Contributor Name	**Prof. Dr. Luca Mari**
Date Submitted	**18/04/2002**
Word Length	**2913 (with abstract, headings and list of figures)**

BMID **Block name (Your block is in bold)**

Super Cluster 2 – Defining the Measurement System Error and Uncertainty

154	Common sources of error in measurement systems
155	**Explanation of key error and uncertainty concepts and terms**
156	General characterisation of systematic and stochastic errors
157	Human error
158	Errors in analogue signal systems
159	Errors in digital signal systems
160	Uncertainty determination
161	Error models, error budgets and their calculation
162	Error minimisation in measurement systems

Retain this section between the red lines in your submission.
It is needed for the editorial and review processing. This section is not published.

EXPLANATION OF KEY ERROR AND UNCERTAINTY CONCEPTS AND TERMS

PROF. DR. LUCA P. MARI

UNIVERSITÀ CATTANEO, ITALY

ABSTRACT

In the formal expression of any measurement result the measurand value must be stated together with an estimation of its quality, that reports all the non-idealities affecting the measurement procedure with respect to both its definition and its empirical accomplishment. Traditionally accounted for in terms of errors, such a quality estimate is evaluated and formalized as a measurement uncertainty, that can be assigned by suitably combining the available objective and subjective information according to a standard formal procedure. This procedure is briefly discussed and a practical example of its application is shown.

KNOWLEDGE LISTING

1. Measurement results and their quality
2. The concept of error
3. The concept of uncertainty
4. Characterizing a measurement with its uncertainty
5. The expression of measurement results and their uncertainty
6. The procedure for assigning the measurement uncertainty: an example

1: MEASUREMENT RESULTS AND THEIR QUALITY

Measurement is a peculiar means of acquiring and formally expressing information about empirical systems. It is aimed therefore at setting up a bridge between the empirical world and the linguistic/symbolic world, the domains of the systems under measurement and measurement results respectively.

A fundamental evidence is that these two realms exhibit extremely different characteristics. Empirical systems are embedded in the space-time universe, and this generates their space and time dependency: any system is only partially isolated from its environment and its dynamics forces to distinguish between the system itself and its temporal versions, i.e., the system states. On the other hand, symbolic entities such as numbers are coextensive with their definitions (in a sense: they *are* their definitions), so that they are always identical to themselves. Paradigmatically, noise exists in the empirical realm, not in the symbolic one; real numbers exist in the symbolic realm, not in the empirical one.

Whenever the two realms interact with each other, as measurement does by means of the mediation of quantities, these diversities (1) require to introduce a concept of quality of the symbols (in our case measurement results) chosen as representatives for empirical states and (2) are the cause of several issues affecting such a quality.

The typical operative context of measurement, that can be presented as follows:

Figure_1_near_here

shows that the required empirical results ("the output") can be in principle obtained as the transformation of the same empirical states ("the input") by a direct manipulation (an "empirical procedure") or a transduction to information entities, to be processed and finally transduced back to the empirical realm (an "informational procedure"):

Figure_2_near_here

The benefits of informational procedures are commonly recognized (basically due to the fact that it is much easier to deal with symbols than with empirical things), but they depend on the faithfulness of measurement results as representative entities for the corresponding empirical states.

Such a faithfulness, and therefore the quality of measurement results, is limited in consequence of causes related to:

- the model of the system under measurement: incompleteness, if not even faults, in the definition of the measured quantity (the {measurand}), as in the case of an ill-characterized system dynamics or an only partial identification of the quantities influencing the measurand;
- the operative accomplishment of the measurement procedure: poor {repeatability}, or {stability}, or {selectivity} of the adopted measuring system (see also mm_66), if not even faults in its usage.

The unavoidable presence of such flaws is the reason requiring us *to state any measurement result by expressing in symbols a measurand value together with an estimation of its deemed quality*.

2: THE CONCEPT OF ERROR

It is a well-known fact that the repeatability of measurements can be increased by:

- improving the measuring system in its empirical characteristics;
- reporting the results with a reduced number of significant figures,

i.e., by adjusting the sensing device or modifying the symbolic expression respectively:

Figure_3_near_here

The repeatability of a measurement, and in more general term its quality, is therefore a relative characteristic, to be evaluated in reference to the goals for which the operation is performed and the available resources (in epistemological terms this can be thought of as a confirmation that a concept of absolute, or complete, precision is simply meaningless).

It is amazing in this perspective to note that the indication of the estimated quality of the results became customary in physical measurement only in the late XIX century, and however several decades after the {Theory of Error} provided by Gauss at the beginning of that century. A plausible reason of this can be recognized in the commonly (in the past) assumed hypothesis that measurable quantities are characterized by a perfectly precise {“true value”}. The choice to adopt the concept of {error} to model and formalize a less-than-ideal quality of measurements originates from this hypothesis: any discrepancy between the measuring system outputs and the measurand true value should be taken into account as an error, and correspondingly dealt with (see also mm_156).

However:

- an error can be recognized as such only if a corresponding “right entity” exists;
- errors can be corrected only if their corresponding “right entities” are known;
- true values, that play the role of such “right entities” in the case of measurement, are in principle unknown (otherwise measurement itself would be useless…) and cannot be operatively determined.

