Chapter 5 – Metrology of Large Parts

H. Philip Stahl, PhD

1.0 Introduction

As discussed in the first chapter of this book, there are many different methods to measure a part using optical technology. Chapter 2 discussed the use of machine vision to measure macroscopic features such as length and position, which was extended to the use of interferometry as a linear measurement tool in chapter 3, and laser or other trackers to find the relation of key points on large parts in chapter 4. This chapter looks at measuring large parts to optical tolerances in the sub-micron range using interferometry, ranging, and optical tools discussed in the previous chapters. The purpose of this chapter is not to discuss specific metrology tools (such as interferometers or gauges), but to describe a systems engineering approach to testing large parts. Issues such as material warpage and temperature drifts that may be insignificant when measuring a part to micron levels under a microscope, as will be discussed in later chapters, can prove to be very important when making the same measurement over a larger part.

In this chapter, we will define a set of guiding principles for successfully overcoming these challenges and illustrate the application of these principles with real world examples. While these examples are drawn from specific large optical testing applications, they inform the problems associated with testing any large part to optical tolerances. Manufacturing today relies on micrometer level part performance. Fields such as energy and transportation are demanding higher tolerances to provide increased efficiencies and fuel savings. By looking at how the optics industry approaches sub-micrometer metrology, one can gain a better understanding of the metrology challenges for any larger part specified to micrometer tolerances.

Testing large parts, whether optical components or precision structures, to optical tolerances is just like testing small parts, only harder. Identical with what one does for small parts, a metrologist tests large parts and optics in particular to quantify their mechanical properties (such as dimensions, mass, etc); their optical prescription or design (i.e. radius of curvature, conic constant, vertex location, size); and their full part shape. And, just as with small parts, a metrologist accomplishes these tests using distance measuring instruments such as tape measures, inside micrometers, coordinate measuring machines, distance measuring interferometers; angle measuring instruments such as theodolites, autocollimators; and surface measuring instruments including interferometers, stylus profilers, interference microscopes, photogrammetric cameras, or other tools. However, while the methodology may be similar, it is more difficult to test a large object for the simple reason that most metrologists do not have the necessary intuition. The skills used to test small parts or optics in a laboratory do not extrapolate to testing large parts in an industrial setting any more than a backyard gardener might successfully operate a farm.

But first, what is a large part? A simple definition might be the part's size or diameter. For optics and diffuse surface parts alike, the driving constraint is ability to illuminate the part's surface. For reflective convex mirrors, large is typically anything greater than 1 meter. But, for refractive optics, flats or convex mirrors, large is typically greater than 0.5 meter. While a size definition is simple, it may be less than universal. A more nuanced definition might be that a large part is any component which cannot be easily tested in a standard laboratory environment, on a standard vibration isolated table using standard laboratory infrastructure. A micro-switch or a precision lens might be easily measured to nanometer levels under a microscope in a lab, but a power turbine spline or a larger telescope mirror will not fit under that microscope and may not even fit on the table.

2.0 Metrology of Large Parts

The challenges of testing large parts are multiple, and they typically involve one or more of the following: infrastructure; gravity sag; stability (mechanical/thermal) and vibration; atmospheric turbulence or stratification; measurement precision and spatial sampling. But, these challenges can be overcome by good engineering practice and by following a structured systems engineering approach. No matter how small or how large your testing or metrology task is; the following simple guiding principles will insure success:

1. Fully Understand the Task
2. Develop an Error Budget
3. Continuous Metrology Coverage
4. Know where you are
5. 'Test like you fly'
6. Independent Cross-Checks
7. Understand All Anomalies

These rules have been derived from over 30 years of lessons learned from both failures and successes. As a validation of these rules, they have been applied with great success to the in-process optical testing and final specification compliance testing of the James Webb Space Telescope (JWST) OTE mirrors (Figure 1). [Ref 1,2]

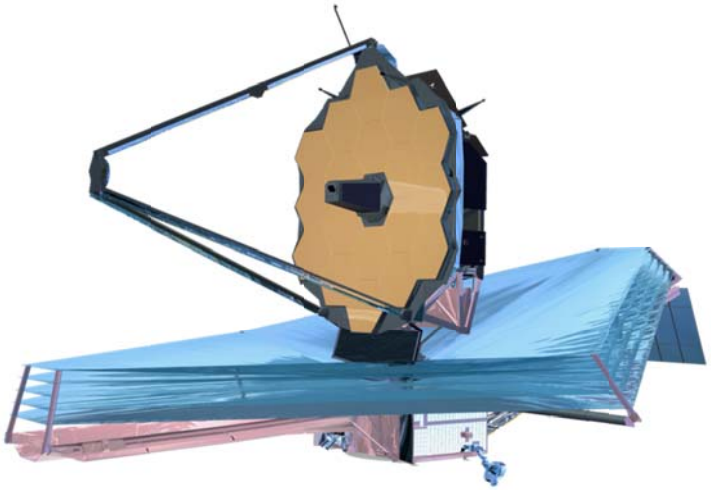


Figure 1: James Webb Space Telescope 6.5 meter primary mirror consists of eighteen 1.5 meter segments.

2.1 Fully Understand the Task

The first step to insure success is to make sure that you fully understand your task. Who is your customer? What parameters do you need to quantify and to what level of uncertainty must must know their value? Do you have the tools and infrastructure to perform the task? And, who is your manufacturing interface?

Before accepting any testing task, study your customer's requirements and understand how they relate to the final system application. Then develop a preliminary metrology plan for how you will quantify each required parameter. This metrology plan should identify the test method to quantify each parameter, the tools and infrastructure required to execute the test and a preliminary estimate of the test uncertainty. We will explore test uncertainty further in the next section. Summarize all requirements and how they will be quantified into a simple table which can be shared with your customer and your manufacturing methods engineer. Make sure that your customer agrees that what you will quantify satisfies their requirements and the manufacturing methods engineer agrees that they can make the part based upon the data you will be providing. Figure 2 shows the final cryogenic temperature requirements for each JWST primary mirror segment assembly (PMSA).

Parameter	Spec	Tol	Units	Verification	Validation
Clear Aperture (Edge Specification)	1.4776 (5)	Min (Max)	mm^2 (mm)	Measure edges at ambient using Tinsley HS Interferometer	Measure area at cryo using XRCF CoC Interferometer
Scratch-Dig	80-50	Max		Ambient Visual Inspection	Independent Visual
Conic Constant	-0.99666	+/- 0.0005		Measured at cryo and defined by null geometry for XRCF CGH CoC test	Ambient test at Tinsley, compare CGH CoC test with auto-collimation test
Radius of Curvature	*	+/- 0.15	mm	Set at XRCF using ADM	ROCO Comparison
Prescription Alignment Error					
Decenter	*	≤ 0.35	mm	Cryogenic test at XRCF, defined by residual wavefront error relative to CGH CoC test and fiducial alignment	Ambient test at Tinsley, compare CGH CoC test with auto-collimation test
Clocking	0	≤ 0.35	mrad		
Piston	N/A			Ambient CMM measurement at AXSYS	Ambient CMM measurement at Tinsley
Tilt	N/A				
Total Surface Figure Error:					
Low/Mid Frequency	20	Max	nm rms	Cryo-Test at XRCF	Cryo-Test at JSC
High Frequency	7	Max	nm rms		
Surface Roughness	4	Max	nm rms	Ambient Chapman measurement at Tinsley	NONE

Figure 2: Primary Mirror Segment Assembly (PMSA) final cryogenic optical performance requirements; the test to verify that each requirement is met and the validation cross-check test for each requirement.

