

# Metrology for Industry for use in the Manufacture of Grazing Incidence Beam Line Mirrors

DE SC0011289

Final Technical Report:

## Explanation of Specular Reflection Deflectometry

Specular Reflection Deflectometry (SRD) is very similar to the classic Hartmann Test of optical surfaces, with the improvement that the Hartmann screen has been expanded so the entire surface of the Unit Under Test (UUT) is sampled. In the Hartmann Test a point source of light is placed near the center of curvature of the UUT. A screen with equally spaced holes in a grid pattern is placed over the UUT so the light reflected back toward the point source is broken up into uniquely identifiable patches of light. This reflected array of light patches was captured on a photographic plate and the displacement of the light patches was an indication of the slope errors in the UUT. Figure 1 shows a schematic diagram of the classical Hartmann Test.

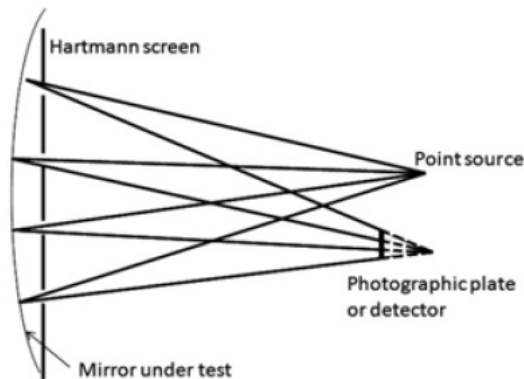


Fig. 1 Layout of the classic Hartmann Test

The Hartmann Test was the principle method used to test astronomical telescope mirrors until the advent of interferometry. The Hartmann Test was used on Mt. Palomar, the largest telescope of its time. However the Hartmann Test has some serious drawbacks for mirrors that require higher precision. First, the mask covers large portions of the mirror so it is not possible to test the entire mirror. Second, the holes in the mask are relatively large so data from each hole is an average of the slope over the entire hole. This makes it impossible to know what the high spatial frequency error in the surface look like.

As we said Specular Reflection Deflectometry is similar to the Hartmann Test. A diagram of the SRD method is shown in Fig. 2. Comparing the two figures it is easy to think of the SRD setup as a reverse of the Hartmann Test. The camera aperture is analogous to the point source of the Hartmann Test and the monitor screen is like the photographic plate. In SRD a pattern on the monitor screen is reflected toward the camera but the only light that reaches the camera detector array is the light that passes through the aperture. This means that knowing the relative positions of the screen, the UUT, and the camera, it is possible to determine the, normal to the mirror surface, at every point on the surface.

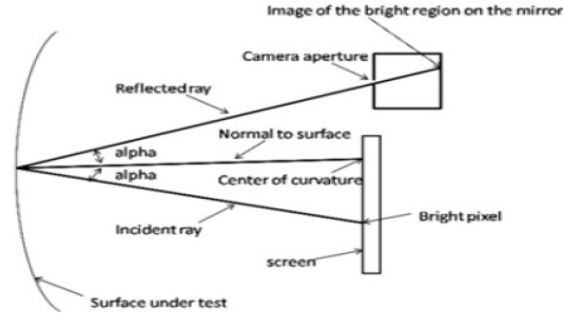


Fig. 2 Diagram of the Specular Reflection Deflectometry Test

The math for using SRD is somewhat complex but it turns out that there is no need to get into the math because if the pattern on the monitor screen is an evenly spaced sinusoidal intensity pattern the image of this pattern looks exactly like the fringe pattern that one would get doing interferometry of a perfect plane surface. This is exactly the place we want to be to determine how well SRD works. We can think of the monitor screen producing an ideal fringe pattern of a plane surface. (It must be kept in mind that we are measuring slope and it takes two fringe patterns in orthogonal directions for complete slope measurement).

How slope information is derived from the fringe pattern on the screen

While the math is a bit complex there are some fundamentals that should be understood. It is possible to measure slopes by putting spots on the screen like a Hartmann Test and measuring their locations relative to their ideal positions. Instead, SDR leverages fringe processing and uses the sinusoidal fringe pattern on the screen as a method of encoding distance along the screen. For example, if we use sixteen pixels to exactly describe one fringe cycle and we know the screen pixels are spaced with a 1 mm pitch then for every one full fringe on the detector we know we are seeing a distance of sixteen mm across the screen. If the camera aperture is one meter from the screen the angle from start to finish of that single fringe is  $16/1000 = 0.016$  radians, or about  $1^\circ$ .

