PRINCIPLES OF AUTOMATIC COMPENSATION

Laser Measurement System
Application Note 156-3

HEWLETT hp PACKARD

Measuring the variable speed of light improves laser distance measurement

In laser interferometry, you can't take comfort in the constancy of the speed of light, because it changes with air temperature, pressure, and humidity; but the variations can be measured and compensated for automatically

by Jonathan D. Garman and John J. Corcoran, *Hewlett-Packard Co., Palo Alto, Calif.*

☐ Everyone knows that the speed of light is a universal constant. Scientists have been refining their estimates of it for over a hundred years, till today its value is known to a part in ten million. It's known, too, that light travels at the speed limit of the universe and that this is axiomatic to relativity and much of modern physics. All of which lends an aura of finality and irrefutability to the concept, and makes it easy to overlook the fact that the speed of light is a constant only in a vacuum.

In laser applications, such as interferometer systems, it's important to remember that the speed of light is *not* constant in air, and in fact varies with air temperature, air pressure, and humidity. If it is not carefully monitored, the accuracy with which a laser interferometer system can measure distance will be severely degraded.

Since it is difficult to measure the speed of light directly, it makes more sense to measure the atmospheric parameters—temperature, pressure, and humidity—and factor them into the calculation. In other words, the best place for a photon speedometer to start from is an electronic weather station. This approach also makes it easy to compensate automatically for errors caused by the thermal expansion of the object being measured.

Laser interferometers measure changes in distance by means of a technique developed by Michelson in the last century. A light source is split into two beams and directed to a fixed and a movable mirror (Fig. 1). The reflected beams recombine at the splitter, where they interfere and form a dark or light spot—a "fringe"—depending on the relative phase of the two waves. If both mirrors are stationary, the fringe pattern at the beam splitter (interferometer) will not change. But when the movable reflector traverses one half wavelength, one full cycle of fringe brightness occurs at the interferometer. This can be detected by a photodiode, and the resultant pulse stored in a counter.

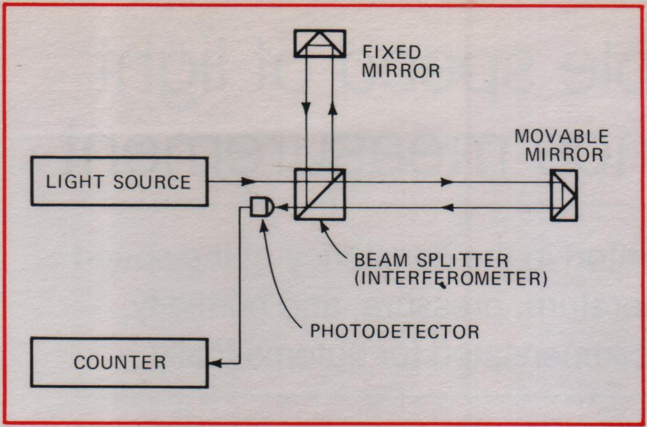
Counting half wavelengths

The number of pulses stored in the counter at any time represents the number of half wavelengths of motion the reflector has moved from some arbitrary reference point where the counter contents were set to zero. To convert this raw fringe count into useful form, it must be multiplied by the length of one half wave. Consequently, the length standard of a laser interferometer is the wavelength of the laser light in the air.

A laser is a frequency standard—not a wavelength standard. It emits light of a constant and well-known frequency. Since the product of wavelength and frequency for any electromagnetic radiation equals the propagation velocity of the wave, variations in the speed of light produce proportional changes in a laser's wavelength. Measurement of the speed of light, therefore, determines the laser wavelength also.

A typical stabilized laser has a frequency stability of about one part in 10^7. To make full use of the measurement accuracy this implies, it would be necessary to continuously measure the speed of light (and thus the laser wavelength) to the same part-in-10^7 accuracy. But in a field instrument measuring the speed of light directly with this degree of refinement is impractical. The most carefully controlled laboratory experiments have only recently reached such a level of accuracy.

Consequently, in designing an automatic compensator for laser interferometry, it is advisable to use the indirect approach and measure the index of refraction of the air (the ratio of the speed of light in a vacuum to its speed in air). This approach takes advantage of the rela-



1. Interferometer principles. Each half wavelength of motion of the movable mirror produces a full cycle of fringe brightness when it is reflected back to the photodetector.

tive ease of finding the precise index of refraction for air, and makes use of the already well-known vacuum speed of light.