Developing a metrology plan for large parts is complicated by the scale of the required infrastructure. For example, while one can easily transport an 8 cm mirror, an 8 meter class mirror with a 16,000 to 20,000 kg mass requires special transport, lifting and handling fixtures, as well as metrology mounts. But, in practice, any part which cannot be safely lifted by two persons also requires special fixtures and should be considered a large. Safety applies both to the technicians doing the lifting and to the part being lifted. And, sometimes the value of a part is such that it requires special lifting and handling equipment regardless of its size. Furthermore, infrastructure is more than just lifting and handling fixtures. It includes industrial scale work spaces with appropriate temperature, humidity and cleanliness controls; computer coordinate measuring machines and test towers; and grinding and polishing machines. Figure 3 shows an illustration of the Itek Autocollimation Test Facility used for the Keck Telescope 1.8 meter mirror segments. The mirror segments radius of curvature was 24 meters for a total air path of 48 meters. The distance from the mirror under test to the fold flat was approximately 12 meters. [Reference 3] Figure 4 shows the Steward Observatory Mirror Lab (SOML) test tower which stands 24 meters tall and has a mass of 400 tons. [Reference 4] Finally, grinding and polishing equipment is important because their capabilities drive metrology requirements such as spatial sampling, test wavelength and measurement precision.

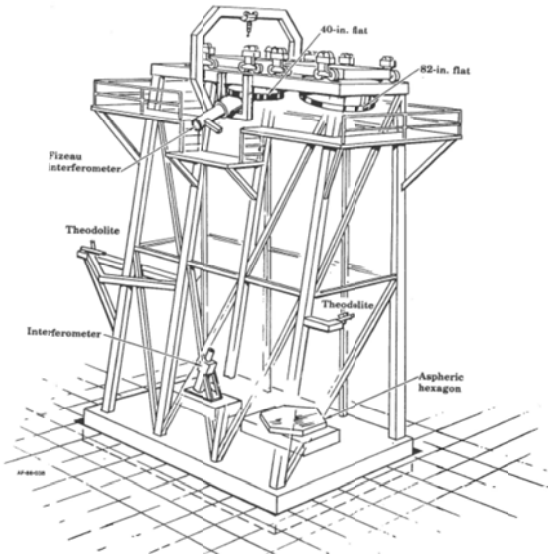


Figure 3: Itek Autocollimation Test Facility. Each Keck segment was tested in over a 48 meter air path. (Figure courtesy of Itek Optical Systems) [Reference 3]

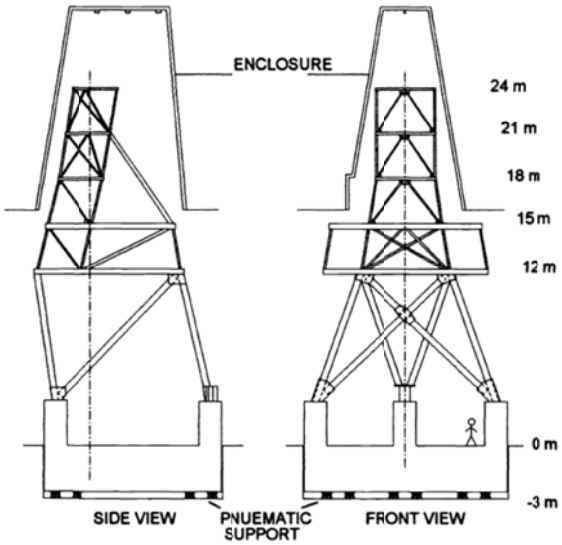


Figure 4: Steward Observatory Mirror Lab test tower. Entire 400-ton concrete and steel structure is supported by 40 air-filled isolators. Drawing by E. Anderson. [Reference 4]

2.2 Develop an Error Budget

The second and most important step is to develop an error budget for every specification and its tolerance. An error budget has multiple functions. It is necessary to convince your customer that you can actually measure the required parameters to the required tolerances. It defines which test conditions have the greatest impact on test uncertainty. And, it provides a tool for monitoring the test process. An error budget predicts test accuracy and reproducibility (not repeatability) of the metrology tools. If the variability in the test data of any element of the error budget exceeds its prediction, then you must stop and understand why. Finally, all elements of the error budget must be certified by absolute calibration and verified by independent test. Figure 5 shows the JWST PMSA high-level error budget for each of its major requirements.

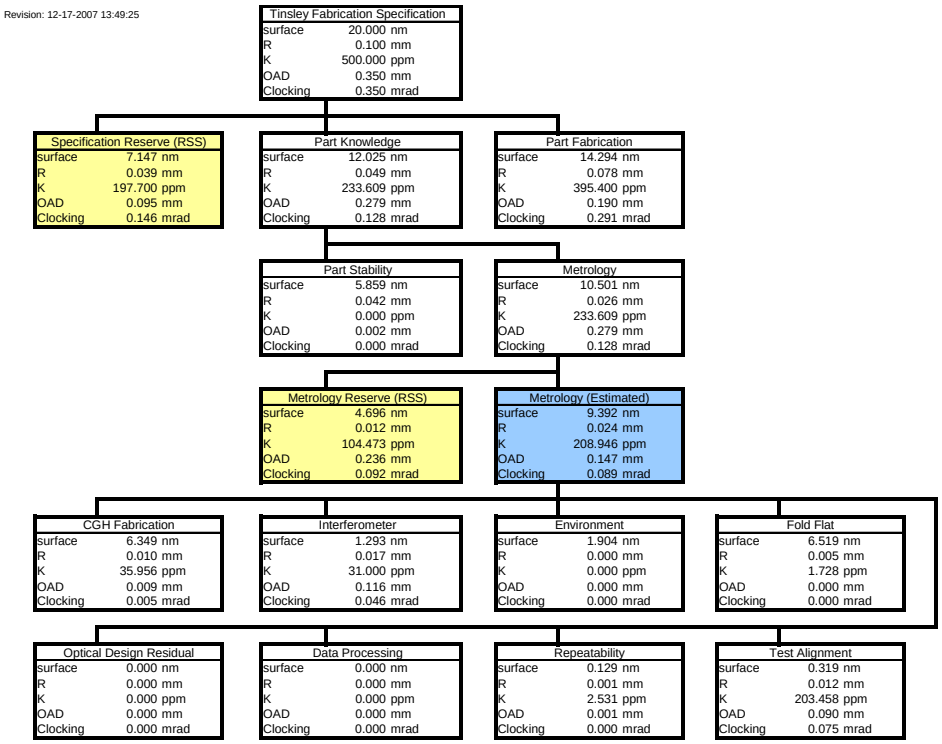


Figure 5: Each JWST PMSA specification had a separate error budget, i.e. surface figure, radius of curvature, conic constant, decenter and clocking of the prescription on the substrate. For every item in this figure, there was a highly detailed error budget.

