

Three Dimensional Metrology

R. Hocken, J. A. Simpson, B. Borchardt, J. Lazar, C. Reeve, and P. Stein, NBS, Washington, D. C.
Submitted by: R. Young, NBS, Washington, D. C./(1)

We present the results of research into the three dimensional measurement process using a classically designed measuring machine. This machine has been retrofitted with laser interferometers to provide a stable metric and is controlled by a minicomputer. Machine motions are programmable in a high level interactive language. Data links are provided to a larger computer for sophisticated data processing.

We have pursued the objective of creating, with the lasers, a machine independent coordinate system based at a point. Measurements made in this reference frame are transformed into the coordinate system of the measured object using the techniques of rigid body kinematics. The error terms inherent to the mechanical design (yaw, pitch, straightness, etc.) are measured over the machine volume, 48" x 24" x 10", on a cubic lattice of spacing two (2) inches. These error terms are stored as matrices and used to correct the data during a measurement. A measurement history on these error terms is being compiled. Real time instrumental drifts due to temperature and other external effects are removed using cross referenced measurement algorithms. Errors that cannot be assessed by calibration, such as axis non-orthogonality, are obtained by measuring the object in different angular positions within the measurement volume. This technique, which we call multiple redundancy, allows the assessment of all metric errors which do not commute with the finite rotation matrix.

Introduction

For several years we at NBS have been concerned with the problem of 3-dimensional measurement as part of our program to provide 3-dimensional calibrations of certifiable accuracy. Our work has been concentrated upon retrofitting a "classical" measuring machine. The machine is a Moore 5-Z(1). We have adapted it for computer control using an interdata 70 minicomputer, added Hewlett-Packard laser interferometers to provide the scales, and added a cooling system to reduce the temperature rise when the machine is in operation. Other features include a temperature controlled environment and a high-speed (1200 baud) data link to a large computer (Univac 1108).

The machine has a measurement volume of 48" x 24" x 10" (120 x 61 x 25 cm). Before retrofitting it had a worst case accuracy of several thousandths of an inch (0.003 cm) along a diagonal of the measurement volume. A schematic diagram of this machine is shown in Figure 1. The workpiece (gage) is mounted on the table which moves 48" (120 cm) to provide the X measurement. Y motion is obtained by movement of the large carriage across the bridge. The Z slide is mounted on the Y carriage. These two movements provide the 24" (61 cm) and 10" (25 cm) displacements respectively. Axes movement is controlled by stepping motors attached to lead screws. The three carriages are mounted upon traditional double-V ways, the X and Y slides with roller bearings and the Z slide with plain ways.

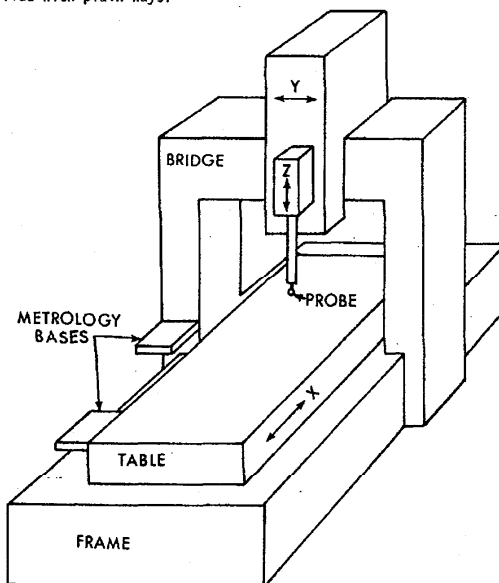


Figure 1 - Schematic diagram of Moore 5-Z

In an attempt to upgrade the performance of this machine, we have developed an extensive formal analysis which to the best of our knowledge is unique to our laboratory(2). The purpose of this paper is to describe this formalism in a broad sense and to illustrate it with specific examples.