Clearly we need to be able to tell to a much finer scale than individual fringes, or full cycles. To do this we use methods from interferometry, or quadrature detection methods used with linear grating scales. The simplest way of doing this is to obtain four images of the screen, one each with the phase of the fringe shifted by a quarter cycle, or four pixels in our example. In the camera, at each detector pixel the four intensities on the screen are recorded and the phase of the screen fringe pattern is found from

$$\phi = a \tan \left[ \frac{I_2 - I_4}{I_3 - I_1} \right]$$

From interferometry experience with an eight bit camera this intensity data gives the phase to, on the order of, 0.001 cycles, or  $16 \mu$  meter in our example. Now with a one meter distance to the camera we have sensitivity of about  $16 \mu$  radians peak-to-valley. In fact we can do better than this in theory

because we can use eight pixels per cycle and there are displays with pixels whose pitch is 0.137 mm giving a fringe pitch of 1.096 mm or angle of 0.0011 radians and taking 0.001 of that is 1.1  $\mu$  radians, and using a rule of thumb of 5:1 for P-V conversion to RMS would be 0.2  $\mu$  radians RMS. Simply going to a two meter base line gets you there. This is why we were optimistic in our proposal that we could get to at least 0.1  $\mu$  radians RMS.

As we found there are a number of practical considerations that kept us from reaching this goal but they may not be insurmountable. Before discussing these practical matters we will talk about the experiments we performed and the results we recorded.

## Experiments performed

### Direct viewing of the monitor screen

Since all the measurements are based on viewing fringes in reflection from the UUT, we felt it made sense to simply view the screen directly to see how well we could sense the phase, and to see the effect of focus on the results. The monitor was a high resolution computer monitor with a pixel pitch of 0.137 mm. We found we could produce a fringe pattern using only four pixels per fringe but these were not well developed fringes. The minimum practical number of pixels per fringe was eight. Also, the intensity linearity of the monitor was poor and even after trying to correct the linearity we were constrained to about five bits of dynamic range on the monitor out of a theoretical eight bits.

We looked at a 536 x 536 camera pixel patch of the monitor from a distance of about one meter. This size patch was a nominal 512 x 512 with extra room at the edges to do low-pass filtering of the pixel data without edge effects interfering with the 512 x 512 data patch. Since there is a continuous pattern of sinusoidal fringes from left to right the arctangent function will go from  $-\pi/2$  to  $+\pi/2$  every fringe producing a saw tooth pattern in phase. This pattern is then “unwrapped” by adding  $\pi$  to the phase every time the phase drops back down to  $-\pi/2$ . Doing this gives a linearly increasing phase from left to right to first order. In Fig. 3 we show the steps in this process from the sinusoidal fringes to the wrapped phase and on to the unwrapped phase.

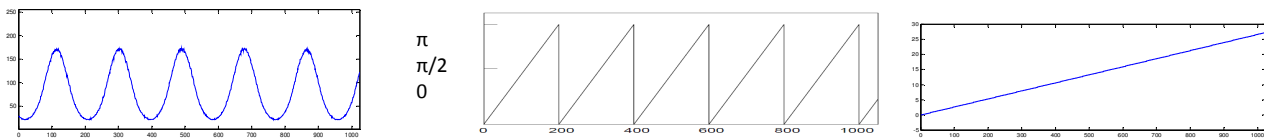


Fig. 3 Evolution of intensity data to unwrapped phase, left, eight-bit sinusoidal intensity, middle, wrapped phase from phase shifted intensity pattern and arc tan algorithm, right, unwrapped phase

The somewhat less than eight-bit dynamic range intensity data displayed on the monitor screen as captured by the camera is shown in the left hand frame of Fig. 3, versus camera pixels. After the four sets of fringe intensity patterns, each shifted by a quarter fringe period, have been used in the arctangent algorithm we get the wrapped phase in the middle frame where the values go from 0 to  $\pi$

versus camera pixel number. Once the phase has been unwrapped by adding  $\pi$  to each discontinuity in the wrapped phase pattern we get a monotonic increase in phase from near 0 to a little more than 5 times  $\pi$  fringes. Assuming as in our original example there are sixteen pixels per fringe and 1 mm pixels, the  $5\pi$  fringes represents a distance of  $5 \times 16 = 80$  mm across the monitor screen.

If all aspects of looking at the screen were “perfect” there should be a strictly linear match between the fringe intensity pattern on the screen and the phase derived from the images captured and processed from the camera data, but the world is not perfect and there will be some non-linearity seen once the phase data has been fit to a straight line, or detrended. In Fig. 4 we show the square set of unwrapped phase data on the left and the same data detrended on the right. This is where we got one of our first surprises. It is clear that something odd is happening in the middle of this frame of data, and the first guess would be an error in the pixel placement on the monitor screen. That was not the case.

