HEWLETT-PACKARD JOURNAL

The Logic State Analyzer—Displaying Complex Digital Processes in Understandable Form

Sixteen consecutive 12-bit words from any part of a data stream are captured and held for display by a new Logic State Analyzer, greatly simplifying the task of monitoring program flow when troubleshooting complex digital processors.

by William A. Farnbach

A S LOW-COST MSI AND LSI components make feasible the use of more sophisticated digital processors in a broadening variety of applications, there is a growing need to be able to perceive what is happening in a digital process. With the increasing power of digital systems that is developing, the amount of data that must be understood to debug or check out a digital machine increases at a dismaying rate.

It is no longer practical to start troubleshooting by asking electrical questions such as, "Do we have too much clock skew somewhere? Is there a transient on a reset bus? Could one of the switches have bounced too long?" There are just too many possibilities to eliminate them one at a time. The problem must first be isolated by answering functional questions, as, "Which loop is it stuck in? Which way does it branch at 036_8? How did it get it to 'halt'?"

Recently described in these pages was a new instrument (Model 5000A Logic Analyzer) that can help answer these questions by capturing and displaying digital data in serial form, like data circulating in a calculator, data transmitted between remote terminals and a computer, or serial data stored on discs.[1] Described here is a new instrument (Fig. 1) that captures and displays digital data in parallel form, such as the output of a random-access memory or the states of a counter. The tabular display of consecutive words provides an overall view of system operation that makes it far easier to trace information flow than would be possible with an oscilloscope or other traditional test instrument.

Tabular Display

The new instrument, called the Model 1601L Logic State Analyzer, displays 16 consecutive 12-bit words as "1's" and "0's" in 16 rows, just as logic data or truth tables are presented in textbooks. The instru-

ment has 12 data inputs using a new probe system that can connect all 12 inputs to closely-spaced pins on digital IC's (Fig. 2). The contents of the displayed words are determined by the placement of the corresponding probes. If all 12 inputs are not to be used, a column blanking control allows the unused data channels to be blanked from the display.

An additional probe input, the "clock" input, is provided for externally generated pulses that are used to strobe the data lines. Each time a pulse edge occurs at this input, the values present on the data lines are strobed into the Analyzer's memory. Clock pulses from the system under test at any repetition

Fig. 1. *Model 1601L Logic State Analyzer system displays program flow in a digital processor in an easy-to-comprehend tabular form, greatly easing the task of finding troubles in a digital system. The Model 182C Oscilloscope mainframe supplied with the system is a standard unit that can be used with other plug-ins. Conversely, the Logic State Analyzer plug-in can be used with other mainframes in the 180 family.*

rate up to 10 MHz are usually applied to this input, but flags, data valid, subcycle counter, or any other signals may also be used.

The words displayed are referenced to a preset trigger word. The Analyzer can be set to display the trigger word and the 15 words following, or it can display the trigger word and the 15 words that preceded it. The trigger word can also start a digital delay generator so the instrument captures and displays the 16 words that follow a preset number of clock cycles. This makes it possible to page through a long digital sequence, as long as 100,000 words.

The trigger word, which can be brightened to show its position in the data stream, is selected with a bank of 12 front-panel switches that correspond to the 12 data inputs. The switches have three positions: HI, OFF and LO. When analyzing positive logic, HI corresponds to a "1" and LO to "0". With negative logic, HI is "0" and LO is "1". OFF means "don't care."

Successive States

An example of how the instrument is used is shown in Fig. 3. This shows the state flow in a BCD counter. The probes corresponding to the four least significant digits are attached to the Q0, Q1, Q2, and Q3 outputs of a type 9310 counter, the corresponding trigger switches are all set to "0" (the others are set to OFF), and the column-blanking control is set to blank the eight most significant digits. The display is then set to start with the trigger word, which is shown at the top of the display.

3
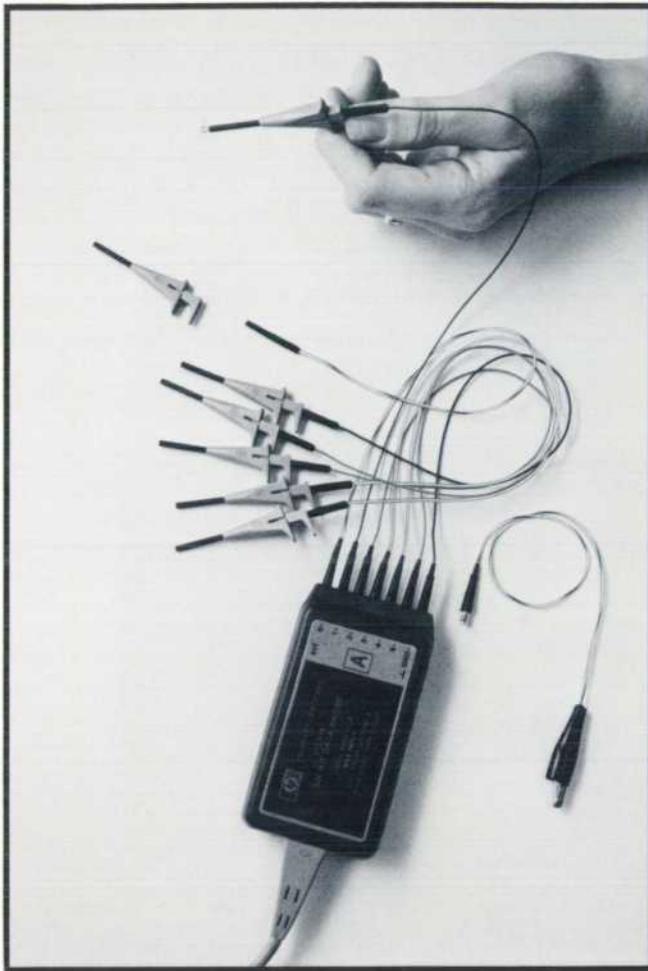
Fig. 2. *New probing system enables probes to be attached to adjacent pins on standard DIP IC packages that have as little as 0.02 inch (0.5 mm) clearance between the package and the leads. "Grabbers" can be removed, leaving a 0.025 inch (0.6 mm) socket for inserting test clips or wire-wrap pins. Twelve signal probes and two ground probes are gathered into two groups of seven to ease handling.*

Fig. 3. *State flow in a BCD counter. The trigger word at the top is brightened indicating here that the CRT displays the trigger and the 15 states that follow the trigger. The numerical scales, silk-screened on the back side of the contrast filter, assist in identifying states.*

One of the most useful characteristics of the Logic State Analyzer is its ability to pick out any sequence in a microprocessor, capture the sequence, and retain it for display. Fig. 4 shows the turn-on sequence in the HP Model 34703A Multimeter[2]. The trigger word set on the Logic State Analyzer is the initial check test state, 167_8, in the turn-on sequence. When the Multimeter is turned on, the Logic State Analyzer shows that the Multimeter's digital controller follows the correct state flow down to the wait loop, 163_8-122_8, where it waits for a step command.

Grouping the Digits

The "1's" and "0's" on the display are easy to read and interpret. To enhance display readability, the displayed bits are arranged vertically in groups, with a minor space between pairs of words and a major space after every fourth word. A front-panel pushbut-ton allows a choice of horizontal groupings, into columns of 3, for convenience in reading octal words, or into columns of 4 for BCD or hexadecimal words (Fig. 5). Another pushbutton inverts the displayed patterns, replacing the "1's" with "0's" and the "0's" with "1's" for display of negative-true logic.

The CRT display is also helpful in identifying erratic states. These appear as "1's" superimposed on "0's". An erratic state that exists only 5% of the time can be identified this way.

Operating Modes

Data flows in digital machines are seldom repetitive—they must be captured on the first pass or lost. The Logic State Analyzer can operate in a single-sample mode, in which a captured sequence is displayed continuously until the RESET pushbutton is pressed, initiating a new acquisition cycle. This mode is especially useful in the END DISPLAY trigger mode, where the instrument displays the trigger word and the 15 words that precede it, a useful technique for finding out how a processor arrived at a particular state. The instrument also operates in a repetitive mode where it updates the display whenever trigger requirements are met. In addition there is a free-run mode where it ignores the trigger switches and initiates a data-acquisition cycle each time a display cycle is completed. This mode is useful for double-checking connections to the circuit under test—active probes properly connected will show superimposed ones and zeros; those not making good contact will show either ones or zeros steadily.

Input Logic Requirements

The instrument samples the input data on either
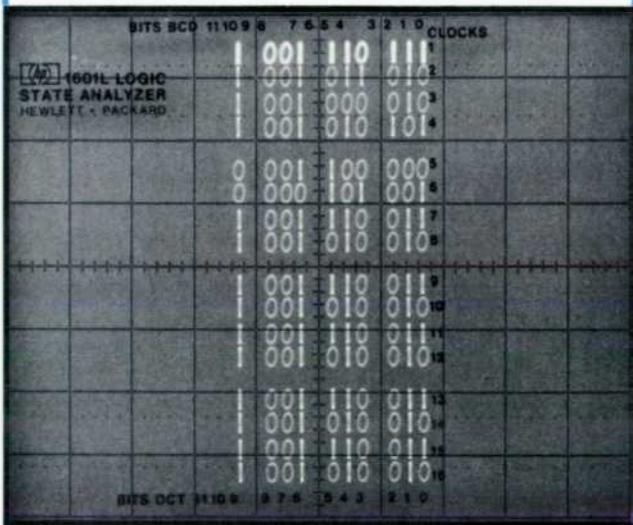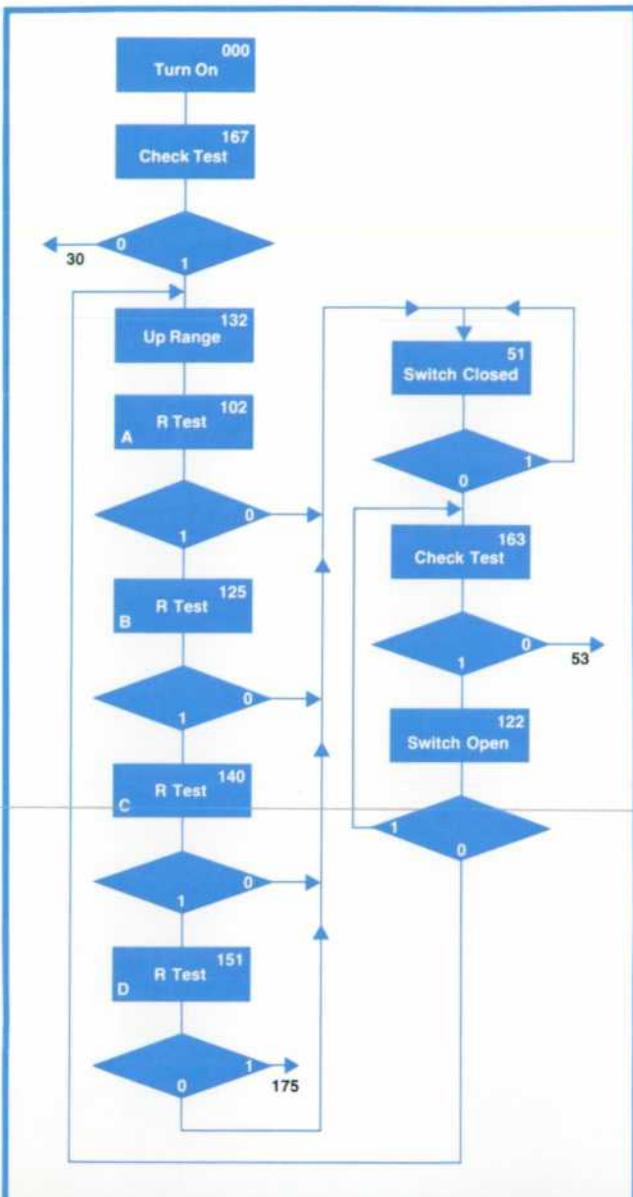
Fig. 4. *Program flow in the turn-on sequence of a multimeter is shown by the diagram and the Logic State Analyzer display.*

the positive-going or negative-going edges of the clock input, as selected by a front-panel pushbutton. The threshold level can be preset to +1.45V, the usual TTL threshold level, or it can be varied over a range of −10 to +10V (both clock and data input thresholds are set by the same control).

