Error Considerations in Measuring Systems

In Chapter 1, the uncertainties in the indicated value of a measurement were examined in a cursory manner. This chapter examines in greater detail the errors in measurement systems that cause these uncertainties, the meaning and interpretation of these errors, and methods of reducing or circumventing certain errors.

ERROR CLASSIFICATIONS

Measurement error is the quantitative difference between the true value of the measurand and the value indicated by the measuring system. It may be expressed either as an absolute difference or on a relative scale. Each component of the measuring system has sources of error that can contribute to measurement error. These error contributions can be classified as *static errors, instrument loading errors,* or *dynamic errors.*

Static Errors

Static errors result from the physical nature of the various components of the measuring system as that system responds to a measurand input that is time-invariant during the mea-

surement. In general, the sources of static error in measuring system include intrinsic imperfections or limitations in the hardware and apparatus (compared to ideal instruments), external influences on the physical properties of the apparatus, inexactness in the calibration of the system, and displaying the output of the measuring system in a way that requires subjective interpretation by an observer. One or more of these error sources is present in each component of an electromechanical measuring system, transducer, signal conditioning, and readout.

Loading Errors

Loading errors result from a change in the measurand after the measuring system or instrument is connected. The loading error of a particular system may have both static and dynamic components. In electromechanical measuring systems, loading error is almost always due to the physical nature of the sensing element of the transducer. For example, the Bourdon tube in a pressure transducer introduces additional volume into the pressurized system under test, thereby changing the rate of pressurization slightly. An accelerometer mounted to a vibrating body adds to the effective mass, slightly changing the natural frequency and the acceleration forces on the body under test.

The effects of instrument loading are unavoidable in most cases, and must be determined specifically for each measurement and measurand. Loading errors may often be the single greatest uncertainty in a physical measurement. Therefore, the measuring system user should select the transducer carefully, with a particular view to choosing one whose sensing element will minimize instrument loading error in the particular measurement involved.

Dynamic Errors

Dynamic errors result from the inability of a measuring system to respond faithfully to a time-varying measurand. Usually, the dynamic response of an electromechanical measuring system is limited by one or more of the following factors:

1. Inertia, damping, or other physical constraints of the transducer.
2. The cutoff or bandpass frequencies of filters incorporated into the signal conditioning, or gain-bandwidth limitations.
3. Inertia, damping, or friction in the readout or display system.
4. In carrier-operated transducers, the ratio of carrier frequency to the highest modulation frequency, and the modulation index.

In ordinary applications, the most likely limitation is the reponse of internal low-pass filters in the signal conditioning equipment.

For cyclic or periodic variations in the measurand input, the dynamic error of the measuring system is characterized by the *frequency* and *phase response* (Bode criteria) of the system. For random or transient inputs, the *time constant* or *response time* describes the dynamic error. In either case, the dynamic characteristics of the measuring system must be known before the system can be used to measure time-varying inputs.

STATIC ERROR

The foregoing discussion points up the fact that there are many possible sources of error in the transducer, signal conditioning circuitry, and readout device that comprise an electromechanical measuring system. Experience indicates that the major concern is the static error introduced by the components of the measuring system.

Types of Static Error

Static error is a combination of three types of errors: reading errors, environmental errors, and characteristic errors. *Reading errors* include *parallax*, *interpolation* error, and *optical resolution*, also known as *readability* or *output resolution*. In certain types of digital displays, there is a plus-or-minus

one count limitation on the reading, which also is a reading error. Reading errors apply exclusively to the readout or display device and have no direct relationship with other types of error within the measuring system.

Several types of reading errors can be reduced or eliminated by relatively simple techniques. For example, parallax can be virtually eliminated by the use of a mirror behind the readout pointer or indicator. Interpolation error can be reduced and optical resolution can be increased by using a magnifier over the scale in the vicinity of the pointer. And, of course, a digital readout or display eliminates most of the subjective reading errors usually made by the observer.

Environmental errors result from external influences on a measuring system. The most common are temperature, pressure, humidity and moisture, nuclear radiation, magnetic or electric fields, vibration or shock, and periodic or random motion. Usually, the effects of environment on each component are independent. Thus, the environmental errors of each component of the measuring system make a separate contribution to the static error. For this reason, the number of environmental variables that could affect the measurement should be minimized.

