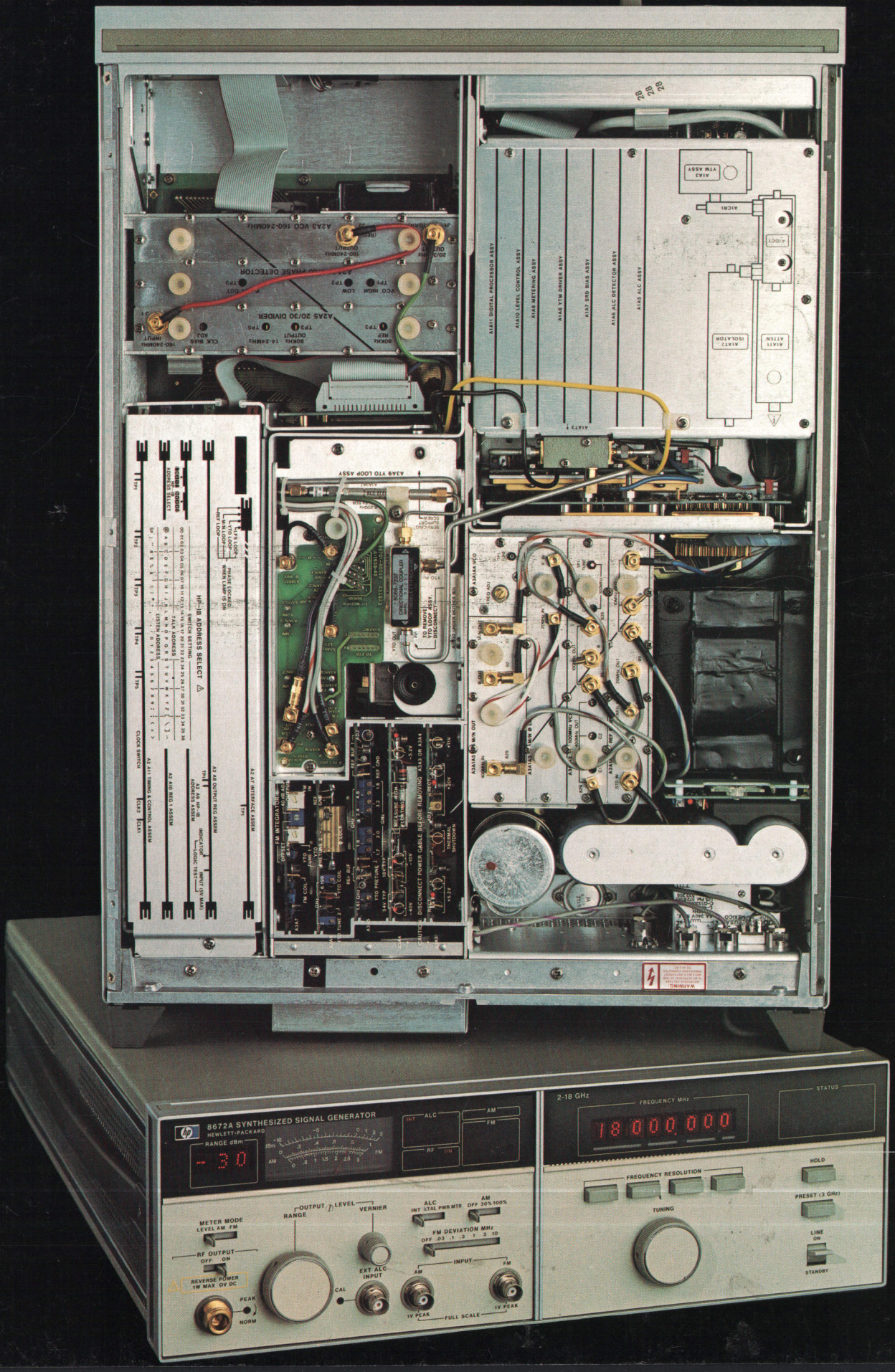
NOVEMBER 1977

HEWLETT-PACKARD JOURNAL



Expanding Synthesized Signal Generation to the Microwave Range

Here are two broadband, programmable, high-spectral-purity microwave signal sources, a 2-to-18-GHz synthesized signal generator and a 2-to-6.2-GHz synthesizer. Both are single, compact, 13-cm-high instruments.

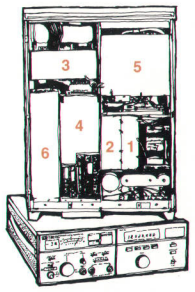
by James L. Thomason

FREQUENCY SYNTHESIZERS are signal sources that produce a selected output frequency by some type of arithmetic operation on a stable reference signal, which is usually derived from a quartz crystal oscillator. The synthesizer's major advantages are the accuracy and stability of its output frequencies. A signal generator, on the other hand, is a signal source that produces a selected output frequency directly from a tunable oscillator. It usually has a calibrated output level and flexible modulation capabilities, and these are its major advantages. The synthesized signal generator, a combination of these two types of signal sources, offers many of the advantages of both.

Broadband frequency synthesizers have been available in the VHF/UHF range since 1964[1]. In 1971 the HP 8660 family of instruments introduced the concept of a synthesized signal generator with modulation and output level control.[2] In 1975 the maximum frequency range of the 8660 family was raised to 2.6 GHz and phase modulation was included.[3]

In the microwave range, frequency sources have generally been narrow-band instruments, often covering one communication band or one octave. Wider frequency coverage has been obtained by means of collections of oscillators that can be switched on and off. In 1975 the broadband HP 86290A Sweeper Plug-In was the first to cover the entire 2-to-18-GHz range with a single YIG-tuned oscillator followed by a YIG-tuned harmonic multiplier.[4]

The new 8671A Synthesizer and 8672A Synthesized Signal Generator, Fig. 1, combine frequency synthesis techniques with improved versions of the 86290A components to provide extremely accurate frequency coverage in the 2-to-18.6-GHz range. The two instruments share a common 2-to-6.2-GHz frequency synthesizer and differ only in their frequency coverage and output signal level control capabilities. The 8671A spans 2 to 6.2 GHz with 1-kHz resolution, can be frequency modulated, and provides more than +8 dBm of unleveled power. The 8672A has an internal YIG-tuned multiplier and provides 2-to-18.6-GHz frequency coverage. 8672A RF output power is internally leveled and calibrated from +3 dBm to −120 dBm. It may also be externally leveled. Six metered frequency modulation ranges and two metered



Cover: *Model 8672A Synthesized Signal Generator works like a well rehearsed orchestra: four phase-locked loops (1-4) and an RF output assembly (5) are conducted by a digital control unit (6). The four loops are the reference loop (1), the M/N loop (2), the low-frequency synthesizer (3), and the YTO loop (4).*

In this Issue:

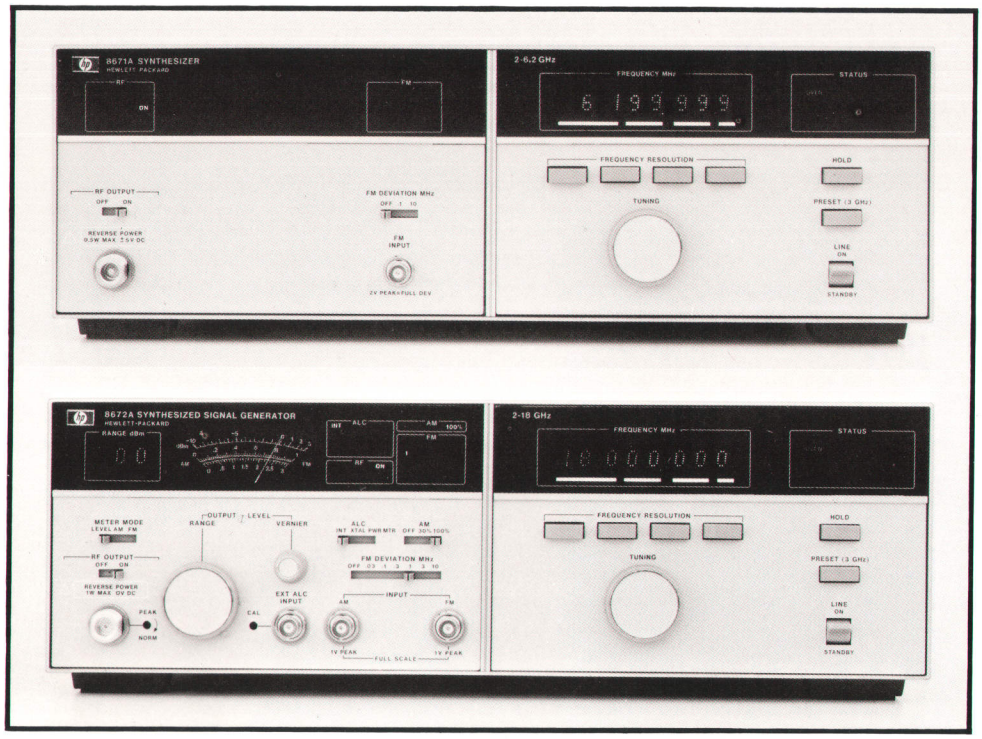


Fig. 1. *Model 8671A Synthesizer (top) generates precise frequencies between 2 and 6.2 GHz. It has 1-kHz resolution and minimum output power (unleveled) of +8 dBm. It can be frequency modulated. Model 8672A Synthesized Signal Generator (bottom) has 1-kHz resolution from 2 to 6.2 GHz, 2-kHz resolution from 6.2 to 12.4 GHz, and 3-kHz resolution from 12.4 to 18.6 GHz. Its RF output power is internally leveled and calibrated from +3 dBm to −120 dBm. It has AM and FM capability. Both instruments are HP-IB compatible.*

amplitude modulation ranges are provided for FM rates to 10 MHz and AM rates to 100 kHz. Both the 8671A and the 8672A are HP-IB (IEEE-488-1975) programmable. All control settings are indicated on the front panel in either local or remote control mode.

Organization and Operation

Internally, the 8671A Synthesizer and the 8672A Synthesized Signal Generator are organized in three sections (Fig. 2):

- the digital control unit (DCU)
- the frequency synthesis section
- the RF output assembly.

These sections are discussed briefly here and in more detail in the articles that follow.

The DCU takes information from the front-panel switches and tuning control in local mode, or the HP-IB interface in remote control mode, and stores this data. Then it calculates the required frequencies of the oscillators in the frequency synthesis section, and the internal switch settings (FM sensitivity, RF output power, etc.) for the rest of the instrument.

The frequency synthesis section of the instrument consists of four phase-locked oscillators and an oven-controlled 10-MHz crystal oscillator for long-term stability. To improve spectral purity, a 100-MHz voltage-controlled crystal oscillator is phase-locked to the 10-MHz crystal oscillator and the reference frequencies used in the instruments are all derived from this oscillator.

The M/N phase-locked loop and the 20-to-30-MHz low-frequency synthesizer loop translate the stability of the crystal oscillators to the 2-to-6.2-GHz range. A YIG-tuned oscillator (YTO) is phase locked to the M/N and low-frequency synthesizer loops and provides the output of the frequency synthesis section.

The RF output section of the 8671A consists of a peripheral mode isolator in the RF path and two sensitivity settings for FM. In the 8672A, the frequency range is extended to 18.6 GHz by means of a YIG-tuned multiplier (YTM) and the circuits required to drive it. In the 2-to-6.2-GHz range, the YTM is tuned to the same frequency as the YTO and the multiplier diode is biased on continuously so the circuit functions as a tuned filter. In the 6.2-to-12.4-GHz and 12.4-to-18.6-GHz ranges, the YTM is tuned to the second and third harmonics of the YTO, respectively, and the diode is biased to perform its harmonic multiplier function. The power amplifier provides a high-level signal to the YTM for maximum efficiency and output power.

RF power level is controlled by the automatic level control (ALC) circuit and the output attenuator. Level adjustment from +3 dBm to −10 dBm is done by the ALC loop, and an additional 110 dB of attenuation is available from the attenuator in 10-dB steps. Unspecified higher output power may be achieved in the overrange mode.

Facilities for wide-range amplitude and frequency modulation are built into the output section of the 8672A. Allowable AM rates are 10 Hz to greater than 100 kHz, and modulation percentage is independent of output level. Both the 8671A and the 8672A can be frequency modulated at rates from 50 Hz to 10 MHz.

Applications of a Microwave Synthesized Signal Generator

Model 8672A Synthesized Signal Generator is designed to satisfy many applications in automatic test, communications, spectrum surveillance, and research and development.

Test Systems

During the design of the 8672A a recurring problem was the lack of a microwave source that had the accuracy and spectral purity required to test the prototype instrument. This was further complicated by the fact that the 8672A has 9,366,666 possible output frequencies! As soon as possible, 8672A prototype units were used to test each other, and desktop computers were added so that testing could be more thorough and test times could be reduced. Finally, several 8672As were included in desktop-computer controlled test systems for research and production.

The 8672A is particularly well suited for production test applications. Many microwave instruments and components must be tested very accurately over broad frequency ranges. Complete programmability allows the testing to be controlled rapidly and precisely. Low-power-level measurements, such as receiver sensitivity or attenuator accuracy, require narrow bandwidths to separate the signal from the noise that is always present. These measurements benefit from the 8672A's excellent frequency stability and accuracy, which are needed to keep the signals within the bandwidth of the receiver. Measurements of the spectral purity of microwave oscillators or signals can be made by comparison with a reference signal. The spectral purity of the 8672A makes it a good local oscillator or reference signal for heterodyning other signals to lower frequencies to make these measurements

Communications Systems

Communications systems generally require local oscillators with superior long-term stability and spectral purity for receivers and transmitters. In addition, the frequency resolution and control must permit easy selection of all required channels. Frequency sources with the same criteria are required to test these systems. In satellite communications, where the up-link and down-link transmission frequencies are often separated by several hundred megahertz or more, the 8672A can perform as a local oscillator or test stimulus generator in either part of the system. Many systems require backup equipment in case of failures. Costs can often be significantly reduced by using a single multiband synthesized signal generator as the backup for several instruments.

Spectrum Surveillance

Another application for a broadband synthesized signal source such as the 8672A is in the area of spectrum surveillance and signal simulation. Here the ability to switch frequencies rapidly while maintaining accuracy and spectral purity is paramount. Used as the local oscillator for a receiver, the programmable synthesizer makes it possible to scan the required band quickly and examine signals of interest to the user. As a signal stimulator, the synthesized signal generator can rapidly test the response of surveillance or other communications equipment to many different signals and frequencies, with or without modulation.

Research and Development Laboratory Uses

A primary goal in the design of the 8672A was to create an easy-to-use microwave signal generator for laboratory bench use. Frequency tuning is accomplished by turning a front-panel knob; all stops and band selection have been eliminated. Instead of coarse and fine tuning knobs, four buttons control the resolution of the single rotary pulse generator used for tuning. Unlike other digitally controlled instruments, the 8672A does not lose its memory after a power interrupt. The advent of desktop computers as instrument controllers has made programmability a real asset to the designer. The 8672A is programmed with single character codes, and provides status feedback to the controller for error detection and analysis. With the large number of HP-IB programmable instruments already available, the designer can quickly configure a desktop-computer controlled system including one or more 8672As to greatly enhance measurement capability.

Spectral purity and frequency accuracy are maintained in the FM mode because the system remains phase locked and the loop bandwidths are not altered.

A companion instrument, Model 11720A Pulse Modulator (see page 6), makes it possible to pulse modulate the output of the 8672A for applications requiring RF pulses.

Lower Frequencies and Higher Resolution

When frequency coverage below 2 GHz is needed, the 8671A/72A may be used with another generator to provide frequencies as low as 1 MHz. Also, because finer frequency resolution may be required than is provided by the 8671A/72A, provision has been made to substitute an external 20-to-30-MHz signal for the low-frequency synthesis loop. Fig. 3 shows a system in which an 8672A is combined with an 8660 Synthesized Signal Generator, two coaxial switches, a relay actuator, and a controller to provide 1, 2, or 3-Hz resolution from 1 MHz to 18.6 GHz. The result is a fully automatic source, with a single output connector, that provides calibrated wide-range output level along with calibrated AM, FM, and phase modulation.

Serviceability and Reliability

An important consideration in the design of a sophisticated instrument such as the 8671A/72A is that the mean time between failures (MTBF) must be long enough to allow effective use of the instrument or any system in which it may be included. The instrument must also be capable of being serviced rapidly so down time is minimized when repairs are necessary.

During the early design stage, a great deal of attention was paid to the thermal, current, and voltage

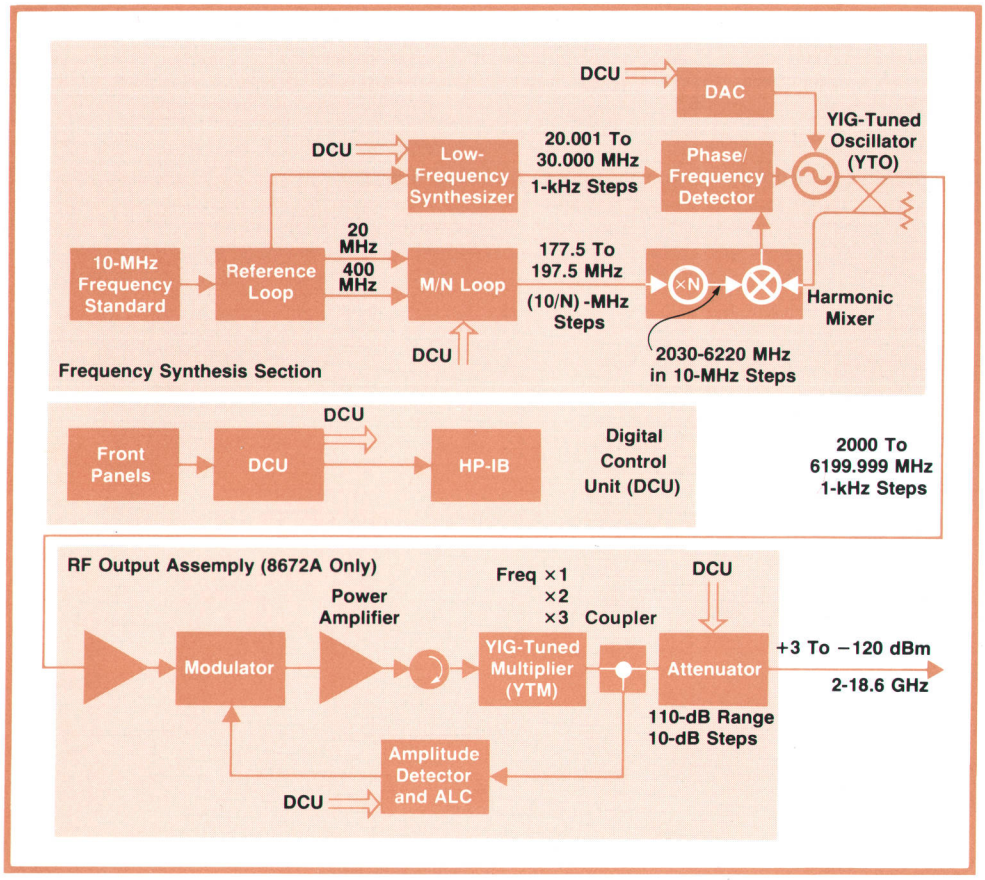


Fig. 2. *Models 8671A and 8672A have the same frequency synthesis section and digital control unit. Model 8672A's RF output assembly contains a YIG-tuned multiplier to extend the frequency range.*

ratings of all components. Temperature profiles of the entire instrument were taken at major prototype and pilot run checkpoints to assure that adequate air flow and heat sinking were provided to all components.

Heat sinks range from simple commercially available thermal dissipators on metal-can transistors and ceramic integrated circuit packages to a specially designed "Swiss cheese" heat sink for the series-pass transistors in the power supply regulators. As a result, the internal temperature rise is less than 10°C above outside ambient, air temperatures inside castings are less than 5°C higher than in the rest of the cabinet, and series-pass elements in the 200-watt power supply exhibit case temperatures less than 30°C above the outside ambient temperature.

All the power supplies are designed with internal current limiting and overvoltage and undervoltage protection, and are individually fused. This prevents minor power supply malfunctions or accidental shorts or overloads during troubleshooting from becoming major secondary problems. To avoid failures that might result from overheating, a thermal shutdown circuit turns the dc power off in the event of excessive heat buildup inside the instrument. The circuit is designed to trip at +85°C and restart at 55°C.