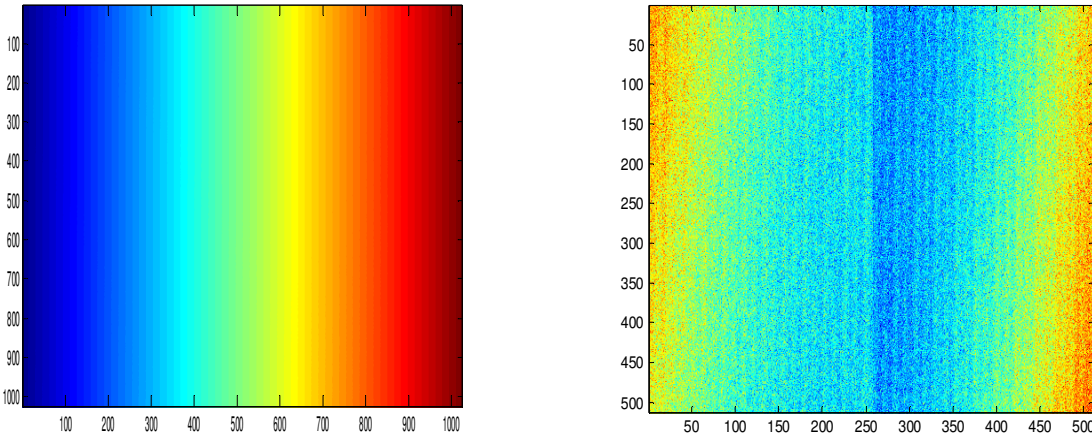


Fig. 4 Unwrapped phase and the same but detrended by removing the best fit plane from left to right

The distinct phase shift in the middle of the frame is due to the camera. The camera we chose for this work was a 4 Mp camera with roughly  $3 \mu$  meter pixels, because we wanted to gather lots of data with high spatial frequency content. It turns out that to read the data out of the camera quickly the detector chip is divided into four sectors vertically and they are read out in parallel but the clocks are apparently not perfectly synchronized giving the apparent shift in phase. In the end, this shift or jump in phase amounts to about  $6 \mu$  radians in slope error because while we are talking about phase the data across the screen are really representative of slope once the distance from screen to camera is factored in.

The other thing that is clear from the detrended phase map is that there is a fine scale error associated with the fringe pattern itself. Depending on the choice of algorithm used to derive phase from the shifted fringe intensity patterns, there is more or less of this fine scale pattern. We did not take the time to study this much but we got the worst results using the four bucket algorithm described above, and used a more sophisticated algorithm to reduce most of our data. Even so, there is some residual systematic error associated with the fringe pattern and we believe it is due to quantization error mainly in the monitor output with some contribution from the camera digitizer.

Finally the map shows high left and right edges and a low center giving the overall map a cylindrical shape. We believe this is due to distortion in the camera lens or due to an apparent shift in the camera pupil with field angle. These effects can be calibrated out of any particular test set up. Fig. 5 shows a slice of eleven rows of phase data taken from the map in Fig. 4.

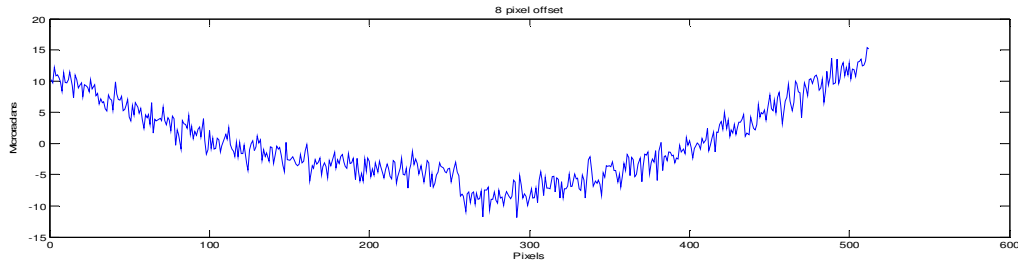


Fig. 5 Profile of phase data converted to slope from Fig. 4

The step in the middle of the phase, or slope, data is clearly obvious in the profile in Fig. 5 along with the cylindrical shape of the overall pattern. Also, the pixel to pixel variation in slope is clear and amount to about  $1.4 \mu$  radians RMS while the P-V variation in slope in this Figure is about  $54 \mu$  radians due to a combination of lens distortion and the camera not looking completely normally at the screen.

#### Effect of averaging data

Next we looked at the effect of averaging data. It helps but not to the extent one would hope. In Fig. 6 we show a single row of phase data, eleven rows of data averaged down columns because we are most interested in slopes running the long direction of mirrors, and an average of sixteen sets of eleven row averages. Clearly averaging helps, but there is still substantial high spatial frequency error that does not average out.