Mathematically, one constructs an error budget by performing a propagation of error analysis. First write down the equation which calculates the specification value. Then take the partial derivative of that equation as a function of each variable. Square each result and multiple times the knowledge uncertainty (i.e. variance in data) for the measurement of each variable. Then take the square root of the sum. For example, assume that a requirement R is a function of variables (a,b,c), i.e. R = f(a, b, c). The uncertainty of the knowledge of the requirement R is give by:

$$\sigma_R = \sqrt{\left(\frac{\delta f(a,b,c)}{\delta a}\right)^2 \sigma_a^2 + \left(\frac{\delta f(a,b,c)}{\delta b}\right)^2 \sigma_b^2 + \left(\frac{\delta f(a,b,c)}{\delta c}\right)^2 \sigma_c^2}$$

If the defining equation is a linear sum, then the result is a simple root mean square of the individual standard deviations. But, if the equation is not linear, then there will be cross terms and scaling factors.

When building an error budget use the standard deviation of measurement reproducibility not of repeatability. Repeatability will give an 'optimistic' result. Reproducibility gives a realistic result. Repeatability is the ability to get the same answer twice if nothing in the test setup is changed. Reproducibility is the ability to obtain the same answer between two completely independent measurements. [Reference 5, 6] If one is measuring the reproducibility of the ability to align a part in a test setup, then to obtain two independent measurements one must physically remove the part from the test setup and reinstall it between measurements. If one is measuring the reproducibility of atmospheric turbulence, then all that is required is to make sure sufficient time has passed since the last measurement to insure that the two measurements are not correlated.

From a real-world perspective, reproducibility is much more important than repeatability. The reason is that a part is never tested just once. They components are tested multiple times during fabrication. This is commonly called 'in-process' testing. Therefore, the error budget must quantify the knowledge uncertainty of how well the test results can be reproduced from test to test from day to day and even month to month. For example, on JWST PMSAs were not only moved back and forth between manufacturing and test at Tinsley, but also from Tinsley to Ball Aerospace Technology Corporation (BATC) and the Marshall Space Flight Center (MSFC) X-Ray & Cryogenic Test Facility (XRCF). On JWST, a complete understanding of each metrology tool's test uncertainty was critical. Data from Tinsley, BATC and the MSFC XRCF was required to reproduce each other within the test uncertainty. Certified cryo-data must be traceable from the XRCF where they were tested on their flight mount at 30K to BATC where they were changed from the flight mount to the fabrication mount at 300K to Tinsley where they were polished on their fabrication mount at 300K. Accuracy is the ability to get the true answer. The only way to get an accurate measurement is to perform an absolute calibration to quantify any systematic errors which must be subtracted from the data.

Finally, the most important element of an error budget is contingency reserve. All error budgets must have contingence reserve. No matter how much one thinks about every potential risk one cannot think of everything. No matter how carefully one executes the test plan, something will go wrong. Based on many years of experience, a 33% reserve is recommended. Also, don't wait too long to validate the error budget. On the ITTT program (which became Spitzer) this author was responsible for the secondary mirror. A complete error budget was developed, but some elements were allocations. The secondary mirror was manufactured to a Hindle sphere test (Figure 7) and the optician achieved an excellent result. Unfortunately, the Hindle sphere was not absolutely calibrated until it was time to perform the final certification and, to much horror, it had a trefoil gravity sag mount distortion. And, because the secondary mirror had a three point mount, every time it was inserted into the test it was aligned to the Hindle sphere's trefoil error. As a result, the optician polished in three bumps which exactly matched the holes in the Hindle sphere. Fortunately, there was sufficient reserve in the error budget such that the mirror still met its figure specification; it just was no long spectacular. The moral of the story is to not only validate the error budget early. But also, as much as possible, randomize the alignment from test to test. Sometimes bad things happen from been too meticulous. (This could almost be an 8th rule.)

In constructing an error budget for large parts, the three biggest potential error sources are gravity sag, mechanical stability and atmospheric effects. Of these, gravity sag may be the most important because it can be significant and a metrology engineer's intuition often fails to fully account for its effect. The intuition challenge arises from the fact that gravity sag is non-linear. To first order:

$$Gravity\ Sag \quad \propto \quad \frac{mg}{K} \quad \propto \quad mg\left[\frac{1}{E}\left(\frac{D^2}{T^3}\right)\right]$$

where: m = Mass
g = Gravitational Acceleration
K = Stiffness
E = Young's Elastic Modulas
D = Diameter
T = Thickness

Therefore, for constant a thickness, a 2 meter part is 4 times less stiff than a 1 meter part. If they both have the same mass, then the 2 meter part will have about 4 times more gravity sag; and if they both have the same area density, then the 2 meter part will have about 16 times the gravity sag. Thus, for most small parts, their intrinsic stiffness is such that any bending or shape change caused by gravity is negligible relative to the surface figure specification and

thus can be ignored. But, for large parts, gravity sag can be orders of magnitude greater than the surface figure error being measured. For example, an 8 meter diameter, 300 mm thick, solid glass mirror (which must be fabricated to a surface figure requirement of less than 10 nm rms) has an edge supported gravity sag of approximately 2 mm. Now, one would never make or test such a mirror using edge support, but if they did, this amount of sag would not be a problem if the mirror will be used in the same gravity orientation as it is made and tested, but if during operation it is to be tilted with respect to gravity or if it is going to be used in space, then the sag must be quantified and if necessary removed from the data.

The key to testing large parts is that the metrology mound must simulate the part's 'as-use' gravity orientation or operational support system. The problem is that metrology mounts are not perfectly repeatable. And, the less stiff the part under test, the more its gravity sag might vary from test to test. When testing large parts, it is desirable to design a metrology mount with sufficient stiffness to hold the part under test such that the uncertainty in its gravity sag knowledge is 10X smaller than the surface figure specification. For example, if the mirror surface figure requirement is 10 nm rms, then the metrology mount should support the mirror in a known orientation with respect to gravity with an uncertainty of less than 1 nm rms. To accomplish this task requires a support structure which is both mechanically (and thermally) stable and introduces known predictable stress/strain and force loads into the part under test. As the part size increases metrology mounts and handling fixtures become more complicated.

Mechanical stability and vibration errors must be included in any error budget. Small parts are typically tested on a small vibration isolated table with sufficient stiffness to maintain micrometer level test alignment for arbitrary periods of time. But, large test setups require large structures. And, for structures sometimes 10s of meters in size, it can be difficult to achieve micrometer (and/or micro-radian) alignment stability between components. Furthermore, at such sizes, the structural material's coefficient of thermal expansion can cause the test setup to 'breath' as a function of room temperature. When operating at large scale, test uncertainty is impacted by static and dynamic stability.

Static stability is the ability of the structure to maintain the alignment of the test elements relative to each other for long periods of time. Insufficient static stability manifests itself in systematic or even unpredictable drifting of the test alignment during the measurement period. Static stability is also the ability to repeatedly position the test elements in the aligned state from test to test. Static instability primarily occurs when strain, which is introduced via mechanical pre-load or misalignment or thermal gradients, is released via stick/slip motion. As a rule of thumb, a test setup should be designed such that the ability to repeatedly position the part under test is sufficiently precise that the uncertainty is 10X smaller than the parameter to be measured. Similarly, any error introduced by drift in the test setup should be 10X smaller than the parameter to be measured.

