Method for evaluation of optical stability of the holographic recording environment

P A Wilksch and D T Ziedars
School of Applied Sciences, RMIT University, Melbourne, Australia
Email: philip.wilksch@rmit.edu.au

Abstract. A simple method is described for displaying and recording drifts in the path length difference between two laser beams from a single source. A beamsplitter is used to recombine the expanded beams, producing two similar interference patterns, one on each side of the beamsplitter. A photodiode is positioned in each pattern such that the two detectors sample the pattern at points where the phase difference between beams is itself different by one quarter of a cycle. The outputs of the two photodiodes are amplified and shown on an oscilloscope in X-Y display mode. After suitable adjustments of amplifier gain and photodiode position, any optical drifts cause the oscilloscope spot to move in a circle in which the angular position of the spot directly represents the phase difference between the light beams. This movement can be tracked visually, or recorded and analyzed to produce a time record of phase drifts. The system is self-calibrated, and sensitivities of 1/100 cycle are easily obtained, on time scales ranging from the minimum response time of the photodetectors to as long as desired. Examples are shown illustrating the effects of air currents and floor vibrations on the optical stability of a holographic table. Results are compared with mechanical vibration measurements obtained with an accelerometer.

1. Introduction
The necessity for optical stability when recording holograms with CW lasers is well-known, and holography texts and handbooks carry many hints and tips on reduction of disturbances due to mechanical vibrations and drifts, thermal movements, changes in humidity, air currents, and the like. The proof of a stable setup is the quality and efficiency of the hologram produced, but the time delay between recording and viewing a developed hologram makes a trial-and-error approach to the reduction of disturbances a tedious process. Immediate, real-time means of evaluation of stability are far more useful.

It is sensible to employ a method which uses the same technique as holography itself, namely interference, to evaluate stability. An interference pattern is readily set up between two or more beams arranged in a similar configuration to the one that is to be employed in recording the hologram. There is then a need for a way of recording the stability of the fringe pattern, and quantifying it so that efforts to improve stability can be evaluated objectively.

This paper points out the advantages of an arrangement that uses two detectors to sample irradiance variations at closely-separated points in a fringe pattern, and a processing algorithm that produces a continuous record of phase difference variations between the interfering beams. The techniques themselves are not new, but our experience has shown that they can be implemented easily with inexpensive equipment and can give valuable information on both short-term and long-term stability.
2. Theory and Method

Figure 1 shows a typical arrangement on an optical bench where a laser beam is split into two beams, which, after traveling different paths, are redirected to a common point. In a holographic recording setup, the common point would be the recording medium, and the two beams would be a reference beam and an object beam, after scattering from the object. The stability of the phase difference between the beams at the point of recombination can be investigated by studying their interference fringe pattern.

![Figure 1](image.png)

Figure 1. Recombining two laser beams to produce two interference patterns, one on each side of the final beamsplitter. BS1, BS2: beamsplitters; M1, M2: mirrors; SF1, SF2: beam expander/spatial filters; D1, D2: detectors.

To produce interference fringes with broad fringe spacing, the two beams must be brought into close parallelism. This is accomplished with another beamsplitter positioned so that the two beams make equal angles with the beamsplitter normal, but on opposite sides of the beamsplitter. Further, to spread the beams over a larger area, a beam expander/spatial filter can be used in each beam. It is an advantage if these are positioned equal distances from the point of recombination at the beamsplitter, so that the curvatures of the wavefronts are closely matched. Then by fine adjustments of the spatial filter positions and the beamsplitter tilt, the separation between the interference fringes on the far sides of the beamsplitter can be broadened, until further widening is limited by the influence of aberrations and nonuniformity in the beamsplitter.

According to the general law of interference [1], the total irradiance $S_T$ at any particular point in the fringe pattern is given by:

$$S_T = S_1 + S_2 + 2r(S_1 + S_2)^{1/2} \cos \delta$$  \hspace{1cm} (1)

where $S_1$ and $S_2$ are the irradiances at that point due to the two beams separately, $r$ is the degree of coherence, and $\delta$ is the phase difference between the beams. Equation (1) can be simplified to:

$$S_T = S_{av} + S_p \cos \delta$$  \hspace{1cm} (2)

where $S_{av}$ and $S_p$ are constants applying at that point. The quantity that is of interest here is the phase difference $\delta$ between the two beams. It can be expressed as:

$$\delta = \frac{2\pi \Lambda_{12}}{\lambda_0}$$  \hspace{1cm} (3)