Three separable techniques form the basis for our understanding of machine behavior. We begin by examining our machine geometry using the formal structure of rigid body kinematics(3). We then add to this the methods of temporal modeling and production sampling(4,5). Finally, we use the techniques of multiple redundancy(6,7). The first three sections of this paper will explain briefly each of these techniques with illustrative examples from our Moore 5-Z. The rest of the paper will be devoted to relating these techniques to provide a coherent picture of an actual measurement with this system.

Rigid Body Kinematics

The formalism of rigid body kinematics allows one to perform a complete analysis of any machine structure without resorting to complicated and error prone geometric arguments. The technique consists of choosing the minimum number of ideal reference frames (coordinate systems) necessary to characterize a machine and using matrix transformations to relate coordinates in these chosen frames. Ideally these coordinate systems must have their origins at a point but in practice small kinematically mounted "metrology bases" serve to decouple these systems from the distorted machine geometry.

In Figure 2 we show a schematic representation of our machine viewed from above. For simplicity the vertical axis (Z) is not shown even though our analysis is 3-dimensional. Three right handed coordinate systems are used. (Counterclockwise rotations about an axis are defined as positive.) The systems are:

- The Space System:** This is the primary measurement system. Its origin is on a small metrology base mounted on the column of the measuring machine. Most measurements are made directly in this system with instruments of small random error. Angles measured in this system will be denoted by Greek letters without superscripts and vectors denoted by \underline{X} without any superscript.
- The Table System:** The origin of this system is on a small metrology base, now rigidly attached to the table. The position of this origin in the space frame, when the probe is at position X_i , is denoted X_{0i} . Vectors in this frame are denoted \underline{X}' . Of necessity, 3 parameters will be measured in this system.
- The Object System:** This is the coordinate system of the workpiece, usually a gage. It has its origin at some

preselected point on the workpiece and all reported coordinates are ideally in this frame. Vectors in the object system are denoted by \underline{X}^o . We assume that goal of any measurement is to obtain a set of N vectors \underline{X}_i^o which define positions in this gage and a set of N error vectors $\underline{\sigma}_i^o$ which estimate the errors in these vectors.

A typical set of vectors \underline{X}_i , \underline{X}_{0i} , \underline{X}_i^o and \underline{X}_i^u are shown in Figure 2. The three coordinate systems are always chosen so that, except for infinitesimal rotations, they are aligned.

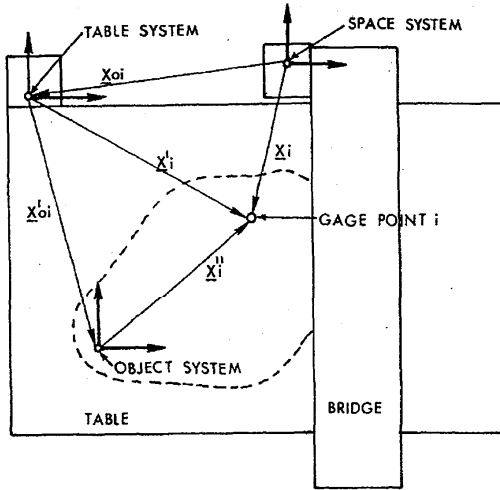


Figure 2 - Schematic diagram of measuring machine viewed from above. Both the table and object system may differ from the space system by infinitesimal rotations.

Any lack of alignment is described by the infinitesimal rotation matrix \underline{R} which is

$$\underline{R} = \begin{pmatrix} 1 & \phi & -\psi \\ -\phi & 1 & \theta \\ \psi & -\theta & 1 \end{pmatrix}$$

This antisymmetric matrix retains only first order terms in the infinitesimal angular rotations ϕ , ψ and θ which are the familiar yaw, pitch, and roll respectively. This matrix will be used to describe the rotations of the three coordinate systems with respect to (WRT) each other. In a well designed measuring machine, the off diagonal terms will be of the order of micro-radians (10^{-6} radians) so that the small angle approximation is quite accurate. The infinitesimal rotation matrices commute, that is

$$\underline{R}_i \underline{R}_j = \underline{R}_j \underline{R}_i, \text{ for small } \phi, \psi, \text{ and } \theta,$$

if we again neglect second order terms in the rotations. Furthermore products of such matrices are also antisymmetric which allows some computational simplification.