Characteristic Errors

The third class of errors contributing to static error are *characteristic errors*. These errors describe the deviation of the output of the measuring system under constant environmental conditions from the theoretically predicted performance, or from nominal performance specifications. *Linearity* errors, *hysteresis*, and *repeatability* errors are characteristic errors present to some degree in each component of a measuring system. Other characteristic errors include *gain* error and *zero offset*, often collectively called *calibration* error.

Similar characteristic errors in each component of the measuring system tend to be additive. Thus, system linearity is usually the sum of the errors in the individual components. The combination and accumulation of errors is discussed later in this chapter.

Repeatability

The single most important factor in any transducer measuring system is *repeatability*. Repeatability is a characteristic of a measuring system whereby repeated trials of identical inputs of the measurand value produce the same indicated output from the system. Repeatability is the only characteristic error which cannot be calibrated out of the measuring system. Thus, repeatability becomes the limiting factor in the calibration process, thereby limiting the overall measurement accuracy. In effect, repeatability is the minimum uncetainty in the comparison between measurand and reference. Outstanding repeatability is a significant feature of a well designed Linear Variable Differential Transformer.

Linearity

A high degree of linearity, or the minimizing of linearity error, is another very important characteristic of most measuring systems. The more linear the measuring system is, the more readily it can be calibrated, and the less uncertainty there will be about a particular output value indicated by the system. Although it is possible to construct a calibration curve for a non-linear measuring system, as long as the system's output is very repeatable, this is tedious, cumbersome, and time-consuming. On the other hand, a simple two-or three-point calibration is sufficient for a repeatable, linear system. Recalibration is also facilitated.

ERROR REFERENCE

Characteristic error is defined as the deviation of the output of a measuring system from the predicted performance, or from nominal performance specifications, under the condition of constant environment. Thus, the "theoretically predicted performance" or "nominal performance specification" is the norm or reference to which deviations are compared. However, the basis for selecting a particular norm is sometimes misunderstood. This is particularly true with regard to linearity.

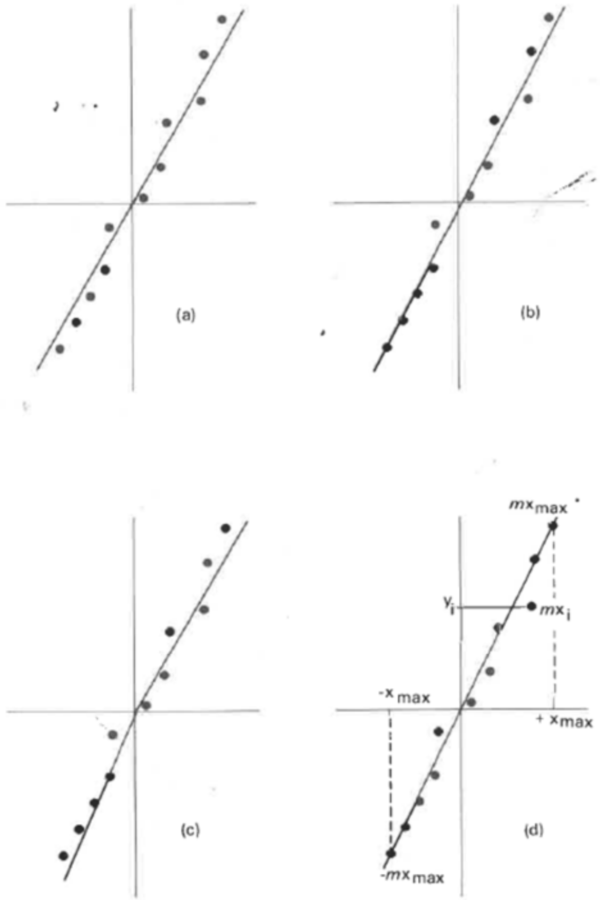

Fig. 14.1 *Graphical representation of reference lines. Line in (a) is forced through zero; (b) is independent of zero; (c) is two-slope line; (d) is least-squares data plot.*

Interpretation of Linearity

Linearity is defined as the maximum deviation of the output of a measuring system from a specified straight line applied to the plot of data points of the output-versus-measurand-input curve. In practice, there are several possible straight lines that can be used as linearity references. Therefore, the exact nature of the reference straight line must be clearly defined before it is possible to interpret or compare linearity specifications for a measuring system or component.