As an aid for the operator when no display occurs, monitor lights (NO CLOCK) indicate whether the clock waveform is above the threshold level (HI), or below (LO), or within ±0.2V of the threshold (both HI and LO on). This arrangement also allows the clock input to be used as a logic probe for single-point measurements.

If proper clock signals are being received, as indicated by both NO CLOCK lights being extinguished, but the trigger word does not occur, then a NO TRIG indicator lights up.

The Logic State Analyzer also has two outputs. One is a repetititon of the clock input, buffered out at TTL levels but with 40 ns delay. The other is a TTL pulse that occurs whenever the selected trigger word occurs. This is useful for triggering an oscilloscope, for example, on the exact digital state desired. Fig. 5 shows how this is used. This shows a BCD counter that is counting incorrectly, resetting on state 89 instead of 99. The Logic State Analyzer is set to trigger on state 88, which triggers the oscilloscope. Examination of the master reset bus with the oscilloscope shows a transient that occurs at the beginning of state 90.

By using a counter to measure the ratio of trigger outputs to clock outputs, it is possible to determine the percentage of time a microprocessor spends in subroutines or the randomness of state occurrences in pseudorandom number generators.

The trigger output can also be used to make other measurements in digital processes. For example, to find out how long a calculator-plotter algorithm spends in a 4-state wait loop, the trigger word is set to one of the wait loop states. In one pass through the algorithm, the counter totals 1538, meaning 1538 passes through the loop or 6150 clock pulses (2 clocks are dropped when the loop exits).

Data Processing

Processing of incoming signals begins in the two pods to which the probes are attached (six probes with each pod). Here, the incoming voltage levels are compared to the instrument's threshold level to classify them as HI's or LO's. Doing this in the pod close to the point of measurement holds the capacitance that each incoming signal must drive to less than 14 pF (input resistance is approximately 40 kΩ).

The incoming clock signal is quantized the same way but in addition it is fed through a slope selector

5
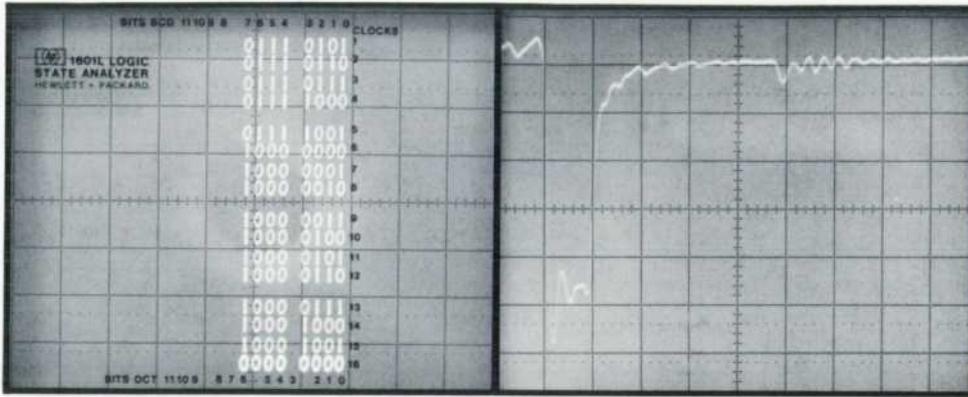
Fig. 5. *Trigger word here is set to end the display, showing that the two-decade BCD counter resets on 89 rather than 99. Resetting the Logic State Analyzer to trigger on 88, and using the trigger output to start an oscilloscope (right), discloses a "glitch" on the reset bus on state 90.*

to an RS latch, a Schottky TTL line driver that generates fast clock transitions that are uniform regardless of the incoming signal rise time. The slope selector (Fig. 6) selects the clock edge that triggers the RS latch.

The outputs of the data probes are transferred to the instrument where they are clocked into a set of 12 D-type TTL flip-flops for temporary storage. The relative delays of the clock and data paths are adjusted so the input has zero "hold time," the time that the data must not change after the triggering clock edge occurs. "Set-up time," the time that the data must be stable before the clock edge occurs, is 35ns, mostly a result of variations in delay between one data path and another. With the provisions for polarity inversion and threshold selection that this instrument has, it can thus be made to respond to digital signals almost exactly like the digital circuits it monitors.

The outputs of the temporary storage flip-flops are passed on to the 16 × 12 memory and to the trigger word comparator, where they are compared to the front-panel trigger-word switch settings.
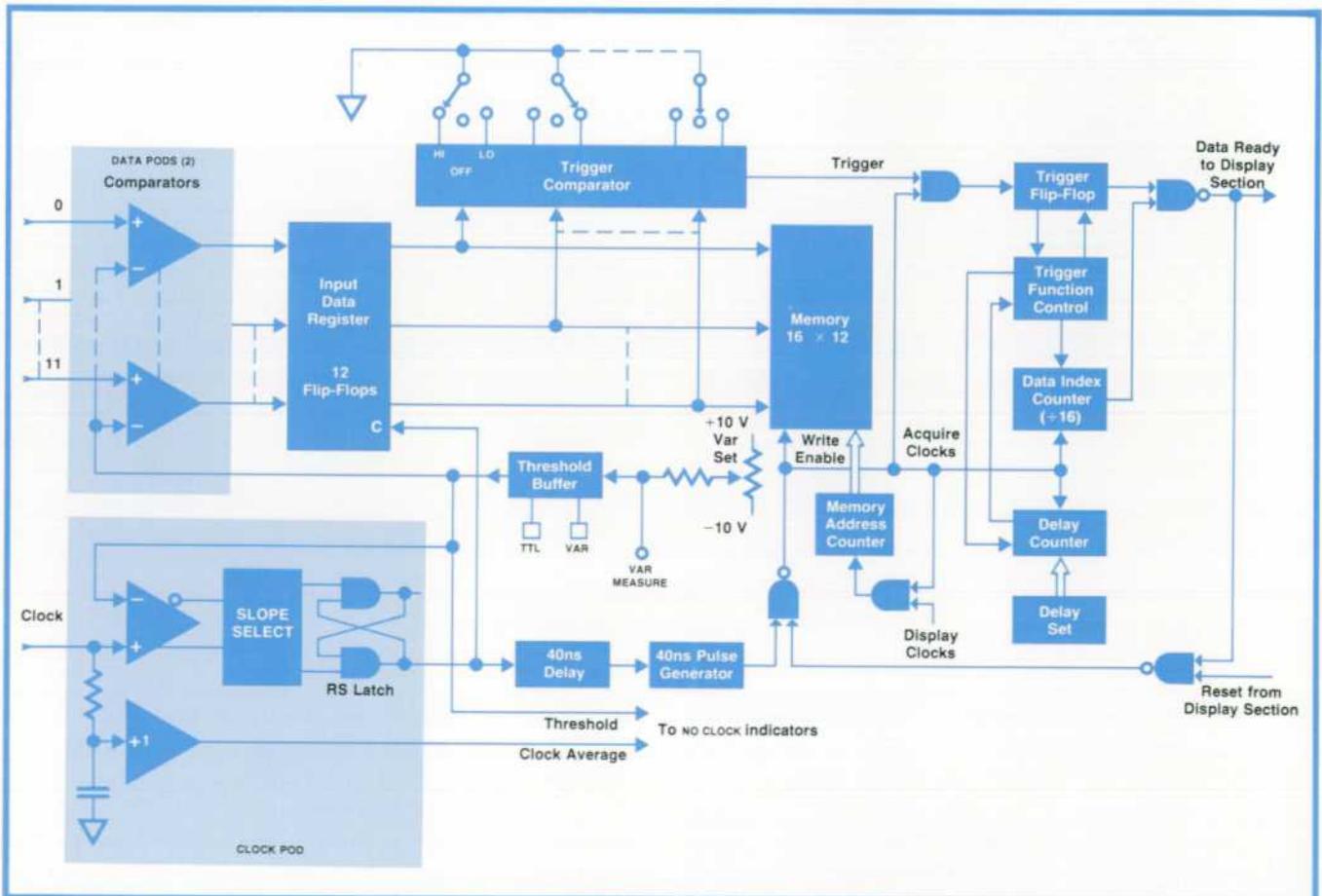
Fig. 6. *Block diagram of the Logic State Analyzers' data acquisition circuits.*

6

The input clock signal is delayed 40ns by a delay multivibrator and then formed into a standard 40-ns pulse by a second multivibrator. The delayed clock is used as the write-enable pulse for the memory and also as the clock signal for data acquisition control.

Acquisition Cycle

Writing of new data into the memory is controlled by the acquisition control circuit. Control consists mainly of stopping data acquisition once the proper data has been stored, the exact operation depending on which trigger mode is selected.

In the start-display mode, data is written constantly into memory. On each clock cycle, the data currently in temporary storage replaces the oldest data word in memory. In other words, the memory acts as twelve parallel, 16-bit, shift registers. This is accomplished by deriving the memory address from a 16-bit counter that is incremented by each input clock.

When a match between incoming data and the trigger word occurs, the trigger comparator sets the trigger flip-flop and enables the data index counter. This occurs on the trailing edge of the write-enable pulse.

Data acquisition continues until the data index counter reaches a count of 15. This turns on the "data ready" line to the display section and gates off the clock pulses in the data acquisition section. Acquisition thus stops, and memory control is transferred to the display section. The memory now contains the trigger word and the 15 data words that immediately followed.