in which $\lambda_0$ is the free-space wavelength of the light, and $\Lambda_{12}$ is the optical path difference between the two beams, measured from the point of division at the first beamsplitter to the point of recombination at the second. The optical path, in turn, is given in general by:

$$\Lambda = \int n(t) dl$$  \hspace{1cm} (4)

where $n$ is the refractive index of the intervening medium, and the integral is over the path between beginning and end points mentioned above. Hence $\Lambda_{12}$ is sensitive to changes in the geometric path length as well as fluctuations in refractive index.
From a recording of the irradiance $S_T$ at a single point, only the cosine of $\delta$ can be determined, so there is ambiguity in translating changes in $S_T$ into terms of changes in $\Lambda_{12}$, particularly when such changes extend over several wavelengths. It is not possible to tell in which direction (increasing or decreasing) the change is taking place, and there is uncertainty whether or not a reversal of direction might have occurred at any point where $\cos \delta = \pm 1$.

Nevertheless, when one looks at an extended fringe pattern, it is obvious in what direction $\Lambda_{12}$ is changing, because the fringes will drift in one direction or the other, and the relation between the direction of fringe drift and the increase or decrease of $\Lambda_{12}$ can be determined simply by imposing a small change of known direction on the beam path, such as by light pressure on one of the mirrors. One could record movements of the fringe pattern with a digital camera, but one would still be faced with the task of analyzing the recorded images to extract quantitative measurements of optical path variations from them.

The classic technique for recovering phase from sinusoidal fringe patterns [2] is quadrature detection, where two detectors are positioned so that, if the phase difference at one of them is $\delta$, then the phase difference at the second is $\delta - \pi/2$ rad. In (2), the irradiance at the second detector will then have a contribution proportional to $\cos (\delta - \pi/2) = \sin \delta$. With knowledge of both $\sin \delta$ and $\cos \delta$, the phase difference $\delta$ can be determined unambiguously, at least over a range of $2\pi$ rad.

This technique could be implemented by situating two detectors side by side in the fringe pattern so that their spatial separation corresponds to one quarter of the fringe separation (or any odd multiple of one quarter). Provided that the fringe separation does not change as the path lengths change, the detectors will remain in quadrature. In most cases this condition will be satisfied, because the spatial details of the fringe pattern depend only on details of the optics between the spatial filter pinholes and the detector plane, and this part of the optical system can be made compact and stable.

However, depending on the size of the detectors, it may be difficult or impossible to situate two of them close enough to sample points within one fringe spacing. One solution could be to use optical fibre pickups to collect the light. But there is another option already available in the optical configuration of figure 1, and that is the “twin” fringe pattern formed on the other side of the beamsplitter. This pattern is a mirror image of the other, and is also contrast-reversed, having bright fringes where the other has dark fringes (if the beamsplitter is non-absorbing). It is not difficult to situate one detector in each fringe pattern, and then fine-tune their positioning so that they are in spatial quadrature. A method for doing this is described in the next section.

3. Implementation

3.1. Apparatus

Measurements have been made on two pneumatically-supported optical tables in the holography laboratory at RMIT. The light source for the interference measurements was a 10 mW helium-neon laser (Melles Griot model 05-LHP-991). An arrangement similar to that in figure 1 was set up, with the beams covering a large area of the table. The two path lengths were matched to within 20 millimeters to ensure good coherence. Spatial filter assemblies were used to spread the beams from points about 100 mm before the second beamsplitter.

Figure 2 is a photograph of the typical appearance of the interference fringes. The system has been aligned to spread the fringes as far apart as possible. The residual fringe pattern shows saddle-shaped contours due to the astigmatism introduced when a spherical wavefront passes through a parallel-sided piece of glass (the beamsplitter) at an angle. A detector is then placed at the centre of each pattern where the fringes are widely separated.

The detectors were two silicon photodiodes (UDT 10DP). These have a circular sensitive area of 1.0 cm$^2$, but each was mounted in an enclosure with an adjustable iris diaphragm in front of the detector, so that a small area of the interference pattern could be sampled. Typically the aperture diameter was between 2 and 5 mm. The photodiode currents were amplified and converted to voltages with Terahertz Technologies PDA-750 photodiode amplifiers.
The two amplified signals were fed to the two inputs of an oscilloscope operating in X-Y mode. Because both signals have a component which varies sinusoidally with changes in $\Lambda_{12}$, the oscilloscope spot is observed to trace out a figure that is in general an ellipse, as the path difference changes. If the two detectors are sampling the same point in the fringe pattern, the figure will be a straight line, as the two sinusoidal components will be in phase with each other. If the detectors sample points that are one quarter of a fringe apart, the figure will be an ellipse symmetrical about the vertical and horizontal axis, or a circle if the amplitudes of both sinusoidal components are the same. The pattern varies between these extremes as the detector is displaced, as shown in figure 3.