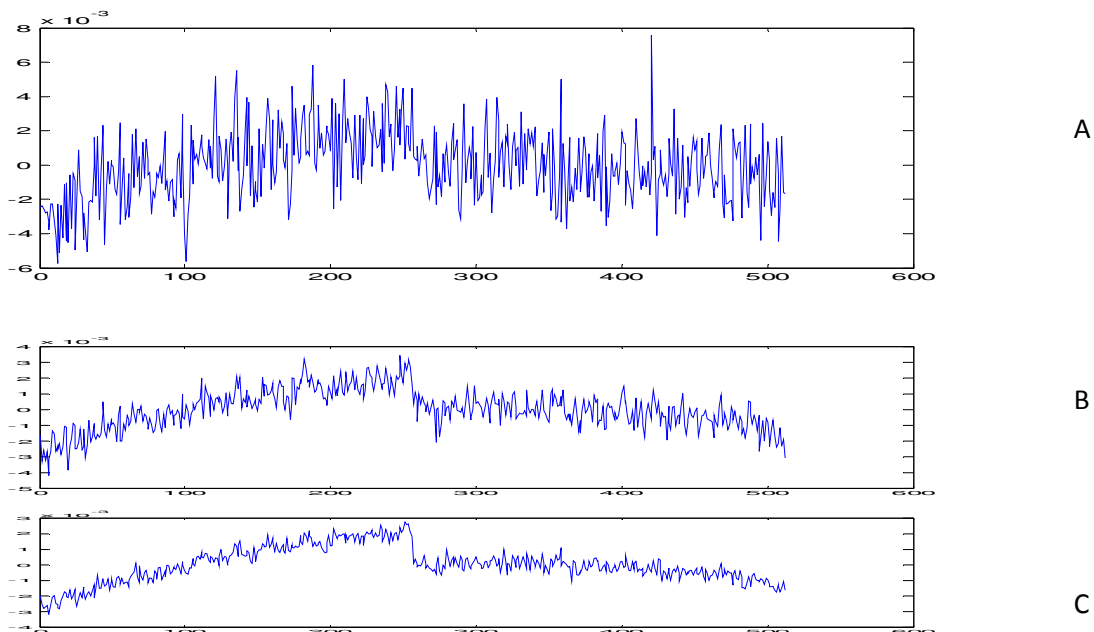


Fig. 6 Effects of averaging; A, one row of phase data, B, eleven rows averaged along columns, C, sixteen sets of eleven row averages. The heights of the graphs have been adjusted so the vertical scales are the same visually.

In Fig. 6 the step in the middle of the data is hidden in the high spatial frequency noise. Averaging eleven rows of data clearly bring out this step, and the additional averaging further reduces the noise. On the other hand, if a low order polynomial curve is fit to the data even after effectively averaging 176 individual measurements the residual RMS is about  $0.75 \mu$  radians RMS, substantially above the desired goal.

If we look at the standard deviation of the polynomial coefficients from the low order fit, they are somewhat smaller than the residual data cited above but still on the order of  $0.5 \mu$  radians RMS, again a bit far from the goal. This though does give a feel for how well the technique will work.

Test of a 300 x 50 mm plane silicon mirror along with a shear test

To get a better feel for sensitivity and repeatability prior to a shear test we looked at just a 536 pixel square region on the camera of the screen with the camera 1 meter from the screen. We used sixteen pixels per fringe so that one fringe was  $2192 \mu$  radians wide with this test arrangement. These data sets were fit to a set of the first five Legendre polynomials in both x and y giving a total of 25 polynomial terms. Legendre polynomials were chosen as a basis set because they are orthonormal over rectangular patches and thus match the geometry of both the camera detector array and the shape of synchrotron mirrors.

We did no averaging but used four frames of data in each experiment. In all we took a first set of data, then turned the lights off above our test setup, then masked the screen so no stray light from the unused parts of the screen could get to the camera, and finally changed from a thirteen to four bucket algorithm to process the phase data. Fig. 7 shows a graph of the Legendre polynomial coefficients of the first and last experiment where the ordinate is  $\mu$  radians. Clearly nothing of substance changed.

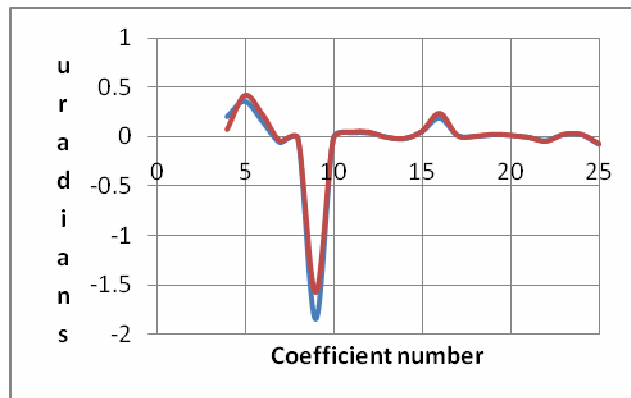


Fig. 7 Legendre polynomial coefficient change from first (red) to last (blue) experimental condition

Next we inserted a 1 mm thick microscope slide close to the camera and on a rotary stage. We wanted a controlled method of displacing the apparent location of the screen pixels. Since the piston term

shows where the phase on the screen starts if we plot the piston term (less the large constant part) and the theoretical shift for rotation of the plate we get the results in Fig. 8. The curves are virtually identical.

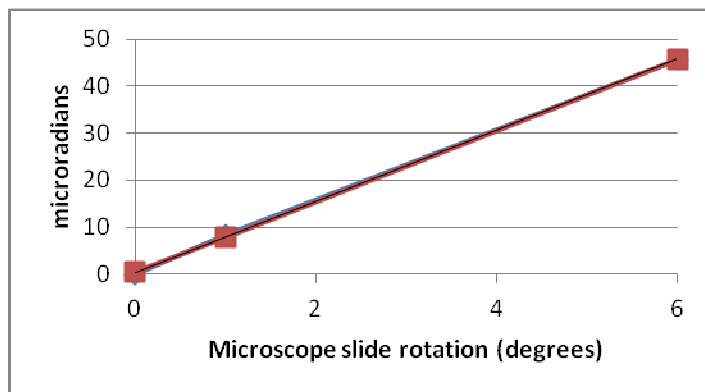


Fig. 8 Pixel displacement in  $\mu$  radians vs. angle for measured and theoretically calculated

To show that this is not just a fluke, Fig. 9 shows the change in the other coefficients due to the rotation. All the other coefficients with the exception of #9, remain essentially constant while 9 changes linearly with angle becoming the largest with the largest rotation of the microscope slide. We did not model what the 9<sup>th</sup> coefficient would be as the slide was rotated but the term was of the correct sort to have changed, a term proportional to  $x^2$ , and the direction of motion. Note that these coefficients are different than those above because the microscope slide introduced some slope changes due to its own errors.