The start-delay trigger mode is essentially the same, except that the trigger flip-flop enables the delay counter. When the delay counter totals the preset number of clock pulses, it enables the data index counter. Data acquisition continues for 15 more clock cycles then stops. The memory now contains the word that occurred on conclusion of the preset number of clock cycles after the trigger word, and the 15 words that followed.

In the end display mode, the sequencing of the trigger flip-flop and the data index counter is reversed. The data index counter begins counting clock cycles as soon as the acquisition cycle begins. When it reaches a count of 15, it is disabled but the trigger flip-flop is enabled. This prevents a trigger from being accepted until the 16th clock pulse after acquisition begins, thus ensuring that the memory does not retain data from the previous acquisition cycle.

Although the data index counter is disabled, the memory continues to accept data until a trigger word occurs. Occurrence of a trigger word sets the trigger flip-flop, stopping acquisition at once since the data counter has already reached full count. Because the trigger flip-flop is set on the trailing edge of the write-enable pulse, the memory now contains the trigger

word and the fifteen words that preceded it.

In the free-run mode, the trigger flip-flop remains enabled at all times so the acquisition cycle stops as soon as 16 words are loaded into memory.

Generating the Display

Once data has been acquired in any of these modes, memory control is transferred to the display section (Fig. 7).

This section has two counters: the 12-state horizontal counter and the 16-state vertical counter. These select the bit in memory to be displayed and position the CRT beam by way of digital-to-analog converters.

The count in the horizontal counter is applied to a one-of-twelve selector that selects the bit in a particular word in memory that is to be displayed. Incremented by an internally-generated 50-kHz clock, the counter steps through all its 12 states before resetting. While resetting, it increments the memory address counter and the vertical counter.

Since the vertical counter steps in synchronism with the memory address counter, it is not involved directly in word selection. The outputs of the four flip-flops in the vertical counter are weighted and added to obtain an analog voltage for positioning the CRT beam vertically. Weightings were chosen to space the words in groups of four with subgroups of two to enhance readability. The relative weightings are:

A	(least significant bit)	1
B		2.1
C		5
D	(most significant bit)	10

To enable the vertical columns to be arranged in groups of either three (OCT) or four (BCD), as selected by the front-panel BYTE switch, the output of the horizontal counter addresses a read-only memory. This selects a 4-bit word corresponding to the position of the bit to be displayed. This word is applied to a digital-to-analog converter to derive the CRT beam horizontal deflection voltage.

Characters are generated by superimposing 100-kHz sinewaves on the deflection voltages. The sinewave to the horizontal deflection system is shifted about 90° by a capacitor and reduced slightly in amplitude so an ellipse is written. To trace a "1", the sinewave is switched off the horizontal deflection by the 1-of-12 horizontal selector. If the polarity of the display is to be inverted, the switch turns off the sinewave for a LO, rather than switching off for a HI.

The 50-kHz clock pulses are derived by dividing down the 100-kHz signal. During the first half of each clock pulse, the CRT is blanked while the CRT beam is being moved to the new position. During the second half, the beam is unblanked and the character is written.
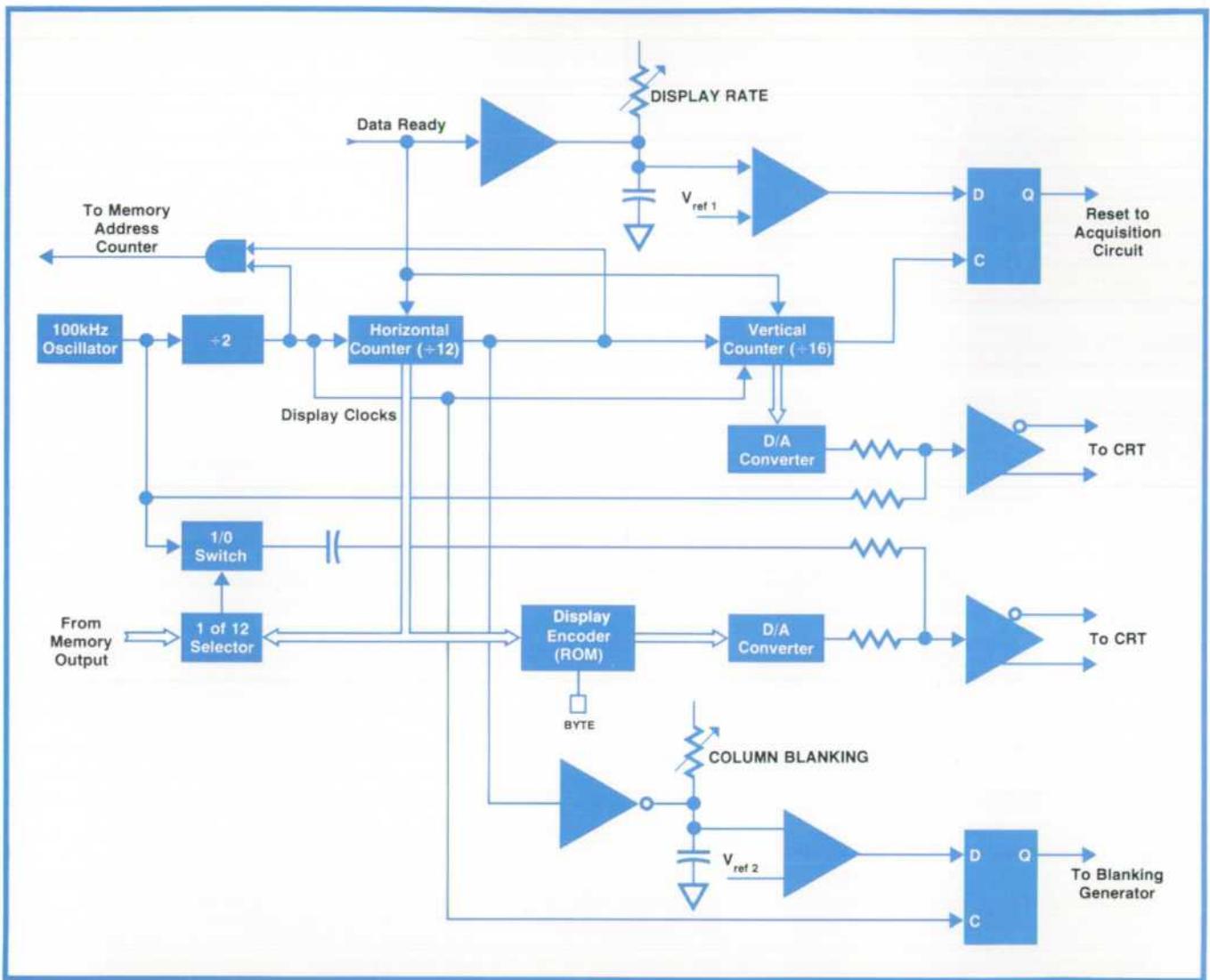
Fig. 7. Block diagram of the display circuits.

Display Cycles

When the instrument is operated in the single cycle mode, the vertical and horizontal counters are reset at the end of a display, and another display cycle is initiated. This continues until the RESET button is pressed, at which time memory control is transferred immediately to the acquisition control circuit, initiating a new data acquisition cycle. The TRIGGER OUTPUT, however, outputs a pulse whenever trigger requirements are met, regardless of what the instrument is doing.

In the repetitive mode, display cycles repeat for a time determined by the front-panel DISPLAY RATE control. This control, which has a range of 40ms to 5s, adjusts the slope of the display rate ramp. At the end of each display, this ramp is compared to a reference voltage and if it exceeds this voltage, memory control is transferred to the acquisition control circuit.

Column blanking is controlled by a voltage ramp that is compared to a reference. The result of the comparison is applied to the D input of a flip-flop clocked by the 50-kHz internal clock. The output of the flip-flop, which goes to the blanking circuit, will not change until the ramp exceeds the reference.

Characters are displayed as long as the ramp is below the reference level (characters are written from right to left on the display). When the ramp exceeds the reference, the character on the next clock pulse and those that follow will be blanked. Blanking digits rather than resetting the horizontal counter to shorten the displayed word ensures that changing the blanking control does not affect CRT intensity.

The ramp is reset at the time the horizontal counter resets. The input to the column-blanking flip-flop is thus below the threshold at the start of the next row so the first digit cannot be blanked. This prevents the display from being completely blanked when the COLUMN BLANKING control is at its extreme position.
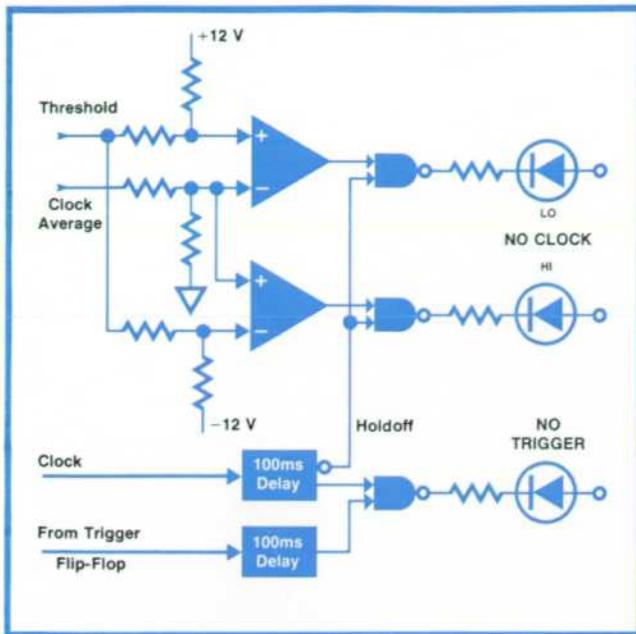
Fig. 8. *NO CLOCK and NO TRIGGER indicators are inter-locked so the NO TRIGGER indicator can turn on only when clock pulses are present but the trigger word does not occur.*

Operator Alerts

The NO CLOCK and NO TRIGGER indicators (Fig. 8) are controlled by the data acquisition control circuit. These are held off by a monostable multivibrator as long as clock pulses set the multivibrator. Anytime that clock pulses are absent for more than one-tenth second, the multivibrator resets, enabling the indicator gates.

The other inputs to the gates come from voltage comparators that compare the low-pass filtered clock input to the threshold voltage. The HI indicator can light up if the average level of the clock signal is above a level that is about 0.2V below the threshold. The LO indicator can light up if the clock average is below about 0.2V above the threshold. If the clock average is within ±0.2V of the threshold, both indicators will light.

The NO TRIGGER indicator is controlled by the trigger flip-flop. Each time the trigger flip-flop is set, it sets a monostable multivibrator. The multivibrator resets 0.1 s after the last trigger pulse, enabling the NO TRIG indicator through a gate. The other input to the gate is from the NO CLOCK delay multivibrator so in the absence of clock pulses, the NO TRIG indicator is inhibited.