The instruments are designed so that, in many cases, careful analysis of the front panel, the front-

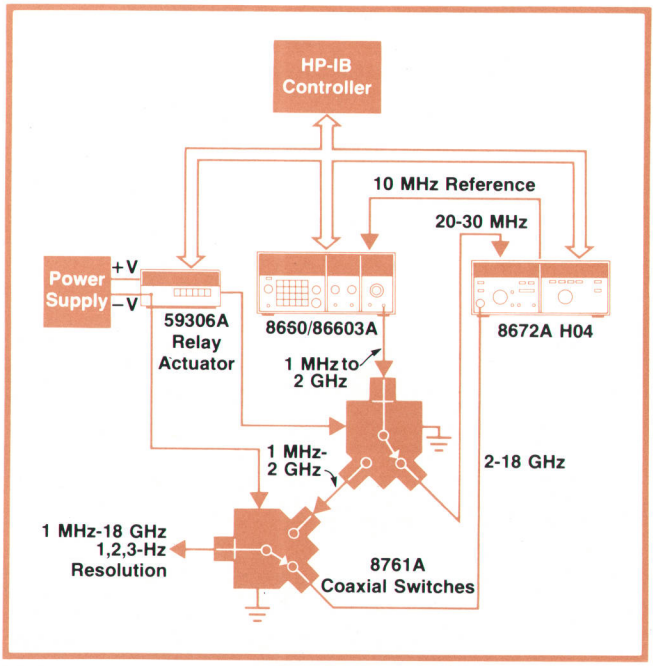


Fig. 3. *To extend its frequency coverage down to 1 MHz, the 8672A can be combined with an 8660 Synthesized Signal Generator as shown here. Frequency resolution is 1, 2, or 3 Hz from 1 MHz to 18.6 GHz.*

A Fast 2-to-18 GHz Pulse Modulator

Model 11720A Pulse Modulator (Fig. 1) is a high-performance 2-to-18-GHz instrument suitable for adding pulse modulation capability to many microwave signal sources, such as the 8671A Synthesizer and the 8672A Synthesized Signal Generator. It employs a novel method of achieving fast pulse transition times while maintaining a good impedance match over the 2-to-18-GHz frequency range. It also achieves low loss and high on/off ratio. Typical rise and fall times are less than 5 ns. Loss is typically less than 4 dB to 12 GHz and 7 dB to 18 GHz. Typical on/off ratio is greater than 90 dB.

Fig. 1. *Model 11720A Pulse Modulator*

The 11720A uses a series-shunt modulator that has one series diode and four shunt diodes (Fig. 2). In this type of modulator the output RF is turned off by turning on the shunt diodes and simultaneously turning off the series diode. The input RF is then terminated by the 50Ω resistor in parallel with the diode. The output RF is turned on by biasing the series diode on and the shunt diodes off. The 50Ω resistor is then shunted by the low impedance of the modulator's forward-biased series diode.

This system avoids the severe reflections that can occur in a shunt modulator system because of the large mismatch produced by the shunt diodes when they are turned on. In a shunt modulator the input and output impedances are essentially zero when the output RF is turned off. The resulting reflections are seen as perturbations on the leading and trailing edges of the pulse and can seriously degrade performance.

Many series-shunt modulator systems have slower transition times than the shunt variety because of the series diode, but this is not the case with the 11720A. The shunt diodes are driven by an IC driver in the conventional manner while the series diode is driven through a bias T by a separate driver. This driver delivers fast, properly timed pulses to the series diode to reduce the transition time to that of a shunt modulator system.

The sequence of developing an RF output pulse is as follows. On the leading edge of the TTL pulse input, the output of driver 1 goes negative. This negative voltage begins to turn off the modulator's shunt diodes and turn on the series diode. The positive pulse out of driver 2 then rapidly increases forward bias on the series diode to force it into full conduction. The RF output pulse now reaches its full steady-state value. When the input pulse returns to zero, the driver 1 output swings positive. This turns on the modulator's shunt diodes and starts to turn off the series diode. The negative pulse out of driver 2 now rapidly reverse-biases the series diode via the bias T, forcing it out of conduction, so the RF output is turned off.

Acknowledgments

Thanks are extended to group leader Ron Kmetovicz for the product concept and for his guidance throughout the project, to Charles Cook for the product design, and to Young Dae Kim for his development of the microwave components.

-Ronald Larson

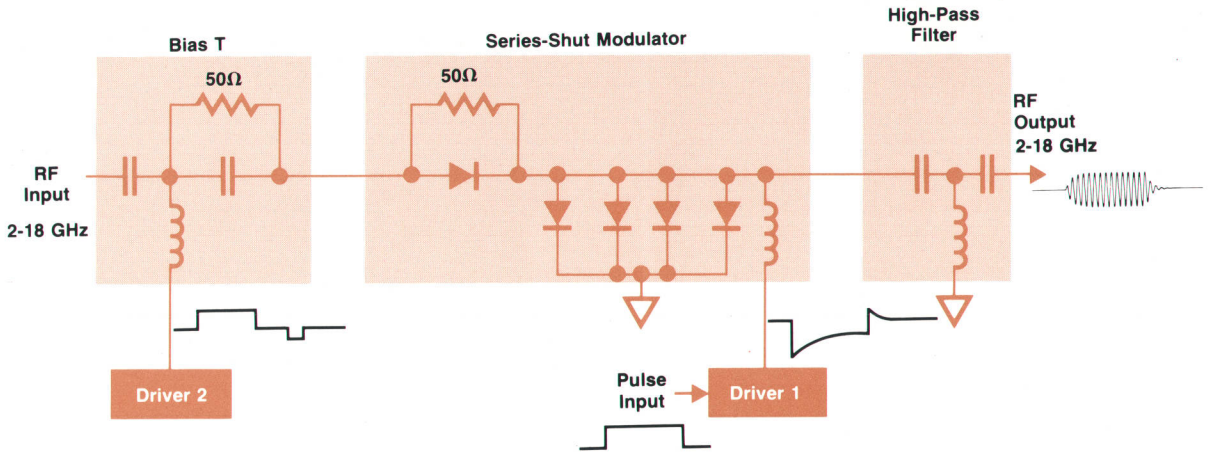
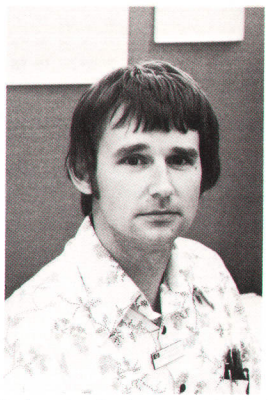


Fig. 2. *A series-shunt configuration, using one series diode and four shunt diodes, provides good impedance matching and fast transition times.*

panel status indicators, and the internal LED indicators can quickly locate a problem to a particular module. The front panel status indicators notify the user of error conditions such as unlock, oven cold, frequency out of range, overmodulation, and unleveled. These indicators can also be reported via the HP-IB to a system controller. Inside there are LED indicators on all the dc power supplies to indicate

their "on" state, a thermal shutdown LED indicator, and LED "lock" indicators for all the phase-locked loops. A CMOS logic probe is built into the DCU to help determine whether the correct digital information is being transmitted to the rest of the instrument.

Complex circuits like phase-locked loops consisting of voltage-controlled oscillators, mixers, dividers, and phase detectors can sometimes be difficult to troubleshoot to the correct subassembly. In the 8671A/72A the simplest phase-locked loop, the VCXO loop, involves two printed circuit boards, and the most complex, the YTO loop, involves four boards and five microwave subassemblies. All of these loops have been designed for troubleshooting to the correct subassembly or printed-circuit board without removing any circuits. RF input, output, and feedback signals are tested by pulling off snap-on connectors (SMB type) and by using commonly available instruments such as counters and spectrum analyzers to verify signal frequencies and levels. VCO tuning voltages are brought outside castings on feedthrough capacitor terminals to facilitate troubleshooting. Access to the YTO loop assembly (two boards inside a casting and five microwave assemblies) is gained by removing three screws and one coaxial connection (SMA type) then remounting the entire assembly on the chassis sheet metal for active troubleshooting.



James L. Thomason

Jim Thomason was the original project manager for the 8671A/8672A project. With HP since 1967, he's served as design engineer, project engineer, project manager, and section manager on a variety of microwave instruments and systems. He was born in Houston, Texas. He received a BA degree in 1962 and a BSEE degree in 1963 from Rice University, and an MSEE degree in 1967 from the University of Colorado. Before joining HP he worked in the petroleum and aerospace industries for four years. Jim is married, has two daughters, and lives in Santa Rosa, California. His hobbies include photography, backpacking, and camping.

Acknowledgments

Early contributors to the project include Doug Clifford and Larry Martin, who designed synthesizer loops, Pete Bice and Mike Marzalek, who did digital

ABRIDGED SPECIFICATIONS

HP Model 8671A Synthesizer

FREQUENCY CHARACTERISTICS
FREQUENCY RANGE: 2.0-6.2 GHz (6.199999 GHz).
FREQUENCY RESOLUTION: 1 kHz
TIME BASE: Internal 10 MHz (<5×10^{-10}/day aging rate) or external 5 or 10 MHz.
SWITCHING TIME: <15 ms to be within 1 kHz.
HARMONICS: <−15 dB.
SINGLE-SIDEBAND PHASE NOISE (1 Hz BW, CW mode):

Offset from F_c	10 Hz	100 Hz	1 kHz	10 kHz	100 kHz
SSB level	−58 dB	−68 dB	−78 dB	−89 dB	−109 dB

SPURIOUS (CW MODE):
Non-Harmonically Related: <−70 dB
PowerLine Related and Spurious Due to Fan Rotation Within 5 Hz Below Line Frequency, and Multiples:

Offset from F_c	<300 Hz	300 Hz to 1 kHz	>1 kHz
	−50 dB	−60 dB	−65 dB

OUTPUT CHARACTERISTICS
POWER (UNLEVELED): +8 dBm (min.), 15 to 35°C.
FLATNESS: <6 dB total variation across full frequency band.
SOURCE IMPEDANCE: 50Ω, <1.7 SWR.
FREQUENCY MODULATION
PEAK DEVIATION (MAX): 10 MHz or $f_{mod}\times5$, whichever is smaller.
RATES (3 dB BW): 50 Hz to 10 MHz typical.
GENERAL
REMOTELY PROGRAMMABLE FUNCTIONS: frequency, FM, RF on/off.
PROGRAMMING FORMAT: HP-IB.
POWER: 100, 120, 220, or 240 V, +5, −10%; 48-66 Hz; 300 VA max.
WEIGHT: net, 24 kg (53 lb); shipping 29.5 kg (65 lb).
DIMENSIONS: 613 mm D × 425 mm W × 146 mm H, (24.1 in × 16.8 in × 5.8 in).
PRICE IN U.S.A.: $17,600.

HP Model 8672A Synthesized Signal Generator

FREQUENCY CHARACTERISTICS
FREQUENCY RANGE: 2.0-18.0 GHz (with overrange to 18.599997 GHz)
FREQUENCY RESOLUTION: 1 kHz to 6.2 GHz, 2 kHz to 12.4 GHz, 3 kHz to 18.0 GHz.
TIME BASE: internal 10 MHz (<5×10^{-10}/day aging rate) or external 5 or 10 MHz.
FREQUENCY SWITCHING TIME: <15 ms to be within 1 kHz, 2-6.2 GHz; 2 kHz, 6.2-12.4 GHz; 3 kHz, 12.4-18 GHz.
SPECTRAL PURITY
HARMONICS, SUBHARMONICS AND MULTIPLES (≤18 GHz): <−25 dB.
SINGLE-SIDEBAND PHASE NOISE (1 Hz BW, CW MODE):

Frequency Range (GHz)	Offset from Carrier 10 Hz	100 Hz	1 kHz	10 kHz	100 kHz
2.0-6.2	−58 dB	−68 dB	−78 dB	−89 dB	−109 dB
6.2-12.4	−52 dB	−62 dB	−72 dB	−83 dB	−103 dB
12.4-18.0	−48 dB	−58 dB	−68 dB	−79 dB	−99 dB

SPURIOUS (CW AND AM MODES):
Non-Harmonically Related: <−70 dB, 2.0-6.2 GHz
<−64 dB, 6.2-12.4 GHz
<−60 dB, 12.4-18.0 GHz
Power Line Related and Spurious Due to Fan Rotation Within 5 Hz Below Line Frequency and Multiples:

Frequency Range (GHz)	Offset from Carrier <300 Hz	300 Hz to 1 kHz	>1 kHz
2.0-6.2	−50 dB	−60 dB	−65 dB
6.2-12.4	−44 dB	−54 dB	−59 dB
12.4-18.0	−40 dB	−50 dB	−55 dB

OUTPUT CHARACTERISTICS
OUTPUT LEVEL (15 TO 35°C): +3 to −120 dBm.
TOTAL INDICATED METER ACCURACY (15 to 35°C)

Frequency Range (GHz)	Output Level Range 0 dBm	−10 dBm	−20 dBm	−30 dBm and below
2.0-6.2	±1.75 dB	±2.25 dB	±2.45 dB	±1.75 dB ±0.3 dB per 10 dB step below 0 dBm range
6.2-12.4	±2.0 dB	±2.5 dB	±2.7 dB	±2.0 dB ±0.3 dB per 10 dB step below 0 dBm range
12.4-18.0	±2.25 dB	±2.85 dB	±3.05 dB	±2.25 dB ±0.4 dB per 10 dB step below 0 dBm range

FLATNESS (15°C TO 35°C): ±0.75 dB, 2.0-6.2 GHz; ±1.00 dB, 2.0-12.4 GHz; ±1.25 dB, 2.0-18.0 GHz.
OUTPUT LEVEL SWITCHING TIME: <20 ms.
SOURCE IMPEDANCE: 50Ω, <2.5 SWR.
AMPLITUDE MODULATION
AM DEPTH (FOR RF OUTPUT METER READINGS ≤0 dB, 15 TO 35°C):
0-75%, 2.0-6.2 GHz.
0-60%, 6.2-12.4 GHz.
0-50%, 12.4-18.0 GHz.
SENSITIVITY: 30%/V, 100%/V ranges. Max input 1 V peak into 600Ω.
RATES (3 dB BW): 10 Hz-100 kHz.
DISTORTION (RATES ≤10 kHz, RF output ≤0 dB, 15 TO 35°C): <3% at 30% depth
FREQUENCY MODULATION
PEAK DEVIATION (MAX): the smaller of 10 MHz or $f_{mod}\times5$, 2.0-6.2 GHz
$f_{mod}\times10$, 6.2-12.4 GHz
$f_{mod}\times15$, 12.4-18.0 GHz
SENSITIVITY: 30, 100, 300 kHz/V and 1, 3, 10 MHz/V ranges. Max input 1 V peak into 50Ω.
RATES (3 dB BW):
30, 100 kHz/V Ranges: Typically 50 Hz to 10 MHz.
300 kHz/V and 1, 3, 10 MHz/V Ranges: Typically 1 kHz to 10 MHz.
DISTORTION: <12% for rates <3 kHz decreasing linearly with frequency to 5% at 20 kHz; <5% for 20 kHz to 100 kHz rates.
GENERAL
REMOTELY PROGRAMMABLE FUNCTIONS: frequency, output level, AM, FM, RF on/off, ALC source.
PROGRAMMING FORMAT: HP-IB.
POWER: 100, 120, 220, 240 V, +5, −10%; 48-66 Hz; 300 VA max.
WEIGHT: net 27 kg (60 lb); shipping 32.5 kg (72 lb).
DIMENSIONS: 613 mm D × 425 mm W × 146 mm H, (24.1 in × 16.8 in × 5.8 in).
PRICE IN U.S.A.: $26,900.

HP Model 11720A Pulse Modulator

FREQUENCY RANGE: 2 to 18 GHz.
ON/OFF RATIO: >80 dB.
INSERTION LOSS: 2 to 12.4 GHz: <6 dB.
2 to 18 GHz: <10 dB.
RISE AND FALL TIMES: <10 ns.
MAXIMUM RF INPUT POWER: +20 dBm.
MINIMUM RF PULSE WIDTH:[1] <50 ns.
PULSE WIDTH COMPRESSION: <20 ns.
MAXIMUM PULSE REPETITION RATE: >5 MHz.
MAXIMUM DELAY TIME: <60 ns.
VIDEO FEEDTHROUGH: <50 mV peak-to-peak.
OVERSHOOT, RINGING:[2] <0.2.
PULSE INPUT
NORMAL MODE: >3 V (on), <0.5 V (off).
COMPLEMENT MODE: <0.5 V (on), >3 V (off).
IMPEDANCE: 50 Ω nominal.
DAMAGE LEVELS
RF INPUT: ac: 2 watts (+33 dBm).
dc: 40 volts.
PULSE INPUT: ±6 V peak from ≥50 Ω source.
+6 V peak, −0.5 V peak from <50 Ω source.
CONNECTORS:
RF (IN AND OUT): Type N Female.
PULSE INPUT: BNC Female.
GENERAL
OPERATING TEMPERATURE: 0°C to +55°C.
RF LEAKAGE:[3] Meets radiated and conducted limits of MIL-I-6181D.
POWER: 100, 120, 220, 240 V; +5%, −10%. 48-440 Hz; 50 VA max.
PRICE IN U.S.A.: $2500.

[1]Off time must be ≥140 ns.
[2]Overshoot and ringing may be reduced by operating at <10 dBm RF input and >+15°C ambient temperature.
[3]For pulse repetition rates <1 MHz.

MANUFACTURING DIVISION: STANFORD PARK DIVISION
1501 Page Mill Road
Palo Alto, California 94304 U.S.A.

control unit design, and Mike Ammirato and Bob DeVries, who did mechanical design. Overall contributors include Dan Derby, who was responsible for the industrial design, and Charlie Sallberg, Randy White, and Cliff Wing, who did the environmental testing. Ron Larson designed the circuits for the 11720A Pulse Modulator. Marc Saunders and Bob Rands provided marketing information that helped define the instruments. Brian Unter served as program manager in the later stages of the project, and Rit Keiter and John Page provided valuable advice and counsel as we progressed. Many others were involved in the project, and I am grateful to all of them for their contributions.

References

1. V.E. Van Duzer, "A 0-50 Mc Frequency Synthesizer with Excellent Stability, Fast Switching, and Fine Resolution," Hewlett-Packard Journal, May 1964.
2. J.C. Shanahan, "Uniting Signal Generation and Signal Synthesis," Hewlett-Packard Journal, December 1971.
3. J.A. Hall and Y.D. Kim, "Synthesized Signal Generator Operation to 2.6 GHz with Wideband Phase Modulation," Hewlett-Packard Journal, July 1975.
4. P.R. Hernday and C.J. Enlow, "A High-Performance 2-to-18-GHz Sweeper," Hewlett-Packard Journal, March 1975.

Frequency Synthesis in a Microwave Signal Generator

by Kenneth L. Astrof

THE SPECTRAL PURITY of a signal generator is among its most important characteristics. High spectral purity is a combination of long-term frequency stability, low phase noise, and low spurious responses. The 8671A Synthesizer and the 8672A Synthesized Signal Generator have this combination to a high degree, along with good frequency resolution, and therefore represent significant contributions to the state of the art in microwave signal sources.

Factors affecting the spectral purity of any synthesizer are the noise performance of the frequency reference, and the noise contributions of the frequency control, conversion, and generation elements (amplifiers, phase detectors, dividers, multipliers, and the oscillators themselves). There are also undesired discrete signals, such as power line related spurious and non-harmonically related spurious signals. It is possible for these discrete effects to be so large that they mask any claimed benefits of low phase noise.

An important objective for the 8671A and 8672A was that the instruments' excellent phase noise performance should not be degraded by spurious signals. Figs. 1 and 2 show measured, specified, and typical 8671A/72A spectral purity. The noise is total single-sideband noise, including AM noise from the amplifiers, multiplier, and ALC system, for two instruments measured together. However, the noise is predominantly phase noise from the oscillators and phase-locked loops.

Synthesizer Section Organization

The block diagram of the frequency synthesis section of the 8671A/72A is shown in Fig. 3. A 10-MHz frequency standard and four phase-locked loops, including three voltage controlled RF oscillators and a microwave YIG-tuned oscillator (YTO), are used to generate a 2-to-6.199999-GHz signal with 1-kHz frequency resolution while maintaining the high spectral purity required in a microwave signal generator.