This problem is compounded when the reference line is "forced" through the origin of the data plot even if the zero output point is not located exactly at the origin. This is illustrated graphically in Figure 14.1. The reference line of Figure 14.1*a* is forced through the origin, whereas the line of Figure 14.1*b* is fitted as closely as possible to the data points, without regard to the origin. The compromise position of the line in Figure 14.1*a* results in greater deviations of the data points than in Figure 14.1*b*. Thus, for the same collection of actual data points, the linearity specification referenced to the line in Figure 14.1*b* will seem to be better than that referenced to the forced-zero line of Figure 14.1*a*.

The finite output at zero-measurand input for the system plotted in Figure 14.1*b* can usually be offset by electrically zero-shifting the output of the signal conditioning equipment. By this technique, a user who rates his system linearity without forcing the reference line through zero (origin) has superior performance to that of a user of the identical system having a forced-zero reference line.

If a user does not require a bipolar output from his system, the reference lines of Figure 14.1*c* can be used. Here, one reference line is used for positive values, and another for negative values, with a common point at the origin. In general, the slope of each line will not be identical. This two-slope method has the advantage that the reference line is fit to only half as many data points over a shorter interval. Thus, it matches the data points more closely than either of the two previously discussed methods. If output from only one side of null is required, this linearity specification could be better than that

referenced to the line in Figure 14.1*b*, even though the same measuring system is used to produce the same data points.

Reference Straight Lines

The discussion above did not depend on how the reference lines fit the data points, except in relation to the origin. Several different reference lines can be used, each differing in the manner it fits the data points. The most common reference lines are "terminal line", "end-point line", "best-fit straight line" and "least-squares line". A *terminal line* is drawn from the origin to the data point at full-scale output. An *end-point line* is drawn between the end points of the data plot, usually without regard to the origin if the output is bipolar. A *best-fit straight line* is a line midway between the two closest parallel straight lines that enclose all the data points. The *least-squares line* is the line for which the sum of the squares of the deviations of the data points from the line or curve being fit is minimized.

Least-Squares Line

Of the four common reference lines cited above, experience and statistical theory both favor the least-squares line. The least-squares line is a truly best-fit curve in the sense that it comes as close as possible to each data point on a plot of output versus measurand input according to the least-square error criterion, i.e., to minimize the sum of the squares of the deviations of the data points from the curve being fit. The method of least squares and its application to curve fitting are covered in detail in most texts on statistics. Suffice it to say that the least-squares line for a transducer measurement system can be found from established slope equations.

The slope, m, of the least-squares best-fit straight line for a collection of data points, Figure 14.1*d*, is given by:

$$m = \overline{X_i Y_i} \Big/ \overline{X}_i^2 = \sum_{i=1}^{n} X_i Y_i \Big/ \sum_{i=1}^{n} X_i^2$$

where:

- n - number of test data points
- X_i - displacement values of the data points
- Y_i - output voltage ratio values of the data points
- $\overline{X_i Y_i}$ - average product of the X and Y coordinates of each data point
- $\bar{X}_i^2$ - average of the squared values of the X coordinates of all data points

When the slope, m, of the least-squares line has been determined, it is possible to check the maximum deviation of linearity from the least-squares line by:

$$\%\,\text{linearity} = \frac{Y_{i\,max} - m X_i}{m X_{full\ range}} \times 100\%$$

where:

- $Y_{i\,max}$ – output voltage ratio at point of maximum deviation from best fit straight line.
- $m X_i$ – point on best fit straight line corresponding to maximum deviation point.

The manual computation of least-squares linearity is laborious. However, mini-computers now available can be readily programmed to do this computation. An on-line calibration system using a mini-computer is illustrated in Figure 14.2. A sample print-out is shown in Figure 14.3.