**Figure 3.** A series of recordings of scaled and normalised detector signals $R_1$ (horizontal axis) and $R_2$ (vertical axis) plotted against each other, as the path difference changes. The position of one of the detectors is shifted incrementally in the sequence (a) to (i). The pattern changes in shape from circular through elliptical to linear and back again as the detector is shifted across the fringe pattern.
To map out the shape of the ellipse, it is necessary to introduce a perturbation to the optical path difference between beams. This is easily done by drumming with one’s fingers on the table top. The position of one of the detectors can then be adjusted until the ellipse is symmetrical about X and Y axes. The gain and offset of each channel are adjusted to produce a circular trace, centred on the oscilloscope screen. The sense of rotation of the oscilloscope spot (clockwise or anticlockwise) tells in which direction the path difference is changing. One whole rotation of the spot corresponds to a path difference change of precisely one wavelength. The system is thus self-calibrating.

To produce a permanent record of drifts, the signal outputs from the detectors must be recorded. Inexpensive digital oscilloscopes are now available which can store a time sequence of voltages and read it out to a computer. Alternatively, USB data logger units can record voltage signals over longer periods, so that long-term drifts can be studied. The series of numbers thus generated is easily converted into a record of optical path variation, as follows.

3.2. Analysis

First, each data series $S$ is inspected to find its maximum and minimum signal values, $S_{\text{max}}$ and $S_{\text{min}}$. From these, we determine $S_{\text{av}} = (S_{\text{max}} + S_{\text{min}})/2$, and $S_p = (S_{\text{max}} - S_{\text{min}})/2$. Then for each reading within the series we calculate the quantity

$$ R = (S - S_{\text{av}})/S_p $$

which from (2) is expected to equal $\cos \delta$. It varies over the range $-1 \leq R \leq +1$.

Suppose that the phase difference at detector 1 is $\delta$, and at detector 2 is $\delta - \pi/2$, and let the corresponding series of values of $R$ be $R_1$ and $R_2$. Then we expect that:

$$ \delta = \arctan \left( \frac{R_2}{R_1} \right) $$

provided that the underlying signals of $S_1$ and $S_2$ are obtained simultaneously from the two detectors.

Equation (6) gives a result over the range $-\pi \leq \delta \leq +\pi$. For convenience of interpretation, it is useful to convert this to a measurement in units of cycles rather than radians. Define a quantity:

$$ D = \frac{1}{2\pi} \arctan \left( \frac{R_2}{R_1} \right) $$

which varies over the range $-\frac{1}{2} \leq D \leq +\frac{1}{2}$. Now it only remains to deal with the discontinuity that results when $D$ crosses the boundaries of this range. Provided that successive samples differ in actual phase by less than half a cycle, this can be done by adding or subtracting one whole cycle whenever $D$ changes by more than $\frac{1}{2}$. We keep track of the number of whole cycles of phase change with an integer $M$, initially set to zero. At each sampling step $i$, we calculate the change in $D$, namely $\Delta D = D_i - D_{i-1}$. Then we set

$$ \Delta M = \begin{cases} +1 & \text{if } \Delta D < -\frac{1}{2} \\ -1 & \text{if } \Delta D > +\frac{1}{2} \\ 0 & \text{otherwise} \end{cases} $$

and put $M_i = M_{i-1} + \Delta M$. Finally, the "unwrapped" phase $P$, expressed in cycles, is given by

$$ P_i = M_i + D_i $$

If the two detectors are not aligned exactly in quadrature, and the shape of the oscilloscope trace is slightly elliptical rather than circular, a small nonlinearity will be introduced into the relation between the calculated $P_i$ and the actual phase difference $\delta$, over each cycle of phase change. The discrepancy can be corrected by using the measured shape of the ellipse [3], but the correction will not be significant except in very precise applications. In any case, the fact that an integer change in $P$ corresponds to a whole cycle of phase change continues to be exactly true.
4. Results

4.1. Long-term drift

Figure 4 shows a series of readings of phase difference taken at one second intervals over a time period of 100 seconds. At the commencement of the series (between 10 and 30 seconds on the time axis), a gradually-increasing pressure was applied to the table top, and then removed, causing the phase to change over a few cycles. From this sequence, the parameters $S_{\text{max}}$ and $S_{\text{min}}$ were determined for each of the two detector signals, and subsequently used in the calculation of the phase difference.