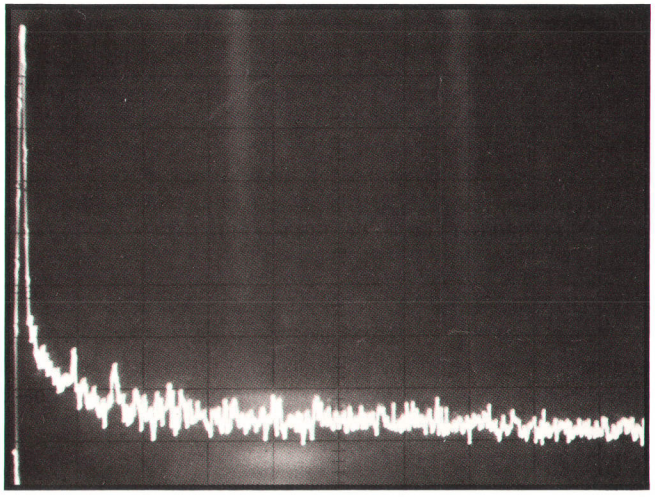


Fig. 1. *Model 8672A Synthesized Signal Generator total single-sideband noise measured at 6.1 GHz with 3-Hz bandwidth. Noise is primarily phase noise but includes line related spurious and AM noise. The 6.1-GHz carrier is the large peak at the left, and the right side of the screen is about 2 kHz above the carrier frequency. Vertical scale: 10 dB/div.*

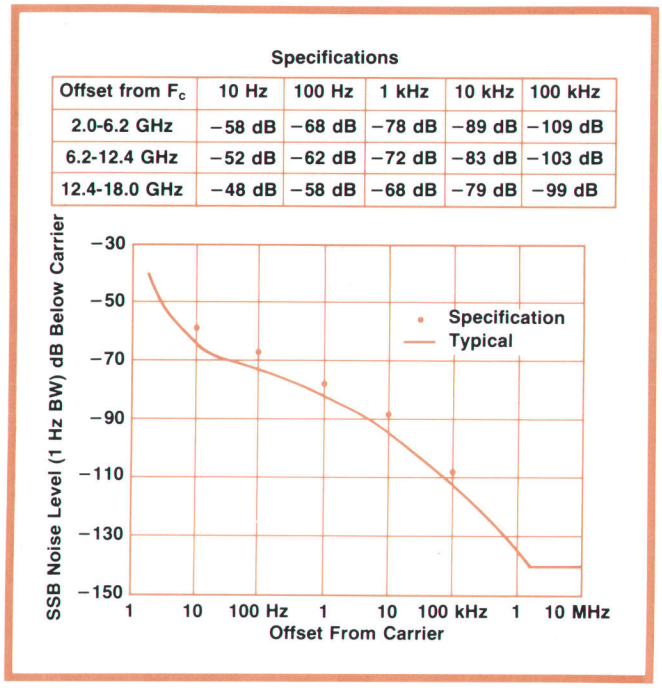


Fig. 2. *Specified and typical 8671A/8672A single-sideband noise in a 1-Hz bandwidth. Graph shows typical performance of two 8672A's at a carrier frequency of 6.1 GHz. Noise increases by 6 dB for 6.2 to 12.4 GHz and 10 dB for 12.4 to 18.6 GHz.*

The principle of operation is to phase lock the YTO output frequency to the outputs of two narrow-band low-frequency synthesizers, the M/N loop and the LFS loop. The M/N loop generates a signal between 177.5 and 197.5 MHz:

$$F_{M/N} = 200 - (M/N) \cdot 10 \text{ MHz}.$$

This is multiplied in a harmonic mixer to produce a broad comb of widely spaced frequencies. By design, the desired Nth comb frequency is an integer multiple of 10 MHz and moves in 10-MHz steps.

The YTO is coarse-tuned to the desired output frequency by a digital-to-analog converter (DAC), and a portion of the YTO output is mixed with the desired Nth comb frequency. The resulting intermediate frequency is phase locked to the output of the low-frequency synthesizer (LFS) loop, which tunes from 20.001 to 30.000 MHz with 1-kHz steps. The output frequency is

$$F_{YTO} = N \cdot F_{M/N} - F_{LFS}$$

$$= 200N - 10M - (20.001\text{-}30.000) \text{ MHz}.$$

Thus this signal is tuned from 2000 to 6199.999 MHz with 1-kHz frequency resolution. In the 8672A, this signal is then multiplied by 1, 2, or 3 by the YIG-tuned multiplier to provide the 2-to-18.6-GHz output in three bands.

The digital control unit generates the required M and N numbers, the data programming signals for the low-frequency synthesizer and the DAC, and the band information for the YTM, based on the frequency information displayed on the front panel.

Reference Frequencies

The frequency standard for the instrument is a vibration-isolated, temperature controlled, 10-MHz crystal oscillator with 5×10^{-10}/day frequency stability and high spectral purity, including low phase noise and spurious responses. An external 5-MHz or 10-MHz frequency standard with 0-dBm signal level may also be used if either greater stability or common use of a single reference is desired. Any number of 8671A/72A synthesizers may share a common frequency standard. Each unit has a rear-panel buffered output derived from the reference loop voltage controlled crystal oscillator (VCXO).

The 10-MHz standard frequency goes to the reference phase-locked loop, which serves as a cleanup loop for the frequency standard. The 100-MHz VCXO in this loop is optimized for low phase noise and spurious responses beyond a 200-Hz offset from the carrier. Within 200 Hz of the carrier the good temperature stability and low aging rate, phase noise, and spurious response of the frequency standard are transferred to the 100-MHz VCXO. The 100-MHz signal is quadrupled to produce a spectrally pure 400-MHz signal that is used in the M/N loop as the basic microwave frequency reference for the instrument. The 100-MHz VCXO output is also divided down to provide 10 and 20-MHz reference frequencies for the rest of the instrument.

Low-Frequency Synthesizer Loop

The low-frequency synthesizer provides the instrument's 1-kHz frequency resolution over any 10-MHz range by repetitively cycling between 20.001 and 30.000 MHz in 1-kHz steps. If the loop had been designed as a simple divide-by-N phase locked loop then the circuit would have involved a 1-kHz frequency reference (10 MHz ÷ 10,000) and a frequency-counter-type divide-by-N circuit, where N would vary from 20,001 to 30,000. The loop output frequency would then be N×1 kHz, or 20.001 to 30.000 MHz in steps of 1 kHz.

Two circuit refinements are used in the 8671A/72A synthesizer to achieve the required phase noise and frequency resolution in a single-loop phase-lock system (see Fig. 4). The first refinement is the conventional technique of increasing the reference frequency to the loop phase detector by a factor of eight and then dividing the VCO output by the same factor, thus retaining the original frequency resolution. The phase detector reference frequency that must be filtered from the VCO tuning line to avoid spurious

sidebands is increased from 1 kHz to 8 kHz, so the loop bandwidth can be increased by the same ratio, thereby increasing the frequency range over which negative feedback can be used to improve the oscillator's performance. This improvement is not without cost, since the VCO frequency range, previously 20.001 to 30.000 MHz, is increased to the much higher range of 160.008 to 240.000 MHz. Higher-speed, higher-cost ECL digital ICs must now be used for the divide-by-8 function and the first stage of the divide-by-N counter is a ÷10/÷11 ECL prescaler with a high-speed T^2L controller that switches between ÷10 and ÷11 by swallowing pulses.

Arbitrarily increasing the loop bandwidth in a phase-lock system will improve the phase noise performance of the VCO close to the carrier where the noise is high. However, if N in the divide-by-N counter is so large that VCO noise sidebands are attenuated below the divider or phase detector noise floor, then the familiar noise pedestal problem of many phase-locked loops appears. The second circuit refinement, a technique called fractional division, overcomes this problem by allowing non-integer division. In the LFS this technique makes it possible to divide by 2000.1, 2000.2, 2000.3, ..., instead of the integers 20,001, 20,002, 20,003, ... For example, to divide by 2000.1, the circuit divides by 2000 nine times and by 2001 once, for an average over ten divisions of $(9\times2000+1\times2001)\div10=2000.1$. The reference frequency is raised from 8 to 80 kHz because the phase detector sees ten zero crossings (ends of count) every 125 μs instead of the one zero crossing associated with dividing by 20,001. Thus the average divide-by-N number is decreased by a factor of ten. **Also, the LFS signal-to-noise ratio improves by 20 dB because the system noise floor remains constant while the dc output of the phase detector increases by the same factor of ten as the input zero crossings.**

M/N Loop

The M/N loop provides a high-spectral-purity signal that, after multiplication to microwave frequencies, serves as the wideband reference frequency for the instrument. The multiplication factor ranges from 11 to 32 as the YTO frequency changes from 2 to 6.190 GHz. Multiplication by 32 creates a 30.1-dB increase in phase noise and spurious responses of the YTO output relative to the output of the M/N loop before multiplication. The specified single-sideband phase noise in a 1-Hz bandwidth at 10 kHz offset from the carrier for the 8671A/72A is 89 dB below the carrier (89 dBc/Hz). With the output frequency at 6.190 GHz and N = 32, the M/N output single-sideband phase noise must be 119 dBc/Hz to meet specifications. Allowing a 10-dB margin between component and instrument specifications and noting that the divide-by-2 circuit following the M/N VCO improves the M/N output phase noise by 6 dB, then the M/N VCO must have single-sideband phase noise of 123 dBc/Hz. The resonator element in this oscillator is a foreshortened coaxial cavity.

To achieve this level of performance in a voltage controlled oscillator that is varactor tuned from 355 to 395 MHz requires careful design and manufacturing procedures. To maintain low resonator losses and thus high Q, all resonator parts are made of gold-plated copper or brass except the resonator housing which is alodined aluminum. Resonator taps for the oscillator and output transistor are designed for minimum loading to maintain a high loaded Q for the oscillator circuit. The oscillator employs a low-noise oscillator transistor and two high-quality microwave varactors in parallel. The varactors are selected for high circuit Q (low RF loss) and low diode noise (internal noise can be caused by leakage current, 1/f noise from diode surface effects, and noise from diode rectification current caused by large RF swings on the varactor). To maintain this level of performance throughout the manufacturing cycle requires not only careful design but also sophisticated assembly and testing procedures. Each M/N VCO is factory tested for phase noise performance, output power level, and tuning sensitivity before it is installed.

In addition to the careful design of the M/N VCO, several other design innovations can be seen in the M/N loop. The values of the M and N numbers used in the digital dividers were selected to reduce the overall tuning range of the M/N VCO. The tuning sensitivity of the M/N VCO varies as a function of frequency. This is balanced within a factor of two by the change in the feedback divide number (M = 8 to 27), thereby keeping loop gain and loop bandwidth nearly constant. Finally, fractional division, which was originally devised for use in the M/N loop, is used to reduce the effective M divide number by a factor of four to realize a 12-dB improvement in the loop signal-to-noise ratio.

YTO Loop

The YTO loop serves as the final summing loop for the instrument and transfers to the RF output the frequency accuracy, stability, resolution, phase noise, and spurious performance of the M/N loop, the low-frequency synthesizer loop, the reference loop, and the internal frequency standard. The use of a wideband low-noise YTO as the final output device makes it possible to use a carefully optimized phase-locked loop to transfer the spectral purity of the low-frequency references to the RF output, while the far-out (>20 kHz) phase noise and broadband noise floor of the YTO offer far superior performance to that of a multiplied or up-converted synthesizer technique. This is a major performance advantage of an indirect synthesizer (one that uses phase-locked loops and

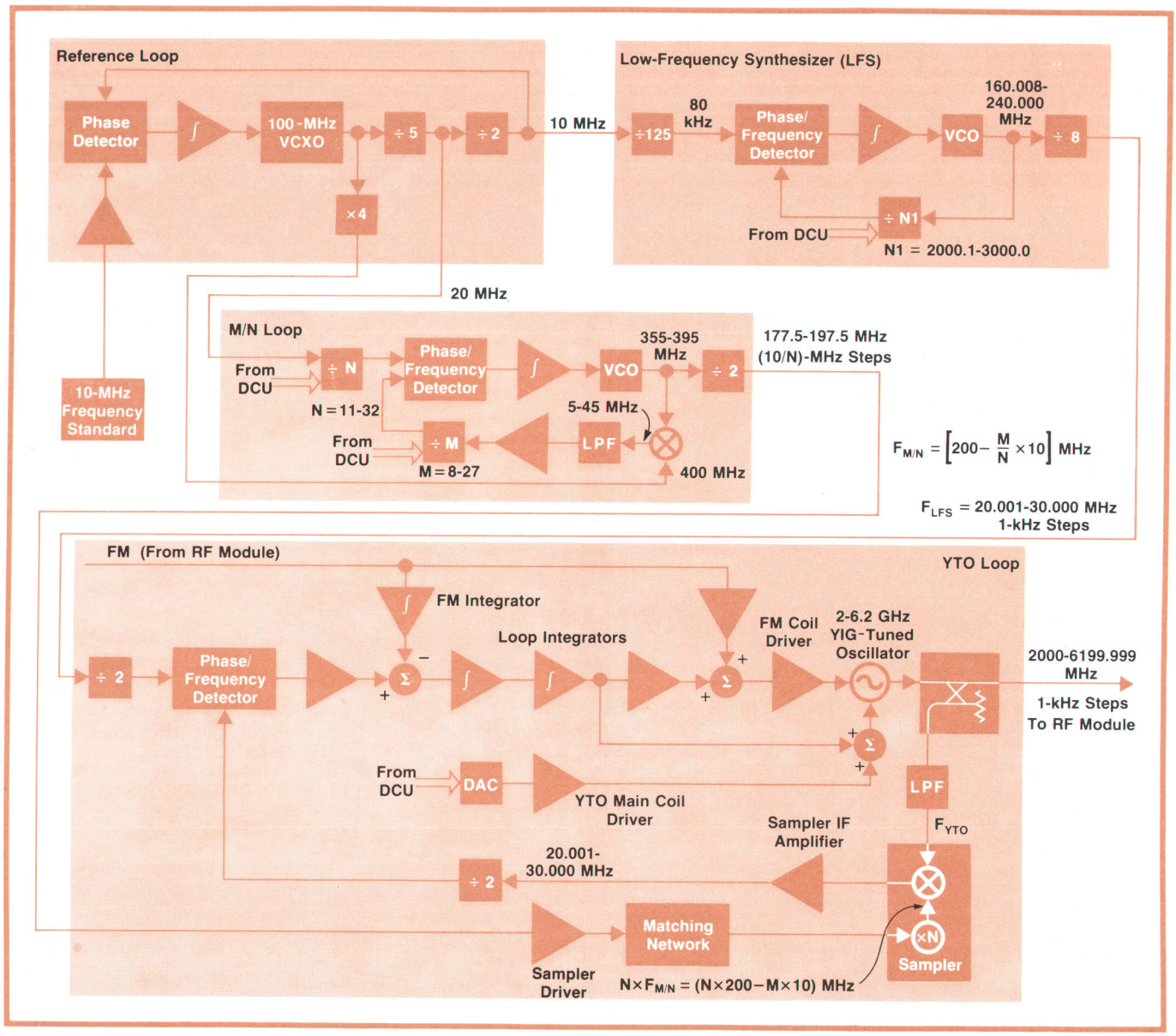


Fig. 3. *Block diagram of the 8671A/8672A frequency synthesis section. Four phase-locked loops transfer the stability of a 10-MHz crystal frequency standard to the output YIG-tuned oscillator (YTO).*

variable frequency oscillators) over a direct synthesis technique (adding and multiplying fixed reference frequencies). Because the far-out phase noise performance of the instrument depends on the YTO, it has been carefully designed to perform well in this regard (see box, page 12).

YIG-tuned oscillators are magnetically tuned devices and are normally designed with high and low-sensitivity electromagnetic coil windings. In the 8671A/72A the main coil or high-sensitivity winding is used to coarse-tune the YTO. It has 40-MHz/mA tuning sensitivity, power dissipation of 1.1 watts, and a relatively slow tuning rate. The tickler coil or low-sensitivity winding (200 kHz/mA) has 10-MHz frequency response and is used for frequency modulation and phase-lock purposes.

Reducing pretune circuit noise can be accomplished merely by placing a large capacitor across the YTO main coil. However, a synthesizer must also have fast switching speeds. The speed-up circuit shown in Fig. 5 is designed to allow large tuning changes quickly, but low-pass filter the tuning voltage for small pretune error signals (i.e., noise). The DAC and YTO main coil driver use this circuit to achieve low noise performance while maintaining fast lock acquisition for any frequency step greater than about 5 MHz (see Fig. 6).

The sampler used to down-convert the YTO output

An Improved 2-to-6.2-GHz YIG-Tuned Oscillator

The 2-to-6.2-GHz YIG-tuned oscillator used in the 8671A/72A is an improved version of the YIG-tuned oscillator used in the 86290A RF Plug-In.[1] It is a high-performance microelectronic component designed for low manufacturing cost and high reliability. Achieving these objectives required special attention to three different areas: a circuit design emphasizing low phase noise, a mechanical design that stressed high performance at low cost, and assembly and test techniques that would reduce the time required and still provide the necessary performance verification.

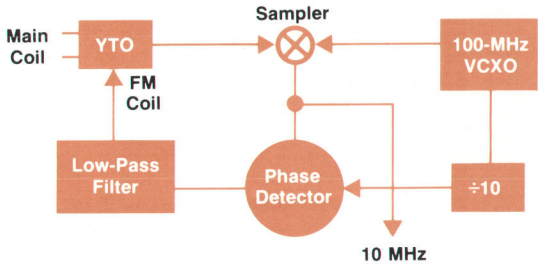


Fig. 1. *System to measure 8671A/8672A YTO phase noise. The 10-MHz signal is down-converted to 100 kHz, passed through a 16-dB carrier notch filter, and measured by a spectrum analyzer.*

Circuit Design

The YTO design uses a YIG sphere and HP 35820 bipolar transistors in a thin-film circuit on a sapphire substrate. The oscillator stage uses inductive feedback in the transistor base circuit. The YIG sphere serves as the high-Q tuning element in the emitter circuit. A two-stage amplifier provides buffering and the desired output power.

Selection and application of the YIG sphere are critical to the performance of the oscillator. The sphere has a diameter of 0.66 mm, a saturation magnetization of 550 gauss, and a typical line width of 0.8 oersted. To obtain good frequency stability the sphere is oriented to a 100 crystallographic axis, mounted onto a sapphire rod, rotated during assembly close to a zero-temperature-coefficient axis, and maintained at 85°C by a self-contained heater.

To achieve low phase noise it is necessary to present the proper collector match across the band. RF currents must be low enough not to excite noise sidebands in the resonator and high enough so the noise floor is far enough below the carrier level. This was assured by computer modeling the oscillator stage to determine the optimum matching impedance.

Testing

Some of the phase-locked loop circuitry of the 8671A/72A was used in a system to measure the phase noise of the YTO. A block diagram of the system is shown in Fig. 1. The oscillator is phase-locked to a harmonic of an extremely clean 100-MHz crystal oscillator (VCXO) signal with a 10-MHz offset. The loop bandwidth is 1 kHz. The difference frequency of 10 MHz has the VCXO's phase noise characteristics for offsets less than 1 kHz from the carrier, and the YTO's characteristics beyond that. This 10-MHz signal is down-converted to 100 kHz, passed through a 16-dB carrier notch filter, and displayed on an 8556A Spectrum Analyzer system. The system is capable of measuring carrier-to-noise ratios of 85 dB in a 1-kHz bandwidth 10 kHz from the carrier.

YTOs are tested by a desktop-computer controlled HP-IB (IEEE 488-1975) system. The system steps the YTO in 100-MHz steps, checks and fine tunes for lock with a 100-MHz VCXO, and proceeds to measure power output, FM noise, harmonics, and linearity (if desired), then plots the measurements. 43 points between 2 and 6.2 GHz are plotted in less than two minutes. The system not only cuts test times considerably, but provides all the data needed to determine device trends and performance margins.

Mechanical Design

To reduce cost and improve ease of assembly and repair, earlier oscillator package designs were modified with emphasis on looser part tolerances, precision assembly fixtures, adjustable or replaceable elements using threaded fasteners, and O-ring seals for hermeticity. For example, the original 2-to-6.2-GHz YTO center body, or package, was assembled using a series of brazing and soldering processes for the circuit shelf, the RF connector barrel, the glass-to-metal seals, and the RF bypass capacitors. The design was changed to allow one-stage assembly of all except the RF connector barrel, using 218°C solder and precision locating fixtures. Part tolerances are looser and part costs and assembly times have been improved. The connector barrel and its mating hole were threaded to ease later assembly and repair. The YIG sphere itself and its support rod are held in place by a mechanical clamp that allows easy adjustment of the sphere during testing and easy replacement if necessary. The clamp, a small but complex device, provides precise alignment for the YIG support rod and thermally insulates the rod to allow heating to a constant temperature. The rod heater, which is also part of the clamp structure, contacts the rod when the clamp is tightened. The screws that provide the clamping force also hold down the substrate carrier for precise positioning and good thermal contact from circuit to heat sink. One of the screws also fastens the FM coil support. The clamp parts are precisely molded from a creep-resistant, glass-filled thermosetting plastic.