Dynamic stability is vibration and it can be driven by either seismic or acoustic sources. Small test structures tend to be very stiff and have first mode frequencies which are much higher than the measurement period. If the vibration is at least 10X higher than the data acquisition rate, then their effect will average to zero – with a small reduction in data 'contrast' due to blurring. [Reference 7] But, large structures can have first mode frequencies which are on the order of 10s to 0.10s of Hertz. For example, the SOML test tower moves as a rigid body with a resonance of about 1.2 HZ and an internal first mode of 9.5 Hz. [Reference 4] Motions in these frequency bands can introduce significant measurement errors. To minimize these errors, it is necessary to minimize the amplitudes of their motions. This is done by vibration isolating the test structure from the ambient environment. One way (as shown in Figure 3) is to bury in a sand pit a very thick concrete slab on which the test structure is setup. The sand dampens vibrations from being propagated from the building into the test structure. As shown in Figure 4, the sand can be replaced via pneumatic supports. A third approach is to build large support legs which are physically attached to the building with pneumatic supports at the top from which the test structure hangs.

Regardless of the approach used, it is virtually impossible to eliminate all vibrations. Therefore, additional means are needed to minimize their impact. The Hubble Space Telescope program mitigated vibration errors by acquiring and averaging many short exposure measurements. [Reference 8] Short exposure measurements 'freezes' the vibration error. And averaging reduces the error contribution to zero because vibration is Gaussian normal (i.e. has a mean value of zero), but it only works if enough measurements are acquired over a long enough time (i.e. over several periods of the vibration) to yield a statistically significant zero mean average. Another approach is to optically or structurally connect the test components such that the vibrations are synchronized. If every test element sees the same vibration such that there is no relative motion, then there are no measurement errors. The Keck

segments were tested in the presence of significant vibration by employing a common path technique. The Twyman-Green reference beam was transmitted along side of the test beam, reflecting from the auto-collimating fold flat three times and a small flat physically attached to the segment under test (Figure 6). [Reference 3] Another trick is to synchronize vibration between test components by structurally connect them. Figure 7 shows the Hindle Sphere test setup used to test the Spitzer secondary mirror. A 2x4 board is connecting the Hindle sphere with the Fizeau interferometer phase modulator. On JWST, it was necessary to characterize the primary mirror segment assemblies (PMSAs) at 30K. This was done by testing them horizontally inside the MSFC x-ray cryogenic test facility (XRCF) with the optical test equipment located outside the chamber at the 16 meter center of curvature (Figure 8). State of the art commercial temporal phase measuring interferometers could not measure the mirrors to the required precision because low frequency structural bending introduced 0.5 mm of piston motion between the PMSAs and the test equipment. To solve this problem, MSFC funded the development of the 4D Vision Systems PhaseCAM instantaneous phase measuring interferometer (Figure 9). [References 9 & 10]

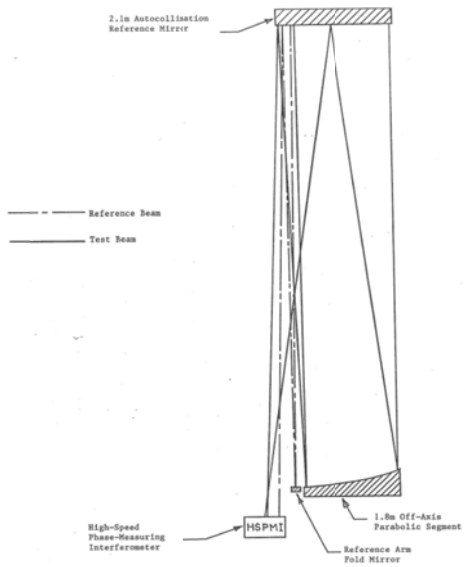


Figure 6: In the Itek Auto-collimation Test Facility, vibration errors between the Keck segments and the auto-collimating flat were minimized by physically attaching a small flat mirror to the Keck segment and bouncing the reference beam off of the auto-collimating flat. [Reference 3]

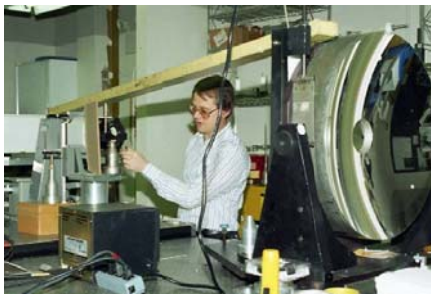


Figure 7: Hindle Sphere test setup to measure the Spitzer Telescope secondary mirror. A 2x4 is used to structurally connect the interferometer phase modulation head and Hindle sphere to minimize vibration induced relative motion. (Photo Courtesy of Goodrich Corporation)

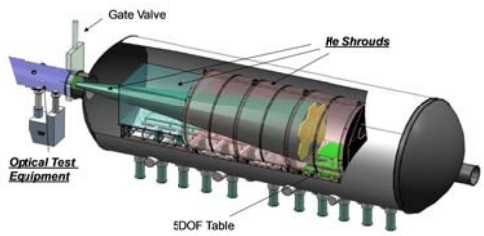


Figure 8: JWST Primary Mirror Segments were tested at meter center of curvature. Because the center of curvature optical test equipment was on a different isolated concrete slab than the

Figure 9: 4D Vision Systems PhaseCAM instantaneous phase measuring interferometer.

Atmospheric turbulence and atmospheric stratification are also important error budget elements. These effects may be easier to understand because they can be seen. Anyone who has ever driven down a hot highway and observed the shimmering thermal boundary layer has an intuitive understanding of its affect. Or, anyone who has stuck a hand into an optical test beam has seen how rising heat distorts the fringes. Thermal variation causes measurement errors because the refractive index of air varies as a function of temperature. A simple illustration of how this can be a problem is if a small pocket of cooler air (which is more dense and with a higher index) moves across an optical surface, it appears as a 'hole' in the surface figure. A more accurate explanation is that optical rays traveling through different parts of the atmosphere with different temperatures experience a differential optical path length error. But, turbulence flow is difficult to model and is another area where an optical metrologist's intuition is frequently inadequate. The challenge for large optics is that, for a constant F/# component, air path volume increases as the cubed power of aperture diameter. Also, while mechanical vibrations are typically periodic, turbulence is chaotic.

Stratification occurs when air forms layers of different temperature, typically cold on the bottom and hot on the top, but temperature inversions are also possible. Normally, one sees this effect in air that is still or not moving, but it can also occur in laminar flow (which is defined as parallel flow with no lateral mixing). Because refractive index varies as a function of temperature, light going though the colder layers has a longer optical path length than light going through the warmer layers. Thus, based on the geometry of how the light traverses the layers, wavefront errors can be introduced by the atmospheric stratification. If linear stratification occurs in a parallel optical beam it introduces a tilt error which can be ignored. But, if linear stratification occurs laterally (perpendicular to the optical axis) in a diverger/converging beam, it acts like a tilted plate and introduces an astigmatic wavefront error. If linear stratification occurs axially along a diverging/converging beam, it acts like a gradient index lens and introduces power (or focal length change) and a small amount of spherical wavefront (or conic constant) error. An analysis of the Gemini 6.5 meter F/11.25 primary mirror predicted that a 0.5 C top to bottom gradient would produce a 2 ppm (parts per million) conic constant error and a 0.3 ppm radius error. [Reference 4] In general, it is best to avoid stratification. An interesting exercise for the reader is to setup a center of curvature test in the laboratory. Take and save a measurement then 'tent' the test and wait for stratification to occur. Then take another measurement and subtract the first. For best results, use a mirror that is larger than 0.5 meters.