Now suppose at time i the machine is moved so that the probe is at point i on the work piece. Let \underline{R}_i denote the rotation of the table system WRT the space system at time i and \underline{R}_i^o denote the rotation of the object system WRT the table system at the same time. The superscripts indicate in which system the measurements are made. Then the vector \underline{X}_i^o is related to the vector \underline{X}_i^u by the simple equation

$$\underline{X}_i^o = \underline{R}_i^o (\underline{X}_i^u - \underline{X}_{0i}^u),$$

where \underline{X}_{0i}^u is the vector to the origin of the object system in the table system. Similarly the vector in the table system, \underline{X}_i^u is related to the vector in the space system \underline{X}_i by

$$\underline{X}_i^u = \underline{R}_i (\underline{X}_i - \underline{X}_{0i}).$$

Combining these relationships enables us to simply express vectors in the object, \underline{X}_i^o , in terms of vectors in the space frame, \underline{X}_i . That is:

$$\underline{X}_i^o = \underline{R}_i^o (\underline{R}_i (\underline{X}_i - \underline{X}_{0i}) - \underline{X}_{0i}^u).$$

It may appear at first glance that the introduction of the table system is unnecessary. This would be true if the table did not move (as in some machines) or if it were an ideal rigid body, which it is not.

Equation 5 gives us a prescription for obtaining coordinates in the body from measurements made in the other systems as well as rigidly specifying what geometrical characteristics we need to measure. The compactness of the notation and strict adherence to sign conventions can prevent errors that often occur when one attempts to do the same problem using analytic geometry. We will defer further discussion of this equation to a later section.

Measurement Modeling

The second technique we use is related to the visualization of measurement as a production process with a product, numbers, whose quality may be controlled by the methods of statistical sampling. The goal here is to model those aspects of instrument behavior which influence the quality of these numbers and to check this model with quality control techniques.

In a large measuring machine many factors conspire to degrade the quality of measurement. Besides those geometrical factors displayed in Section 1, one of the most troublesome of these factors is a temporal dependence of the measurements. The causes of this time dependence are most likely changing temperature distributions in the machine and drifts in the associated machine electronics. A complete analytic model of such behavior is currently impossible, but an empirical model based on observational data is not. A very limited set of intuitively plausible assumptions makes such an empirical model possible.

Inherent in our discussion in Section 1, though not explicitly stated, was the assumption that the workpiece (object) was an ideal rigid body. Here we make that assumption explicit with the further corollary that, besides being a perfect rigid body, the object is characterized by a fixed set of coordinates of the gage points which, in the object system, are functions only of the workpiece temperature. That is, the numbers computed from equation 5 are invariant in time except for a random error component which is assumed small and independent of the magnitude of \underline{X}_i^u . Call this error vector $\underline{\sigma}_i$. Then, if on a repeated measurement, the coordinates obtained for a point on the gage differ from the previous value by more than $\underline{\sigma}_i$, the change may have been caused by a drift in some instrument parameter. One can then assume that the drift was linear in time and correct intervening points accordingly. Again this procedure is best illustrated with an example.

Suppose the workpiece is a gage with N gage points which may be located with the same algorithm. The random error in this algorithm, $\underline{\sigma}_i$, may be accessed by repeated location of the same gage point during a time short compared with the time of a complete measurement of the gage (called a "run"). Our technique then is to choose one particular gage point, called the repeated point, and K other points called check points. A run begins with a measurement of the repeated point. This point is remeasured at random intervals throughout a run, and then measured as the last measurement. Between remeasurements of the repeated point the check points are also remeasured. The constraint is then imposed that measurements of the repeated point should agree within the estimated random error $\underline{\sigma}_i$ (i.e. the gage itself does not change).