William Farnbach
Bill Farnbach joined Hewlett-Packard in 1967 upon getting his MSEE degree from Brigham Young University, where he had also earned his BSEE degree. Initially he worked on the 1424A and 1425A Sampling Time Bases and then the 1817A 12.4 GHz sampling head for the 1815A TDR system. Prior to taking on the 1601L program, he was project leader for the 1810A Sampling Oscilloscope plug-in. In his spare time, he enjoys fishing or a good fast game of racket ball.

References

1. R. Adler, M. Baker, and H. Marshall, "The Logic Analyzer: A New Instrument for Observing Logic Signals," Hewlett-Packard Journal, October 1973.
2. R. Gardner, A. Dumont, and S. Venzki, "A Greater Range of Capabilities for the Compact Plugin Digital Multimeter," Hewlett-Packard Journal, March 1973.

SPECIFICATIONS
HP Model 1601L Logic State Analyzer
(Includes 1601A LSA plug-in, 182A mainframe, clock and data probes)

Clock and Data Inputs
REPETITION RATE: 0 to 10 MHz.
INPUT IMPEDANCE: 40 ±3k ohms shunted by <14 pF.
INPUT BIAS CURRENT: <30 µA.
INPUT THRESHOLD: TTL, fixed at approx +1.5 Vdc; variable, ±10 Vdc.
MINIMUM INPUT SWING: 0.4V +10% of threshold V p-p.
MINIMUM CLOCK PULSE WIDTH: 25 ns.
MINIMUM SETUP TIME: time data must be present prior to clock transition, 35 ns.

MINIMUM HOLD TIME: time data must be present after clock transition, 0.

Display Rate
VARIABLE: from <40 ms to >5s.

Clock and Trigger Outputs
HIGH: >2V into 50 ohms (line driver interface).
LOW: <0.4V into 50 ohms (line driver interface).
PULSE DURATION: approx 40 ns (RZ format).

General
WEIGHT: 31.75 lb (14.4 kg).
POWER: 115V or 230V ±10%, 48 to 440 Hz, 200 VA max.
DIMENSIONS: approx 8 in W × 13 in H × 19 in D (20 × 34 × 49 cm).
PRICE IN U.S.A.: Model 1601L Logic State Analyzer, $2650.
    Model 1601A Logic State Analyzer Plug-in with probes but without 182A mainframe (can use any 180-series mainframe as display), $1700.
MANUFACTURING DIVISION: COLORADO SPRINGS DIVISION
    1900 Garden of the Gods Road
    Colorado Springs, Colorado 80907

A Laser Interferometer That Measures Straightness of Travel

This unique "straightedge" combines the high accuracy of the optical flat with the long range of the laser to achieve submicron resolution over ranges up to 100 feet.

by Richard R. Baldwin, Barbara E. Grote, and David A. Harland

IT IS BECOMING increasingly important in the machine tool industry to determine the precise geometry as well as the positioning accuracy of machine slides. The problem is particularly acute for multi-axis machine tools, where the method most often used for single-axis machines, the Abbé Comparator Principle[1], cannot be followed. Present methods allow accurate measurement of machine slide positioning, but can measure geometry only in terms of angular quantities such as pitch and yaw. The only reliable tools available for determining straightness of travel have been the short straightedge and the optical flat.* These devices are adequate in cases where

* A glass plate carefully lapped and polished for maximum flatness.

the total machine slide travel is two feet or less, but it is not feasible to produce physical standards to accommodate travel longer than this.

Traditionally, attempts have been made to overcome this problem by resorting to an optical or mechanical substitute such as a mobile fluid, a stretched wire, an alignment telescope, or an alignment laser. Unfortunately, such attempts have proved much less accurate than the short straightedge or optical flat, which have remained the most accurate straightness standards for the last century.

The Straightness Interferometer

The straightness interferometer, as shown in Fig. 1,
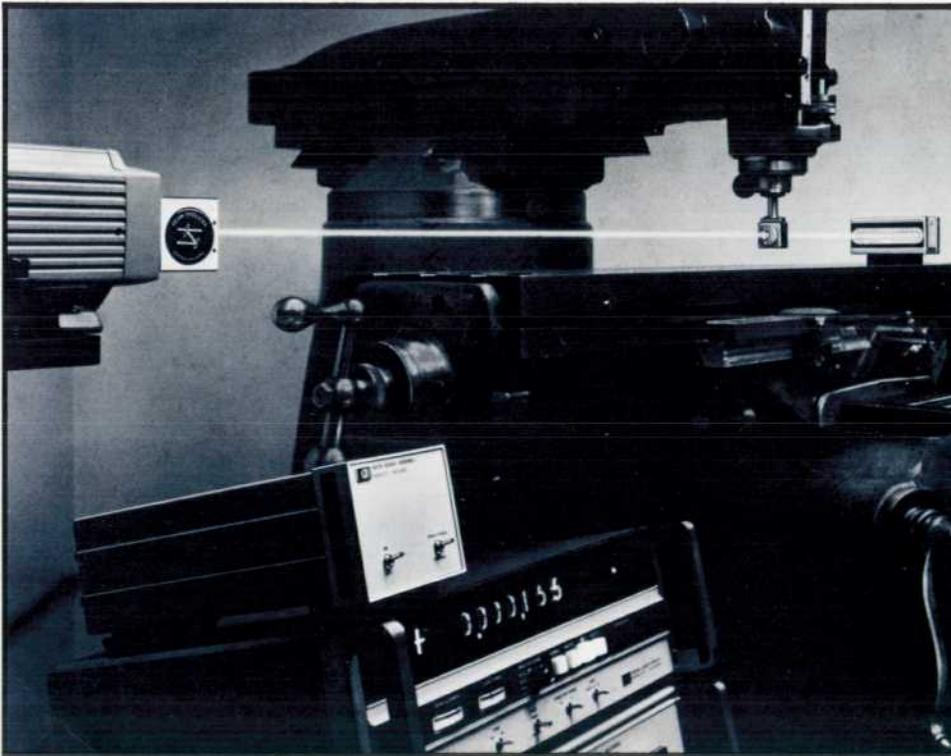
Fig. 1. *The straightness interferometer, an optional accessory to the 5526A Laser Interferometer System, measures straightness, parallelism, and perpendicularity of machine-tool and measuring-machine coordinate motions over a range of approximately 100 feet or 30 meters. Basic accuracy is five microinches per foot of travel.*

provides a more direct answer to the straightness problem: use small optical flats to measure straightness over a long travel. Using this approach the straightness interferometer has achieved an accuracy of five microinches per foot of travel for travel up to one hundred feet, without requiring state-of-the-art optical flats. The interferometer is presently available in two versions: a short-range instrument that has one-microinch resolution and measures straightness over ten feet of travel, and a long-range instrument that has ten-microinch resolution and measures straightness of travel over one hundred feet of travel. Both have inherent accuracy of five microinches per foot.

Besides high accuracy, the straightness interferometer shares another advantage with the optical flat. The accuracy of both devices can be improved by a technique known as "reversing the straightedge" in which a second straightness measurement is taken with the reference straightedge reversed so that it faces in a direction opposite that of the first measurement. By averaging the two measurements, the error of the interferometer, or optical flat, can be reduced to the limit of repeatability of the device under test.

Principle of Operation

The principle of operation of the straightness interferometer is shown in Fig. 2. The interferometer consists of a Wollaston prism[2] which acts as a beam-splitter, and a reflector comprised of two plane mirrors that are cemented together as a rigid assembly precisely matching the angle of divergence of the Wollaston prism. The reflector is mounted on a fixed surface with its axis aligned with the coordinate axis whose straightness is being measured. The prism is mounted on the slide that moves along the measured axis.

If the Wollaston prism is illuminated by light from a two-frequency laser[3] containing frequency components f_1 and f_2, polarized as shown in Fig. 2, then the frequency components of the laser beam will be separated into two beams which diverge at a small angle ϕ. This is because the composite refractive index

Fig. 3. *Optical path difference in the Wollaston prism for f_1 and f_2 is $P_2 - P_1 + 2z\sin\phi/2$, where z is the distance from the bottom of the prism and P_2 and P_1 are optical path lengths for z = 0. A displacement E of the prism in the z direction causes the path difference to change by $2E\sin\phi/2$. This change is measured by the interferometer.*

of the prism is different for the two planes of polarization of the two frequency components. The diverging beams are then reflected from the plane mirror assembly back through the Wollaston prism, which recombines them into a single beam.

During recombination, the two beams interfere within the Wollaston prism so the single beam returning to the laser contains relative phase information proportional to the optical path difference of the two interfering beams. Any motion of the Wollaston prism in a direction parallel to the angular bisector of the plane mirror assembly will not cause any change in the optical path difference of the two interfering beams, since both beams will be lengthened or shortened by the same amount. However, if the Wollaston prism moves in a direction perpendicular to the angular bisector of the plane mirror assembly an optical path difference will be generated within the Wollaston prism.

The reason for this change in optical path length can be seen in Fig. 3. If we consider a plane wavefront of one frequency, say f_1, to be normally incident on the Wollaston prism from the left, it will leave the prism at an angle $\phi/2$ as shown in Fig. 3a. It follows

Fig. 2. *The straightness interferometer causes the two differently polarized frequency components in the laser beam to diverge. After reflection from two precisely plane mirrors, the two components recombine and interfere in the interferometer. Any transverse motion of the interferometer relative to the mirror bisector causes an optical path change that is detected and displayed. The system is insensitive to motion along the mirror bisector and to rotation of the laser.*

from the principle of equal optical paths[4] that the total optical path from any point on the incoming wavefront A-A′ along the direction of a particular ray to the wavefront B-B′ must be equal to the total optical path from A-A′ to B-B′ for any other ray. However, because the outgoing wavefront is inclined to the prism by the angle $\phi/2$, the optical path through the prism is not constant, but is equal to $P_1 - z\sin\phi/2$, where P_1 is the optical path length through the prism for the lowest ray of frequency f_1, z is the height of the incident ray of frequency f_1 whose path length is to be determined, and $\phi/2$ is the angle of inclination of the emerging wavefront B-B′.

Similarly, as shown in Fig. 3b, the optical path length through the prism for any ray of frequency f_2 would be equal to $P_2 + z\sin\phi/2$, where P_2 is the optical path length through the prism for the lowest ray of frequency f_2, z is the height of the incident ray of frequency f_2 whose path length is to be determined, and $\phi/2$ is the angle of inclination of the emerging wavefront. Therefore, the difference in optical path within the Wollaston prism for the two frequency components f_1 and f_2 is $(P_1 - P_2) + 2z\sin\phi/2$ for any ray entering the prism at height z.