Turbulence is caused by the convective flow of warmer/cooler air pockets moving through ambient air (or lateral mixing and eddy current mixing at air temperature boundaries). Because refractive index varies as a function of temperature, pocket to pocket (or across boundaries) temperature differences manifests themselves as measurement errors (caused by differential optical path length variations). These fluctuations can be distributed laterally as well

as axially along the test beam. These pockets can be large and moving slowly, or (with increased mixing) they can be slow and moving rapidly. This size and rate of motion is described by diffusion, the greater the mixing or the more turbulent the flow, the shorter the diffusion length.

Ideally, the best test environment is an atmosphere with no temperature variation. In such a case, even if there was significant air flow, there would be no optical turbulence. But, such an environment is difficult to achieve. Typical air handling systems are good to 1C. The Hubble program solved the atmospheric turbulence problem by testing the primary mirror at center of curvature in a vertical vacuum chamber (Figure 10a). [Reference 8]

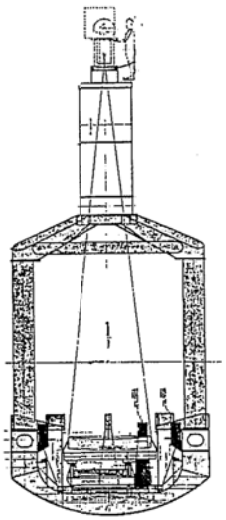


Figure 10a: The Hubble primary mirror was inside a vacuum chamber to eliminate atmospheric turbulence as a source of measurement error. [Reference 8] Figure 10b: Hubble primary mirror being loaded into the vertical test chamber [photo courtesy of Goodrich Corporation].

When it comes to optical testing in air, there many different opinions. Some think that the best approach is to maximize turbulent mixing to minimize the size of pockets (diffusion length). Others believe that you should stop the mixing and test as soon as the air becomes quiet but before it becomes stratified. Some believe that the air should flow along the optical axis while others believe that it should flow perpendicular to the optical axis. This author recommends perpendicular flow with maximum turbulent mixing. The problem with axial flow is that pressure gradients can form in front of the mirror and eddy current vortices can be produced around the edge. The best test environment that this author has even experienced was a 10 m by 20 m room whose air flowed from one end to the other, was exchange approximately every 5 minutes and was controlled to 0.01 C. [Reference 11] The next best test environment was the Ball Aerospace Technology Corp (BATC) Optical Test Station (BATC) for the JWST PMSAs (Figure 11). [Reference 12] Each PMSA was tested at center of curvature in a thermally insulated test tunnel. Thermally controlled air was flowed down the tunnel with fans producing vertical mixing.

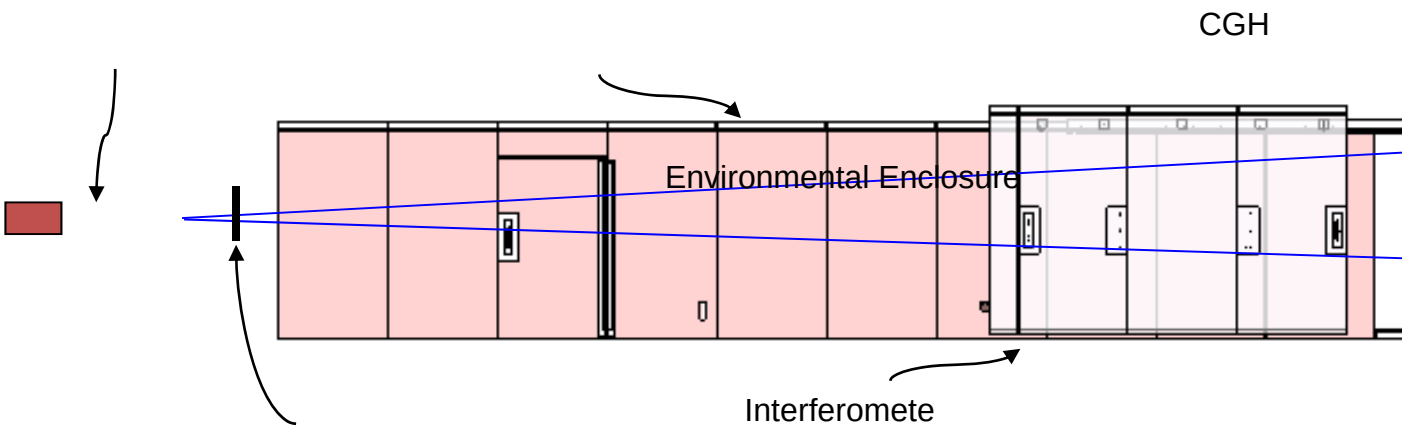


Figure 11: Ball Optical Test Station (BOTS) for ambient environment testing of JWST Primary Mirror Segment Assemblies. [Reference 12]

An important fact to understand about testing in ambient atmosphere is that turbulence is not statistically random. It does not average to zero. Rather, atmospheric turbulence is chaotic with a diffusion length. Thermal pockets are 'correlated' with each other axially and laterally. Therefore, one cannot eliminate atmospheric turbulence errors simply by taking lots of short exposure measurements and averaging (as one does for vibration). And, according to the ergodic principle, the temporal variation along an optical path has the same statistical properties as the spatial turbulence. Thus, two measurements separated in time by less than the diffusion time are correlated and, therefore, averaging them will not yield a 'zero' error. Rather, averaging correlated measurements yields a low order error. The only way to eliminate atmospheric turbulence effects is to average measurements which are acquired at time intervals longer than the diffusion or correlation time. And, the only way to obtain short diffusion times is a highly mixed, highly turbulent atmosphere.

2.3 Continuous Metrology Coverage

The old adage (and its corollary) is correct: 'you cannot make what you cannot test' (or 'if you can test it then you can make it'). The key to implementing these rules is simple. Every step of the manufacturing process must have metrology feedback and there must be overlap between the metrology tools for a verifiable transition. Failure to implement this rule typically results in one of two outcomes, either very slow convergence or negative convergence.

Overlapping metrology coverage requires tools which can precisely measure large dynamic ranges, for a range of surface textures during different fabrication processes, and over a range of different spatial frequencies. Regarding measurement precision and range, it is much easier to measure a 1 meter radius of curvature to a precision of 10 micrometers than it is to measure a 10 meter radius of curvature to a precision of 10 micrometers (or even to 100 micrometers). The metrology tools designed to make such precision measurements can have range limitations. Also, as distances become greater, all of the previously discussed problems such as mechanical stability and atmospheric turbulence affect precision. Another, well know but subtle effect, is the Abbe sign error if the radius measurement is not being made directly on the optical axis of the component. Fortunately, the dimensional tolerances for large optics are frequently more relaxed than for small optics.