An empirical formula is available for expressing the refractive index of air in terms of air pressure, air temperature, and partial pressure of water vapor in the air.[1] For typical indoor atmospheric conditions, the refractive index is around 1.0003, or a variation of three parts in 10^4 from the vacuum value. For a wide range of environments, this value can be expected to change about one part in 10^4. The empirical formula allows this range to be split into a thousand parts, so that the refractive index can be determined to a part in 10^7. The net result is that an instrument with roughly 0.1% accuracy and suitable environmental sensors can determine refractive index (and thus laser wavelength) over a wide range of conditions to a part in 10^7. This four-orders-of-magnitude reduction in accuracy requirements makes the indirect approach very attractive.

Computing the wavelength

A reduced form of the formula for refractive index, which holds for dry air and wavelengths near 6,328 angstroms (the wavelength of a helium-neon laser), is

$$N = 1 + (104.79 \times 10^{-6})P/T \qquad (1)$$

where N is the refractive index, P is the air pressure in millimeters of mercury, and T is the air temperature in degrees Kelvin. The formula merely states that the refractive index is a linear function of the ratio of air pressure to air temperature. More simply, refractive index is a linear function of air density.

The speed of light (and thus the laser wavelength) is proportional to the inverse of the refractive index. Since the refractive index is only slightly different from 1, the inverse can be found by subtraction rather than division. Within a part in 10^7, Equation 1 becomes:

$$1/N = 1 - (104.79 \times 10^{-6})P/T \qquad (2)$$

As air density increases, therefore, the speed of light decreases, agreeing with intuition. For example, with P = 760 mm and T = 300° K, 1/N = 0.9997345, which is the ratio of air wavelength to vacuum wavelength for the stated conditions.

Some laser interferometers—such as the Hewlett-Packard model 5525B—contain a calculator which uses 1/N directly to compute distance,[2] which is given by the product of fringe count, vacuum wavelength, and inverse refractive index. Since the three most significant 9s in 1/N do not change for any reasonable environment, they are hard-wired into the calculator. Only the last four digits, which resolve the wavelength to a part in 10^7, need be supplied by the electronic compensator.

Several techniques are available for solving the equation for 1/N electronically. In designing a compensator for the 5525B, the constraints were that the selected pressure sensor (aneroid capsule) and the temperature sensor (thermistors) are inherently analog while the output to the interferometer system had to be digital. The linear designer's approach might be to use an analog divider to form the ratio of P and T, an operational amplifier to perform the difference operation, and an a-d converter to produce the final result. The digital designer's approach might be to a-d convert P and T individually, then perform the division and subtraction operations using a small arithmetic unit. But both these techniques have disadvantages—for example, there is imprecision in the former and complexity in the latter.

Another, perhaps less obvious, approach was adopted. A flexible, precise, dual-ramp integration technique was used to perform the division, subtraction and a-d conversion operations all in a single step (see panel, "How the dual-ramp integrator works").

Equation 2 only holds for dry air. If water vapor is present the refractive index is changed and a correction must be made to 1/N. The correction required is directly proportional to the partial pressure of water vapor in the air. In this compensator, direct measurements are made of humidity and air temperature. Water vapor pressure is then derived from these two quantities in a second cycle of the dual-ramp integrator, and the appropriate correction to 1/N is simultaneously made.

Practical limitations

Wavelength errors of a part in 10^7 are introduced by quite small errors in the measurement of the important environmental parameters. Air temperature, for example, must be known to 0.1°C to achieve this level of accuracy. But temperature measurements at different points in the same room often differ by as much as several degrees, and it is unrealistic to expect the same temperature to occur everywhere along the beam's path.

Uncertainty in the measurement of air temperature can easily therefore introduce a wavelength uncertainty of several parts in 10^7. An error in measuring air pressure of just 0.01 in. of Hg produces a wavelength error of a part in 10^7. Yet a mercury barometer must be very carefully maintained to achieve this accuracy, and only the best aneroid barometers approach it.

If long vertical distance measurements are undertaken, another problem enters—air pressure drops about 0.01 in. of Hg for each 10-foot increase in altitude. Humidity (or water vapor pressure) has a smaller effect—a part in 10^7 wave-length error results from roughly a 5% error in determining relative humidity at normal room temperatures. But the accurate measurement of humidity also presents some problems.