Now if the prism is displaced in the z direction by a distance E, the optical path difference between the two frequency components will change by an amount $2E\sin\phi/2$. In other words, if the Wollaston prism does not move along the angular bisector of the plane mirror reflector in a straight line, out-of-straightness will be measured as an optical path difference of $2E\sin\phi/2$ where E is the error in straightness and ϕ is the divergence angle of the Wollaston prism.

It is important to notice that the two plane mirrors shown in Fig. 2 are aligned parallel to the emerging wavefronts shown in Fig. 3. Therefore the optical path to the mirrors from A-A′ in Fig. 3 is the same for all rays of a given frequency and the interferometer is not sensitive to motion of the laser beam along the z direction provided the Wollaston prism remains fixed with respect to the mirrors.

The interferometer measures any *relative* motion between the angular bisector of the plane mirror reflector and the Wollaston prism. It doesn't matter whether the Wollaston prism moves, the bisector moves, or both move. The mirror bisector axis in effect represents the straightedge along which the interferometric straightness measurement is made. It is *not* the laser beam that is used as the straightness reference.

The difference between the long-range and short-range versions of the straightness interferometer lies in the Wollaston prism and the plane mirror reflector. Externally, the two versions are the same; however, the beams exit from the long-range interferometer at one-tenth the divergence angle of the short-range version, and the reflector angle is changed accordingly.

How Straight Is It?

The straightness of the interferometer axis is a direct function of the difference in flatness of the two plane mirrors that serve as the interferometer reflector. For example, suppose one of the mirrors is flat but the other is convex by a small amount δ as shown in Fig. 4. Even though the Wollaston prism, or straightness interferometer, were to travel in a straight path with respect to the mirror axis, the optical path length of f_2 would be shortened by an amount equal to δ with respect to the optical path length of f_1, and this would be interpreted by the interferometer as "out of straightness" of the interferometer travel.

The straightness interferometer therefore relies on the flatness of the reflector mirrors as its straightness reference. Since the mirrors are in fact small optical flats, the straightness interferometer represents the application of a pair of small optical flats to the measurement of straightness over a large travel as mentioned previously. Fortunately, extremely accurate optical flats are available. Laboratory-grade flats are typically accurate to within one microinch per foot. The straightness interferometer can tolerate devia-
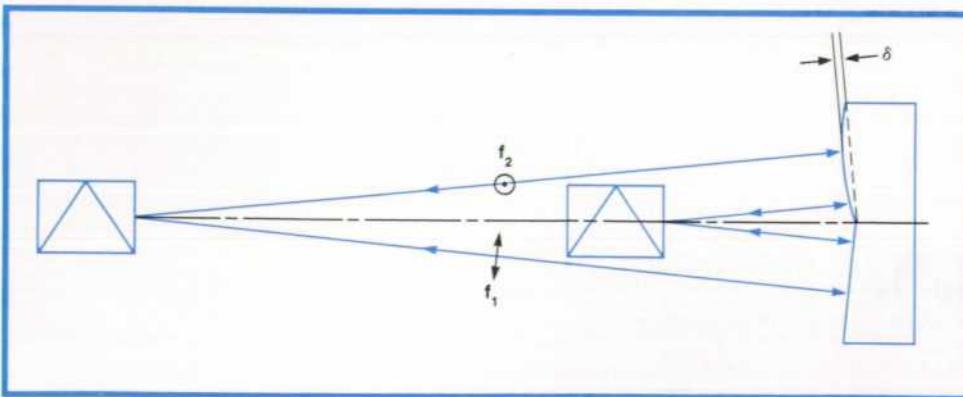
Fig. 4. *Straightness interferometer accuracy is a function of the flatness of the plane mirrors, which are small optical flats. These optical flats–not the laser beam–are the ultimate straightness reference for the system. The system can tolerate a modest flatness error (δ) of about a microinch over each two-inch mirror before the straightness error exceeds five microinches per foot of travel.*

Machine Tool Calibration with the Expanded Laser Measurement System

The straightness interferometer is the latest in a series of re-mote interferometer developments for the HP 5526A Laser Measurement System.[1,2] It adds an important class of measurements—those of the geometry of a coordinate system—to the capabilities of the system.

Up to now the HP Laser Interferometer has been able to make accurate measurements of linear position plus two super-imposed angular motions, pitch and yaw. But the straightness of the coordinate axes and the relationships between them has been assessable only with non-interferometric instruments which, in general, have limited range, resolution, or linearity.

The alignment laser is better than other non-interferometric in-struments, but accuracy and resolution claims of a few tenths of thousandths of an inch are difficult to substantiate as a result of its extreme sensitivity to air turbulence. This effect is more se-vere for alignment lasers than for the HP straightness interfero-meter because the alignment laser depends on the spatial sta-bility of the light beam whereas the straightness interferometer is sensitive only to differential changes in optical path length. In practice, alignment lasers are limited to resolutions of about 0.001 inch. This limits their suitability for certification purposes to situations where specifications are given only to the nearest 0.01 inch.

Other drawbacks of non-interferometric instruments for mea-suring geometric errors are the length of time required to make a calibration, dependence on operator skill and, in most cases, no digital output to facilitate on-line digital data reduction and analysis.

Measures Squareness and Parallelism

The linear, angular, and straightness interferometers are capable of evaluating five out of a possible six degrees of freedom of a rigid body, roll being the exception. Fig. A illus-trates the six degrees of freedom.
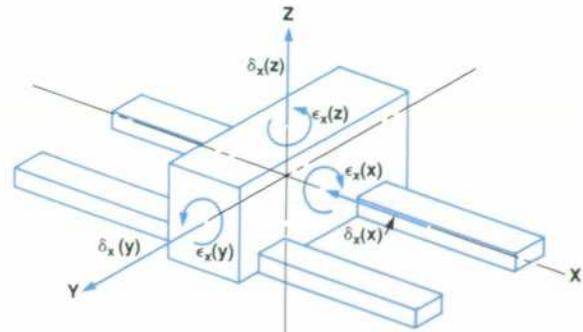
Fig. A. *The six degrees of freedom of a machine tool table.*

All machine tools and measuring machines depend on the movement of a tool or probe along at least one axis. Al-though the importance of each of the six components of mo-tion varies with the type of machine, all such devices re-quire six measurements to define their performance com-pletely.

Many machines have two or more coordinate axes, and this introduces the need to maintain geometric relation-ships. In the case of boring, milling, drilling and other simi-lar machines there are two or three orthogonally related axes. For machines where the workpiece rotates (the single largest category of machine tools) parallelism between tool and spindle axes is of fundamental importance.

The straightness interferometer is able to make both per-pendicularity and parallelism measurements. This is done by comparing two consecutive straightness calibrations of adjacent axes, introducing either a 90-degree or a 180-de-gree rotation into the straightness reflector axis between traverses.
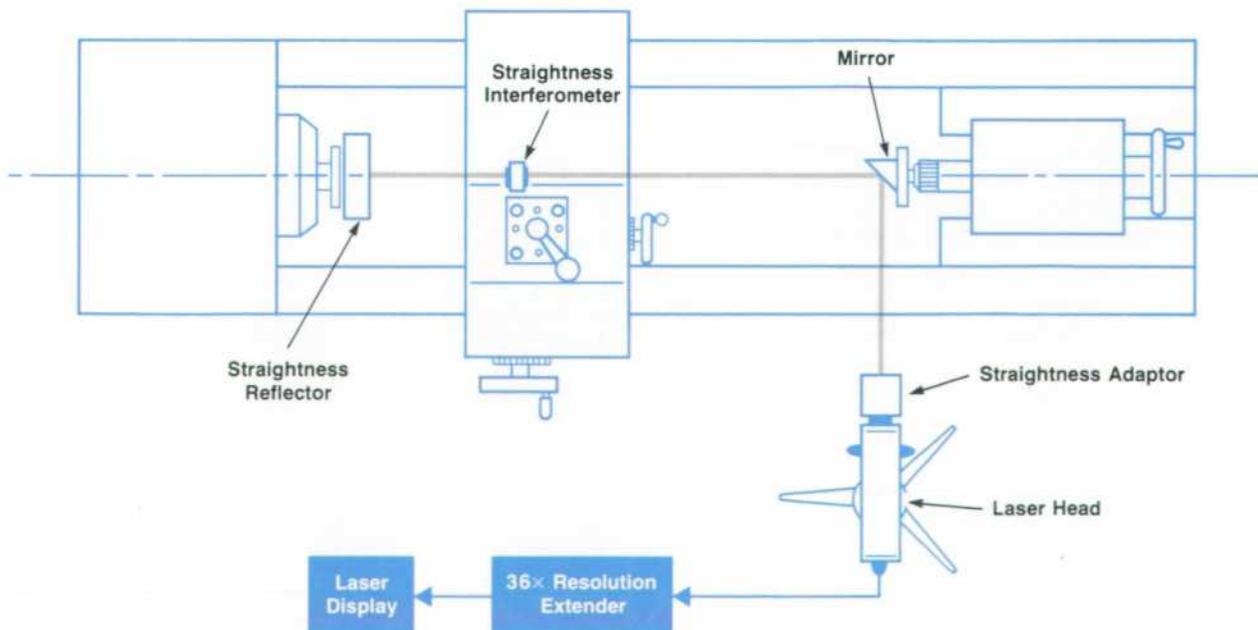
Fig. B. *Plan view of setup for straightness and parallelism calibration of a lathe.*

13

For perpendicularity a separate right-angle reference is required. It can either be a pentaprism, which causes the extended axis of the straightness reflector to be viewed with a 90° deflection, or a precision indexing table, which physically rotates the reflector assembly after the first axis has been measured.

For parallelism no separate device is requred. Rotation of the spindle/reflector assembly about the spindle axis itself defines the spindle axis. Fig. B shows how this principle is applied to lathe calibration.In this case the straightness reflector is chucked in the spindle with its long side in the horizontal plane. The reference plane,defined as the bisector of the angle between the two reference mirrors, is projected vertically. The straightness interferometer is fixed in the toolpost. Deviations in a horizontal plane about the bed axis create optical path differences in the straightness interferometer that are registered in the laser diplay as microinches of out-of-straightness.

An advantage of the straightness interferometer over conventional straightness references is immediately apparent. A physical straightedge or mandrel has mass and elasticity. Sagging therefore limits its useful length in cantilevered fixtures. The straightness reflector, on the other hand, measures only 4.5 by 2 by 2 inches whether the range of straightness measurement is six inches or 100 feet.

After a satisfactory definition of the relative travel between the tool axis and the reflector axis has been recorded, the spindle/reflector assembly is carefully rotated through 180°. By making a second traverse with respect to this redefined reference plane, the out-of-parallelism, α, can be determined by:

$$\alpha = (\alpha_1 + \alpha_2)/2$$

where

α_1 = angle between mean axis of travel of carriage and axis of straightness reflector.

α_2 = angle between mean axis of travel of carriage and axis of straightness reflector after 180° rotation.