Large parts go through a variety of manufacturing processes, from machining to rough grinding to fine grinding to polishing and figuring. Each process has a different surface texture and different precision and dynamic range requirements. Typically, coarse metrology is done via a profilometer for machining and grinding operations and an interferometer for polishing and figuring. The problem comes in making the transition from grinding to polishing. Coordinate Measuring Machines (CMMs) are great for machining and rough polishing. They have large dynamic ranges and work well with 'mechanical' surfaces, i.e. surfaces which are not smooth enough to reflect light. The primary issue for large optics is getting a CMM with a sufficiently large measurement volume. A secondary issues is that the larger the measurement volume, the more difficult it is to obtain high precision. And, high precision is what drives the overlap problem. A CMM with a 0.100 mm rms measurement uncertainty cannot provide a good

metrology hand-off to optical interferometry. To achieve good overlap with optical interferometry requires knowledge of the surface shape under test to an uncertainty of approximately 0.010 mm or 10 micrometers rms. Traditionally, this gap has been filled with infrared interferometry [Reference 13], but improvements in CMM precision will eventually allow for direct transition to optical interferometry. (CMMs capable of 8 to 10 meter are expensive. So the choice may be aperture dependent.)

For JWST, Tinsley developed overlapping metrology tools to measure and control conic constant, radius of curvature, prescription alignment and surface figure error throughout the fabrication process. During rough grinding this was accomplished using a Leitz Coordinate Measuring Machine (CMM) (Figure 12). The CMM was the primary tool used to establish radius of curvature and conic constant. While these parameters can be adjusted in polishing, it is much easier to set them during grinding. During polishing, metrololgy was provided by a Center of Curvature (CoC) interferometric test. Ordinarily, optical fabricators try to move directly from CMM to optical test during fine grinding. But, given the size of JWST PMSAs and the mid-spatial frequency specification, this was not possible. Bridge data was provided by a Wavefront Sciences Scanning Shack Hartmann Sensor (SSHS) (Figure 13). Its infrared wavelength allowed it to test surfaces in a fine grind state. And, its large dynamic range (0 to 4.6 mrad surface slope), allowed it to measure surfaces which were outside the interferometer's capture range. The SSHS is an auto-collimation test. Its infrared source is placed at the focus for each PMSA prescription (A, B or C) to produce a collimated beam. An infrared Shack-Hartmann sensor is then scanned across the collimated beam to produce a full aperture map of the PMSA surface. The SSHS was only certified to provide mid-spatial frequency data from 222 to 2 mm. When not used, convergence was degraded. Figure 14 shows an example of the excellent data agreement between the CMM and SSHS.

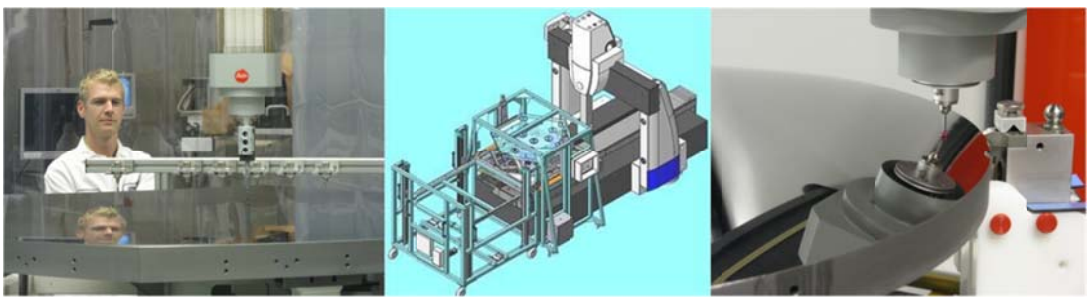


Figure 12: Leitz Coordinate Measuring Machine (CMM) was used at Tinsley during generation and rough polishing to control radius of curvature, conic constant and aspheric figure for Primary Mirror Segment Assemblies, Secondary Mirrors and Tertiary Mirror.

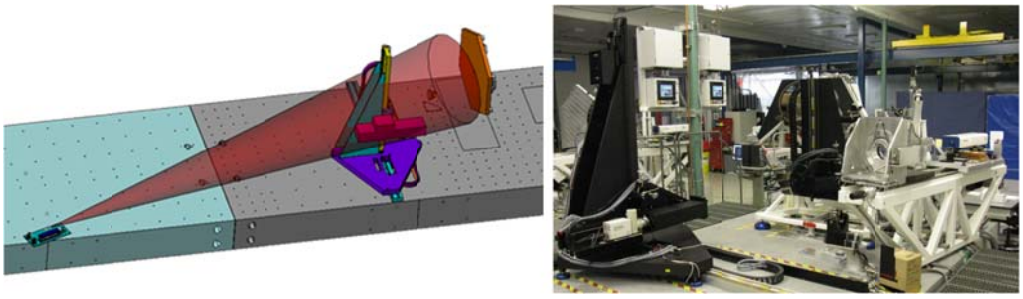
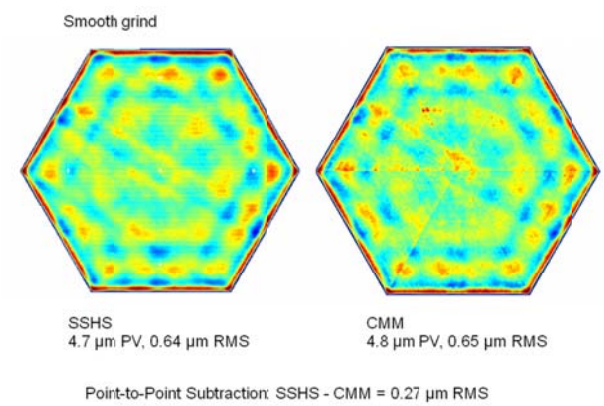
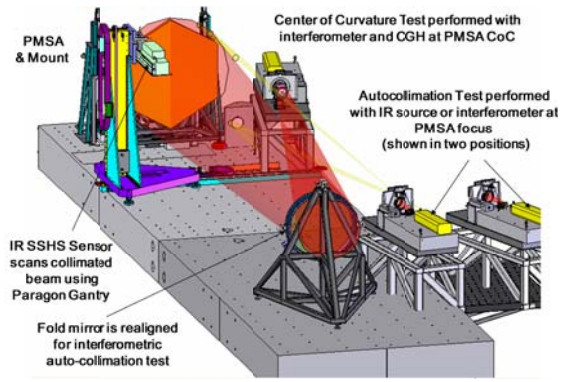


Figure 13: Scanning Shack Hartmann Sensor (manufactured by Wavefront Sciences) is an auto-collimation test. A 10 micrometer source is placed at focus and a Shack-Hartmann sensor is scanned across the collimated beam. There are three different source positions for the three PMSA off-axis distances. Photo on right shows the sensor (white) mounted on the Paragon Gantry (black).





6 DOF Test Stand and Mirror

In additional to dynamic range and fabrication process stage, spatial sampling metrology overlap is also important. As the part becomes more and more perfect, it is necessary to control smaller and smaller features. High resolution spatial sampling is needed to drive the polishing process. It is especially important if the optical component is an asphere. A common fabrication process for aspheric optics and for large optics is small tool computer controlled polishing. But, the size of the tool which can be used is limited by the spatial sampling of the metrology data. [Reference 14] If one has an 800 pixel interferometer taking data on a 0.8 meter component, then one has 1 mm spatial sampling. According to the Shannon Sampling Theorem, this should be sufficient to correct 2 mm spatial period errors, but in practice it is only good enough for 3 to 5 mm spatial frequency errors. Extrapolating to larger apertures, an 800 pixel interferometer taking data on an 8 meter mirror has 10 mm spatial sampling which can control 30 to 50 mm spatial frequencies. Depending upon the mirror's structure function specification, i.e. its required surface figure vs spatial frequency, such a spatial sampling may or may not be sufficient. Additionally, segmented telescopes have edge requirements. On JWST, the polished optical surface needed to meet its specification to within 7 mm of the physical edge. While the JWST center of curvature interferometer had a projected pixel size of 1.5 mm and should have been able to resolve a 4.5 to 7.5 mm edge, it could not.