If they do not, a linearized drift correction is computed and applied to all data. A simple statistical test is applied to confirm that this correction, computed from the repeated point measurements only, reduces the standard deviation of the check points to the predicted σ_p .

The model this algorithm defines is quite simple. Though the machine behavior is governed by many uncontrolled variables, over a time short compared with the machine's thermal time constant, the behavior of the machine is a linear function of time. Extensive experimentation has shown that this is indeed true, an observation that will be proven in a later section. The process of statistical comparison at the check points, analogous to production sampling for part quality, continually checks the adequacy of the model.

Multiple Redundancy

Our third technique is perhaps the most powerful and yet the most difficult to elucidate. It is based on the realization that properly chosen redundant measurements, made with an imperfect instrument can contain sufficient information to also measure and remove the known imperfections in the measurement system itself. In three dimensional measuring machines the type of redundancy we have found most suitable is that of remeasuring the workpiece at different angular orientations WRT the measuring machine. The data then allows us to remove errors due to scale imperfections and axis nonorthogonality. Again we illustrate this procedure with a simple example.

Suppose we had an object with three gage points that we measure in a nonorthogonal coordinate system as depicted in Figure 3. In such a system X coordinates are defined by distances to the Y axis (rather than projections on the X axis) and Y coordinates are defined similarly. Vectors in the non-orthogonal system,

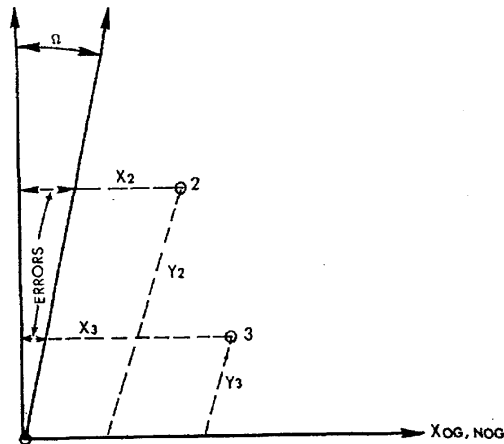


Figure 3 - A three point gage as measured in a non-orthogonal coordinate system.

X_{NOG} , will be related to vectors in the orthogonal system, X_{OG} , with the same origin by

$$X_{NOG} = \underline{A} X_{OG} \text{ where } \underline{A} = \begin{pmatrix} 1 & -\alpha \\ 0 & 1 \end{pmatrix},$$

and α is assumed small. The X axes of the two systems are chosen to be aligned for convenience. This choice is arbitrary. Now let us suppose we rotate the object by a finite angle Δ , which corresponds to a rotation of the coordinate systems, and remeasure the vector to the same gage point. The new vector, denoted with the superscript N, is related to the old vector in the orthogonal system by the simple finite rotation matrix \underline{B} . That is

$$X_{OG}^N = \underline{B} X_{OG} \text{ where } \underline{B} = \begin{pmatrix} \cos \Delta & \sin \Delta \\ -\sin \Delta & \cos \Delta \end{pmatrix}.$$

The coordinates of this new vector in the non-orthogonal frame are simply

$$X_{NOG}^N = \underline{A} X_{OG}^N = \underline{A} \underline{B} X_{OG}.$$

It is easy to see that equation 8 does not correspond to a simple rotation of the original vector in the non-orthogonal frame.

Such a vector, call it X_{NOG}^B , would be given by

$$X_{NOG}^B = \underline{B} X_{NOG} = \underline{B} \underline{A} X_{OG}.$$

The two vectors, X_{NOG}^B obtained by a simple rotation of the non-orthogonal coordinates and the measurements of the rotated object X_{NOG}^N will be equal if and only if

$$\underline{A} \underline{B} = \underline{B} \underline{A},$$

ie, if the matrix which describes the imperfection of the measurement system commutes with the finite rotation matrix. For this case it is easy to show that

$$\underline{A} \underline{B} \neq \underline{B} \underline{A} \text{ if } \Delta \neq n\pi \text{ radians, } n = 0,1,2,\dots$$

Furthermore it can be shown that with just 3 points, (one of which is defined as the origin) measured twice as in this example, it is possible to compute both the coordinates of the points and the angles Δ and α . There are 8 measurements and only 6 unknowns so the system is just slightly over constrained. For a real measurement where there are many gage points and measurements at several angles the measurements are termed "multiply redundant" and the angles and coordinates determined by large least squares fits.