The mini-computer can be programmed to do more than the routine calculations of an on-line calibration, however. The data points from this on-line calibration can be used to derive the coefficients of a second or third degree polynominal whose curve even more closely satisfies the least-squares criterion for curve fitting than a least-square straight line. This polynomial is then stored in the computer.

When a measurement is taken, the reading is passed through the polynomial equation. The result is a reading of improved accuracy.

The equations for deriving the coefficients of these higher order polynomials are beyond the scope of this book, but can be found in several of the references on statistics listed in the bibliography, Appendix A.

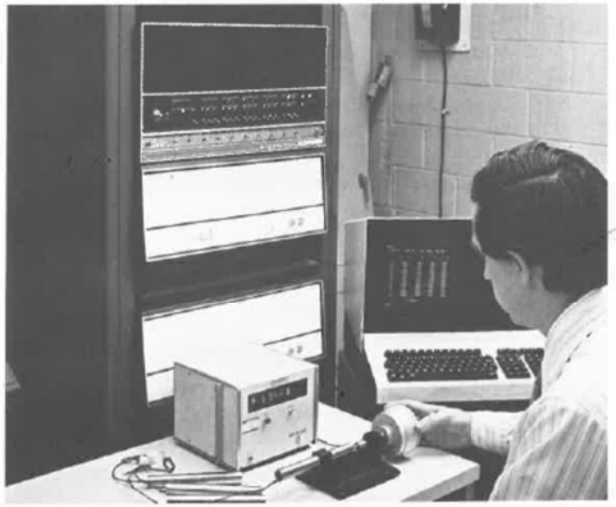

Fig. 14.2 *On-line calibration equipment connected to mini-computer.*

S/N: TST LVDT: 35 - : STD - 10/6/71 STA. NO. 1

P	X	YA	YC	D
01	10000	9996	9987	-0009
02	8000	7996	7992	-0004
03	6000	5992	5994	+0002
04	4000	3992	3996	+0004
05	2000	1995	1998	+0003
06	1000	0997	0999	+0002
07	0500	0496	0499	+0003
08	0200	0200	0199	-0001
09	0000	0001	0000	-0001
10	-0200	-0199	-0200	-0001
11	-0500	-0497	-0500	-0003
12	-1000	-0998	-1000	-0002
13	-2000	-1997	-1999	-0002
14	-4000	-3995	-3997	-0002
15	-6000	-5992	-5995	-0003
16	-8000	-7993	-7993	0000
17	-10000	-9991	-9990	-0001

L = 00.05

Fig. 14.3 *Typical computer print-out of least-squares linearity.*

LVDT Linearity

Many of the points made in the foregoing discussion about reference lines are applicable to the interpretation of the linearity specification for an LVDT. Thus, if the reference line is forced through zero (origin), the linearity can become compromised, particularly if there is not a very close match between the secondary windings during manufacturing. Minor secondary unbalance can result in a two-slope reference line that might have a linearity 0.1% of full range for each line, but 0.5% of full range compared to a single reference line over the entire range.

For applications where the LVDT is used on only one side of null, the linearity may be perfectly acceptable, but the same LVDT, used over its full range, would appear to have unsatisfactory linearity. By special winding techniques, it is possible to balance the secondaries to achieve linearities of better than 0.1% of full range.

Even without special balancing techniques, however, it is possible to have typical linearity specifications of 0.25% of full range if the reference line is not forced through the origin. This is usually acceptable for LVDT measuring systems, because the zero output can be restored electrically. In the case of a DC-LVDT, the zero point is not forced, because its null is defined as the core position for zero output.

Another important fact is that an LVDT used at less than full-scale displacement will exhibit correspondingly improved linearity. Thus a typical LVDT that has a single-side-of-null linearity of 0.1% of full range has a linearity of 0.05% of full range when used up to half-scale displacement. This same comment applies to RVDTs used at less than full angular range.

ERROR ACCUMULATION

RMS Static Error

The total static error of a measurement system can be measured in terms of the root-mean-square (rms) of the com-

ponent characteristic errors if the following conditions are fulfilled:

1. The component characteristic errors are independent.
2. The component errors are of the same order of magnitude.
3. The distribution of the errors is normal (Gaussian).