On JWST, grinding and polishing feedback was provided by a custom built optical test station (OTS) (Figure 15). The OTS is a multi-purpose test station combining the infrared SSHS, a center of curvature (CoC) interferometric test with a computer generated hologram (CGH) and an interferometric auto-collimation test. This test simultaneously controls conic constant, radius of curvature, prescription alignment and surface figure error. The CoC test pallet contains a 4D PhaseCAM, a Diffraction International CGH on a rotary mount and a Leica ADM. The ADM places the test pallet at the PMSA radius of curvature with an uncertainty of 0.100 mm which meets the radius knowledge requirement. Please note that this uncertainty is an error budget built up of many contributing factors. Once in this position, if the PMSA were perfect, its surface would exactly match the wavefront produced by the CGH. Any deviation from this null is a surface figure error to be corrected.

2.4 Know Where You Are

It might seem simple, but if you don't know where a feature is located on the part, you cannot correct it. To solve this problem you must use fiducials. There are two types of fiducials: Data Fiducials and Distortion Fiducials. Data fiducials are used to define a coordinate system and locate the measured data in that coordinate system. Sometimes this coordinate system is required to subtract calibration files, other times it is required to produce hit maps. Distortion fiducials are used to map out pupil distortion in the test setup. Many test setups, particularly those with null optics can have radial as well as lateral pupil distortion. Distortion can cause tool mis-registration errors of 10 to 50 mm or more.

Fiducials can be as simple as a piece of tape or black ink marks on the surface under test or as sophisticated as mechanical 'fingers' attached to the edge protruding into the clear aperture. Tape fiducials are acceptable for simple

reproducibility or difference tests or to register a calibration alignment. But, they are not recommended for computer controlled process metrology. In these cases, fiducials define your coordinate system and need to be applied with a mechanical precision of greater accuracy than the required prescription alignment to the substrate. Additionally, because the interferometer imaging system might invert the image or because fold mirrors in the test setup might introduce lateral flips, an asymmetric pattern is highly recommended. A good pattern to use is one with fiducials at 0, 30 (or 120), 90, and 180 degrees. The 0/180 degree fiducials produce a central axis for the data set. The 90 degree fiducial defines left/right and the 30 degree fiducial defines top/bottom. Additionally, for test setups with null optics, pupil distortion can be a problem. In these cases, distortion fiducials are required. One option is to place multiple fiducial marks along a radius. For null tests with anamorphic distortion, a grid of fiducial marks is recommended. Finally, if one has a clear aperture requirement, make sure to place fiducial marks inside and outside of the required clear aperture distance, this way it can be certified whether or not the requirement is achieved.

Another problem is software coordinate convention. Most interferometer analysis software assumes that the optical (Z axis) positive direction points from the surface under test towards the interferometer, such that a feature which is higher than desired is positive. But, many optical design programs define the positive optical axis to be into the surface. The problem occurs because both programs will typically define the Y-axis as being up, so it is critical to understand which direction is +X-axis. The problem is further complicated when interfacing with the optical shop. To avoid doubling the height or depth of a bump or hole because of a sign error, or adding a hole or bump to a surface because of a coordinate flip or inversion, a good metrologist must know the coordinate system of every computer controlled grinding and polishing machine in the optical shop.

On JWST, the CoC null test simultaneously controls the PMSA conic, radius, figure and prescription alignment. The key is knowing where the prescription is on the substrate and knowing where the prescription is in the test setup. Prescription alignment (off-axis distance and clocking) is controlled by aligning the PMSA into the test setup with an uncertainty which is smaller than the decenter and clocking tolerances. PMSAs are manufactured in Observatory Coordinate Space as defined by 'Master Datums' on the back of each substrate. The optical surface figure is registered to the mirror substrate and to the observatory coordinate system via data fiducials placed on the front surface of each mirror. The CMM is primary in establishing compliance with prescription alignment. Starting with the master datums, the CMM defines 'transfer' fiducials on the side of the mirror. Then, the CMM establishes the data fiducials based on these secondary fiducials. Figure 16 shows fiducialized mirrors being loaded into the MSFC XRCF for cryogenic testing. Some of the mirrors have only the data fiducials. Others of the mirrors have both data fiducials and distortion fiducials (2D grid of dots). Distortion fiducials are necessary to compensate for anamorphic distortion introduced by the CGH.

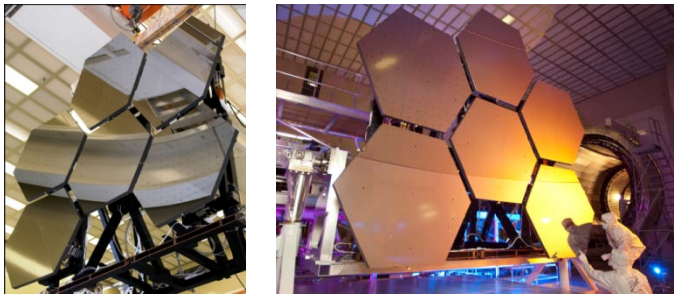


Figure 16: PMSA mirrors with Data and Distortion Fiducials are ready for loading into the MSFC XRCF.

2.5 Test like you Fly

'Test like you fly' covers a wide range of situations, and of course, for ground applications this rule could be 'Test like you use'. Whenever possible, the part should be tested in its final mount, at its operational gravity orientation and at its operational temperature. While gravity is typically not a problem for small stiff optics, it can be a significant problem for large optics. Any optical component going into space needs to be tested in a 'zero-g' orientation. This is typically accomplished by either averaging a cup-up/cup-down test to remove the concave/convex gravity sag contribution, or by averaging a horizontal multiple rotation test to remove mount induced bending. [Reference 15] Gravity sag can be every significant for very large ground based telescopes. In this case, the best approach is to test them in their final structure (or a suitable surrogate) at an operational gravity

orientation. The one thing that a good metrologist should avoid is agreeing to test a very low stiffness mirror without a final support system. The reason is that it will be virtually impossible to achieve a stable, repeatable measurement. With such mirrors, simply picking it up and setting it back down on the metrology mount might result in unacceptable shape changes. Finally, it is important to test a part under its intended atmospheric pressure and temperature conditions. If a lightweight mirror intended for use in vacuum does not have proper venting paths, it can result in a damaged mirror. And, a mirror intended for use at a cryogenic temperature can have very large coefficient of thermal expansion (CTE) induced figure changes. In such cases, it is necessary to characterize these changes and generate a cryogenic 'hit' map to 'correct' the surface figure for 'at-temperature' operation.

Because JWST mirrors were fabricated at room temperature (300K) but will operate in the cold of space (< 50K), it is necessary to measure their shape change from 300 K to 30K, generate a 'hit-map', and cryo-null polish the mirrors such that they satisfy their required figure specification at 30K. After coating, all mirrors underwent a final cryo-certification test of conic constant, radius of curvature, prescription alignment and surface figure error. These tests were performed at Marshall Space Flight Center (MSFC) in the X-Ray and Cryogenic Test Facility (XRCF) show in Figure 17. Additionally, because JWST operates in the micro-gravity of space but is manufactured in the gravity of Earth, it is necessary to removed gravity sag from the measured shape. This is accomplished using a standard 6 rotation test.