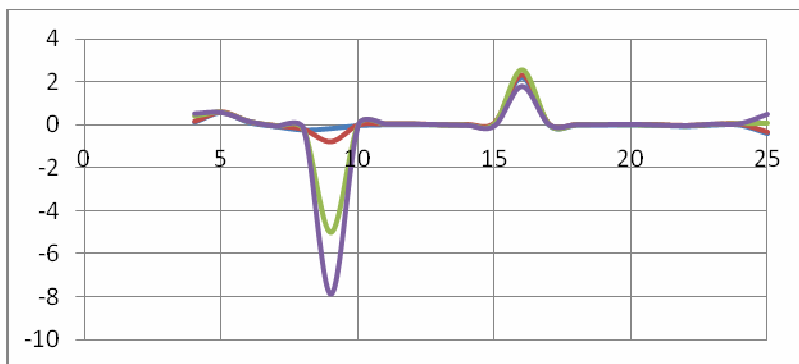


Fig. 9 Change in Legendre coefficients with change in angle of microscope slide, 0 (blue), 1° (red) and 6° ( green)

Finally we returned the window to the 0° position but translated it in the x direction 1 mm. Now the coefficients are decidedly different with the difference showing up largely in the  $x^2$  and  $x^3$  coefficients. The change is due to the errors in the microscope slide. This is proof positive that the SRD Test also works in transmission. It is also proof that we can see differences when the UUT is shifted relative to

the camera and screen on the order of  $1\ \mu$  radian and that the changes seen make sense given the type of perturbation to the test setup.

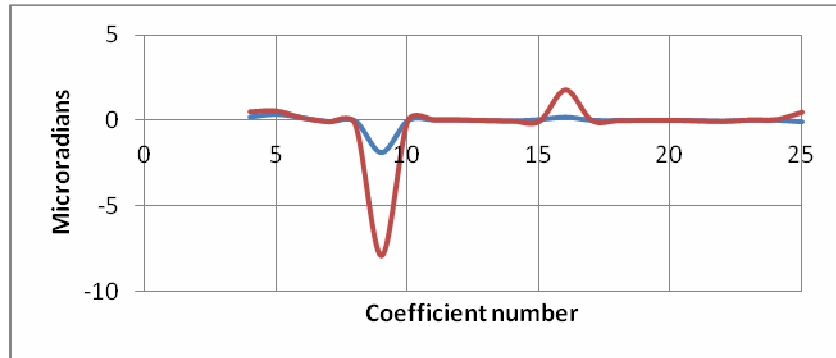


Fig. 10 Change in coefficient value with a 1 mm translation of the microscope slide, initial position (blue), 1 mm shift (red)

Finally, without changing anything we went back and used the thirteen bucket algorithm instead of the four bucket we had been using for most of these examples. Again, there is no change when doing the low order fitting even though we had seen an obvious four cycle ripple in the four bucket data that the thirteen bucket algorithm removed, see Fig. 11.

While the polynomial fitting can only represent low order errors in slope we believe the representation is robust and could prove useful. It is tedious to extend the order of the fit but there is no fundamental problem in doing so. As many as 800 terms have been used to this end at the University of Arizona. Further, for long, thin mirrors only a few terms need to be fit in the narrow direction leaving more terms for the longer direction to map at a higher spatial frequency.

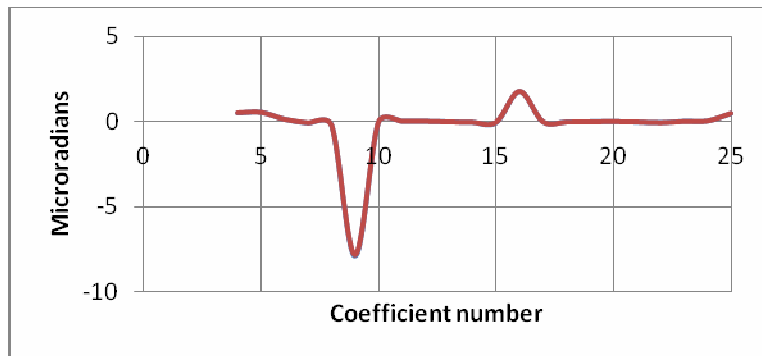


Fig. 11 Coefficients using both four and thirteen bucket phase calculation algorithms

Finally, we took data from the entire 300 x 50 mm plane silicon mirror. Fig. 12 is a close up of the mirror standing on end. This was done because it was easier to adjust angles of screen and mirror than if the mirror were sitting horizontally. The high contrast fringes over the whole mirror surface are clearly visible. Fig. 13 shows the whole test setup with the camera (in the lower right with the red dot) focused on the mirror. The computer monitor that controls the display and camera is to the left, and the monitor with the fringes is seen from the rear. The picture gives a feel for the relative simplicity of



the SRD Test and that most of the costly part of the test is in the software, not the hardware. The advantage of this is that once a solid software package is written that cost can be amortized over many different hardware and test configurations.