The generalization of the above analysis to three dimensions is messy but straightforward. The class of errors that may be computed and removed is exactly the same, that is, those metric errors that do not commute with the finite rotation matrix. Other common errors at this type, besides nonorthogonality, are scale errors. In fact, with sufficient measurements, only one good scale is required for a three dimensional measurement.

It should be noted that though the analysis in Section 1 is made in orthogonal coordinate systems it is equally valid in slightly non-orthogonal systems. This is because the infinitesimal rotation matrix \underline{R} commutes with the non-orthogonality matrix \underline{A} (to first order in the angles).

The technique of multiple redundancy shares with the technique of measurement modeling the assumption of gage stability. In a sense both are self calibrating algorithms in that, through a closed series of measurements, errors in instrument performance may be assessed and removed.

Now with this background and terminology, we are ready to discuss the process we use for three dimensional measurement on our Moore 5-Z.

Machine Calibration

In order to actually make a measurement using the techniques outlined in the preceding sections several more steps are required. The first of these is to actually write out, by components, the geometrical observation equation, equation 5. One then must identify those terms which must be measured, the system they must be measured in, and devise techniques for their measurement. An inspection of the equation reveals that at least 15 independent measurements are required, three (3) for each rotation matrix, and three (3) for each vector, even for this simple model which assumes that the X,Y, and Z positions of the probe can be accurately determined directly in the space frame. Some reduction in this number can be obtained by observing certain machine characteristics. For instance, we assume the table of our Moore 5-Z to be a quasi-rigid body in that it remains rigid WRT the X,Y plane. This assumption merely reflects the

fact that distortions of the table are larger in the direction of gravity (Z) and gravity induced changes in X,Y distances only cosine errors. Three measurements, X'_{0i} , Y'_{0i} and ϕ'_i are eliminated by this means.

Table 1 - The geometric parameters needed in 3-dimensions

Symbol	Description	Measurement Coordinate System	Method
1) Y_0	Y straightness of X table	space	stored
2) Z_0	Z straightness of X table	space	stored
3) ϕ_0	Yaw of X table	space	stored
4) ψ	pitch of X table	space	stored
5) θ	roll of X table	space	stored
6) XSP	X straightness of probe	space*	stored
7) YSZ	Y straightness of Z axis	space*	stored
8) ZSY	Z straightness of Y axis	space*	stored
9) ψ'	Pitch of object	table	on-line
10) θ'	roll of object	table	on-line
11) Z_0	Z position of object	table	on-line
12) X_0	Basic table position	space	on-line
13) Y_0	Basic Y-carriage position	space	on-line
14) Z	Basic Z carriage position	space*	on-line

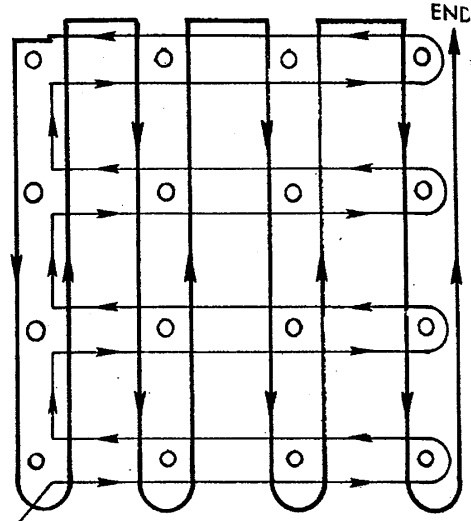
*These measurements are transferred to the space system via an intermediate coordinate system at height $Z = 0$. This system is not described in detail in the text.