Fig. C shows a typical plot.

References

1. J.N. Dukes and G.B. Gordon, "A Two-Hundred-Foot Yardstick with Graduations Every Microinch," Hewlett-Packard Journal, August 1970.
2. R.R. Baldwin, G.B. Gordon, and A.F. Rudé, "Remote Laser Interferometry," Hewlett-Packard Journal, December 1971.
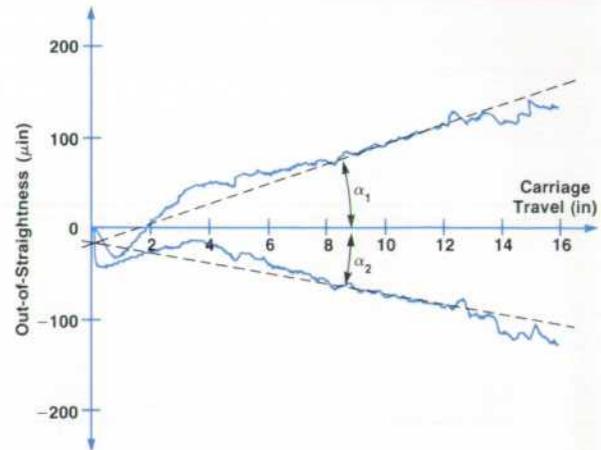
Fig. C. *Typical lathe carriage straightness and parallelism calibration.*

tions from flatness as high as five or six microinches per foot (the actual tolerance is one microinch over each two-inch mirror).

Air Turbulence

The principal factor that limits the accuracy of the straightness interferometer is air turbulence. Air turbulence is caused by local variations in the index of refraction of air. These variations are almost always the result of localized temperature differences, although occasionally a similar phenomenon can be caused by variations in the gas content. Air turbulence causes random variations in the optical path lengths of the two interfering laser beams, which in turn cause a random fluctuation in the interferometer reading.

Although it is difficult to predict the amount of air turbulence present in a given environment, there are ways to minimize its effects. First, the interferometer is affected by air turbulence only in the region between the Wollaston prism and the plane mirrors. Therefore the closer the interferometer is to the mirrors the smaller the air turbulence effects will be. Some "typical" values of air turbulence effects versus interferometer-mirror separation are shown in Fig. 5.

Air turbulence can also be minimized by eliminat-

ing heat sources in the vicinity of the measurement whenever possible. If a heat source cannot be removed during a measurement it is helpful to direct
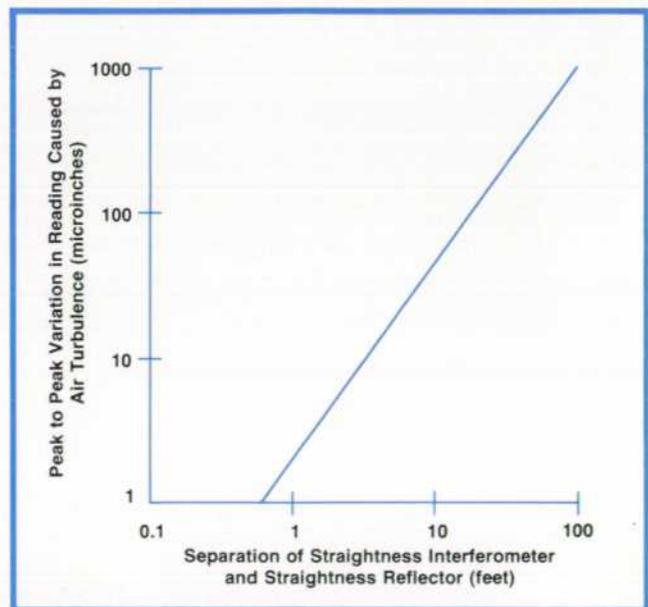
Fig. 5. *Air turbulence limits straightness interferometer accuracy by causing random fluctuations in the display. These can be reduced by placing the reflector closer to the interferometer.*

hot air in its vicinity away from the measurement path by means of a fan. A fan, however, is not as effective as the elimination of heat sources and is not recommended except in cases where extremely poor air circulation is causing thermal gradients, or where the effects of heat sources cannot be removed by any other means.

Because of their random nature, air turbulence effects can also be reduced by averaging a number of measurements. If the uncertainty in a particular straightness measurement is N microinches this uncertainty can be reduced to $N/\sqrt{n}$ microinches if n measurements are averaged.

Other Error Sources

Whenever a new method of making physical measurements is devised, two questions always arise: "How well does the method measure what it is supposed to?" and "Does it try to measure anything it is not supposed to?" Up to now we have been concerned with the first question.

To answer the second question a detailed ray tracing program was run on an HP 2100A Computer. This program imposed undesired translational and rotational motions on different parts of the interferometer system to determine what error would result in the interferometer reading. The ray tracing program was quite comprehensive, involving eighteen independent perturbations for both short-range and long-range interferometer systems. A summary is given on page 17. In this article only those motions which caused a significant effect in the interferometer reading are included.

As predicted earlier, translational motions of the laser beam in any direction had no effect on the interferometer reading. Also, relative translation along any axis perpendicular to the z direction in Fig. 3 between the Wollaston prism and the plane mirror reflector had no effect.

It was expected that such translational motion would have no effect other than that caused by imperfections in the optical geometry of the Wollaston prism or plane mirror reflector. The primary concern was that the interferometer would be sensitive to rotational motions of the laser, Wollaston prism, or plane mirror reflector, particularly if such motion occurred around an axis perpendicular to the plane of the drawing of Fig. 2. If the interferometer is to behave as a proper straightedge, it is imperative that a common point exist somewhere within the system about which small rotations of any element in the system cause no effect on the interferometer reading. Also, it is imperative that the interferometer be insensitive to small rotations about any axis passing through that point. If such a point can be found, then any compound motion of any element in the system can be re-

solved into three translations and three rotations about that point as an axis.

Fortunately such a point does exist, and is labeled N in Fig. 2. It lies on the centerline of the Wollaston prism, about one-quarter inch from the exit face, and it is the straightness of travel of this point with respect to the plane mirror axis that is measured by the interferometer.

Small rotations of the Wollaston prism, the plane mirror reflector, and the laser about an axis perpendicular to the plane of Fig. 2 and extending through this point cause the effects shown in Figs. 6, 7, and 8, respectively. Rotation about this axis has been arbitrarily called "pitch" and the point N of Fig. 2 has been called the "nodal point" of the interferometer system. Since the effects shown in Figs. 6, 7, and 8 are small, they should not cause significant error during a straightness measurement, because large angular pitch is invariably accompanied by large out-of-straightness.

The effects of laser and prism yaw (where "yaw" is defined as rotation about an axis perpendicular to the mirror bisector, passing through point N and lying in the plane of Fig. 2) are shown in Figs. 9 and 10 respectively. The effects shown in Figs. 7, 8, and 9 are in fact smaller than the limit of resolution of the interferometer. Attempts to measure these effects experimentally were unsuccessful, because the interferometer showed no effect at all. For this reason experimental data has been included only in Figs. 6 and 10. All data presented is for the short-range straight-
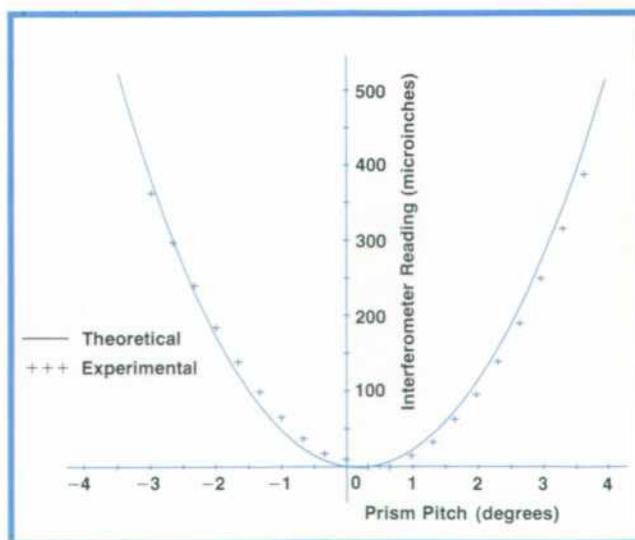
Fig. 6. *Effect on straightness reading caused by pitch of the short-range straightness interferometer (Wollaston prism) about the nodal point (N in Fig. 2). Pitch was plotted over a wide angular range to demonstrate the agreement between theory and experiment. Pitch does not cause significant errors, because large angular pitch is always accompanied by large out of straightness.*
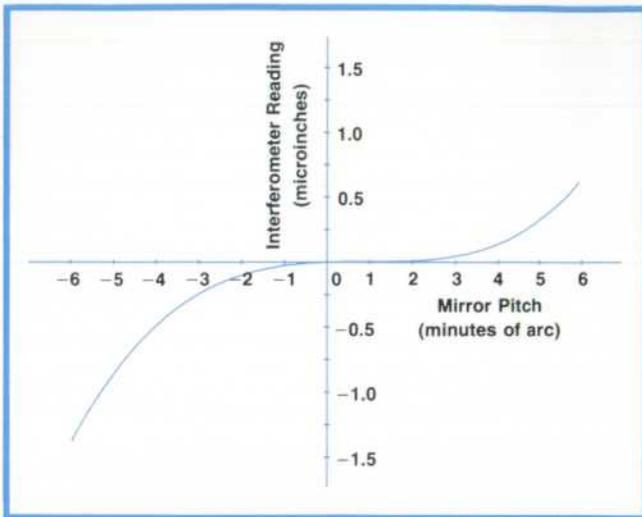
Fig. 7. Effect on reading caused by pitch of the short-range reflector assembly about the nodal point.

Fig. 9. Effect of laser yaw on reading of short-range version. Effects shown in Figs. 7, 8, and 9 were too small to measure experimentally. Thus experimental data is shown only in Figs. 6 and 10.

ness interferometer. The effects caused on the long-range interferometer are smaller, and are less likely to cause error because of the coarser resolution of the long-range instrument.

It should be mentioned that because the interferometer is insensitive to any translational motion of the laser beam, and because rotation of the laser about any point can be resolved into a translation and a rotation about point N, rotational effects of the laser will be minimal regardless of the axis about which it is rotated.

A New Kind of Remote Interferometer

The straightness interferometer is an accessory to the 5526A Laser Interferometer System. Being insensitive to laser motions, it can be remote from the laser, and its addition to the 5526A allows the mea-

surement of five out of six degrees of freedom of a machine tool slide (see page 13).