A final example of the new approach is in sealing the magnets to the center body. Previously, epoxy was used to fasten and seal these joints, and clamps were added for mechanical reinforcement. The drawbacks were tight diameter tolerances (for alignment), slow assembly, and difficult removal and cleanup when repair was needed. The use of O-rings in place of the epoxy allows looser diameter tolerances, thanks to the centering action of the rings. Assembly, removal, and replacement of magnets is now greatly simplified. Also, hermeticity is attained more easily and reliably. The O-ring was chosen for low permeability, flexibililty over a wide temperature range, and chemical compatibility.

Acknowledgment

The phase-locked noise measurement system was originally developed by Larry Martin.

-G. Basawapatna
-J. Nidecker

Reference

1. P.R. Hernday and C.J. Enlow, "A High-Performance 2-to-18-GHz Sweeper," Hewlett-Packard Journal, March 1975. Also contains a section describing how YIG tuning works.

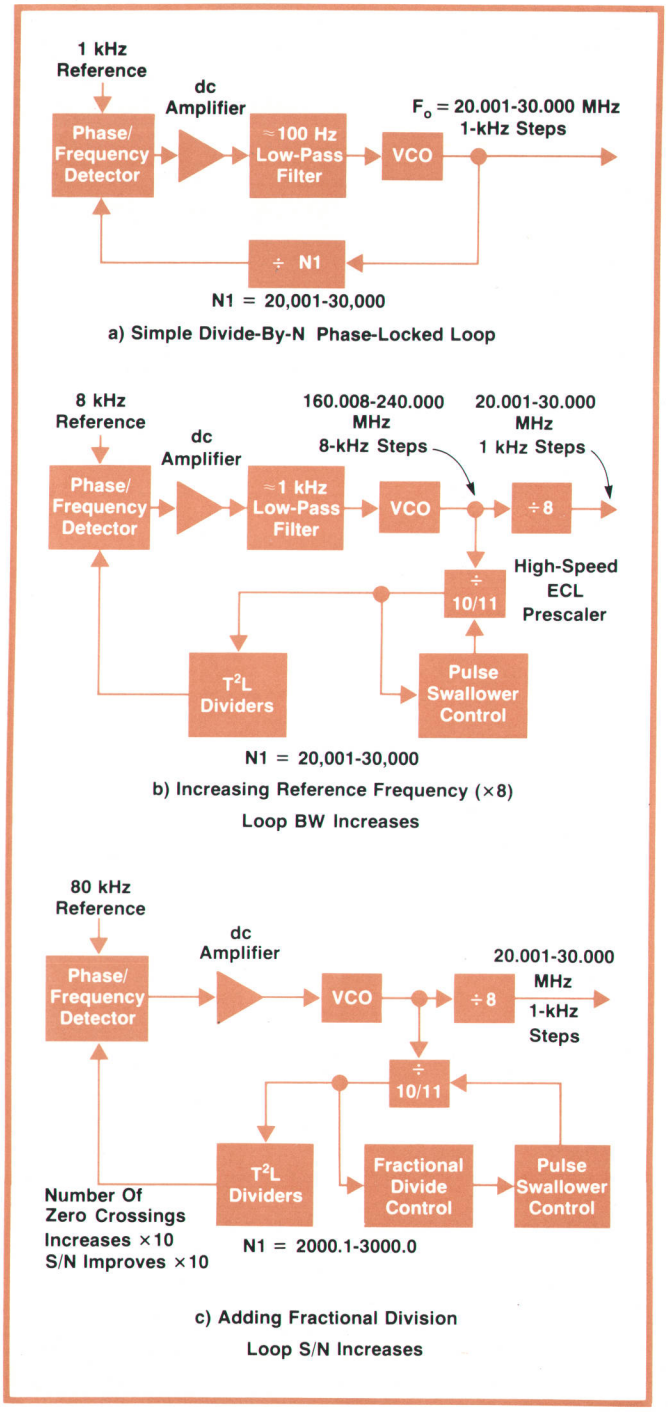


Fig. 4. *Evolution of the low-frequency synthesizer loop. Two refinements to the basic divide-by-N loop improve its phase noise performance by more than 20 dB. Increasing the reference frequency increases the loop bandwidth. Fractional division decreases the value of N.*

to the 20-to-30-MHz range is a sophisticated balanced harmonic mixer. Some of the requirements placed on this component and its associated driver and IF amplifier are as follows:

- Mixer sampling efficiency and IF noise figure must achieve a 110 dB/Hz signal-to-noise ratio over the RF frequency range of 2 to 6.2 GHz
- Feedthrough of the high-level sampling signal (177.5 to 197.5 MHz) and its harmonics, which extend past 8 GHz, must be 80 dB below the RF output signal level.
- Intermodulation products between the sampling signal and the IF signal (20 to 30 MHz), and between the sampling signal and harmonics of the YTO, must also be 80 dB below the sampler IF level.

Other considerations in the YTO loop design were that the phase detector, dc amplifier, integrators, and

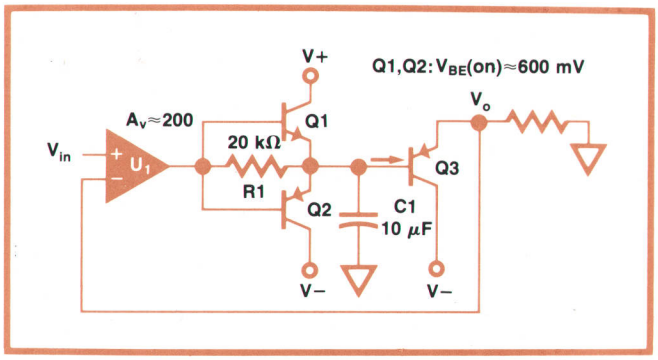


Fig. 5. *Speed-up circuit allows large YTO frequency changes quickly, but low-pass filters the YTO tuning voltage for small error signals (noise). When $|V_{in}-V_o|$ is small, Q1 and Q2 are off and the amplifier bandwidth is about 160 Hz. When $|V_{in}-V_o|>3$ mV, either Q1 or Q2 is on, R1 is bypassed, and C1 charges rapidly. Thus V_o reaches the desired level quickly. Then Q1 (or Q2) turns off and the bandwidth is again 160 Hz.*

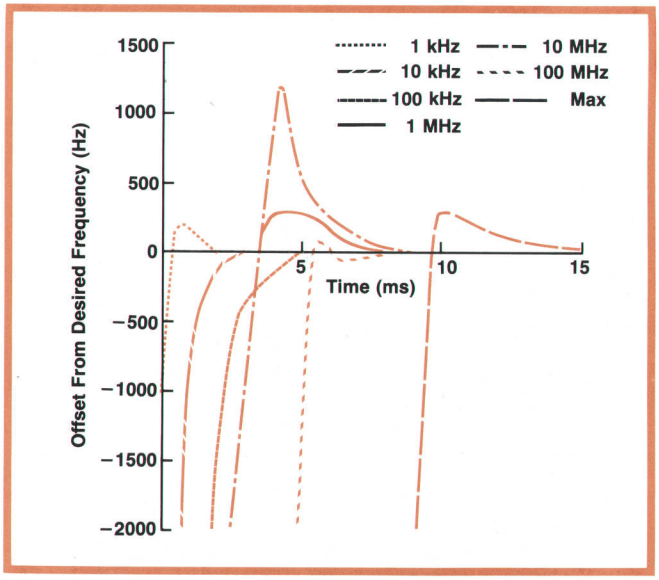


Fig. 6. *Typical switching characteristics of the 8671A/8672A in band 1 (2-6.2 GHz). Each curve is a plot of frequency offset versus time for a change of one unit in the frequency digit shown. (Note: a change from 999 to 1000 kHz is equivalent to a change in the 1-MHz digit.) Specification (band 1): <1 kHz error at 15 ms. For bands 2 and 3, multiply offsets by 2 and 3, respectively.*

Dealing with Microphonic Sidebands

Fan vibration can be a pesty source of line related spurious signals. Fig. 1 shows the microphonic sidebands that appeared on the microwave output of an early prototype 8672A Synthesized Signal Generator when it was operated with a high-slip fan at a 50-Hz line frequency. The sidebands were 30 Hz from the carrier and 36 dB below it. They were caused by the vibrational energy that reached the 10-MHz frequency standard. Fig. 2 shows these accelerations.

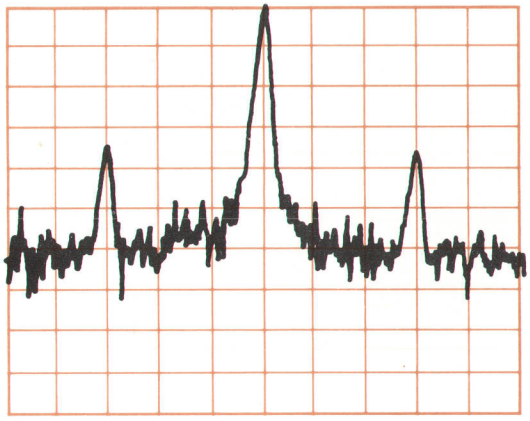


Fig. 1. *Microphonic sidebands caused by a high-slip fan. Horizontal scale: 10 Hz/div. Vertical scale: 10 dB/div. Carrier frequency: 6100 MHz.*

Direct observation of the output of the frequency standard would show sidebands 92 dB below the 10-MHz carrier level. The corresponding sideband levels in the microwave output are 56 dB higher. This is because the 6.1-GHz microwave output is generated by the YIG-tuned oscillator, which is phase-locked via the M/N loop to the 10-MHz standard, so in effect any frequency deviation on the standard is multiplied by the ratio of the microwave frequency to 10 MHz, or by 610 in this example. The following equation applies:

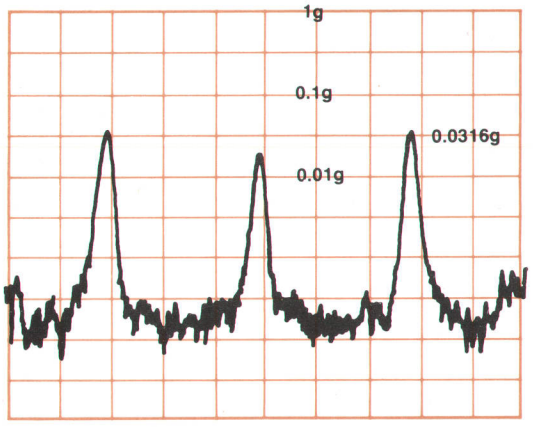


Fig. 2. *The sidebands represent the fan-induced vibrations of the 10-MHz reference oscillator assembly that caused the sidebands shown in Fig. 1.*

$$\left(\begin{array}{c}\text{Sideband Level Relative}\\ \text{to 6.1-GHz Carrier}\end{array}\right) = 20 \log \left[\frac{\left(\begin{array}{c}\text{Peak Frequency Deviation}\\ \text{of 10-MHz Standard}\end{array}\right)}{2 \times \text{Deviation Rate}} \times 610\right]$$

$$= \text{Sideband Level at 10 MHz} + 56 \text{ dB}.$$

Although this kind of fan is not used in the 8672A, it demonstrates the need for taking steps to prevent fan-generated spurious signals. In general, the vibrations a fan can couple into an instrument are not limited to just the slip frequency, such as the 30-Hz rate illustrated in Fig. 1. Vibration rates up to 1 kHz are typical for fans operated from a 50-Hz or 60-Hz power line. Fig. 3

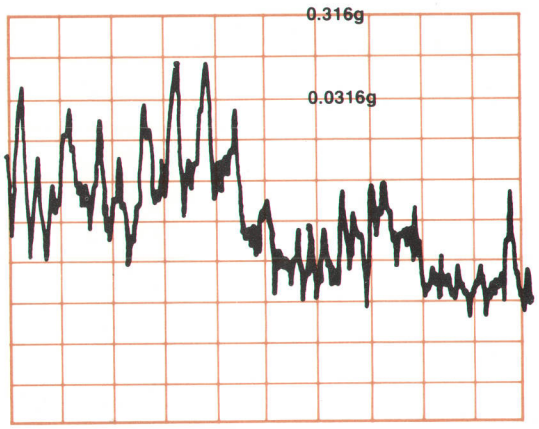


Fig. 3. *Typical radial vibration components of a freely suspended 50/60-Hz fan similar to those used in the 8672A.*

shows the radial acceleration components generated by a fan of the type used in the 8672A. These fans are balanced and individually tested to assure low vibrational acceleration levels.

Flexible mounting is the usual technique for preventing higher-frequency fan vibrations from reaching the instrument chassis. Natural rubber and neoprene mounts are readily available for this purpose. However, there is a potential problem: the fan and its flexible mounts resonate when fan vibrations coincide with the natural frequency of the system, which is given by $\sqrt{K/M}$, where K is the effective spring constant of the mounts and M is the fan mass. The resulting vibration level transmitted to the instrument chassis can be 15 to 20 dB greater than it would be with a rigidly mounted fan. Care must be taken to insure that improved performance at frequencies above resonance is not gained at the expense of degraded performance at resonance.

In the 8672A, the resonant frequency of the fan and its mounts is 58 Hz in the radial direction and 116 Hz in the axial direction. Consequently, vibration rates above 116 Hz are effectively decoupled from the chassis. At the two resonant frequencies, the only subassembly requiring extra attention is the 10-MHz reference standard. This oscillator's microphonic susceptibility is specified at less than 2×10^{-9} Hz/Hz/g. If a worst-case, fan-induced chassis vibration of 0.06 g at 58 Hz were coupled directly to the oscillator assembly, the resulting sideband on a 6-GHz output would be 40 dB below the carrier, slightly lower than the sideband in Fig. 1, and at 58 Hz. Using flexible mounts for assembly and selecting the mounts so that resonance occurs below 30 Hz, the potential −40 dB sideband is lowered to at least −60 dB.

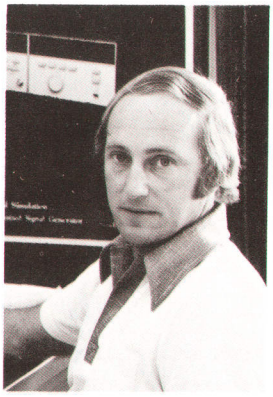
Provided that the source is stable enough, measuring vibration-induced spurious signals on a microwave source is relatively easy now that low-frequency spectrum analyzers with 1-Hz IF bandwidths are available (for example, the HP 3580A). The 8672A's long-term stability is 5×10^{-10}/day, so two of these synthesizers can be heterodyned together to generate a signal containing the same vibration sidebands that are present on the microwave output, but at a frequency within the range of a low-frequency spectrum analyzer.

-Carl Enlow

FM coil driver should also have signal-to-noise ratios and spurious responses consistent with the above objectives, and that the instrument should be capable of calibrated frequency modulation inside the phase-locked loop bandwidth without degrading the overall phase noise or spurious responses. How the FM requirements are met is discussed in greater detail in the article that follows.

Acknowledgments

Many talented people contributed to the frequency synthesis section of the 8671A/72A. Howard Booster worked on the D-to-A converter, the power supplies, and several electronic tools used for production testing of the instrument. Bob Dildine had responsibility for the YTO phase-locked loop and its FM circuitry. Carl Enlow worked on line related spurious and microphonics throughout the instrument. Bob Jewett performed the R&D characterization of non-harmonic spurious signals and wrote the production software to test for them. Ganesh Basawapatna and John Nidecker designed the low-noise YTO. Siegfried Linkwitz designed the harmonic mixing sampler, and Larry Nutting designed the M/N VCO. Ron Trelle did the product design, including the thermal and vibrational analysis of the instrument structure. Dave Wait provided production engineering support and design help on the VCXO phase-locked loop. Production introduction support was provided by Jeff Gould, Richard Bauhaus, Jim Gardner, Ken Kinser, Sam Zuck, Don Meador, and Val Peterson. In addition to those mentioned, many others contributed to the team effort required to introduce this product. My thanks to all of them.



Kenneth L. Astrof

Ken Astrof, project manager for the frequency synthesis section of the 8671A/8672A, has been doing microwave development for HP since 1966. A native of Vancouver, British Columbia, Canada, Ken attended Queen's University in Kingston, Ontario, graduating in 1964 with a BSc(EE) degree. Continuing his studies after joining HP, he received his MSEE degree in 1969 from the University of Santa Clara, California. Ken is married and has two sons and a daughter. What with landscaping his new home in Santa Rosa, California, studying for his MBA degree, and helping coach his sons' soccer team, he finds his schedule quite full, but whenever he can he likes to get away for some camping, swimming, or skiing.

Signal Generator Features for a Microwave Synthesizer

by Bradley C. Stribling

THE STABILITY AND SPECTRAL PURITY of synthesized frequencies are required for many applications. However, for a synthesizer to be truly useful as a laboratory instrument and for the broadest scope of applications, it needs certain signal generator features. These include precision leveling, modulation, and convenience of operation. These and other capabilities are provided by the digital control unit and RF output section of the 8672A Synthesized Signal Generator.

Tuning Controls

The 8672A is primarily a source of precision variable-frequency signals. Therefore, an easy-to-use tuning control was considered a must. However, since the 8672A consists of four phase-locked loops, each requiring its own special code to generate the proper frequency, and since it is a multiband instrument, thus requiring a division and roundoff algorithm, a certain amount of circuit sophistication is required to achieve simple operation. The necessary control functions are vested in the digital control unit (DCU).

For frequency selection, a single rotary tuning control is used. While it retains the feel of a mechanically tuned generator and permits tuning from 2.0 GHz to

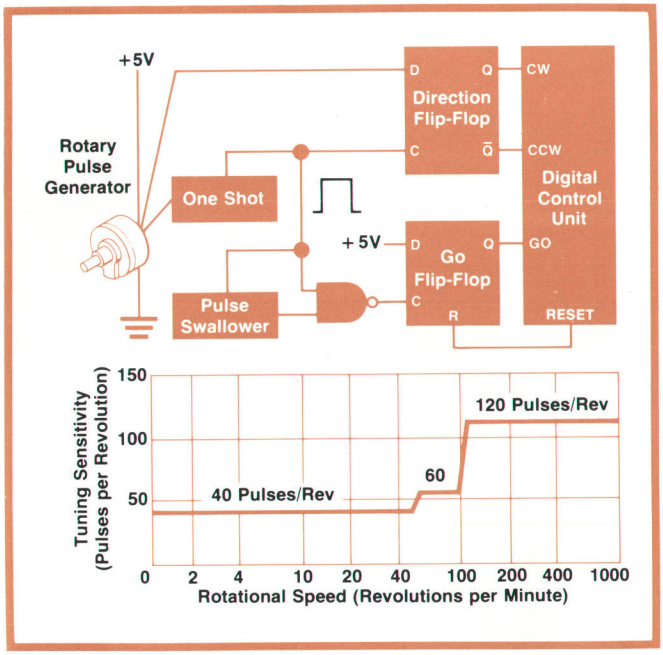


Fig. 1. *The 8672A Synthesized Signal Generator's single rotary tuning control permits tuning from 2 to 18.6 GHz without band switching, yet provides resolution as fine as 1 kHz. The tuning control is a rotary pulse generator. The adaptive resolution circuit shown here adjusts the control's tuning sensitivity to its speed of rotation, providing rapid tuning between widely spaced frequencies and slower tuning when approaching the desired frequency.*

18.6 GHz in a single operation without range or band switching, it easily affords resolution as fine as 2 kHz. This control is actually a rotary pulse generator, sometimes called an incremental shaft encoder. Its output pulses either increment or decrement a dedicated register in the DCU. The contents of this register are processed to generate the proper programming signals for the loops and the multiplier.

An adaptive resolution feature enhances the convenience of the tuning control. When tuning rapidly between frequencies, the control produces 120 frequency steps per revolution. Approaching the desired frequency, when tuning is slower, the control generates fewer frequency steps per revolution. Fig. 1 shows the adaptive resolution circuit and its function.

Tuning in fine-resolution steps across the entire frequency range would take a long time. To save time, the tuning resolution can be prescribed by one of four pushbuttons. Special light bars underscore the numerals of the digital frequency display to indicate which digits will respond to the tuning action, either directly or via carries. For example, with the finest resolution selected, all numerals from 1 kHz on up are underscored. With the coarsest resolution selected, only the 100-MHz digit and above are underscored, the lower ones being frozen (except that they are affected by the roundoff algorithm in the third band).