Table 1 catalogs those terms remaining which must be measured for an accurate three dimensional measurement. There are 14 in all*, requiring great expense for instrumentation and making the problem of just fitting the instruments on the machine very difficult. Many of the terms are, however, small and if they are sufficiently smooth functions of measuring machine position and are also sufficiently reproducible, they can be measured at regular intervals and used as corrections to the basic X,Y,Z metric. This is the course we chose to pursue.

We concluded that the generalized error term, E, must be a function of the machine position (X,Y,Z) and of such variables as the machine temperature distribution $T(X,Y,Z)$ and the table loading $L(X,Y)$. The load term is the most difficult of these effects to handle. We are currently standardizing the table loading to avoid large changes in its structure. Changes in machine geometry due to self-loading are built into the analysis as functions of the coordinates.

We divide the measurement volume of the machine into a 2 inch (5 cm) cubic array with 1950 lattice sites. Each of the stored error terms is to be measured at all lattice sites. Our algorithm for doing this is, however, basically two dimensional. The instrument is set up for measuring the error term and probe moved through the measurement volume in the X,Y plane at constant Z. This path is depicted in Figure 4.

*The fourteen (14) terms become necessary due to the addition of a 4th coordinate system used to transfer the Z axis measurements into the space system. Since the methodology of this process is the same as described in Section 1 and its addition to the algorithm only an additional complexity, this system (in which only two parameters are measured) has been deleted from the discussion.



START

Figure 4 - The measurement path for generation of a machine error surface at constant Z.

After an initial standardized warm up exercise, the machine is moved until the probe is at $X = 0, Y = 0$ in the X,Y plane (lattice site 1,1,1 for $Z = 0$), and the interlaced path is swept out as indicated. The error term(s) being measured is, by the end of the run, measured 4 times at each lattice site. Temporal average over a 30 sec interval is performed for each of these measurements. The resulting set of data (1300 points) is divided into what we call cycles (one back and forth run in either the X or Y direction) and a linear temporal drift correction applied within each cycle. A least squares fit is used to "tie in" the whole surface and the standard deviation of the error surface thereby assessed. The stability and reproducibility of the drift correction is discovered by remeasuring the whole surface. Since each run takes some 16 hours, the test that two or more runs should agree is indeed stringent. The extent of this reproducibility is best illustrated by examining the data. In Figure 5, we show typical error surface. This surface is a measure of the roll of the table (at its origin) as a function of X and Y machine positions at $Z = 0$. The total roll is about 2 sec (as measured in the space frame) and Figure 5 is a result of combining 6 runs (7800 data points, each averaged for 30 sec). The resulting standard deviation of this surface is about ± 0.01 sec. Similar surfaces have been prepared for the error terms listed as stored in Table 1. The standard deviations of the angular error surfaces are of the order of hundredths of a second of arc and of the straightness errors of the order of microinches ($1 \mu\text{inch} = 25 \text{ nm}$). Our work so far has shown that the five table errors ($\phi, \psi, \theta, Y_0, Z_0$) are independent of the Z probe position, which considerably simplifies the analysis.

The excellent reproducibility of these error surfaces leads us to have considerable faith in both the geometrical integrity and temporal predictability of our machine. That is not to say that the drift is always the same, but rather by careful choice of a measurement algorithm such behavior may be reliably removed. We are now ready to describe a typical measurement of a gage at NBS.