Fig. 12 Fringes reflected off 300mm x 50mm Silicon plane mirror



Fig. 13 Overall test SRD set up for Si plane mirror

Once the fringe data was taken and fit to the first 25 Legendre polynomials we get a low spatial frequency slope map in the long direction of the mirror as shown in Fig 14.

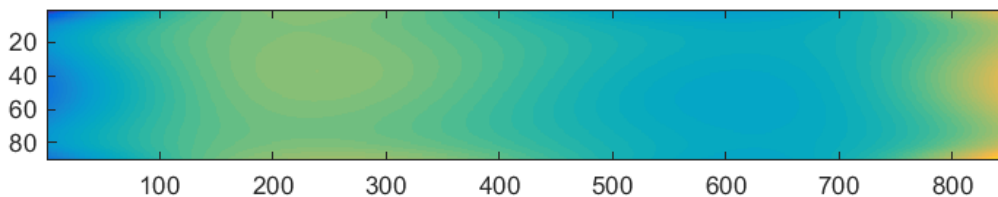


Fig. 14 Slope contour map in the x, or long, mirror direction as represented by low order Legendre polynomials (the numbers represent pixels)

What is shown in Fig. 14 is just the x slope as fit to a double set of the first 5 in x and in y Legendre polynomials, that is, the 25 polynomials show the y as well as the x variation in just the x direction slope. If the fringes had been rotated 90° a similar map could have been produced showing the slope variation in the y, or narrow, direction of the mirror. The map covers virtually the entire surface; a few data points at either end were removed as bad data but accounted for less than 1 mm at either end of the mirror.

The peak-to-valley error in this map is about 4,300  $\mu$  radians but is almost entirely made up of alignment and distortion error in the test set up. This is why it is important to be able to understand

what types of errors as represented by the low order terms represent what types of alignment and distortion errors. On the other hand this background of alignment and distortion error shows the value of performing shearing to understand the true figure of the UUT.

The mirror was shifted upwards 20 mm and another set of data taken. Over the common area in both data sets the data were subtracted to see the change due to the shear. A map of this difference is shown in Fig 15 and has a peak to valley of about 250  $\mu$  radians. It should be said that is the first and only time we did the test and have not had a chance to double check the results.

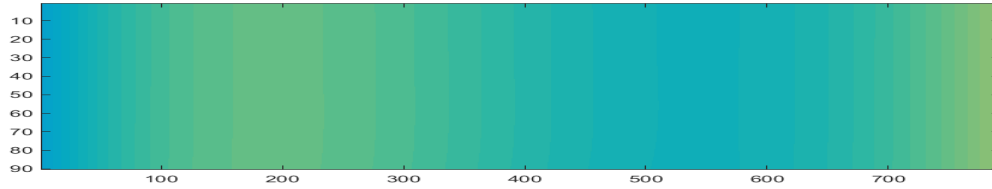


Fig. 15 Difference map between original mirror data and that shifted by 20 mm. The color map was scaled to same peak-to-valley as Fig. 14

If we look at the change in the low order Legendre coefficients as in Fig. 16 we see that there was a change of significance in only 3 coefficients, all of which are a function of  $x$  only, that is  $x^2$ ,  $x^3$  and  $x^4$  with the largest change in  $x^3$ . If there were no higher order errors, when integrated the  $x^3$  slope behavior would lead to a surface error with a  $x^4$  dependence. Because we have not paid attention to the scaling it is difficult to say what this means in terms of the actual height variation.

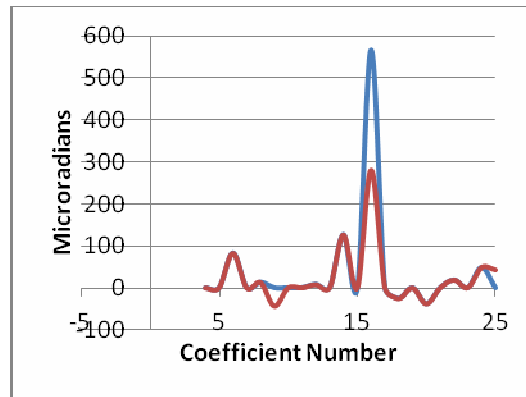


Fig. 16 Change in Legendre coefficients before and after shifting mirror 20 mm

The very promising aspect of the shear data is that only those coefficients that might reasonably be expected to change were the ones that did change. All other remained constant. We feel this shows we are moving in the correct direction in the data analysis.