There is also a HOLD button that electrically disables the tuning control to prevent inadvertent mistuning. A 3.0-GHz PRESET button provides a convenient starting point for selecting new frequencies.

Since the frequency is digitally specified, some means of protecting it during accidental loss of power is required. The frequency register, a low-power MOS latch, is protected by a small NiCad battery pack, which is normally trickle-charged from the power line. In the absence of power, frequency memory can be retained for about one year.

Level Control

Accurate control of the output level is accomplished by three mechanisms. First, an internal automatic level control (ALC) assures output power flatness and stability. Second, a continuous vernier can vary the output over a 13-dB range, with the effects monitored on a front-panel meter. Third, a 10-dB step attenuator adjusts the final output, with its settings displayed on a front-panel numeric readout. These features combine to give accurate control of the output power level from −120 dBm to +3 dBm.

The ALC loop (see box, page 17) determines the instrument's output level stability and flatness with frequency, and its vernier and AM capabilities. A number of novel approaches are used. For example, the vernier linearly controls the output level expressed in dBm. In most instruments, the vernier varies the output voltage linearly, so setting a prescribed power level in dBm is very touchy at low levels, and the meter scale is very crowded there. The 8672A circumvents this by shaping the level detector's output to make it a logarithmic function of the output voltage. This yields the desired control characteristic.

In many applications it is desirable to level the power delivered to the load, instead of just to the output connector, to reduce the effects of reflections and cable losses. This is done by connecting a suitable power sensor at the load, such as a coupler and a crystal detector or power meter, and feeding the resulting sense voltage back to the signal generator to accomplish level correction (see Fig. 2). To this end, the 8672A provides for external ALC. The input circuitry of the ALC loop automatically adapts to accommodate either polarity of crystal or power meter, and can accept the bandwidths of all standard HP power meters. The instrument can be calibrated from the front panel so that an externally leveled setup retains good metered output accuracy.

Amplitude Modulation

Most high-quality instruments that provide AM

A High-Performance Microwave Power Leveling Loop

Accurately calibrated output power was one of the design goals for the 8672A. Another was wideband linear AM capability. Both of these features depend on the automatic level control (ALC) circuitry in the RF output module. A simplified diagram of the 8672A ALC loop appears below.

A portion of the RF output is fed to the detector by means of a broadband directional coupler. The dc output of the detector, point A in the diagram, is held constant to a value determined by the reference and AM inputs. This is accomplished by feedback action via the PIN modulator to compensate for variations in oscillator output, amplifier gain, YTM conversion efficiency, and various mismatches. With the detector output held constant, the output level accuracy depends on the RF flatness of the detector, the coupler, the attenuator, and the connecting hardware following the coupler. Detector temperature coefficient is a serious problem and is compensated in the ALC circuitry.

The leveled output is affected by harmonic content because the detector responds to harmonics as well as the fundamental. The YTM attenuates harmonics to at least −30 dBc, so their effect is negligible.

Only a small amount of RF power is coupled to the detector, so the detector operates in its square-law region. Thus the voltage at point A is not a linear function of either the RF voltage or the RF power expressed in dB. The detector logarithmic amplifier converts the detector's output to a voltage at point B that is linear with RF output expressed in dB. Thus the loop's reference input, which comes from a digital-to-analog converter, can be linear in dB, simplifying the digital control of output power.

The external leveling input goes to an absolute value converter (gain = ±1, depending on input polarity), then to a logarithmic amplifier. Therefore, any square-law detector or linear power meter can be used for external leveling without affecting the accuracy of the front-panel level indicators.

The AM input is also logarithmically shaped and fed to the loop summing point (C) with appropriate gain. This permits linear, calibrated AM, independent of RF output level and modulation depth.

Constant Gain for Wideband AM

To maintain a high AM bandwidth it is necessary to hold the ALC loop gain fairly constant. There are several sources of gain variation. PIN modulator gain is proportional to oscillator output voltage as well as being a strong nonlinear function of the drive current. The power amplifier is driven close to saturation. The YTM is very nonlinear on the multiplying bands, and the detector has a square-law characteristic.

Variations caused by the detector are corrected by the detector logarithmic amplifier. Its output is a constant 6 mV/dB independent of RF output level. This amplifier operates over an 80-dB range of input voltages while maintaining a minimum bandwidth of 1 MHz. Its temperature coefficient is less than 2 μV/°C and its bias current is less than 10 nA. The amplifier itself is a discrete design with a gain-bandwidth product of 500 MHz.

The PIN modulator transfer characteristic is of the form

$$\log\left[\frac{P_{out}}{P_{in}}\right] = K - n\log(I + I_o)$$

where I is the modulator drive current. If I_o is removed, the graph of this function is a straight line on log-log paper. The

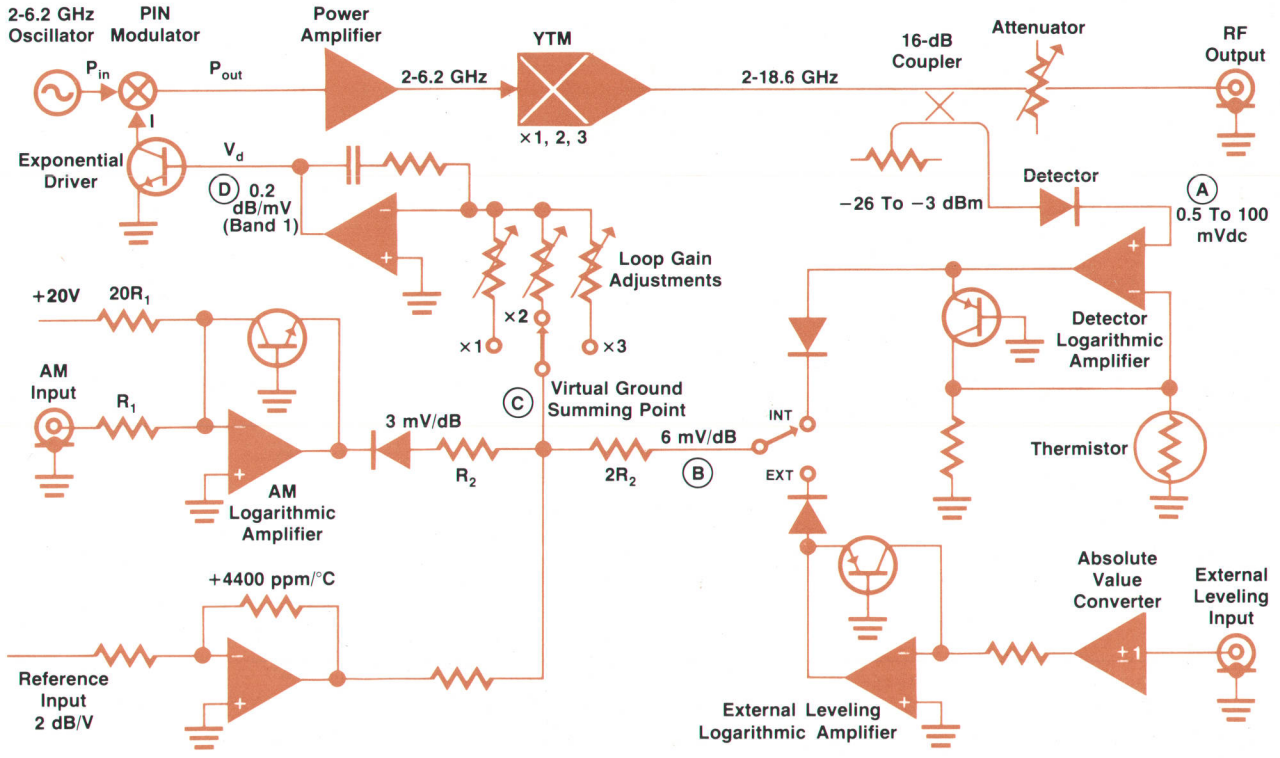


Fig. 1. *Simplified diagram of the 8672A ALC loop.*

slope, n, changes slightly with frequency. With the amplifier and YTM included, the line is not so straight, because n also changes with power level. The use of self-biasing on the step recovery diode in the YTM reduces variations in n to less than 2:1 over any one band. Changing the harmonic band of the YTM changes n considerably more.

If the PIN modulator is driven from an exponential current source

$$I = I_1 e^{V_d/V_o},$$

then the modulator transfer function becomes

$$\log \frac{P_{out}}{P_{in}} = K_1 - n \frac{V_d}{V_o}.$$

Thus the modulation characteristic can be expressed as N dB/mV, relatively constant over power and frequency within one band. Band switching simply results in a gain change, easily handled by switching gains elsewhere. The gain from point D to point B is the product of the modulator characteristic, N_1 dB/mV, and the detector logarithmic amplifier characteristic, N_2 mV/dB, or N_1N_2, a dimensionless constant. This allows a minimum AM bandwidth of 500 kHz on the fundamental band and about 300 kHz on the multiplying bands, with no gain adjustments required for varying operating conditions.

Temperature Compensation

The detector uses a low-barrier Schottky diode. Compared to older point-contact types, it has improved resistance to physical shock and lower video resistance (given by kT/qI_s, where I_s is the diode saturation current). Lower video resistance improves the detector bandwidth, but has a detrimental effect on temperature coefficient. The diode junction sees only a portion of the open-circuit RF voltage because of a voltage divider consisting of the RF source resistance, some resistors on the detector's substrate, and the diode video resistance. Since I_s doubles every 20°C, the division factor changes with temperature. The resulting temperature coefficient is a function of both temperature and power level, and varies from 0.03 dB/°C to 0.2 dB/°C.

The resistance of a thermistor is an exponential function of temperature, as is I_s. A properly specified thermistor in a voltage divider network will match the coefficient of the detector, leaving a small residual drift that is a function of power level but linear with temperature. Referred to the output of the detector logarithmic amplifier, this appears as gain variation with a temperature dependence of +1100 ppm/°C. A logarithmic amplifier with its offset drift cancelled exhibits a gain drift of +3300 ppm/°C, and this adds to the residual detector drift.

The reference voltage feeding the summing point (C) is derived from a network with a TC of +4400 ppm/°C, thereby matching the drift of the detector and logarithmic amplifier. The overall temperature dependence, allowing for component tolerances, is typically within ±0.015 dB/°C for output power up to +5 dBm and circuit temperature from 10°C to 70°C.

-Stephen Sparks

capability accomplish the modulation by impressing the AM signal on the ALC loop reference. This has the effect of raising and lowering the output level in response to the input signal, thereby achieving amplitude modulation. This is also done in the 8672A, except that the modulating signal is pre-shaped before being impressed on the reference. There are two reasons for this. First, as in all signal generators, the output level detector has certain nonlinearities, which falsely represent the power level and distort the AM. Second, since the 8672A level is controlled linearly in dBm, the detector characteristic is pre-shaped as described earlier, so the output level fluctuates exponentially in response to variations on the reference. Pre-shaping the AM signal in the 8672A not only achieves low-distortion linear AM, but also constant AM percentage as the output level is changed.

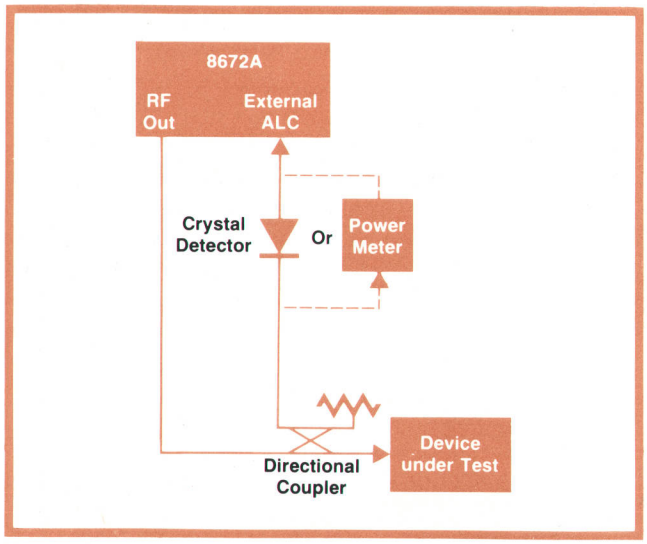


Fig. 2 *8672A RF output power is internally leveled. An external ALC input makes it possible to level the power actually delivered to the load. A coupler and a power meter or detector serve as the power sensor at the load.*

The 8672A provides two ranges for AM operation, 100% per peak volt and 30% per peak volt. The front panel meter monitors the amount of modulation.

Achieving wideband AM in a signal generator requires maintaining the ALC loop gain as constant as possible. This is a demanding task when there are devices in the loop with nonlinear dynamic characteristics. In the 8672A, the YTM characteristic is made more linear by self-biasing. The PIN modulator in the loop is driven from an exponential current source, and the resulting composite characteristic is precisely what is needed to account for the logarithmic characteristic of the shaped detector.

Fig. 3 shows the AM frequency response of the 8672A.

Frequency Modulation

A common annoyance when frequency modulating

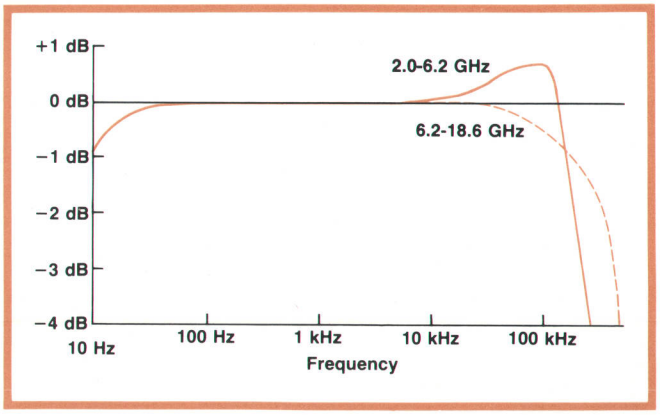


Fig. 3. *8672A AM flatness. Specified 3-dB bandwidth for amplitude modulation is 10 Hz to 100 kHz.*

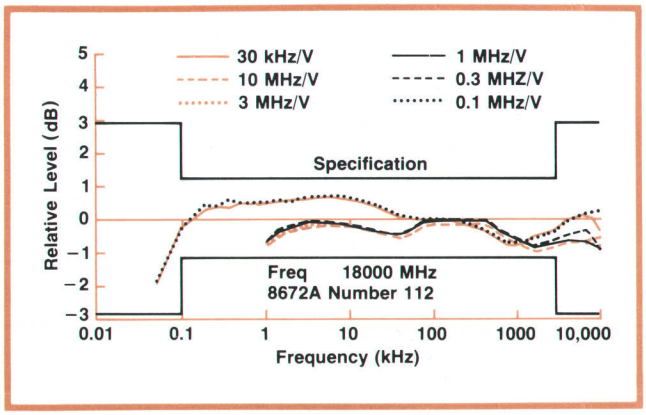


Fig. 4. *8672A FM flatness relative to 100 kHz. Both the 8671A and the 8672A can be frequency modulated at rates up to 10 MHz with peak deviations as high as 10 MHz, depending on the FM rate.*

many signal generators is their inherent frequency drift in the FM mode. Synthesizers can get around this problem by modulating while phase-locked. But this approach also has its drawbacks. Frequency modulation at rates well above the phase-locked loop's bandwidth is no problem. But if the modulation frequency is less than the loop's bandwidth, the loop will act to eliminate the perturbation, which is precisely the loop's purpose. Since in most synthesizers the loop bandwidths are intentionally wide, the useful FM range is substantially impaired.

The 8672A circumvents this problem with some innovative design. For low frequencies, the FM signal is integrated, inverted, and then summed with the phase-detector's output. The result is cancellation of the loop error voltage generated by the desired FM, thereby allowing low-rate FM inside the phase-locked loop bandwidth.

Operation of the 8672A's FM system is described in the box on page 20. Fig. 4 shows the 8672A's FM frequency response.

Remote Operation

To be truly versatile in today's technology, an instrument must be easily interfaced for remote programming in computer-based automatic measurement systems or easily-configured desktop-computer controlled test setups. Remote programming also eases diagnosis and testing of the instrument itself.

The 8672A is equipped with the HP interface bus (IEEE-488). It can be interfaced quickly with a desktop computer or other controller, along with digital power meters, counters, DVMs, and so on, to implement a flexible and sophisticated test system.

Programming the 8672A is designed to be as simple as possible. Each internal function (such as a frequency digit, the output attenuation, or the modulation function) has its own internal address. These are generally the letters A through P. Programming is accomplished by the controller's transmitting the address followed by the desired value (a BCD or hexadecimal digit). Where a large number of addresses are to be programmed, for example, in setting all digits of the output frequency, the 8672A automatically sequences through the addresses until told to stop by a specific address code.

The 8672A can also output certain information onto the bus. This information is mainly status or fault data. For example, the instrument may be interrogated as to whether it is out of lock, unleveled, or programmed out of range. This can be important to a test system that must wait until a certain condition (such as phase-lock) exists before making its measurement. It can also avoid misleading data should a failure occur. Five fault conditions are included in the status byte (Fig. 5). If one or more of these occurs, a

Bit Number	8	7	6	5	4	3	2	1
Decimal Value	128	64	32	16	8	4	2	1
Function	Crystal Oven Cold	RSV Request Service	Out Of Range (Frequency)	RF Off	Not Phase Locked	Level Uncal	FM Over-Mod	+10 dBm Over Range

Fig. 5. *HP interface bus (IEEE 488) programmability makes it easy to configure systems. The 8672A can place various status indicators on the interface bus. If any fault condition occurs a service request is automatically generated.*

A Calibrated 50-Hz-to-10-MHz FM System

FM in the 8672A is applied to the FM coil of the YTO while it is phase-locked. At the FM input one volt peak gives maximum deviation on each of the six FM ranges, which are 10, 3, 1, 0.3, 0.1, and 0.03 MHz per volt peak. Deviation varies linearly with the modulation input between zero and one volt peak.

Fig. 1 is a block diagram of the FM system and YTO phase-locked loop. The noise performance of the synthesizer is preserved while frequency modulating by maintaining the YTO loop bandwidth at 20 kHz. When modulation is applied to the YTO FM coil at frequencies lower than 20 kHz, it is reduced by the feedback action of the loop. To prevent this reduction of modulation, a signal equal in magnitude to the error signal from the phase detector is used to cancel the phase detector error signal. This technique results in essentially flat FM response above and below the YTO loop bandwidth.

The error voltage, V_e, from the phase detector is proportional to the phase difference of its two inputs, that is, $V_e = K_D\phi$, where ϕ is the phase of the YTO with respect to the reference signal and K_D is the phase detector gain. The phase of the YTO signal with FM is:

$$\phi = \int [\omega_c + \omega_d(t)]dt$$

where ω_c is the carrier frequency and $\omega_d(t)$ is the instantaneous deviation, usually expressed as $A\sin\omega_m t$ for sinusoidal modulation. The deviation is proportional to the modulating voltage, $V_{FM}(t)$:

$$\omega_d(t) = K_{FM}\, V_{FM}(t).$$

The error voltage from the phase detector is then:

$$V_e = K_D\int [\omega_c + K_{FM}V_{FM}(t)]\, dt = K_D \int \omega_c dt + K_D K_{FM}\int V_{FM}(t)dt.$$

The first term is the error voltage that keeps the loop locked and the second term is the error voltage caused by the applied FM. If a voltage proportional to the integral of the applied FM signal is used to cancel the phase detector error voltage produced by the FM, the loop is transparent to FM. In the 8672A, the FM signal is integrated, inverted, and summed into the YTO loop just after the phase detector.

The band attenuator at the FM input divides the modulation signal amplitude by a factor of 1, 2, or 3 to compensate for the deviation multiplication that occurs in the YTM, so deviation is independent of the YTM multiplication number. For example, in the 6.2-to-12.4-GHz frequency band, the YTM multiplies its input frequency by two. Deviation is multiplied by the same factor, so the modulation amplitude is divided by two to maintain constant deviation independent of microwave output frequency.

The FM amplifier has adjustable gain for calibrating the system. It also provides isolation between the band attenuator at its input and the range attenuator at its output. Both attenuators are resistive dividers switched by CMOS analog switches.

The FM range attenuator allows calibrated and metered FM at sensitivities that normally would not show up on the meter. This attenuator is divided into three main parts. The first part, located after the FM amplifier, has 0, 10, 20, or 30 dB of attenuation. The second part consists of 40 dB of switchable attenuation located at one input to the FM coil driver. The third part provides a switchable 40-dB gain reduction at the output of the FM integrator. Locating the 40-dB pad at the output instead of the input of the FM integrator improves the signal-to-noise ratio in the 0.1 and 0.03-MHz ranges.