The Complete Measurement Algorithm

The complete measurement algorithm for measurement of a three (or two) dimensional gage is schematized in Figure 6. The algorithm begins with positioning the gage on the table. No attempt

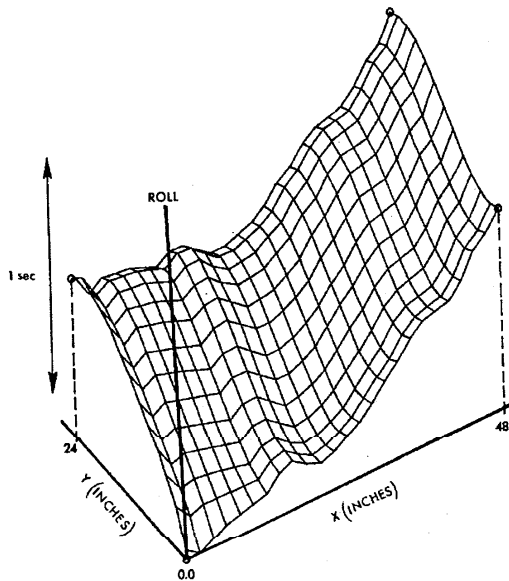


Figure 5 - The roll of the table system WRT the space system on the MOORE 5-Z. This surface is the result of 5 runs combined and has a standard deviation of the order of 0.01 second of arc.

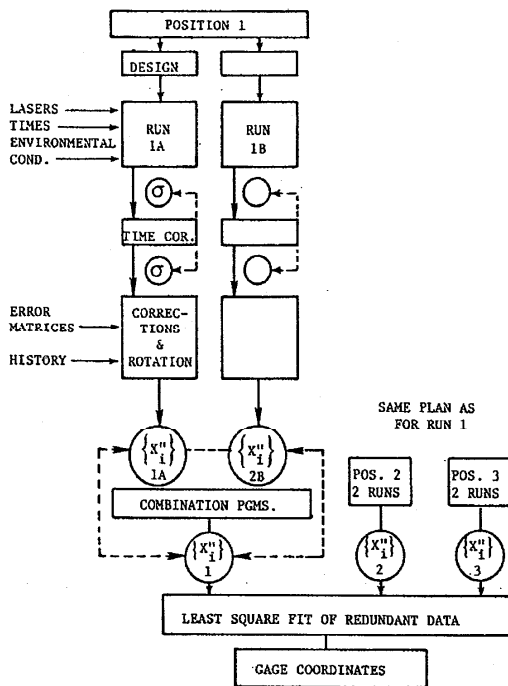


Figure 6 - Measurement flow chart for a multi-dimensional calibration. Each set of 2 runs at a single position has a flow diagram like that shown for position 1. Dashed lines indicate where statistical checks are performed. The last least squares fit of the redundant data is of course a stringent statistical test.

is made to align the preferred (by the user) gage coordinate system with the machine system. Furthermore once the original position (position 1 on the Figure is established) the possible positions for 2 and 3 are constrained since rotations of $m\pi$ ($m = 1, 2$) yield no new information, in fact $\pi/4$ and $\pi/2$ rotations are preferred. The operator then designs a typical run algorithm (as described in Section 1) which contains the repeated point and the check points. One such path, for the NBS 2-0 ball plate, is shown Figure 7. An attempt is made to randomize the path so that distances measured will not be correlated with time. The machine encoders are set at prescribed values to allow retrieval of the error matrices previously described. The plate is then measured, recording the laser readings, encoder readings and time of each gage point location. Atmospheric conditions and average gage temperature are recorded at the beginning and end of each run.

The resulting data set is then transferred to a large computer (UNIVAC 1108) where the following computations are performed. First the time correction is computed and applied using the standard deviation of the check points to test the applicability of the model. Next the corrections for the machine geometric distortions are made with the error matrices stored in the computer. Finally the laser wavelength is corrected for atmospheric conditions and the gage size corrected for thermal expansion (all results are reported for 20 °C).

The resulting set of coordinates, called a single run, and their standard deviations are stored for future analysis. Such a run may take from 3 to 6 hours.