The straightness interferometer is adapted to the 5526A as shown in Fig. 11. A plane-mirror/beamsplitter assembly called a straightness adaptor directs light from the straightness interferometer to the photodetector, and a 36× resolution extender enables the 5526A display to read straightness errors directly and in the correct units. Since the straightness interferometer measures an out of straightness error E as an optical path change $2E\sin\phi/2$, the angle ϕ was chosen such that $2\sin\phi/2 = 1/36$ for the short-range interferometer and 1/360 for the long-

Fig. 8. Effect of laser pitch on reading of short-range version.
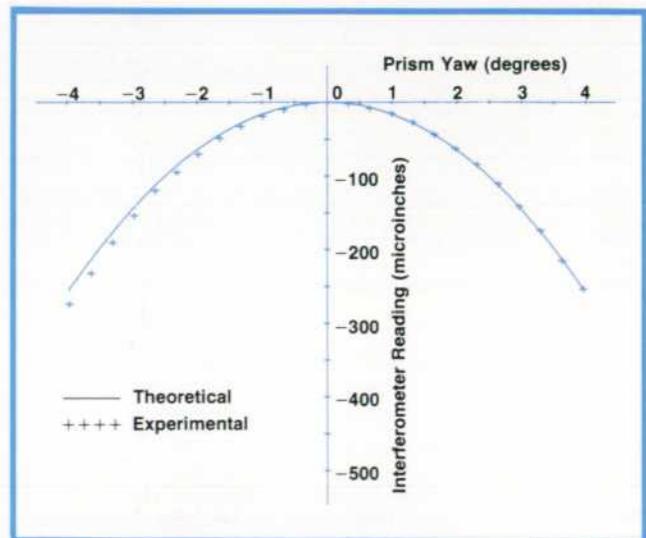
Fig. 10. Effect of prism yaw on reading of short-range version. Effects shown in Figs. 6 through 10 are less significant for the long-range version of the Interferometer.

16

Ray-Trace Program

As part of the Straightness Interferometer development a ray-trace program was written to predict the behavior of three-dimensional optical systems in six degrees of freedom with accuracy greater than 1 part in 10^{10}. The program, which runs on an HP 2100A Computer, can be used to analyze flat-surfaced optic systems of isotropic or birefringent materials. It offers an alternative to traditional pencil-and-paper ray tracing and empirical analysis.

Operation of the ray-trace program is very simple. The user models his system by writing a straightforward FORTRAN IV program, calling upon a library of three-dimensional vector mathematics subroutines. The optical components, their relationships to each other, and the specific surfaces involved in ray-tracing are specified by the programmer. He usually defines some of the parameters as variables whose values will be set later during program execution.

The ray-trace program performs three tasks. First, it creates a set of (x,y,z) points to represent each optical element geometrically. Then it translates or rotates a part of the system. Each optical component can be moved in six degrees of freedom—x,y,z, yaw, pitch, or roll—around a user-defined point. Finally, the program analyzes each system configuration.

The program ray-traces from one point to another through the optical system, following Snell's Law extended to three dimensions. Point P_n on the nth surface can be found given the two previously-traced points P_{n-1} and P_{n-2} on the $(n-1)$st and $(n-2)$nd surfaces. Referring to Fig. A, P_n is the intersection of:

1) the plane formed by the incident ray $P_{n-2} P_{n-1}$ and the surface normal at the $(n-1)$st surface,
2) the object plane at the nth surface,
3) the cone with vertex P_{n-1} formed by Snell's Law at the $(n-1)$st surface.

The quadratic nature of the equations generally results in two solutions for P_n. The program chooses the appropriate one for refraction or reflection. Whenever optical path lengths are calculated, it is sufficient to trace the central ray for a complete description of system behavior, because all rays in a wavefront have equal optical path lengths.
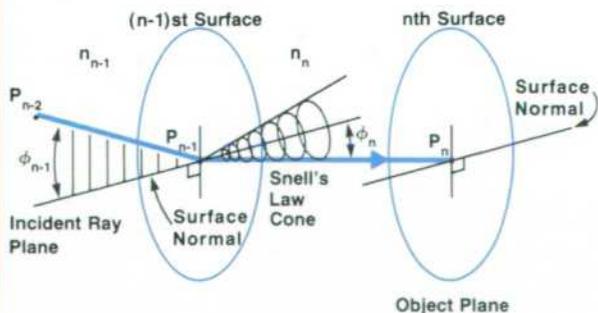
Fig. A. *The ray-trace algorithm follows Snell's Law: If n_{n-1}, ϕ_{n-1} and n_n, ϕ_n are the index of refraction and the angle for the incident and refracted ray, respectively, then $n_{n-1}\sin\phi_{n-1} = n_n\sin\phi_n$.*

Birefringent Materials

Birefringent materials exhibit double refraction. Wave propagation is described by two principal indices of refraction, n_o and n_e, and an optic axis OA. Two types of waves exist, called ordinary, or o-rays, and extraordinary, or e-rays. Wavefronts polarized perpendicular to the optic axis are o-rays and propagate with the same velocity $v = v_o$ in all direc-

tions, behaving as though they were in an isotropic material. Their index of refraction n_o, defined as c/v, where c is the speed of light in a vacuum, is constant. Wavefronts polarized parallel to the optic axis are e-rays and propagate in an ellipsoid of revolution about the optic axis following velocity v_e', which varies from v_o to v_e as a function of the angle between the ray and the optic axis. The instantaneous index of refraction n_e' varies between n_o and n_e in an elliptical manner corresponding to v_e'.

One feature of this program is its ability to approximate n_e' for birefringent three-dimensional refraction at some surface $n-1$ where $n_n = n_e'$, and P_{n-2}, P_{n-1}, n_{n-1}, ϕ_{n-1}, ϕ_n, n_e, n_o and the optic axis in the material between the $(n-1)$st and nth surfaces are known. Snell's Law is assumed to hold with the direction of $\vec{n_e}$ defined as the direction of $\vec{P_{n-1} P_n}$. n_e' is a composite function of itself. Referring to Fig. B, n_e' is a function of angle β and is the simultaneous solution to three equations: the projection of n_e' onto the optic axis, the equation of a circle with radius n_e' coplanar with $\vec{P_{n-1} P_n}$ and OA, and the coplanar ellipse described by n_e and n_o. Angle β is a function of point P_n by simple geometry, and point P_n is a function of $n_n = n_e'$ by the ray-trace algorithm. Thus $n_e' = f(n_e')$. Using iterative mathematical methods, n_e' can be found and birefringent refraction can be performed for uniaxial materials such as quartz.
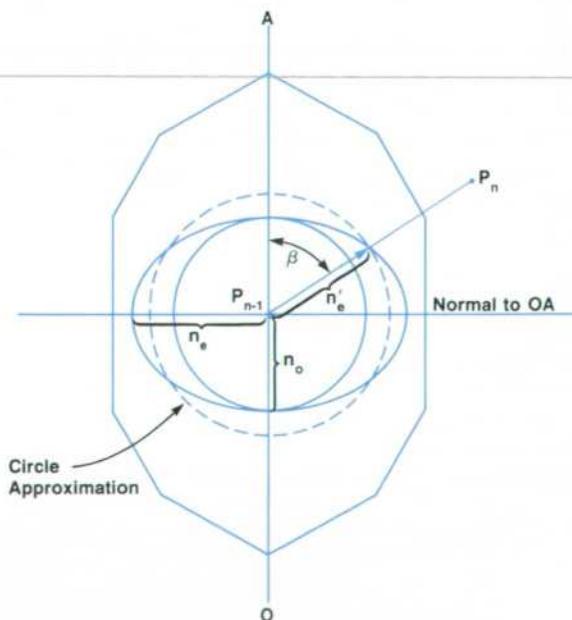
Fig. B. *The extraordinary index in a uniaxial material, such as quartz, is a function of the angle between $\vec{P_{n-1} P_n}$ and the optic axis and can be approximated using a circle method. The ellipse has rotational symmetry about the optic axis OA.*

Programs

Ray-trace programs were written to model the long-range and short-range straightness interferometers, each program requiring 16K of core to run. Fig. C shows the complete construction of the short-range straightness interferometer. Notice that ten points in space define the compensated Wollaston prism. An initial ray-trace through the unmoved system calculates the optical path difference between the interfering rays at normal incidence and establishes the system's "zero" reference for all future measurements.

(continued)
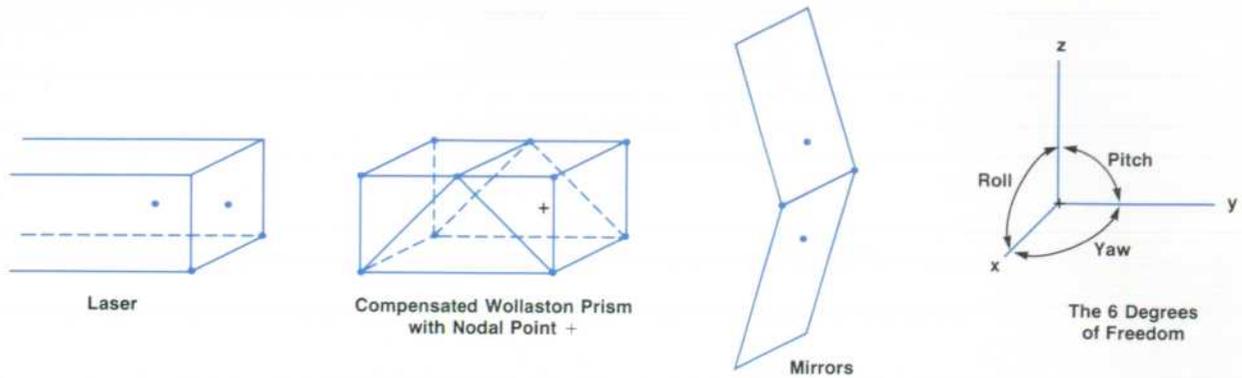
Fig. C. *The short-range straightness interferometer program creates a three-dimensional model of the system, using points in space. All movement will occur about a user-defined point.*

For the straightness interferometer, it is critical that the user find a point such that rotational motions of the system about this point have minimal effect on the optical path difference. This is the system's "nodal" point. The computer locates the place inside the prism that appears to be the source for the two rays exiting toward the mirrors, offering this to the user as the possible point. It would be the correct choice if the extraordinary ray in birefringence propagated spherically, i.e., if $n_e' = n_e =$ constant. However, because n_e' is actually not constant, the choice is not the best and can be improved. The user tries another point by entering it into the program via the teleprinter and running the ray-trace. He repeats this until a suitable point is found. For the short-range straightness interferometer, the computer's choice, when displaced slightly in the y-direction by the user, becomes the nodal point, for it minimizes the interferometer readings for motion in all directions except straightness, or the z-direction.

Printouts of either summarized or detailed ray-trace results are obtainable at the user's request. These lead to system understanding and curves such as Figs. 6 through 10.

range interferometer. The resolution extender forces the display to read E directly if the short-range interferometer is used, or to read 0.1 if the long-range interferometer is used.