Because of these effects, it is perhaps impractical to expect wavelength determination accuracy of much bet-

How the dual-ramp integrator works

The heart of the automatic compensator is a dual-ramp integrator—a precision a-d converter with built-in computational capability. The integrator finds the ratio of the two analog signals required by the wavelength formula. It multiplies the expansion coefficient by the material temperature to correct for material expansion. And it converts individual analog sensor readings for display as digital quantities.

The dual-ramp integrator uses current switches, an integrator, a comparator, and a timer (Fig. A). The comparator threshold is V_c at the integrator output. Two currents, I_1 and I_2, are switched into the integrator during each conversion cycle. The first current, I_1, is turned on to start the cycle, and the time interval, T_1, begins when the integrator output crosses the comparator threshold. At the end of T_1, current I_1 is switched off, and at this point,

$$V_{peak} = V_c + I_1T_1/C$$

I_2 (of polarity opposite to I_1) is now switched into the integrator, and the time required to return the integrator to the comparator threshold—T_2—is measured. Another expression for V_{peak} can then be written:

$$V_{peak} = V_c + I_2T_2/C$$

Setting these two expressions equal leads to

$$I_1T_1 = I_2T_2, \text{ or}$$

$$T_2 = (I_1/I_2)T_1$$

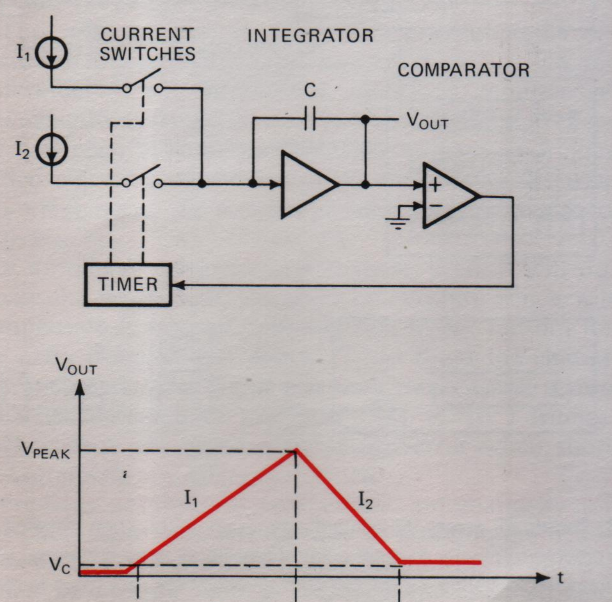
Important features of this technique are

- V_c cancels out, so comparator drifts do not degrade the conversion accuracy.
- If I_1 and I_2 are linearly derived from the same reference voltage, then any drift in that voltage will cancel out.
- If T_1 and T_2 are measured from the same oscillator, then any drift in the frequency of the oscillator will cancel.
- T_2 is proportional to I_1/I_2, permitting a dc ratio measurement to be made if T_1 is a constant. This is used to form the ratio of P to T in the wavelength computation.
- T_2 is also proportional to I_1 and T_1, enabling a multiplication (analog current times digitally derived time) to be made if I_2 is a constant. This is the mode used to correct for thermal expansion. The coefficient of expansion (entered on thumbwheels) sets the time interval T_1 directly, and a current proportional to material temperature (which is zero at 20° C) is switched in as I_1.
- If T_1 and I_2 are constant, the dual ramp cycle becomes a straight a-d conversion of the unknown current I_1. Individual sensor readings are converted in this manner.