FM frequency response of the 8672A is 50 Hz to 10 MHz. Frequency response inside the YTO loop bandwidth depends on how well the FM integrator performs and how well the phase detector error signal resulting from FM is cancelled. Actually, the FM integrator performance is important to several octaves beyond the loop bandwidth, because the error signal is significant until the loop gain has dropped to well below unity. Frequency response in the range of 10 kHz to 10 MHz is further affected by a gradual loss of FM coil sensitivity that amounts to about 6 dB over the three-decade range. This loss is compensated by a frequency shaping network in the FM driver.

The FM coil and its feedthrough capacitor are resonant at approximately 14 MHz. To maintain constant coil current as the FM frequency approaches the resonance of the coil, the coil and

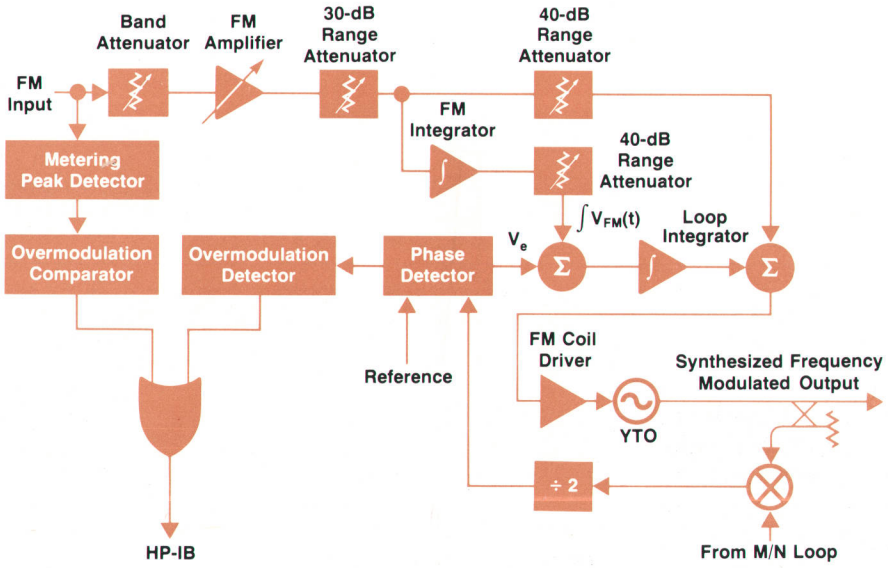


Fig. 1. *Block diagram of 8671A/8672A FM system and YTO loop.*

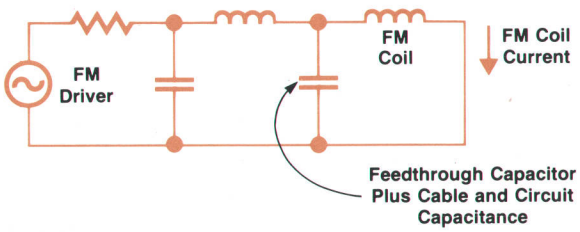


Fig. 2. *To maintain constant FM coil current to the YTO as the FM rate approaches the coil resonance (14 MHz), the coil and its feedthrough capacitor are incorporated into a low-pass filter.*

its feedthrough capacitor are incorporated into a low-pass filter designed to be terminated by a short circuit (see Fig. 2). Since the FM coil is the last inductance before the short circuit, the coil current is the same as the output current and is maintained constant as frequency increases. The filter cutoff frequency is only slightly higher than the highest modulating frequency, helping to reduce harmonic distortion at high modulation rates.

FM frequency response variation is typically less than 1 dB from 100 Hz to 3 MHz and less than 2 dB at 10 MHz. Because of the high sensitivity (10 MHz per volt on the most sensitive range), extremely small line related signals or noise at the input of the FM system result in noticeable sidebands on the output. For example, only 60 nanovolts of stray 60-Hz signal at the input will result in 60-Hz sidebands 40 dB below the carrier. Because of the inevitable ground loops that occur as a result of connecting the 8672A to FM sources, the low-frequency cutoff is specified as 1 kHz for FM deviation ranges greater than 0.1 MHz per volt. Response extends down to 50 Hz for FM deviation ranges 0.1 MHz per volt and below.

The frequency modulated signal from the YTO appears at the input to the phase detector. Therefore the instantaneous phase excursions must not exceed the range over which the phase detector can operate. The phase detector used in the YTO loop has a range of $\pm 2\pi$ radians and therefore limits the modulation index at the phase detector input to 2π. A divide-by-two circuit ahead of the phase detector increases the maximum modulation index of the YTO to 4π. If this modulation index is exceeded, the phase detector output becomes unsymmetrical and the resulting dc offset causes the loop to false-lock at a frequency offset that is an integral multiple of the FM rate. For the loop to recover reliably, the modulation index must be reduced to approximately half the maximum theoretical value. Therefore, the maximum modulation index allowed is half the available range of the phase detector. This assures that transients in the FM signal or transients due to frequency changes of the synthesizer do not cause false locking.

FM deviation metering is accomplished with a broadband peak detector that monitors the input voltage and outputs a dc level proportional to the input voltage. This dc voltage is used to drive the front-panel meter and the overmodulation comparator.

Overmodulation in the 8672A can occur if the maximum deviation exceeds 10 MHz or if the modulation index exceeds the specified value of 5 (which allows some margin below the theoretical value of 2π). The maximum modulation index at the output of the instrument increases on bands 2 and 3 because of the multiplication of the YTM. Two separate detectors are used to indicate overmodulation. The first detector is a comparator that monitors the dc level in the metering circuit and is actuated whenever the input signal exceeds one volt peak. The second detector monitors the phase excursions in the YTO loop and is actuated whenever a modulation index of approximately 2π is reached. Both detector outputs are combined to indicate whether an overmodulation condition exists.

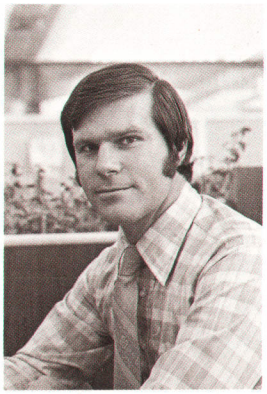
-Robert Dildine
-Ronald Larson

service request is generated.

The status indicators can also be used to speed up operation in an automatic test system. For example, the 8672A's specified switching time between any two frequencies is 15 ms, but it often switches faster than this. The system controller can interrogate the 8672A to determine when switching is complete, instead of having to wait the full 15 ms between frequency steps.

Acknowledgments

The success of the 8672A derived from the dedicated efforts of many talented people. Greatly appreciated are the efforts of Bob Devries and Charles Cook for the mechanical design, Paul Zander for the digital circuitry, John Hasen for the 20-30-MHz loop, Ron Larson for the modulation circuitry, Steve Sparks for the ALC loop design, Russ Mendenhall for YTM driver circuitry, Young Dae Kim for his work on the microwave components, and Dave Whipple and John Cooper for their producibility efforts and testing.



Bradley C. Stribling

Brad Stribling was project manager for the 8672A RF output section and digital control unit. Born in Edmonton, Canada, he received his BS and MS degrees in electrical engineering from the University of California at Berkeley in 1967 and 1969. With HP since 1968, he's had various responsibilities related to the 5360A Computing Counter and the 8660 Synthesized Signal Generator system. He's now section manager for 8640 production. Brad is married, has a son, and lives in Sunnyvale, California. For recreation he goes skiing and backpacking, coaches a youth soccer team, plays bridge, and wheels and deals in real estate.

Personal Calculator Algorithms III: Inverse Trigonometric Functions

A detailed description of the algorithms used in Hewlett-Packard hand-held calculators to compute arc sine, arc cosine, and arc tangent.

by William E. Egbert

BEGINNING WITH THE HP-35,[1,2] all HP personal calculators have used essentially the same algorithms for computing complex mathematical functions in their BCD (binary-coded decimal) microprocessors. While improvements have been made in newer calculators,[3] the changes have affected primarily special cases and not the fundamental algorithms.

This article is the third of a series that examines these algorithms and their implementation. Each article presents in detail the methods used to implement a common mathematical function. For simplicity, rigorous proofs are not given, and special cases other than those of particular interest are omitted.

Although tailored for efficiency within the environment of a special-purpose BCD microprocessor, the basic mathematical equations and the techniques used to transform and implement them are applicable to a wide range of computing problems and devices.

Inverse Trigonometric Functions

This article will discuss the method of generating $\sin^{-1}$, $\cos^{-1}$, and $\tan^{-1}$. An understanding of the trigonometric function algorithm is assumed. This was covered in the second article of this series and the detailed discussion will not be repeated here.[4]

To minimize program length, the function $\tan^{-1}A$ is always computed, regardless of the inverse trigonometric function required. If $\sin^{-1}A$ is desired, $A/\sqrt{1-A^2}$ is computed first, since

$$\sin^{-1} A = \tan^{-1} \frac{A}{\sqrt{1-A^2}}.$$

For $\cos^{-1}A$, $\sin^{-1} A$ is computed as above and then $\cos^{-1} A$ is calculated using

$$\cos^{-1} A = \pi/2 - \sin^{-1}A.$$

$\cos^{-1}$ is found in the range $0 \leq \theta \leq \pi$ and $\sin^{-1}$ and $\tan^{-1}$ are computed for the range $-\pi/2 \leq \theta \leq \pi/2$. The $\tan^{-1}$ routine solves only for angles between 0 and $\pi/2$, since $-\tan A = \tan(-A)$. Thus A may be assumed to be positive and the sign of the input argument becomes the sign of the answer. All angles are calculated in radians and converted to degrees or grads if necessary.

General Algorithm

A vector rotation process similar to that used in the trigonometric routine is used in the inverse process as well. A vector expressed in its X and Y components can easily be rotated through certain specific angles using nothing more than shifts and adds of simple integers. In the algorithm for $\tan^{-1}|A|$, the input argument is $|A|$, or $|\tan\theta|$, where θ is the unknown. Letting $\tan\theta = Y_1/X_1$, $|A|$ can be expressed as $|A|/1$, where $Y_1 = |A|$ and $X_1 = 1$. A vector rotation process (see Fig. 1) is then used to rotate the vector clockwise through a series of successively smaller angles θ_i, counting the number of rotations for each angle, until the Y_2 component approaches zero. If q_i denotes the number of rotations for θ_i then

$$|\theta| = q_o + q_1\theta_1 + \ldots + q_i\theta_i + \ldots$$

This process is described in detail below.

Vector Rotation

To initialize the algorithm, A and 1 are stored in fixed-point format in registers corresponding to Y_1

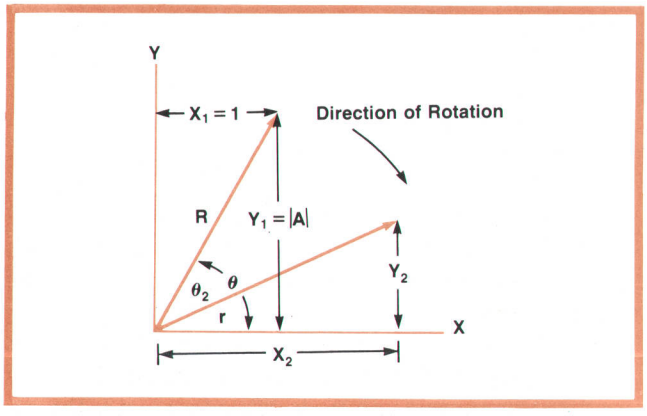


Fig. 1. *Vector rotation.*

and X_1. This is done in such a way as to preserve as many digits of A as possible when the exponent of A differs from zero.

At this point the sign of A is saved and $Y_1 = |A|$. Now comes the vector rotation (see Fig. 1). If the vector **R** is rotated in a clockwise direction, Y_2 becomes smaller and smaller until it passes zero and becomes negative. As soon as Y_2 becomes negative, we know that we have rotated **R** just past the desired angle θ. Thus to find θ, **R** is simply rotated clockwise until Y_2 becomes negative. The amount of rotation is remembered and is equal to the desired angle $\theta = \tan^{-1}|A|$. To rotate **R**, the following formula is used.

$$\frac{X_2}{\cos\theta_2} = X_1 + Y_1 \tan\theta_2 = X_2' \qquad (1)$$

$$\frac{Y_2}{\cos\theta_2} = Y_1 - X_1 \tan\theta_2 = Y_2' .$$

This equation is the same as equation 1 of the article on trigonometric functions,[4] except that the plus and minus signs are exchanged because **R** is rotated in the opposite direction. As before, $\tan\theta_2$ is chosen such that the implementation requires a simple shift and add ($\tan\theta_2 = 10^{-j}$). To find θ, **R** is initially rotated with $\tan\theta_2 = 1$ ($\theta_2 = 45°$). Y_2' soon becomes negative and the number of successful rotations is stored as the first digit of what is known as the pseudo-quotient. Y_2' is then restored to the last value it had before becoming negative and **R** is rotated again, this time through a smaller angle, i.e., $\tan\theta_2 = 0.1$ ($\theta_2 \approx 5.7°$). This process is repeated with the angle of rotation becoming smaller and smaller until five pseudo-quotient digits have been generated.

At the end of each series of rotations, Y_2 is multiplied by 10 to preserve accuracy.

Pseudo-Multiplication

It is now time to shift gears and add up all the small angles represented by the pseudo-quotient digits. There remains a residual angle ,r, represented by the final X_2' and Y_2'. Since the residual angle is small, we would like to say $Y_2' = \sin r = r$. However, this is true only if $X_2' = 1$. Unfortunately, X_2' in this case is the product of all the $1/\cos\theta$ terms resulting from several applications of equation 1. However, Y_2' is this same product times Y_2. Thus $Y_2'/X_2' = Y_2/1$. Therefore, the final Y_2' is divided by the final X_2' and the result is sin r, which for small angles in radians is approximately equal to r, the residual angle.

With the residual angle as the first partial sum, θ is generated by adding the angles represented by the digits of the pseudo-quotient. This is exactly the reverse of the pseudo-division operation in the trigonometric routine. Thus:

$$\theta = q_o \tan^{-1}(1) + q_1 \tan^{-1}(0.1) + q_2 \tan^{-1}(0.01) + \ldots + r \qquad (2)$$

Each coefficient q_i refers to the count in a particular pseudo-quotient digit.

The result of this summation process, also called pseudo-multiplication, is an angle θ that is equal to $\tan^{-1}|A|$, where $|A|$ is the input argument to the $\tan^{-1}$ routine. At this point the original sign of A is appended to θ. For $\tan^{-1}$ this angle is normalized, converted to degrees or grads if necessary, and displayed. Recall that for $\sin^{-1}$, $A/\sqrt{1-A^2}$ was first generated. Thus for $\sin^{-1}$, the result of the $\tan^{-1}$ routine is again simply normalized, converted to degrees or grads if necessary, and displayed. For $\cos^{-1}$, the $\tan^{-1}$ routine returns $\sin^{-1}$. $\cos^{-1}$ is then simply found as $\pi/2 - \sin^{-1}A$.

Summary

In summary, the computation of inverse trigonometric functions proceeds as follows:

1. Calculate $A/\sqrt{1-A^2}$ if the desired function is $\sin^{-1}$ A or $\cos^{-1}$ A.
2. Place $|A|$ and 1 in fixed-point format into appropriate registers, while preserving the sign of A.
3. Repeatedly rotate the vector with A=Y and 1=X clockwise using equation 1 until Y approaches zero. The number of rotations and the amount of each rotation is stored as a pseudo-quotient along the way.
4. Using the pseudo-multiplication process of equation 2, sum all of the angles used in the rotation to form $|\theta|$.
5. Append the proper sign to the answer and calculate $\cos^{-1} A = \pi/2 - \sin^{-1} A$ if required.
6. Convert to the selected angle mode, and round and display the answer.

The calculator is now ready for another operation.

References

1. T.M. Whitney, F. Rodé, and C.C. Tung, 'The Powerful Pocketful': An Electronic Calculator Challenges the Slide Rule," Hewlett-Packard Journal, June 1972.
2. D.S. Cochran, "Algorithms and Accuracy in the HP-35," Hewlett-Packard Journal, June 1972.
3. D.W. Harms, "The New Accuracy: Making $2^3=8$," Hewlett-Packard Journal, November 1976.
4. W.E. Egbert, "Personal Calculator Algorithms II: Trigonometric Functions," Hewlett-Packard Journal, June 1977.

Bill Egbert is a project manager at HP's Corvallis (Oregon) Division.

Viewpoints

Tom Hornak on Fiber-Optic Communications

Much is being said in the technical press these days about fiber-optic communications. By now, it would seem, all electronics engineers are aware of the great advantages that optical fibers have over coaxial cable as a transmission medium, namely, a drastic reduction in weight and size for a given bandwidth, electrical isolation, immunity to electromagnetic interference, and a capability for secure transmission of information. So, in view of the explosive growth of communications—beginning with essentially man-to-man messages and now encompassing transmission of complex information, computation, instrumentation, and control data—the question arises: why are fiber-optic links not in high-volume use?

Before addressing that question, a brief review of the present state of fiber-optic technology is in order. The concept of communicating via a light beam in air has been known for many years, but range, dependability, and general usefulness were limited by atmospheric absorption of light and the need for line-of-sight transmission paths. A way around these limitations was developed over 20 years ago by N.S. Kapany at the Imperial College of the University of London, who proposed the use of glass fibers as optical waveguides for conducting light waves around corners. A few years later, scientists at the Standard Telecommunication Labs in England determined the theoretical limits of attenuation in glass fibers to be in the neighborhood of only 1 dB/km, much lower than that available in glass fibers at the time. However, it was not until 1970, when the Corning Glass Works developed a low-loss glass fiber with an attenuation of 20 dB/km, that large-scale developments in fiber-optic communication systems got under way.

Another necessary ingredient for a practical fiber-optic communications system is a highly concentrated but convenient-to-use light source. During the 1960s, developments in solid-state light sources by such companies as General Electric, IBM, Bell Telephone Labs, Hewlett-Packard and Monsanto—coupled with the availability of solid-state optical detectors—contributed towards the realization of practical optical communication systems. Presently, fiber-optic communication links are being tested by the military and the telecommunications companies for the purpose of exposing all components of the system to a real-life environment, and to gather operational experience. Western Electric's Atlanta and Chicago experiments, GTE's link in California, used in regular telephone service, and other installations here and abroad involve multi-kilometre links using repeaters with transmission rates up to more than 100 Mb/s. (Present systems transmit information in digital form to permit noise-free regeneration at each repeater. Analog transmission, as in multichannel CATV, is used only where wide analog signal bandwidths preclude digitization by inexpensive means.)

Fiber-Optic Basics

An optical fiber is basically a cylindrical waveguide made of a transparent material, glass or plastic, with a diameter ranging from 50 to 400 μm. It is made in such a way that the index of refraction, is maximum at the cylinder axis (about 1.5) and minimum at the circumference (about 1% less). Because of the higher velocity of propagation where the index of refraction is lower, light waves that are not exactly parallel to the fiber's centerline are bent back towards the center as they approach the circumference of the waveguide and thus follow undulating paths along the fiber. A fiber's maximum acceptance angle for light waves increases with greater variation in the refractive index, measured along the radius of the fiber, and ranges up to 20° for fibers with a step change in refractive index.

Attenuation minima exist in glass fibers at light wavelengths around 820 and 1060 nm. The 820-nm wavelength is quite compatible with existing light sources and detectors. Plastic fibers, suitable for low-cost, short-link applications, have a minimum attenuation of approximately 350 dB/km at a wavelength around 655 nm. With contributions from many companies worldwide, a steady improvement in fiber "transparency" is taking place with fibers made of high-purity silica, using processes comparable to those of the semiconductor industry, achieving attenuations approaching theoretical limits.

The bandwidth capability of an optical waveguide is one of its most striking characteristics. With the "carrier" frequency of a light source being greater than 3×10^{14} Hz, attenuation is independent of modulation frequencies up to many GHz. By contrast, attenuation in coaxial cables, typically 6 dB/km at 10 MHz, increases with the square root of frequency, requiring the use of complex equalization networks for wideband transmissions.

The principal bandwidth limitation in commonly used fibers results from the fact that all light rays captured by a fiber at its input do not experience the same delay in traveling to the output. Consequently, a zero-rise-time change in light intensity at the input has a finite rise time at the output. This is known as pulse dispersion and results from two phenomena.