The resolution obtainable with the straightness interferometer is shown graphically in Fig. 12. This represents the straightness of travel of a precision air bearing slide having a total straightness error of about 130 microinches. Although this slide does not represent the state of the art in straightness, it is representative of the type of error generally found in machine tools in use today.

Acknowledgments
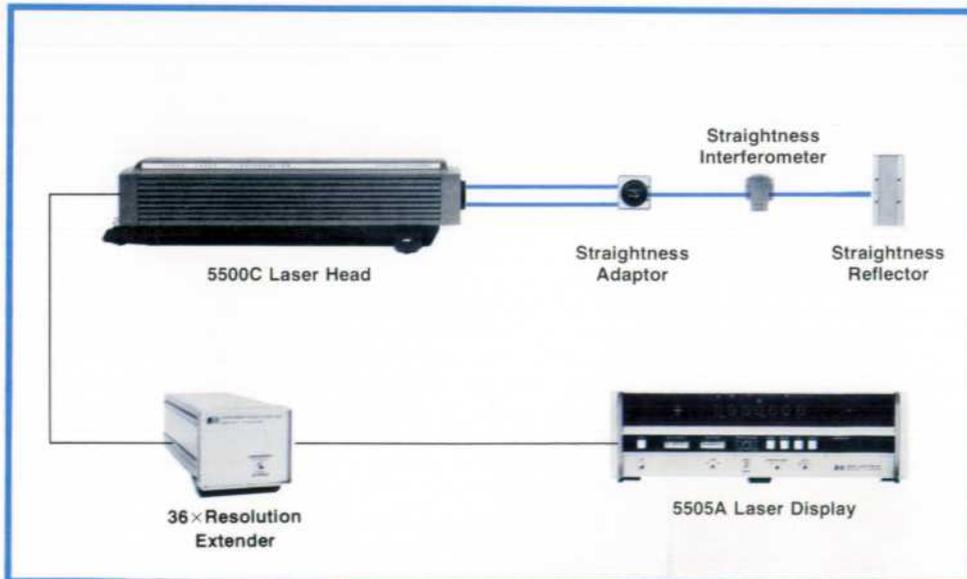
Fig. 11. *5526A Laser Interferometer System with straightness interferometer option. Resolution extender applies a proportionality constant so display reads correct units.*

18
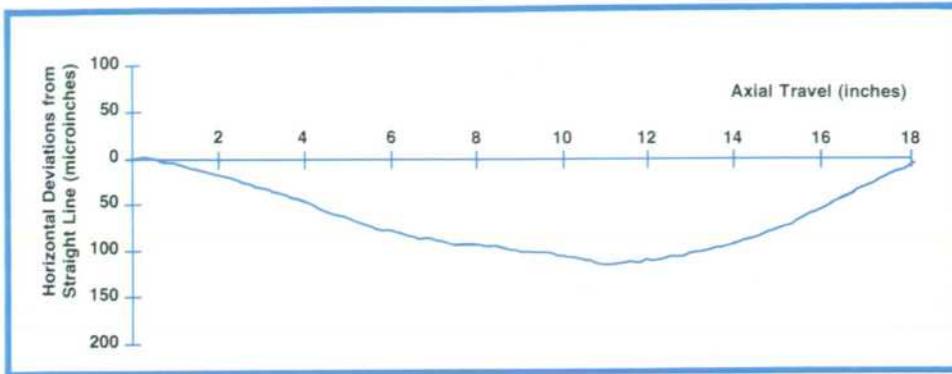
Fig. 12. *Typical straightness calibration of an air platform and precision rail.*

SPECIFICATIONS
HP Model 5526A Laser Interferometer System
Option 30, 31, 32 Straightness Interferometers

(Unless otherwise stated, given specifications apply to both the short and long-range versions.)

ACCURACY:
INCH: ±5 microinches/foot ±1 count in last digit.
METRIC: ±0.4 micrometer/meter ±2 counts in the last digit.
Note: This accuracy can be improved to the limit of repeatability by rotating the Straightness Reflector through 180° and making a second pass (equivalent to reversal of a straightedge).

LINEARITY:
INCH: ±1 count in last digit.
METRIC: ±2 counts in last digit.

CALIBRATION:
SHORT RANGE: ±5% of reading.
LONG RANGE: ±10% of reading.
Note: The actual calibration value is marked on each Straightness Reflector.

RESOLUTION:
SHORT RANGE:
Normal and Smooth Modes:
Inch: 0.00001 inch
Metric: 0.1 micrometer
×10 Mode:
Inch: 0.000001 inch
Metric: 0.01 micrometer
LONG RANGE: One-tenth that of the Short-Range version; e.g., in the ×10 mode, metric units, the last digit has a value of 0.1 micrometer; in the NORMAL mode, inch units, the last digit has a value of 0.0001 inch.

AXIAL RANGE:
SHORT RANGE: 10 feet (3m) ±5%.
LONG RANGE: 100 feet (30m) ±10%.
Note: The minimum distance between the Straightness Interferometer and the Straightness Reflector at which measurements can be made are:
Short Range: 4 inches (100 mm)
Long Range: 3 feet (1 m)

LATERAL RANGE: ±0.1 in (±2.5 mm).

Option 33 Pentaprism
Contact Hewlett-Packard for specifications.

PRICES IN U.S.A.: 5526A Laser Measurement System, $9550.
Option 30 Short Range Straightness Interferometer, $3600.
Option 31 Long Range Straightness Interferometer, $3600.
Option 32 Full Range Straightness Interferometer, $5950.
MANUFACTURING DIVISION: SANTA CLARA DIVISION
5301 Stevens Creek Boulevard
Santa Clara, California 95050 U.S.A.
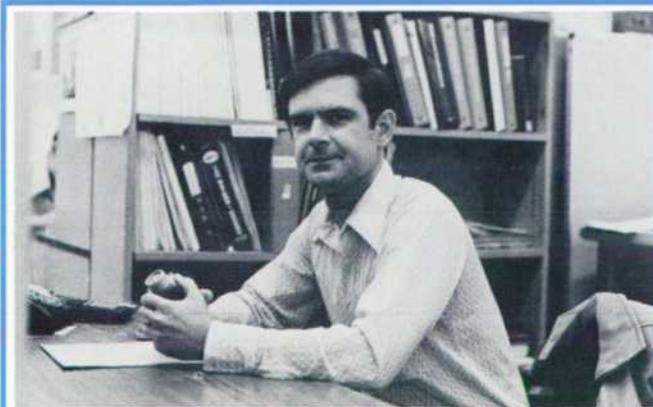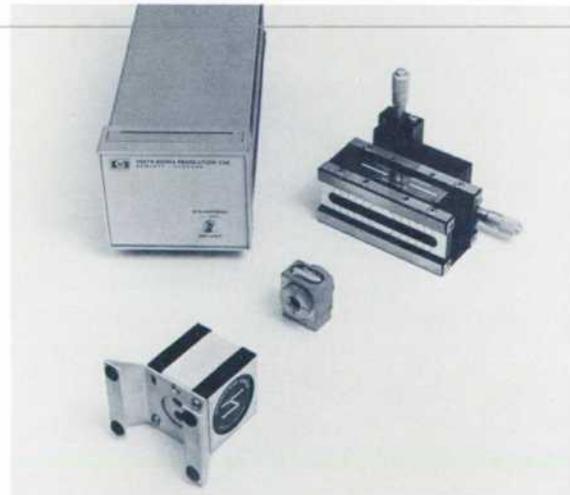
Richard R. Baldwin
Dick Baldwin is project manager for HP Laser Interferometer accessories, including the straightness interferometer. Since 1969, when he joined HP with a seven-year background in metrology and precision machining, he's been responsible for most of the optical design of the HP Laser Interferometer System. He's a member of the Optical Society of America, author of several papers on precision machining and interferometry, and holder of a pair of interferometer patents. A native of San Antonio, Texas, Dick received his BS degree in engineering physics from Ohio State University in 1959. He's married, has two small daughters, and lives in Saratoga, California.

Barbara E. Grote

Barbara Grote is applied mathematician and computer programmer for the dimensional measurement section of HP's Santa Clara Division. She developed the ray-trace program to check the Straightness Interferometer concept. Born in Sacramento, California, Barbara attended Stanford University, graduating in 1970 with a BS degree in mathematics. She joined HP the same year as a programmer in computer-aided integrated-circuit artwork, IC testing, and IC marketing. When she isn't programming, Barbara likes gardening, hiking, watching old movies, and tennis, among other things. She and her husband, who also works for HP, live in Palo Alto, California.
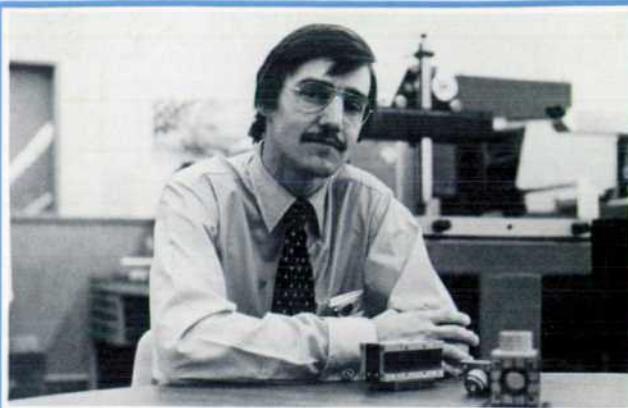
David A. Harland

David Harland is product manager for HP Laser Measurement Systems. A native of the United Kingdom, David came to the U.S.A. in 1966 just after receiving his B.Tech. degree in mechanical engineering from Brunel University. After two years in air-cushion vehicle development, he returned to the campus at Carnegie-Mellon University, graduating in 1970 with an MBA degree. He joined HP in 1970. David is a member of the American Society for Quality Control and the author of papers on air-cushion vehicles and machine tool calibration. He will soon be leaving HP to return to the United Kingdom, where he and his wife plan to explore some new opportunities.

References

1. See, for example, Kurt Rantsch, "Machine Tool Optics," International Research in Production Engineering, American Society of Mechanical Engineers, New York, 1963, p. 629.
2. R.A. Soref and D.H. McMahon, "Optical Design of Wollaston-Prism Digital Light Deflectors," Applied Optics, Vol. 5, No. 3, March 1966.
3. G.M. Burgwald and W.P. Kruger, "An Instant-On Laser for Length Measurement," Hewlett-Packard Journal, August 1970.
4. M. Born and E. Wolf, "Principles of Optics," Pergamon Press, 1959, p. 131.
5. R.R. Baldwin, G.B. Gordon, and A.F. Rudé, "Remote Laser Interferometry," Hewlett-Packard Journal, December 1971.

CHANGE OF ADDRESS: To cha... ...abel (it peels off). Send c...anges ... ...1501 Page Mill Road, Palo Alto, California 94304 U.S.A. Allow 60 days.