Wherever possible, condition 3 should be verified by experimental analysis.

From the foregoing discussion, the total static error (TSE) of a typical system would be:

$$TSE = \sqrt{RE^2 + (LE_1 + LE_2 + LE_3)^2 + EE_1^2 + EE_1^2 + EE_3^2 + CE_1^2}$$

where:

RE - reading errors
LE - linearity errors of individual components
EE - environmental errors
CE - a characteristic error (other than linearity)

Presumably, all of these errors are expressed as a percentage of full scale so the total static error is also expressed as a percentage of full scale.

The rms static error thus determined can be used to evaluate the most probable value of the measurement. To do this, the rms error is equated to the standard deviation, σ. Confidence bands of 1σ (68%), 2σ (95%), or 3σ (99.7%) can then be established. This technique is described more fully in textbooks on statistics. As a rule, allowing for an error three times as great as the rms static error gives a 99.7% probability that the true measurement error is no greater.

Static Error Band

Occasionally, it is desirable to show the combined effect of characteristic errors graphically instead of numerically. In such a case, the usual practice is to plot a *static error band* along the reference line. This band encompasses the maximum deviations of the data points from the reference line. The error band is always plotted over at least one full calibration cycle. Usually, it is plotted over three or more calibration cycles to include the repeatability errors. It is also possible to plot the envelope of the static error band along points corresponding to the 2σ or 3σ values of the deviations, but this is not often done. The concept of static error band is covered more fully in several of the references listed in Appendix A.

TRANSDUCER SYSTEM CALIBRATION

The multiple-component nature of a transducer measuring system makes it imperative to calibrate the system as an integrated whole. Then, the ultimate calibration will take into account the error-producing properties of the sensing element, transducer element, signal conditioning equipment, and readout.

Calibration Technique

A transducer measuring system is usually calibrated by adjusting the transducer, signal conditioner, and readout to assure zero output for zero-measurand input. A known value of the measurand is applied to the input of the transducer and the displayed output of the system is observed. The system *gain* and *calibration* controls are then adjusted to display an output equivalent to the known measurand input.

Ordinarily, the magnitude of the calibrated measurand input is chosen to be the full-scale input value of the transducer. However, it may be less than full scale if the transducer has a greater full-scale capability than the range of measurand values to be measured. For maximum calibration accuracy, the measuring system should be calibrated as close as is practical to the maximum measurand values that are to be measured. The reason for this is that most transducer system errors are by nature a percentage of full-scale output.

The transducer system calibration should be performed under environmental conditions that are as close as possible to those conditions under which actual measurements are to be made. This is particularly true regarding ambient and measurand-media temperatures. The readout equipment is often located remotely from the transducer, so the ambient temperature of the readout and the signal conditioning electronics may be significantly different from that of the transducer. This

environmental difference should be taken into account when calibrating the system.

Calibration Standards

A meaningful system calibration requires that the reference measurand input be known to a much greater degree of accuracy than the desired degree of measurement accuracy. This is because repeatability and similar characteristic errors are inherent in the transducer measuring system. If the calibration uncertainty is small compared to system errors, the measurement accuracy will depend essentially on system errors alone.

Ideally, the calibration standard for the system should be at least one order of magnitude more accurate than the desired measurement system accuracy. An accuracy ratio of 3:1 may be acceptable for some applications, but ratios of 10:1 or more are desirable to insure that the system accuracy depends primarily on the characteristic errors of the system components.

MEASURING SYSTEM ACCURACY

If the calibration standard is sufficiently accurate, the overall accuracy of a measurement system depends on the combined errors due to loading and total static error. If the errors are of similar magnitude, system error will be:

$$SE = LE + TSE$$

where:

SE - system error
LE - loading error
TSE - total static error (rms)

DSPM Industria®
sensori & trasduttori
Via Paolo Uccello 4 - 20148 Milano
Tel +39 02 48 009 757 Fax +39 02 48 002 070 info@dspmindustria.it www.dspmindustria.it