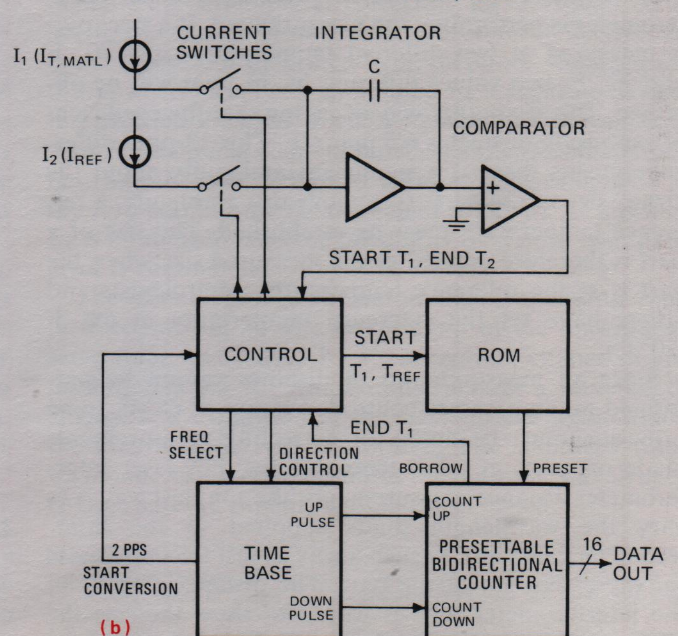
As an example of the operation of the dual ramp integrator, suppose it is desired to display material temperature in degrees centigrade. A number representing the desired time interval T_1 is read from a small read-only memory and preset into a counter (see Fig. B, which is a simplified block diagram of the 5510A automatic compensator). Then the material temperature current is turned on, and when the integrator output crosses the comparator threshold, down-count pulses are routed to the counter. When the counter reaches zero, T_1 is complete, and the material temperature current is turned off.

The reference material temperature number (20.0° C, where the material temperature current is zero) is then read from the ROM and preset into the counter. A fixed current (with the proper sign to return the integrator to zero) is now switched on to begin T_2, and simultaneously a pulse train is routed into the appropriate (up or down) counter input. When the integrator output again crosses the comparator threshold, the counter contains the digital representation of the temperature since the reference number has been changed up or down by an amount proportional to the difference between the actual and the reference temperature.

When multiple terms must be converted separately and then summed, as in the computation of the total compensation factor, the first term is converted normally, and at the end of T_2 the result from the counter is stored in a scratchpad register. For subsequent terms, the scratchpad contents are preset into the counter just before the beginning of T_2, then stored again in the scratchpad at the end of T_2. When all terms have been completed and summed, the scratchpad contents are read into an output buffer register, and conversion is complete.



(a)



(b)

2. Compensation. Errors produced by thermal and other factors are corrected by 5510A automatic compensator. The probes have magnetic bases for easy attachment near the measurement path.

ter than a part per million, except under laboratory conditions. Still other sources of measurement error in laser interferometry can degrade the over-all accuracy to several parts per million, unless corrections are made for them. Two of the most important of these other error sources are the thermal expansion of the object being measured and what is referred to as "deadpath error."

Compensating for thermal expansion

All common materials expand as their temperature is increased. Although this expansion may not seem to change the size of an object significantly, it becomes very visible when a precision yardstick like an interferometer is performing the measurement. If a steel part is measured at two different temperatures, say 70° F and 72° F, two values differing by 13 ppm will be obtained. This is another way of saying that the coefficient of expansion of steel is 6.5 ppm/°F. This 13-ppm difference means there is a significant ambiguity about the "true" size of the part. To remove this ambiguity, a reference temperature must be established. The size of a part is therefore defined as the measured size when the part is at the reference temperature. Metrologists and others have set the reference temperature at 68° F (20° C).

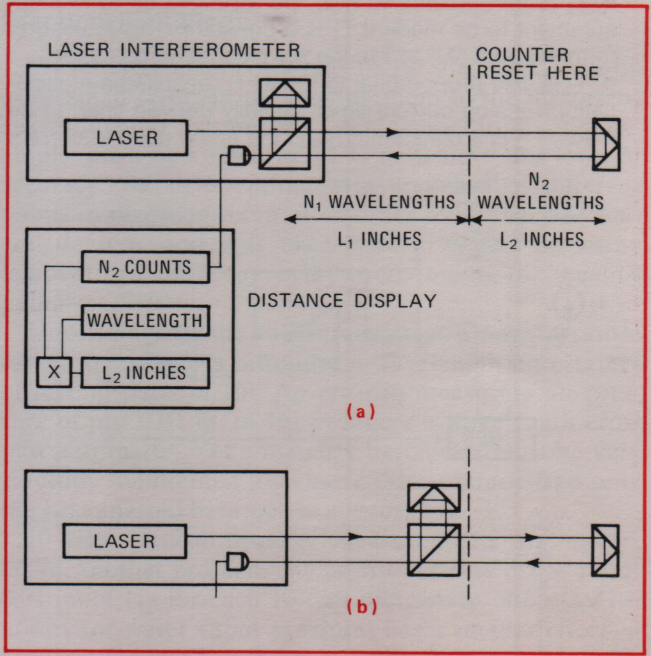
Since all measurements of distance cannot be performed in temperature-controlled rooms at 68° F, some provision must be made for correcting measurements made on parts at other temperatures. For laser interferometer distance measurements, the simplest way is to vary the wavelength digits supplied to the interferometer calculator in such a way as just to compensate for the expansion of the part. The distance display of the interferometer then is forced to show the size the part would be at 68° F (the "true" size of the part) instead of the actually measured, expanded size.