![Figure 4](image)

**Figure 4.** Variation of phase difference as a function of time, during which a phase change was induced by pressure on the table top. The phase is sampled at one second intervals.

Longer sequences of measurements are shown in figure 5, for four different conditions. The sampling period was again one second. Figure 5(a) was recorded with the room air-conditioning operating, and the table top open and unprotected. In an attempt to minimize local draughts, temperature-controlled air is introduced into the room through long, porous "socks" running overhead the length of the room. Nevertheless, random phase fluctuations extending over about half a cycle are apparent. That these are due to air currents is confirmed by figure 5(b), where the air-conditioning was switched off. The random fluctuations are seen to be reduced in amplitude, although they are still significant.

![Figure 5](image)

**Figure 5.** Recorded variations in path difference as a function of time, expressed in terms of number of cycles of phase change. The four traces are for (a) unenclosed table, air conditioning (AC) on; (b) unenclosed table, AC off; (c) enclosed table, AC on; (d) enclosed table, AC off. The phase at the starting time is set arbitrarily to offset the traces.
For the traces in figures 5(c) and (d), the table top was enclosed by walls and a roof made of acoustic insulation batts, preventing air currents from crossing the table. There is little difference between (c), with air-conditioning running, and (d), with the air conditioning off, and the random fluctuations are further reduced to less than one tenth of a cycle. However, all four traces show a steady drift in path difference over the 400-second time interval of the recordings, with some indication that the drift might be more serious when the table is covered. More measurements are needed to identify the cause of this drift.

4.2. **Short-term drift**

To investigate phase jitter over time periods shorter than one second, a digital oscilloscope was used to record the output of the two detectors. The output data consisted of 4096 samples taken over a period of one second, with a sampling rate close to 0.25 ms. A short segment of the processed recording is shown in figure 6.

There is a prominent periodic component in phase perturbation with a frequency close to 100 Hz, and an amplitude of about 0.01 cycle. This was present in most of the records. The ability of the technique to resolve phase changes of this magnitude is well illustrated by this example. To verify that these oscillations were not due to output power fluctuations of the laser, recordings were made of the two beams individually. No 100 Hz component was evident.

![Figure 6: Phase difference variations recorded with a digital oscilloscope, with a sampling period of 0.25 ms. A strong periodic perturbation with a frequency of 100 Hz and an amplitude of about 0.01 cycle is apparent.](image)

4.3. **Accelerometer measurements**

Short-term mechanical vibrations were also investigated with the use of a single-axis seismic accelerometer (Wilcoxon Research model 731A). Records of acceleration over time periods of a few seconds were Fourier-analyzed to produce measurements of the frequency components of the amplitude of displacement at one-Hertz intervals.

Measurements were made on a table next to the one on which the interferometer was mounted. Figure 7 shows the recorded vertical displacement on a logarithmic scale. The upper trace is for the accelerometer mounted on the floor under the table, and the lower trace is for the accelerometer mounted on the table top. Conversion from acceleration to displacement involves multiplying the data by the inverse square of the frequency, and results become unreliable at frequencies below about 10 Hz.

On the floor there are prominent components at 100 Hz and 200 Hz. These are attenuated by over an order of magnitude by the table suspension and damping system, but the 100 Hz vibration is still detectable on the table top. Horizontal displacements were also measured, and were found to be similar to the vertical ones but generally smaller in amplitude.
The accelerometer confirms the presence of the 100 Hz vibrations found with the interferometer. It is thought that they might be caused by motors or power transformers elsewhere in the building, generating acoustic waves that are transmitted through the building structure.

5. Conclusions
A simple method has been used to measure and record phase difference variations between two interfering beams in a holographic setup. With the use of two detectors sampling an interference pattern in spatial quadrature, phase variations can be tracked over both short and long time periods without ambiguity in direction of the phase drift or discontinuities in the phase record. Phase variations can also be displayed in real time for rapid evaluation of the stability of holographic environments. Resolution is of the order of 1/100 cycle or better.

Use of the technique in the holography laboratory has illustrated the importance of good shielding of the optical table from air currents, and mechanical vibrations at 100 Hz have been identified as a source of phase perturbation.

References
[1] Hariharan P 1996 Optical Holography: Principles, Techniques, and Applications (Cambridge, New York: Cambridge University Press)
[2] Creath K 1988 Phase-measurement interferometry techniques Progress in Optics vol 26, ed E Wolf (Amsterdam: Elsevier) pp 349–93
[3] Wu C-M, Su C-S and Peng G-S 1996 Correction of nonlinearity in one-frequency optical interferometry Meas. Sci. Technol. 7 520–4