The process is now repeated and a second such set of coordinates and errors generated. These results are combined with those of the previous run. The standard deviation (rms) of the combined runs is compared to the single run deviation. If there is a

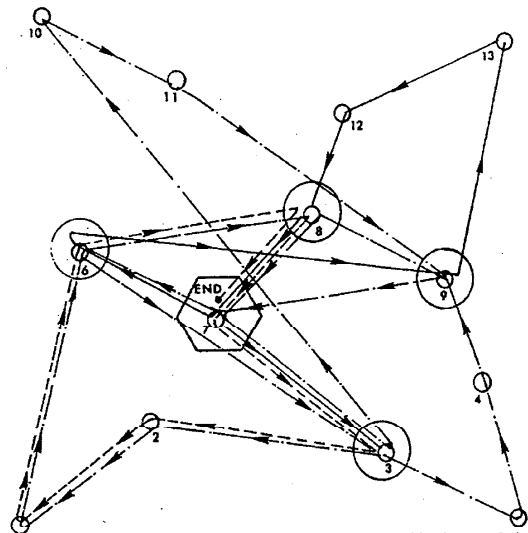


Figure 7 - Typical run path for the NBS 2-0 ball plate. Point 7 is the repeated point and points 3, 6, 8 and 9 the check points.

statistically significant increase in the random error a new run is instituted. This process checks specifically the applicability of the time, atmospheric and thermal expansion correction, as well as giving a good estimate of the gage point location error.

The whole process is now repeated for 2 different orientations of the gage WRT the measurement machine. Rotations of 90° ($\pi/2$ radians) and 45° ($\pi/4$ radians) are particularly sought after since the former exchanges the axes and the latter mixes the axes.

The final results of these measurements are three complete sets of coordinates and errors in three different coordinate systems. Naturally these coordinates look nothing alike as they differ from each other by finite rotations, by scale errors and by machine axes non-orthogonality. These results are used as the input to a large nonlinear least squares fit which treats these numbers as data and the "coordinates", the angles describing the gage orientations and the angles characterizing the machine axes non-orthogonality as parameters. The constraint that the preferred gage coordinate axes should pass through specific gage points is used in this final fit.

Summary

The algorithm described in the preceding text is a recent synthesis of several different approaches. At the time of this writing it is fully implemented in two-dimensions and nearing completion in three. Only the passage of time will allow complete assessment of its usefulness. Particularly we must establish a history on the multi-dimensional error matrices and perhaps a wear model for the updating of these matrices. At this time the hardware for the pitch, roll and Z-straightness of the gage WRI the table is also in the development stage. (These measurements are only required in three dimensions).

We have however already learned a great deal. The short term reproducibility of the error matrices is quite encouraging and has been shown to be valid over periods of a month or so. We have also created several questions that as yet are unanswered. Specifically though we can obtain single run standard deviations

on a large ball plate that are equal to our location error (~ 6 inches, $.75 \mu\text{m}$). The total error (1) of the complete algorithm after multiply redundant measurements is of the order of twice that. This inclines us to believe that some systematics have been excluded from our analysis and we are pursuing their sources. Scale errors are not currently included in our multiple redundancy programs on the assumption that the properly aligned H.P. interferometers provide an accurate and stable metric. Tests of these lasers against an iodine stabilized helium-neon laser show a frequency stability of parts in 10^8 which would seem to support this assumption. We expect that with continued research and study these problems will be solved.

Bibliography

- 1) Brand names are mentioned only in the interest of clarity and in no way constitute an endorsement by the National Bureau of Standards.
- 2) Many other workers have studied three dimensional measuring engines. See, for example, D. L. Lane, UCRL-50679, 1969.
- 3) Many mechanics texts are applicable. See, for example, Classical Mechanics, Goldstein, Addison-Wesley, 7 ed., 1965. and P. A. McKeown, 23rd CIRP General Assembly.
- 4) J. M. Cameron and G. Hales, NBS TN 844, 1974.
- 5) P. E. Pontius, NBSIR 74-545, 1974.
- 6) D. C. Brown, AFMTC-TR-58-8, ASTIA Doc. #134278, 1958.
- 7) C. P. Reeve, NRSIR 74-532, 1974.