The compensator under discussion corrects for thermal expansion in precisely this fashion. A surface-temperature sensor is supplied with each instrument, and thumbwheel switches on the front panel provide for entry of the coefficient of expansion of the material being measured (Fig. 2). The compensator finds the difference between the temperature of the part and 68° F, a result which is multiplied by the coefficient of expansion to get the expansion in ppm. The four digits representing the wavelength are then decremented by the amount of the expansion. This forces a smaller-than-actual distance display, the size the part would be at 68° F. The multiplication and subtraction operations necessary are again performed by the dual-ramp integrator, a single cycle being sufficient to correct the wavelength digits for thermal expansion.

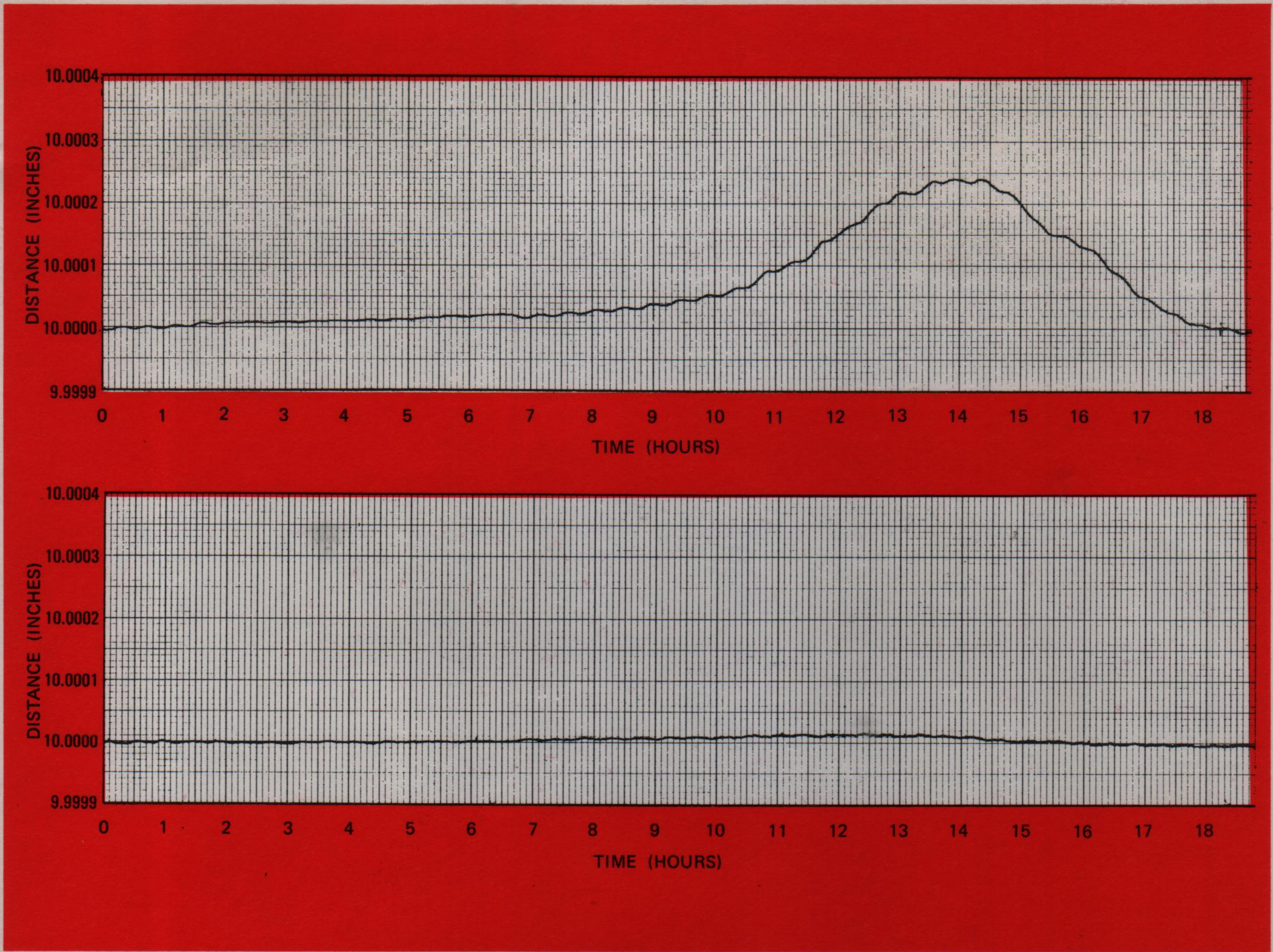
When large parts are being measured, thermal homogeneity cannot be assured. For such situations, the compensator has provision for averaging up to three strategically placed temperature sensors. This consideration is particularly important in the laser calibration of large machine tools.