#### Toroidal Mirrors

One of the big advantages of SRD testing is its intrinsic dynamic range when compared with interferometry. We did not have a sample that was representative of a synchrotron quality mirror, but

we did have a small concave toroid about 30 x 60 mm in aperture with radii of 217 mm x 438 mm. This mirror would have been impossible to measure with an interferometer unless an expensive Computer Generated Hologram (CGH) was created to null the measurement. When this mirror was placed in front of the monitor in a rather arbitrary position we could certainly get perfectly useable fringes from the surface as shown in Fig. 17. The toroidal mirror is about in the center of the picture and the “J” shaped fringes are clearly visible. It is clear that for a mirror this size the monitor should be smaller and the camera lens focal length shorter to get better data. We did not analyze the fringes from this mirror in the interests of time but wanted to show that we could get fringes of equal quality to those on the plane mirror to show there was no fundamental problem with testing toroids or other low f number optics with a SRD setup.

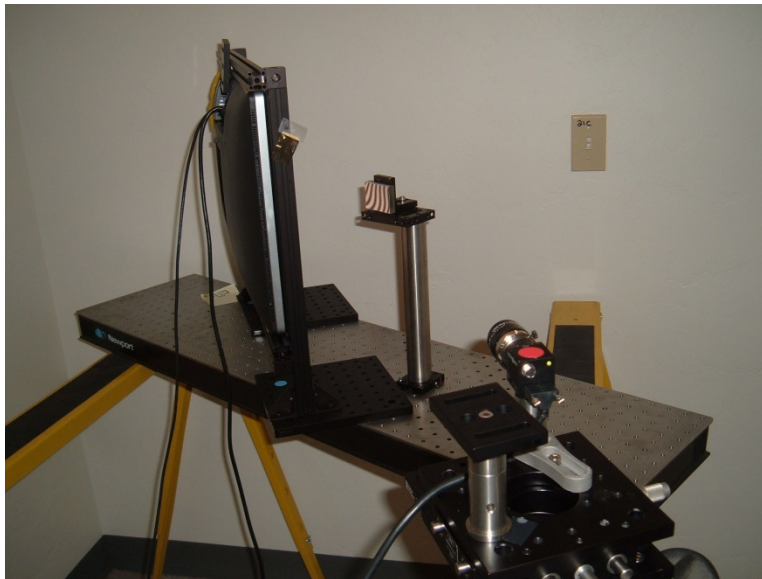


Fig. 17 Test of a small toroidal mirror with the SDR setup. Mirror and fringes are in center of picture

This concludes the experimental part of the study. We have shown that a typical plane synchrotron mirror can be tested to a level of about  $1 \mu$  radian RMS over its entire aperture. We have also shown it is possible to separate the effects of mirror and test by shearing the UUT and that the results are what would be expected. Finally, we have shown that the test also works for toroidal and other fast optics and that the scale of the monitor and camera lens should be adjusted to the size of the part under test.

#### What was learned during this study

- 1) Perhaps the single most important aspect of SRD is that plane or near plane mirrors can be tested with screens, or displays, that are not much larger than the mirror under test. This means that it is perfectly reasonable to test 1.5 meter rectangular mirrors with a 75" class TV display.
- 2) Light is a limiting factor. The high frequency noise is partially due to photon and electron noise. The noise can be reduced by increasing the amount of light available (brighter monitors), which allows

shortening exposure time so more measurements can be averaged (more total photons). This is really a systems problem because we need a large depth of focus to maintain a high modulation of the fringes. This requires a small aperture in the camera, or slow  $f/\#$ , to give the depth of focus but this causes less light to reaching the detector. The best balance of these factors is not easy to figure out but more light from the display makes the systems choices easier and reduces the measurement time.

3) The plane mirror under test should be kept as close to the screen as possible for two reasons; one is the depth of focus issue mentioned above. The closer the mirror is to the screen the less out of focus the fringes will appear when the camera is focused on the mirror as it must be to have high spatial resolution on the mirror. Again, this is somewhat a systems problem depending on the exact test requirements. Second, keeping the mirror close to the screen means the screen can be smaller given the size of the mirror. For example, if the mirror is half the distance from screen to camera, the screen must be twice the size of the mirror to see the whole mirror, while if the mirror is close to the screen, the screen only needs to be 10-20% larger than the mirror.

4) It is easy to use a lens design program such as Zemax to simulate the SRD Test with the understanding that the simulation is modeled backwards. The source of the light, or object, in Zemax is actually the camera aperture and the screen the image plane. A uniform array of rays is traced from the object off the mirror to the screen. For the particular test set up, the array of spots on the screen is what the camera would see if the mirror were perfect. Departures from this spot pattern are the indications of slope errors in the mirror.

5) Because the test setups are easily modeled in Zemax, these departures from a uniform array of spots, or ray intercepts on the screen, can be represented by power series of polynomials. For circular arrays, of mirrors, Zernike polynomials are probably the best choice and are directly available in Zemax. For rectangular mirrors, Legendre polynomials are a better choice of basis functions.