The first, called modal dispersion, occurs because rays of differing angles of entry within the acceptance angle propagate over different path lengths to the output. Modal dispersion, being the largest for a fiber with a step-index profile, increases with acceptance angle, which in turn is determined by the variation in refractive indexes. Modal dispersion can be minimized by use of graded index profile fibers, first introduced by Nippon Electric in Japan. Here, the increase of path length with a ray's increasing angle of entry is compensated by the higher average propagation velocity as the ray's path extends into areas that have increasingly lower refractive index.

The second phenomenon, called material dispersion, is caused by the variation of refractive index as a function of wavelength, which affects propagation velocity. Material dispersion is thus proportional to source line width, which is typically 35 nm for LEDs and 2 nm for solid-state lasers.

Some examples will illustrate the effects of these dispersion mechanisms. Modal dispersion in a step-index fiber is about 35 ns/km, limiting the bit rate of a 1-km link to about 25 Mb/s, and it exceeds material dispersion using either LED or laser sources. In a silica fiber material dispersion with an LED source is larger than modal dispersion with a graded-index profile, and is about 3.5 ns/km. It would limit bit rate to about 250 Mb/s in a 1-km link. With a laser source, however, material dispersion in the same fiber drops to 0.2 ns/km so bit rate is then limited by modal dispersion to about 1 Gb/s in a 1-km link. In a single-mode fiber (one in which the diameter of the higher-index core is comparable to the source's wavelength) bandwidth is limited by material dispersion, even with a laser source. For a 1-km link, maximum bit rate would be 5 Gb/s or, in view of the present practical limits on laser modulation, 1 Gb/s on a 5-km link. However, the active core of a single-mode fiber is only a few micrometres in diameter and the resulting difficulty in splicing and connectoring is a practical obstacle to widespread use of these fibers.

Convenient Size

The size and weight of a fiber-optic cable leads to an important

contribution of this technology to communications: the ease of installation and the savings in duct space. The inherent mechanical strength of thin glass fibers is surprisingly high. Any initial weakness is caused by surface scratches and other flaws along the fiber. Work is being done to better understand the long-term behavior of these imperfections and how the environment affects them. Clean-room techniques and coating the fiber with a protective plastic during the pulling operation lead to mechanical strengths greater than 0.7 GN/m^2, equivalent to the stress on a fiber 100μm in diameter when suspending a 0.5-kg weight. Fibers are usually combined with additional strength members into cables that can withstand the tensions encountered during installation in ducts. The diameter of a fully protected cable is still very small, however, being on the order of 2.5 mm. This is a significant reduction in size and weight compared to a coaxial cable of equivalent bandwidth and attenuation.

It is possible, of course, to combine several fibers into a common protective cable assembly. This was done in the early days of fiber-optic communications to increase the waveguide cross-section so the coupling problem would not be so severe, and also to avoid complete loss of communication in case a single fiber breaks. Because of progress in fiber reliability and coupling technology, this approach is now used mainly for high-risk, short-link applications only. On the other hand, each fiber could be dedicated to a separate communication channel. This is possible because of the almost total absence of cross-coupling between adjacent fibers.

Zero Electrical Conductivity

The dielectric nature of the fiber contributes many advantages. An optical-fiber link maintains practically complete electrical isolation between transmitter and receiver, effectively eliminating ground loops. It is also immune to electromagnetic interference. Both of these properties are advantageous in heavy industrial applications (control systems, data acquisition systems, distributed computer links, etc.).

A fiber-optic link itself does not radiate electromagnetic energy. Therefore it is suitable for carrying signals near sensitive electronic equipment or for securing data transmissions against electromagnetic eavesdropping. Using clandestine taps for eavesdropping is also difficult because the optical power diverted into the tap will be noticed as a drop in signal level at the legitimate receiver.

Sources

Concurrent with developments in fibers, a great deal of development work is being devoted to light sources. Two are commonly used, both based on GaAlAs and both emitting around 820 nm. LEDs are generally used for lower cost applications where bandwidth requirements are below 30 MHz. The light output of an LED is nearly proportional to its drive current, leading to simple drive circuits.

The disadvantages of LEDs are their relatively large response time (on the order of nanoseconds) and linewidth, both leading to bandwidth limitations. Furthermore, LEDs emit light with no expressed directionality so light rays exceeding the fiber's acceptance angle are not utilized. A typical LED for fiber-optic applications generates an output power of 1 mW with a quantum efficiency of around 3%. Only 5 to 25% of the output power is coupled into the fiber.

To match light-emitting area to fiber cross-section, efforts have been directed towards developing high-radiance LEDs in which the current flow, and hence the light generation, is concentrated into an area less than 100 μm in diameter. This leads to current densities as high as 10,000 A/cm^2, requiring elaborate structures with good heat-sinking capabilities to assure long life, such as the LED developed by Burrus of Bell Telephone Laboratories.

The solid-state laser, a high-current-density, multilayer device fabricated by liquid-phase epitaxy, is the preferred source for high-performance links where the emphasis is on distance and/or bandwidth. When the drive current is above a certain threshold value, light generation in a laser is the result of stimulated rather than spontaneous (LED-like) emission. In this region, a typical solid-state laser generates about 5 mW of output power (per face) with a differential quantum efficiency of 30% and a response time of less than 1 ns. Bandwidth is limited more by the driving circuits than by the laser itself.

The solid-state laser emits light in a cone with half angles varying from 10° to 22° from an emitting area of micrometre size, considerably smaller than the typical fiber cross-section. Coupling efficiency, about 50%, is superior to that of a LED.

On the other hand, the threshold current of a solid-state laser is susceptible to ambient temperature and varies from device to device. To avoid destruction of the laser while maintaining stimulated emission, the drive circuit must include an optical feedback arrangement for controlling the laser current. This, added to the higher cost of the laser itself, limits the use of solid-state lasers to high-performance systems where cost is a secondary consideration.

At the Far End

The maximum length of a fiber-optic link is determined by source power, fiber attenuation, and the minimum power required by the receiver, which is a function of receiver noise only because of the fiber's immunity to external interference.

Two types of detectors are in common use. One is a simple PIN diode in which photons generate electrical carriers with quantum efficiencies of typically 80%. The subnanosecond response time and the low bias requirements make the PIN diode easy to use. The main source of noise is usually the post-detector amplifier.

The other detector is an avalanche photodiode. The avalanche process amplifies the photon-induced current about 100 times but also generates noise, so overall improvement in signal-to-noise ratio is about tenfold. The disadvantage of the avalanche diode is the high optimum bias, about 100 V, and its temperature dependence.

For illustration, let us consider a high-grade fiber-optic link operating at 50 Mb/s with a bit-error rate of 10^{-9}, which is equivalent to a peak-to-peak-signal-to-rms-noise ratio of 12:1. The required signal power would be about −44 dBm (40 nW) for a PIN-diode receiver and one-tenth of that for an avalanche-diode receiver. Assuming a laser with 5 mW (7 dBm) output power and 50% (−3 dB) coupling efficiency, the loss in the cable could be 48 dB for the PIN-diode receiver. Allowing −3 dB for splices, this represents a respectable 9-km repeater spacing with a 5-dB/km cable. With an avalanche diode receiver, repeater spacing could be 11 km. Pulse dispersion is not a limiting factor if the cable is a silica fiber with a material dispersion of 0.2 ns/km (because of the laser) and with a graded-index profile that gives a modal dispersion of 1 ns/km.

The Deterrents

So with the major technical problems solved and no significant breakthrough required, why are fiber-optic links not yet in high volume use? The reason lies in the fact that fiber-optic communication is a totally new approach, not a refinement of an existing technique where only a small part of the system would be subject to innovation. This reflects in a lack of standardization in fibers, cable, connectors, wavelengths, and power levels. Without standardization, communication is possible only between equipment of common origin.

Standardization is usually not achieved by decree. It is a natural process that occurs when one design clearly exceeds the others in cost and performance and becomes widely used because of its merits. The fiber-optic equivalents of RG58/U cable and the BNC

(continued on back page)

An NMOS Process for High-Performance LSI Circuits

Fast 16-bit microprocessors, 16K read-only memories, and a variety of special-purpose random-logic chips are the result of an NMOS process that produces high-performance large-scale integrated circuits.

by Joseph E. DeWeese and Thomas R. Ligon

THE MOST DRAMATIC IMPACT of LSI (large-scale integration) in recent years has been on computers and calculators. While the development of hand-held calculators has been highly visible to all, what has happened and continues to happen with their larger and more sophisticated cousins has been little less than profound. While the further integration of TTL and other small and medium-scale integrated circuits has in some cases resulted in much smaller machines consuming less power, more significantly it has led to much greater functional density allowing dramatic increases in performance for a given volume. Thus it is that the newer desktop calculators use more complex and powerful languages, contain more memory, and are able to solve more complex problems in a shorter time than their predecessors. Their new capabilities class them more properly as desktop computers.

Obtaining these advanced capabilities in HP desktop computers was not simply a matter of adapting available off-the-shelf LSI devices to meet objectives. This would have stretched development time since new devices must be made available before product design can be undertaken. With circuit design, LSI process development, and product design proceeding simultaneously on parallel paths, the result can be products based on LSI devices that are years ahead of anything available on the market. The products thus achieve exceptional performance/cost ratios. This is the approach used by Hewlett-Packard.

Developing an LSI process is a major undertaking (a process as used here refers to the procedure by which an integrated circuit is fabricated beginning with the bare silicon wafer). The technologies employed for forming or modifying particular parts of the structure, such as oxidation, diffusion, ion implantation, photolithography (pattern formation), and various types of thin-film deposition, are all complex subjects in themselves. What makes process development particularly challenging is the high degree of interaction that can and almost always does occur between the various steps in forming the device.

Undertaking NMOS

Previously described in these pages was a central processing unit (CPU) built on a single chip for mainframe computer systems by an HP-developed CMOS/SOS process.[1] We would now like to describe an NMOS process that was developed to meet the needs of desktop computers. This process provides the HP Model 9825A Desktop Computer with a CPU that has the speed and power of the HP Model 2100A 16-bit Computers on a hybrid circuit (Fig. 1) small enough to fit in a vest pocket.[2] Two of these hybrids are used in the new Model 9845A Desktop Computer (Fig. 2).

Why MOS? LSI circuits use MOS (metal-oxide-semiconductor) field-effect transistors more often than not because thousands of these devices can be packed on a single silicon chip of reasonable size. While bipolar devices (NPN and PNP transistors) are used in some LSI circuits, their manufacture requires several diffusion steps and the electrical isolation from the common substrate required by each device

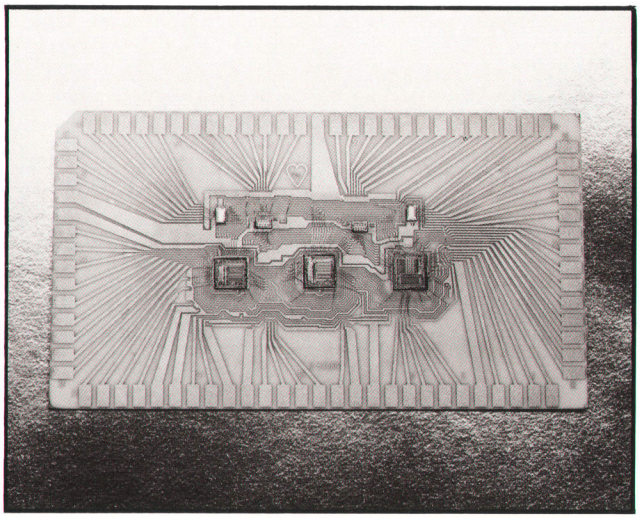


Fig. 1. *Hybrid microcircuit has a CPU chip, an input-output chip, an extended-math chip, and four bipolar interface buffers to form a 16-bit parallel processor that has the speed and power of a minicomputer.*

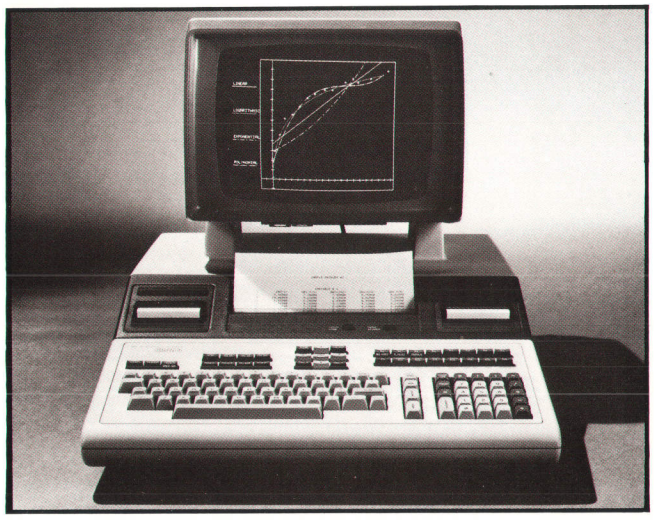


Fig. 2. *Recently-announced Model 9845A Desktop Computer uses two hybrid microprocessors similar to that shown in Fig. 1. These plus NMOS 16K ROMs mounted on 8-chip hybrid substrates give Model 9845A an extremely powerful central processor and an exceptionally large mass memory. Other NMOS LSI chips drive the thermal printhead and control the tape driver.*

has been a serious constraint on dense packing. MOS devices in most cases require only a single diffusion step and they are inherently self-isolating so they can be packed more closely.

There are powerful incentives for packing more devices on each chip, the primary incentive being an economic one. At today's prices the cost of a fully processed, 3-inch wafer might range from around $50 to $100. The wafer may have some hundreds of chips but only a fraction of that number will be sufficiently free of defects to be acceptable. The contents of each chip may range from a few hundred transistors, for simple logic or driving circuits, to well over 10,000 transistors for a microprocessor chip. The point is, the cost of chips may vary by a factor of only 10, depending on yield, process complexity, and so on, so processes that can put thousands of transistors on a single chip achieve substantial reductions in the cost per function.

Using a larger chip to allow more devices per chip may not be cost effective because the chances of a defect occurring on each chip then increases, with a consequent drop in yield. This is further compounded by the reduction in the number of chips per wafer. Given a particular defect environment, the yield Y can be related to active chip area A and defect density D (defects per unit area) by the expression:

$$Y = \left(\frac{1-e^{-AD}}{AD}\right)^2$$

The steep rise in the curves of Fig. 3, which relates device cost to yield Y, explains why the sizes of chips at the present time seldom exceed 25 to 30 mm² in even the most carefully controlled processes. Hence, major efforts are being made to make devices smaller so more can be placed on chips of limited size.

There are additional advantages to reducing device size. Scaled-down devices have less of the area-related capacitance that slows switching times. With more devices on a chip, more of the interconnections occur on the chip rather than between chips, resulting in better system speed and less power loss.

Why NMOS?

MOS devices exist in two basic types. In one, NMOS, the source and drain are doped n-type and the channel between them is p-type. A positive voltage applied to the gate causes the channel to turn n-type, allowing conduction between source and drain. In the other type, PMOS, the source and drain are p-type and the channel is n-type. A negative voltage applied to the gate then enables conduction between source and drain.

NMOS devices are faster because conduction is

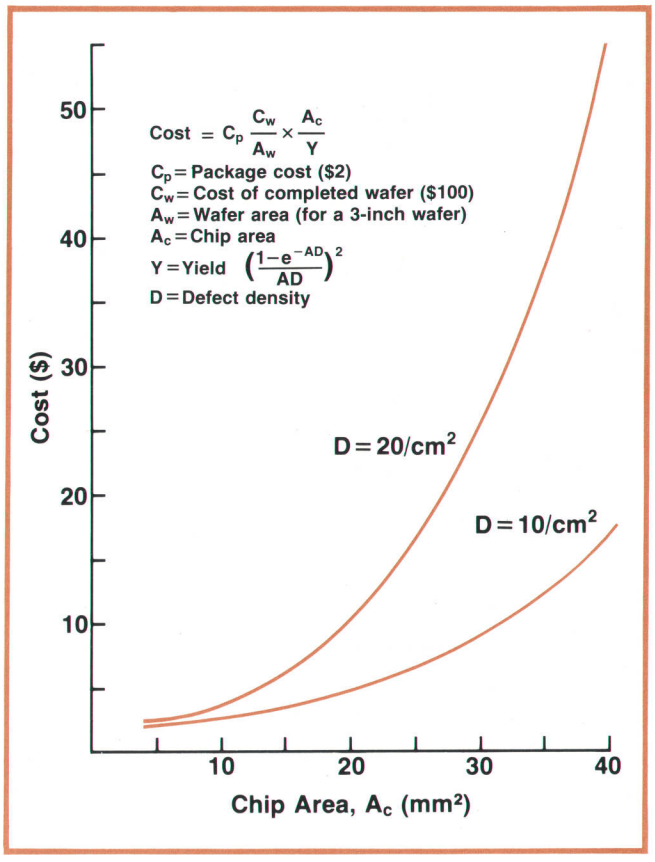


Fig. 3. *Curves plot typical device cost as a function of chip area A for two values of defect density D, showing why very large IC chips are uneconomical. These curves are for a 3-inch wafer with arbitrarily selected values of package and processed wafer costs.*

predominately by electrons whereas conduction in PMOS devices is by holes, which are less mobile. However, NMOS devices have been harder to make. This is largely because mobile ions from contaminants, principally sodium, migrate to the silicon-dioxide/silicon interfaces and tend to invert them (induce turn-on), giving rise to leakage currents.

In the late 1960's, a number of manufacturers were developing MOS memories to meet the anticipated demand that the substantial size and cost advantages of solid-state memories would generate, but none of the available devices met the particular needs of a new family of desktop computers being conceptualized at HP. Thus, the development of an NMOS process was undertaken at HP's Loveland, Colorado, facility for the purpose of manufacturing read-only memories (ROMs) for this family, the HP 9810A/20A/30A series. It was believed that the performance benefits of NMOS justified the additional process complexity.

The problem of contamination-caused leakage currents was largely reduced by the use of "back-gate" bias, a voltage applied to the substrate itself to counteract the potential caused by contaminating ions. As a result, HP was one of the first, if not the first, manufacturer to produce NMOS devices in quantity. The 4K ROMs developed in this facility enabled the 9810A/20A/30A family of desktop computers to have far greater capabilities than any similar products available at the time.

The Next Generation

The central processor in the 9810A/20A/30A machines included a number of off-the-shelf ICs occupying a good-sized printed-circuit board. Looking beyond the 9810A/20A/30A family, it was realized that if all this circuitry could be integrated on a large scale, the next generation machines could have a significant increase in processing speed and capability with a concurrent reduction in cost and size.

Using the experience gained with this first NMOS process, it was decided to develop a new NMOS LSI process that could enable integration of the central processor. The initial objectives were straightforward: twice the density of the current process with at least three times faster operating speeds. These objectives are complementary to a degree since small size means lower capacitance, the principal limit on speed.

The smallest feature size in the process at that time was 7μm. Therefore, to achieve a 50% reduction in device area, dimensions needed to be reduced to about 5 μm. This would also halve the capacitance area.

Obtaining a 30% reduction in linear dimensions was not a simple matter. As active portions of devices are brought closer together, undesired effects quickly arise. Two of these are the "short-channel" effect, in which gate-voltage control of the channel conductivity is degraded, and "punch-through", where conduction occurs between two depletion layers. Compounding these problems is the need to maintain more accurate dimensional control so the same allowable degree of dimensional error can be maintained. Tight tolerances must be obtained at each step in the process, otherwise there can be a surprisingly broad range of device parameters in finished chips, and since the designer must make allowances for the worst case, he sacrifices speed and efficiency if tolerances are too loose.