Reducing deadpath error

The other common source of error in interferometer measurements is referred to as "deadpath" or uncompensated path length error. This type of error can arise when environmental conditions change during a measurement. Assume the measurement shown schematically in Fig. 3a is being made. Let L_1 be the distance from the interferometer to the reference point of the measurement (the point where the counter is set to zero), and let N_1 be the number of wavelengths of light constituting this length. Similarly, let L_2 be the distance



3. Deadpath errors. Significant errors can occur when distance L_1 from interferometer to reset point of counter is large compared with measured distance, L_2 (a). Because its optics are very close to reset point, remote interferometer (b) almost completely eliminates deadpath and the need to compensate for it.



4. It works! Typical before-and-after readings illustrate stability improvement provided by compensation. Upper graph shows change in distance reading without compensation. Lower graph, which was produced at the same time, shows the benefits that arise from including automatic compensation for both wavelength and expansion of material.

being measured and N_2 the corresponding number of wavelengths (the number of pulses stored in the counter). Now assume that the environment changes and the wavelength decreases by 10 ppm.

This will increase the number of wavelengths required to represent the fixed distance (L_1+L_2). The interferometer will detect the increased number of wavelengths as fringe motion, just as if the target reflector had been moved. However, the number of fringes added will be proportional to the total fringe distance (N_1+N_2), not to the number of fringes in the measured path N_2.

If automatic compensation is being used, it will decrease the wavelength standard by 10 ppm, which will compensate for the 10-ppm addition to N_2, but not for the extra fringe counts due to N_1. In short, the change in wavelength scale factor has been tracked, but a zero shift has occurred of magnitude $N_1 \times 10^{-5}$ fringes. Such a shift can cause a significant error when short measurements are being performed.

Previous attempts to compensate automatically for this effect have required the user to estimate the deadpath distance L_1 and enter it on thumbwheel switches. When wavelength changes occurred, the assumed zero shift was computed and then corrected for.

Another type of deadpath zero shift can be caused by the thermal expansion of the material upon which the interferometer and reflector are placed. Its correction requires entry of a coeffecient of expansion. These "open-loop" attempts at correcting for deadpath error are inconvenient, unsatisfying and often ineffective. In the compensator described here, therefore, the need for thumbwheels has been eliminated by a separate device, the Hewlett-Packard 10565A remote interferometer.[3] This small optical element moves the interferometer optics from the laser head to a point as close as desired to the zero point of the measurement (Fig. 3b). As a result, L_1 is minimized, and the possibility of serious deadpath error is eliminated.

The effectiveness of the automatic compensator is illustrated in Fig. 4. The large hump in the uncompensated trace was caused by a temperature change of approximately 2.5° C, which induced both a wavelength change and expansion of the steel surface upon which the measurement was performed. The stabilized trace was made using compensation for both the wavelength change and the expansion of the steel bench. □

REFERENCES
1. Bengt Edlen, "The Refractive Index of Air," Metrologia, Vol. 2, No. 2, 1966, pp. 71-80.
2. J. N. Dukes and G. B. Gordon, "A Two-Hundred-Foot Yardstick with Graduations every Microinch," Hewlett-Packard Journal, August 1970.
3. R. R. Baldwin, G. B. Gordon, and A. F. Rude, "Remote Laser Interferometry," Hewlett-Packard Journal, December 1971.