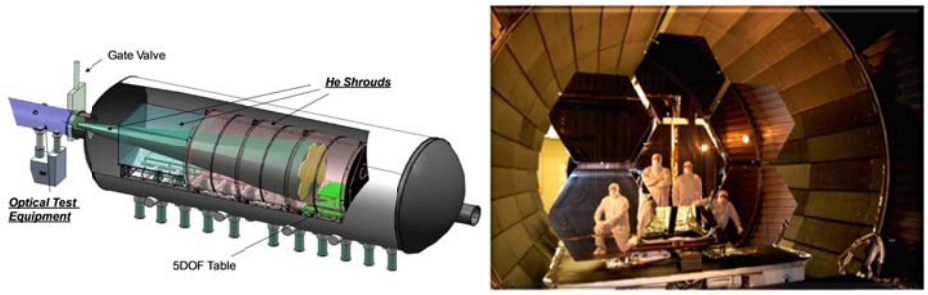


Figure 17: MSFC X-Ray and Cryogenic test Facility (XRCF), with its 7 meter diameter and 23 meter length can test up to 6 JWST PMSAs. Test equipment is located outside a window in ambient temperature and atmospheric conditions.

2.6 Independent Cross-Checks

Probably the single most 'famous' lesson learned from the Hubble Space Telescope is to never rely on a single test to certify a flight specification. Therefore, every component specification must have a primary certification test and a secondary confirming test. Also, it is very important that these confirming secondary tests be performed early in the metrology process. While the metrologist is always going to be under pressure to start in-process testing as soon as possible. And, while the argument will be made that precision is not required during the early fabrication phases, a good metrologist must insist on certifying and confirming the ability of their test setup to achieve the required error budget for each phase of the metrology process.

While technically not an independent cross-check test, it is recommended that a metrologist occasionally depart from their test routine and deliberately attempt to randomize the test. Metrologists tend to be highly structured, process driven individuals – as is required by a profession which measures quantities to nanometers. But, if by chance such an individual is unknowingly introducing an error into their measurement, then by being overly systematic, they will introduce that exact same error into the test every time they conduct the test. Examples of how to vary the metrology process include: deliberately misalign and realign the test setup; perform a settling vibration; take data with different amounts of tilt or defocus; etc.

As summarized in Figure 2, each JWST PMSA requirement has a verification and at least one validation cross-check test. For example, the optical prescription has multiple cross-checks. The prescription is defined during fabrication at ambient temperature using the Tinsley CoC interferometer CGH test and confirmed with an independent auto-collimation test. The PMSA prescription is further tested via an independent ambient test at BATC and the MSFC XRCF 30K test. The prescription receives a final confirmation test at 30K when the entire assembled primary mirror is tested at center of curvature with a refractive null corrector at Johnson Space Center.

2.7 Understand All Anomalies

Finally, of all the rules, this one maybe the most important and must be followed with rigor. No matter how small the anomaly, one must resist the temptation of sweeping a discrepancy under the metaphorical error budget rug. Any time that the actual data uncertainty for a given measured value is larger than its error budget, the reason for this discrepancy must be determined and understood. Do not eat into the contingency reserve because it will be needed at the end of the fabrication process or for the integration, alignment and test (IA&T) process when, if something goes wrong, it is very difficult to fix an error. Similarly, if the actual data uncertainty for a measured value is less than its error budget, one can either adjust the total error budget to create margin for other more difficult parameters or increase the contingency reserve

3.0 CONCLUSION

The discussion above has walked through the challenge of measuring large optics. In doing so, this chapter has defined seven guiding principles that can be applied to any metrology application.

1. Fully Understand the Task
2. Develop an Error Budget
3. Continuous Metrology Coverage
4. Know where you are
5. "Test like you fly"
6. Independent Cross-Checks
7. Understand All Anomalies

Although we have used specific examples from optical testing applications, clearly the issues of error budgets, environmental issues, datum points, cross-checks and understanding anomalies can apply to any part, but particularly to measuring larger parts and structures such as described in the previous two chapters. Large sections on machine tools sag under gravity, girders holding up bridges will change with temperature, and many small errors in an engine will add up to a bad engine.

Many of these issues become most noticeable on large parts being made to high precision. A system like a turbine engine is made so precisely that a large engine able to move a jumbo jet can be easily turned by hand. However, as tolerances keep increasing for all manufacturing, more often than not these considerations will hold true for smaller parts as well. The seven guiding principles therefore can be a valuable tool for any metrology application.

References

1. Stahl, H. Philip, “Rules for optical metrology”, SPIE Proceedings 8011, 22nd Congress of the International Commission for Optics: Light for the Development of the World, 2011

2. Stahl, H. Philip, et. al., “Survey of interferometric techniques used to test JWST optical components”, SPIE Proceedings 7790, 2010, DOI:10.1117/12.862234

3. Stahl, H. Philip, "Testing Large Optics: High Speed Phase Measuring Interferometry," Photonics Spectra, pp. 105 112, December 1989.

4. Burge, J. H., D. S. Anderson, D. A. Ketelsen, and S. C. West, “Null test optics for the MMT and Magellan 6.5 m F/1.25 primary mirrors”, SPIE Proceedings 2199, pp.658, 1994.

5. Stahl, H. Philip, "Phase-Measuring Interferometry Performance Parameters," SPIE.680, 1986

6. Stahl, H. Philip, and Jay A. Tome, "Phase-Measuring Interferometry: Performance Characterization and Calibration," SPIE Vol.954, 1988.

7. Hayes, John Bradford, Linear Methods of Computer Controlled Optical Figuring, PhD Dissertation, University of Arizona Optical Sciences Center, 1984.

8. Montagnino, Lucian A., “Test and evaluation of the Hubble Space Telescope 2.4 meter primary mirror”, SPIE Vol. 571, pp. 182, 1985.

9. Stahl, H. Philip, “Development of lightweight mirror technology for the Next Generation Space Telescope”, SPIE Vol.4451, p.1, 2001.

10. Smith, W. Scott, and H. Philip Stahl, “Overview of mirror technology development for large lightweight space-based optical systems”, SPIE Vol.4198, p.1, 2001.

11. Stahl, H. Philip, John M. Casstevens, and Ronald P. Dickert, "Phase Measuring Interferometric Testing of Large Diamond Turned Optics," SPIE Vol.680, 1986.

12. Stahl, H. Philip, et. al., “Survey of interferometric techniques used to test JWST optical components”, SPIE Proceedings 7790, 2010, DOI:10.1117/12.862234

13. Stahl, H. Philip, Infrared Phase-Shifting Interferometry using a Pyroelectric Vidicon, PhD Dissertation, University of Arizona Optical Sciences Center, 1985.

14. Mooney, James T., and H. Philip Stahl, “Sub-pixel spatial resolution interferometry with interlaced stitching”, SPIE Proc.5869, 58690Z (2005)

15. Evans, Chris J., and Robert N. Kestner, “Test optics error removal”, Applied Optics, Vol.35, No.7, p.1015, 1 Mar 1996.