6) Because the arrays of points can be represented by polynomials it is easy to model the effects of misalignment of the test setup. Particular translations and rotations of mirror relative to the screen will effect particular low order polynomial terms in a specific way. This knowledge then helps sort out misalignment of the test from slope errors in the mirror.

7) Another advantage of representing the slope data by polynomials is that they act as low pass filters and remove the high frequency noise in the test data and show the collective behavior of the slopes.

8) A further advantage of the polynomial representation is that the low order surface height data may found by algebraic manipulation of the polynomial coefficients rather than explicit integration of the slope data. Similarly, the polynomial representation provides a straightforward basis for doing stitching of sub-aperture data.

9) When we first looked at the SRD problem we figured that we could get higher slope precision using smaller displays and shorter focal length lenses looking at small patches of the mirror surface. With further thinking, this is probably a false assumption although it was not tested. As size scales are shrunk for looking at small patches of the surface, the distances between screen, mirror and camera

also shrink, and thus any expectation of achieving less than  $\sim 1 \mu$  radian RMS slope resolution on the small scale is not realistic. If one were interested in height variation over small distances there would be an improvement, but not in slope resolution.

10) One very positive aspect of the SRD is that the test seemed quite insensitive to the environment. Whether the air conditioning was running or not seemed to have little effect on the test results. Further, our test setup was not vibration isolated. It was set up in a seismically quiet location, but we could see little evidence that any residual vibration affected the results. Also, ambient lighting had little to no effect. Whether the lights above our test area were on or off did not affect the results, nor did masking the screen so that stray light from the screen would not enter the camera.

11) Another general comment is that at the level of quantization at which these experiments were done, averaging did not have a large effect. As seen in Fig. 6, certainly averaging 10-11 sets of data is worthwhile, but from here on there is a rather slow improvement in the data. At this point it would probably make no sense to average more than 20 or so sets of data. This gives a sort of bound on data collection and time to do a measurement assuming there was sufficient light to minimize exposure time.

12) A final general comment would be that there does not seem to be any “long pole” that is limiting the angular resolution of SRD. It appears that if the test is to be made better there are a lot of small things that each need to be improved; brighter, more linear screen or display to start with. A twelve bit camera would be another improvement and perhaps a further investigation of phase determination algorithms. From there it would be another round of experiments to see the effects of these upgrades.

Also, there needs to be a good look at what the requirements of the test are in terms of slope resolution as a function of spatial period. For example, if mirrors are being polished using classical methods, changes in polishing will have the greatest affect at the lower order characteristics of the mirror’s figure and the mirror finish at the high spatial frequency end. On the other hand, if deterministic polishing is used, then the test must have the most emphasis on the mid to high spatial frequencies in order to guide the polishing process.

The comments so far have been rather general in nature. We have also found a few very specific items for anyone planning on furthering this work.

13) We specifically wanted to use a high resolution camera to attempt to get the highest angular resolution. This was probably, in hindsight, a poor choice. First, as was shown in Fig. 4-6, there was a jump in the slope data in the middle of the frame. After some investigation it turns out that the camera detector array was readout into four parallel buffers in order to handle the large amount of data (4 Mp). Apparently the clocks controlling the four buffers were not precisely synchronized and this led to the step error in the data. We had first thought it was an error in the screen but the jump remained fixed as the camera was moved relative to the screen.

The second possible short coming of the high resolution camera is that the small pixels each collect fewer electrons each than would larger pixels. This leads to a combination of photon and electron

noise that would probably be lessened with a lower resolution camera such as one with only 1 Mp. The readout of a lower resolution camera would certainly be accomplished with a single buffer, eliminating the synchronization issue.

14) Although we spent very little time on this issue, it is clear now that rather than spending money on a higher resolution camera, a camera with a greater dynamic range would have been a better choice. A twelve bit camera rather than the eight bit camera we used would have helped with the noise issue. It is unlikely though that the twelve bit camera would have helped enough to get to the  $0.1 \mu$  radian RMS level. It is relevant to note that the twelve bit camera will cost several times more than the high resolution camera, because of its need for cooling.

15) Although we only investigated three phase calculating algorithms, it was obvious that some were better than others in terms of high frequency noise. The classic four bucket algorithm had a definite four cycle per fringe ripple. This was probably due to quantization error of our 8 bit data. A thirteen bucket algorithm almost completely eliminated this error, although a look at the right hand example in Fig. 4 shows there is still the same four cycle per fringe coherent vertical structure discernable in the slope mapped, again probably due to quantization error.

16) One thing leading to quantization error was the distinct non-linearity of the display screen. Even after attempting to correct for the non-linearity of 8 bit pixel input signal, to actual intensity of the screen, we were only able to utilize about 4-5 bits of intensity level in our fringe pattern. Certainly an improved screen would improve the input fringe pattern for the test. How much this would improve the end result was not estimated.

In the end, if all the improvements suggested were implemented it would be worthwhile repeating these experiments to see what the collective effect of the improvements are on the precision of the test. Some of the improvements suggested could be simulated such as improved screen linearity and a higher dynamic range camera. If this gave evidence of a possibility of  $0.1 \mu$  radian RMS precision then it would definitely be worthwhile carrying on with new experiments.