For the above reasons, and others, new and more controllable technologies had to be employed. The process that ultimately emerged, known within the company as NMOS II, makes use of the following

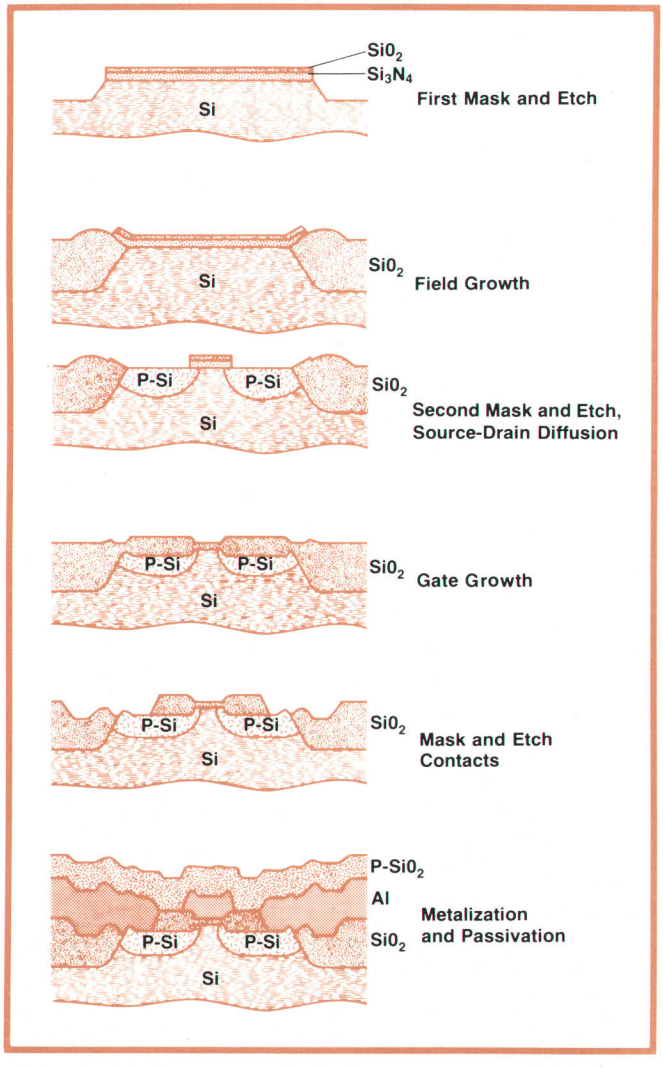


Fig. 4. *Drawings of magnified cross-sections of a silicon wafer show steps in the NMOS II process. The vertical scale is exaggerated for clarity; the depth of the phosphorous diffusion (P-Si) is actually less than 2 μm.*

technologies.

Local Oxidation of Silicon. Rather than oxidize the entire wafer and then etch away the oxide where the devices are to be formed, this technique first covers the wafer with a layer of silicon nitride and then removes it everywhere except where the devices are to be formed (Fig. 4). Silicon-dioxide is then grown on the exposed areas while the silicon nitride prevents the underlying areas from oxidizing. As shown by the cross-sectional drawing of Fig. 4, the growth of the oxide into the wafer forms partial sidewall containment for subsequent diffusions (Fig. 5), providing a margin against punch-through between adjacent diffusions, and it gives more gentle steps on the surface for metal passover. Furthermore, since the first etch is made through a thin layer of silicon nitride rather than through a relatively thick SiO2 layer, better definition of device features is achieved.

Ion implantation (I^2). This technique employs a particle accelerator to implant dopant ions into the silicon in a highly controlled manner (solid-state diffusions, such as that used for the source and drain, require temperatures around 1000°C, which can complicate a process by causing more than the desired effect). I^2 is used in NMOS II primarily for doping the gates of selected FETs so they are lightly depleted. These devices are then used as load resistors with the advantage that they are much smaller than resistors of equal value created by diffusion of a long, narrow area.

Self-aligned gates. This technique minimizes overlap of the gate structure onto the source or drain diffusion areas. Overlap results in a small gate-to-source or -drain capacitance that is magnified by the Miller effect, slowing switching speeds.

A number of techniques are used for self-alignment but a highly effective one was developed for the NMOS II process. It is based on the fact that under the proper conditions phosphorous-doped silicon oxidizes several times more rapidly than undoped silicon. After the source-drain diffusion, the protecting silicon nitride in the gate region is etched away, then the gate oxide is grown. The source-drain oxide forms simultaneously but at about five times the rate. The gate oxide is then centered perfectly between source and drain (see Fig. 4). Subsequent gate-metal placement is relatively uncritical. Furthermore, the feature size can be made smaller since less allowance is needed for aligning the gate to the source and drain.

Hard-surface masks. Until the time that the NMOS II process was being developed, the masks used to transfer patterns to the wafer were made of photoemulsion on glass. These were not well suited to the NMOS II process for two reasons: (1) the edge definition is not sharp enough for the $5 \pm 0.5\ \mu m$ line definition wanted, and (2) repeated clamping and aligning of the masks to the wafers soon damages the emulsion, leading to the generation of defects. Hard-surface masks are not so easily damaged. This was a brand-new technology at the time, so many difficulties were encountered in obtaining the required quality for the NMOS II process. Materials successfully used have been thin films of silicon and of iron oxide on glass.

Cu-Si-Al metalization. Aluminum is the standard material for making surface interconnections, one reason being that it alloys with silicon, providing good, low-resistance connections. On the other hand, with the shallow doping used in the NMOS II process, this alloying can extend through the doped silicon, shorting the contact to the substrate. Using silicon-saturated aluminum reduces this "alloy-through." The addition of copper to the alloy suppresses crystallization, smoothing the silicon-aluminum alloy. The smoothed alloy also covers steps more uniformly.

The Major Problems

The "alloy-through" was one of the most troublesome problems that had to be overcome. The boiling points of silicon and aluminum are quite different, making composition control difficult when using the evaporative technique of depositing thin-film conductors. Conventional sputtering techniques, while allowing easy composition control, are too slow and raise the temperature of the wafers to around 300°C, causing further problems with separation and crystallization of the alloy materials.

A new technique, commonly referred to as planar-magnetron sputtering, was developed. In P-M sputtering, a magnetic field in the plasma region causes electrons to follow spiral paths which results in more efficient ionization. Consequently, there is a large improvement in both deposition rate and a reduction in the substrate heating that was caused by secondary electron currents. P-M sputtering is now generally used throughout the industry.

Another troublesome area involved the problem of leakage. The introduction of phosphorous into silicon dioxide can be a good way to solve this problem because phosphorous is a rather effective "getter" for

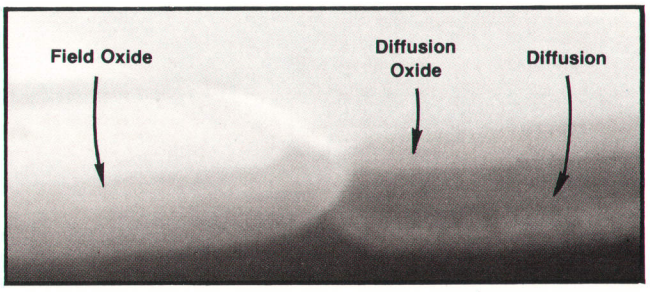


Fig. 5. *Scanning electron micrograph of NMOS II IC shows sidewall containment of diffusion (magnification: 8800×).*

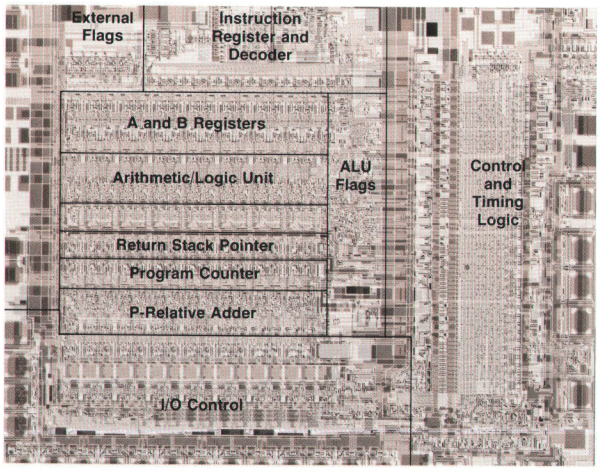
Applications of the NMOS-II Process

The first use of the NMOS II process described in the main text was for mask-programmable 16K ROMs used in the Model 9815A Desktop Computer[1] (DTC), introduced in September 1975. These ROMs are organized in a 2K × 8 structure for operation with a commercially available 8-bit microprocessor. In addition to providing four times as much firmware as the 4K ROMs used in earlier DTCs, the NMOS II ROMs decreased power consumption by turning themselves off when not being accessed by the microprocessor. This dynamic "power-pulse" operating mode decreased ROM power dissipation by a factor of greater than 10 without affecting access time significantly.

Model 9815A

The increased storage capability provided the Model 9815A by the NMOS II ROMs enabled the standard machine to have as part of its standard language many of the math functions and other options that were available only as add-on ROMs for the previous generation DTC's.

The first use of an NMOS II microprocessor was in the Model 9871A Impact Printer,[2] also introduced in 1975. Use of ROM-controlled processor logic gave the flexibility needed to improve performance at low cost. By enabling the printer to perform its internal functions itself, the processor also reduced the amount of communications needed between the DTC and printer.



Layout of BPC chip

This processor, known as the Binary Processor Chip (BPC), is a 16-bit parallel processor capable of executing 59 unique instructions. It has a memory address space of 32K 16-bit words. With NMOS II 16K ROMs organized in a 1K × 16-bit structure, the BPC forms a controlling processor that operates with a 6-MHz clock supplied by an external 2-phase, non-overlapping clock source. The processor controls the acceleration, speed, and deceleration of the motors that position the print wheel and it determines the hammer force according to the density of the character selected. Many features, such as self-test, automatic tabulation, plotting, and character substitution, were added at little cost simply by including them in the firmware.

The first totally NMOS II LSI-based desktop computer was the Model 9825A[3] introduced in January 1976. The heart of this DTC

Model 9825A

is the seven-chip hybrid microprocessor mentioned in the main text. Three of the chips are processors (a BPC, an IOC chip, and an extended math chip) that among them execute 86 instructions including control of all communications between the DTC and peripherals via the I/O port. The three processor chips are isolated from the electronics in the rest of the DTC by four bipolar interface buffers, minimizing the capacitance that the processor chips must drive. This allows the processor to operate at clock rates above 6 MHz.

Firmware is contained in 16K ROMs, organized 1K × 16. Besides saving energy with the power-pulsing (automatic self turn-off) feature, these ROMs have "three-state" (input, output, off) pins that allow multiplexing of addresses and data. The three-state arrangement reduces the overall capacitance on the parallel address-data bus by cycling non-addressed ROMs into a high impedance (low capacitance) state.

In addition to the processor and 16K ROMs, NMOS II is used in the Model 9825A in a random-logic chip that controls the operation of the keyboard, LED display, and thermal printer. This chip controls all operations of these "internal peripherals" and the data communications between them and the microprocessor. It generates key scans, allows for key debounce delay, generates repeat-key timing, and communicates key-code data to the microprocessor. It shifts and clocks data to the printer registers and controls paper advance and the burn timing of the thermal printhead. It shifts display dot character information to the LED display, controls display scan rate, and

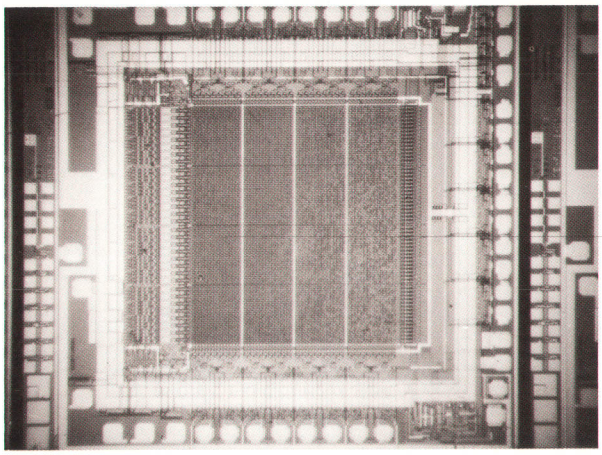
generates the display cursor. The character set used by the display and printer is contained in a mask-programmable ROM on this chip, enabling alternative character sets, such as Katakana, to be generated simply by changing photomasks during chip fabrication.

The most recent use of NMOS II LSI is in the recently announced 9845A Desktop Computer (to be described in a forthcoming issue of the HP Journal). The processor in this DTC was expanded to include two hybrid microprocessors of the type used in the Model 9825A. One processor executes the language processing while the other is dedicated to communications between the mainframe electronics and peripherals. A new BPC chip was designed for these microprocessors with four times the addressing space of the original BPC chip. The new BPC chip also makes use of NMOS power transistors for the processor's clock drivers, accepting a TTL-level clock input and generating a 12V, two-phase non-overlapping clock capable of charging 300 pF to above 10V in less than 30 ns.

The firmware memory of the 9845A has 16K ROMs on 8-chip hybrid substrates. Each of these ROMs has its own power-pulse transistor on the chip rather than using external bipolar transistors for this function as the earlier NMOS II ROMs did. Each of the memory hybrids is isolated from the mainframe electronics by bipolar interface buffers to reduce the capacitance the ROMs must drive, resulting in improved speed of operation.

Another random-logic NMOS II chip controls the operation of the minicartridge tape drive and data transfer between mainframe electronics and the tape, performing tape read-and-write control and detecting interrecord gaps. Another NMOS II chip, taking the place of a previously used bipolar chip, shifts and stores the data to the 80-character, page-width printer-plotter, directly driving the thin-film thermal printhead thin-film resistors with NMOS power transistors that are capable of sinking 150 mA with less than 6 ohms ON resistance.

Thus, it can be seen that the NMOS II process is not limited to memory and microprocessor applications but is branching out to encompass the whole array of desktop computer circuits, giving the designers of desktop computers more capability and greater flexibility in less space for enhanced performance.



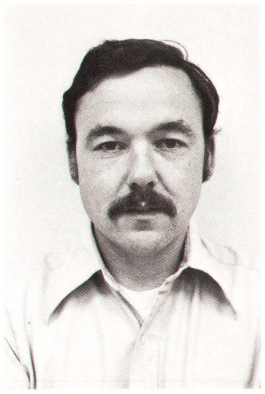
Microphotograph of NMOS II 16K ROM.

References

1. D.M. Clifford, F.T. Hickenlooper, and A.C. Mortensen, "Mid-Range Calculator Delivers More Power at Lower Cost," Hewlett-Packard Journal, June 1976.
2. R.B. Bump and G.R. Paulson, "Character Impact Printer Offers Maximum Printing Flexibility," Hewlett-Packard Journal, June 1976.
3. D.E. Morris, C.J. Christopher, G.W. Chance, and D.B. Barney, "Third Generation Programmable Calculator Has Computer-Like Capabilities," Hewlett-Packard Journal, June 1976.

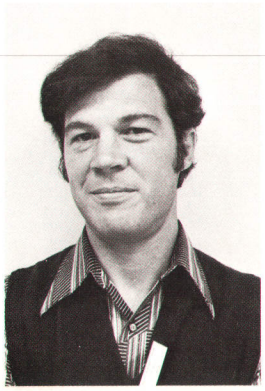
sodium ions. However, the procedures used to achieve other desired features, such as the self-aligned gate, did not allow the easy introduction of sufficient phosphorous into the field oxide. The solution was to dope an oxide layer with phosphorous while the layer is deposited chemically at 400°C as the final passivation layer at the end of the process.

With the more sharply defined features, the shallower dopings, the self-alignment technique, and the use of depletion load transistors, plus refinements in circuit design, circuit speed improved by a factor of 10 with the factor-of-2 area reduction. During the two and a half years the NMOS II process has been on line, it has proved to be relatively high yielding. Besides providing the hardware basis for the 9815A/25A/45A desktop computer family and its many high-performance peripherals, it is also finding its way into other kinds of products, such as the new Model



Joseph E. DeWeese

Joe DeWeese has devoted five years to the NMOS-II process, beginning in 1972 with development of prototype process equipment and going on to development of products, production equipment and process monitors. The past two years, he was project manager for NMOS II LSI product support. He is now a project manager in the R and D Lab. A native of Clinton, Kentucky, Joe earned a BSEE degree at the University of Illinois and is working towards an MSEE degree at Colorado State University. In off hours, he does some woodworking and he also participates in the HP Golf League and intramural basketball. He and his wife have one small boy and a baby girl.



Thomas R. Ligon

After obtaining BS and MS degrees in physics from Oklahoma State University in 1966, Tom Ligon joined Hewlett-Packard, working initially on the development of precision high-frequency wirewound resistors, then the 17-layer printed-circuit ROM used in the Model 9100A Calculator, and finally integrated circuit development. He left HP in 1970 to pursue a private business venture but returned two years later to work on the NMOS-II process development, becoming project manager in 1973. A native of Wewoka, Oklahoma, Tom, his wife and three children live on a 1½ acre tract in Loveland, Colorado, raising fruits, vegetables, chickens, rabbits, and an occasional calf or pig.

8568A Spectrum Analyzer.

Acknowledgments

Major contributors to the NMOS II process were Larry Hall, Jim Mikkelson, Dana Seccombe, and Mark Lundstrom. In the LSI circuit development effort, Gene Zellmer and Howard Abraham were early notables while IC engineering manager Tom Haswell guided the overall effort. Many other people, too numerous to mention here, contributed significantly to the success of the process and products.

References

1. B.E. Forbes, "Silicon-on-Sapphire Technology Produces High-Speed Single-Chip Processor," Hewlett-Packard Journal, April 1977.
2. W.D. Eads and D.S. Maitland, "High-Performance NMOS LSI Processor," Hewlett-Packard Journal, June 1976.

(continued from page 25)

connector have not yet arrived.

Standardization is also a prerequisite of high-volume production. Because of low-volume developmental manufacture, the present costs of fiber, cable, solid-state lasers and LEDs are an order of magnitude higher than that considered economically feasible. It will take one or two years yet before the "gain" in the volume-cost-demand "loop" becomes great enough to create a steep increase in production and a drop in cost, as occurred with integrated circuits.

Then there is the usual reluctance of reliability-minded engineers to abandon highly developed, well proven systems for a totally new technology, with the active light source in the transmitter being particularly suspect. Material defects appear to be the major limiting factor on the life of LEDs and lasers. The specific failure mechanisms are being studied in a number of laboratories throughout the world with assurance of many solutions. Recent reports from Bell Laboratories cite an extrapolated life of up to 10^6 hours for a solid-state laser, corresponding to an uninterrupted service of about 100 years. Even a 10-year life span, more commonly achieved today, should be quite sufficient for reliable operation.

The Future

No doubt about it, wide-scale use of optical communications is coming, not only for telephone transmissions, cable TV, and distributed computer networks, but also for such short-haul applications as pc-board-to-pc-board signal transfer within an instrument. Most transmitters and receivers now are designed for specific bit rates and link lengths. Large-scale integration, and possibly trade-offs in receiver sensitivity and maximum bit rates, will result in compact circuits that will be able to accommodate a wide range of link lengths and bit rates without requiring any adjustments. This will facilitate the economical application of fiber-optics to a wide range of applications.

In the immediate future, we can see further developments in splices and connectors. To avoid undue loss of power, the faces of the adjoining fibers must be of equal size, in line, parallel, and as close to each other as possible. This leads to strict diameter control in fiber production and, given the small diameter of the fiber, requires a precision in the micrometre range for the connectors and splices. At present, the optical power loss in a typical connector is 0.5 to 2 dB, equivalent to a substantial length of low-loss fiber cable. The use of connectors or splices in a system therefore must be carefully planned, quite a different situation than that encountered with communications over metallic wire.

Farther in the future we anticipate developments in single-mode fibers and also in integrated optics, the optical equivalent of semiconductor ICs. These will consist of solid-state components, usually based on thin-film technology, in which optical rather than electrical signals are switched, amplified, modulated, and processed directly without conversion to electrical signals. The impact of this technique will be far-reaching.

Tom Hornak is Manager of the Applications and Engineering Department within the Solid-State Lab of HP's central research laboratories where he is responsible for R and D on LSI circuits and optoelectronic systems. He holds a Dipl. Ing. degree from the Slovak Technical University and a PhD degree from Czech Technical University and worked on radar, instruments and memories before joining Hewlett-Packard in 1968. He has published over two dozen papers on electronics and holds a like number of patents.

Hewlett-Packard Company, 1501 Page Mill Road, Palo Alto, California 94304

HEWLETT-PACKARD JOURNAL

NOVEMBER 1977 Volume 29 • Number 3

Technical information from the Laboratories of Hewlett-Packard Company

Hewlett-Packard Central Mailing Department
Van Heuven Goedhartlaan 121
Amstelveen-1134 The Netherlands
Yokogawa-Hewlett-Packard Ltd., Shibuya-Ku
Tokyo 151 Japan

Editorial Director • Howard L. Roberts
Managing Editor • Richard P. Dolan
Art Director, Photographer • Arvid A. Danielson
Illustrator • Susan E. Wright
Administrative Services, Typography • Anne S. LoPresti
European Production Manager • Dick Leeksma

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

3250392 0TH157HARRAAA 134

MR JULIAN A HARRIS JR
CHAYO ELECTRONICS LTD
COMMUNICATIONS DEPT
1575 NO 20TH AVE
PENSACOLA FL 32503

CHANGE OF ADDRESS: To change your address or delete your name from our mailing list please send us your old address label (it peels off). Send changes to Hewlett-Packard Journal, 1501 Page Mill Road, Palo Alto, California 94304 U.S.A. Allow 60 days.