These assertions imply that the Theory of Error is grounded on metaphysical, empirically inapplicable, bases. Consider the following two statements:

- “at the instant of the measurement the system is in a definite state”;
- “at the instant of the measurement the measurand has a definite value”.

Traditionally they would be considered as synonymous, whereas their conceptual distinction is a fundamental fact of metrology: the former represents a basic assumption for measurement (we are not considering here measurement in quantum mechanics), while the latter is epistemically unsustainable and however operationally irrelevant. Measurement results are symbolic, and not empirical, entities: what in the measurement is *determined*, and therefore considered pre-existing, is the system state, not the measurand value that is instead *assigned* on the basis of the {instrument reading} and the {calibration information}.

3: THE CONCEPT OF UNCERTAINTY

The search of a more adequate framework reached a crucial point about thirty years ago, when it was understood that a common approach for modeling and formally expressing a standard parameter describing the quality of measurement results was a condition to establish a strict co-operation among the national {calibration laboratories}. To build up and maintain a mutual confidence between accreditation bodies and compatibility for their {calibration certificates} required to have the quality of their measurement results evaluated and expressed according to some harmonized protocol. To this goal the *International Committee for Weights and Measures* (!CIPM!), started a project together with several international organizations involved in standardization (ISO, IEC, OIML, …): its final result is the {*Guide to the Expression of Uncertainty in Measurement*} (!GUM!), first published in 1993 and later introduced as a Standard by each of such organizations. While originally intended for calibration laboratories, the GUM is presently to be considered as the basis for expressing the results of any measurement performed in accordance with an international Standard.

According to the GUM, the uncertainty of a measurement result is "a parameter, associated with the result of a measurement, that characterizes the dispersion of the values that could reasonably be attributed to the measurand". Apart from this rather classical definition, the most important innovation of the GUM stands in its recognition that the uncertainty of measurement results can be evaluated according to two distinct and complementary methods:

- some uncertainties, designated as “of {type A}”, are computed as suitable statistics of experimental data, usually obtained as repeated instrument readings;
- some other uncertainties, designated as “of {type B}”, are instead estimated on the basis of the observer’s personal experience and the available a priori information, and therefore express a {degree of belief} on the possible measurand values.

The recognition that even measurement, an operation traditionally deemed as the paradigm of objective information acquisition, requires the introduction of some subjective evaluation is of capital importance. Therefore the shift from “error” to “uncertainty” is far more than a terminological issue, and witnesses a conceptual transition from an ontological position to an epistemic one: according to the GUM standpoint, to establish the quality of measurement results is an issue related to the *state of knowledge* of the measurer, and therefore “absolute quality” cannot be reached simply because some {*intrinsic uncertainty*} is always part of the measurement system.

The possible {sources of uncertainty} that the GUM itself lists are exemplar at this regards: together with the “variations in repeated observations of the measurand under apparently identical conditions”, the reason usually recognized for random variability, and some causes related to instrumental issues such as “approximations incorporated in the measurement procedure” and “instrument resolution or discrimination threshold”, the GUM identifies several epistemic sources, and among them the incomplete definition of the measurand and the imperfect realization of its definition (see also mm_154).

4: CHARACTERISING A MEASUREMENT WITH ITS UNCERTAINTY

To accomplish a measurement process three distinct activities must be sequentially performed:

- *acquisition*: by means of a sensing device the measurand is transduced to a quantity suitable for direct access by the measurer (e.g., the angular position of a needle with respect to a reference scale), possibly through the mediation of an "intermediate" quantity (a typical role for electrical quantities) to drive processing and presentation devices;

Figure_4_near_here

- *evaluation*: the access to the transduced quantity (i.e., the instrument reading) concludes the empirical part of the operation; by gathering and processing the available information (the transduced quantity itself, together with everything is known on the measurement system: the {measurand definition} and realization, the instrument {calibration diagram}, the values of relevant {influence quantities}, …) the measurer evaluates the measurand value and its uncertainty; this inferential process is based on both objective and subjective information;
- *expression*: the obtained information is expressed in symbolic form according to an agreed formalization.

It should be noted that the same information could be in principle expressed in different forms for different needs, by adopting, typically, a statistical or a set-theoretical formalization (or some generalization of the latter, as in the case of representations based on fuzzy sets: we will not deal with such generalizations here). Consider the traditional indication, $x \pm y$, that admits two distinct interpretations:

- the measurand value is expressed as the scalar x, with y as its estimated uncertainty;

- as the measurand value the whole interval [x–y, x+y] is taken, whose half-width, y, expresses the quality (sometimes called {*precision*} of such a measurement result.

Neither of them is the “right one”: they should be selected according to the specific application requirements. The GUM adopts this approach, and while basing its procedure on the first interpretation recognizes that “in some commercial, industrial, and regulatory applications, and when health and safety are concerned”, it is often necessary to express the measurement results by means of intervals of values.

Measurement results must be therefore assigned according to the goals for which the measurement is performed; they are adequate (and not “true”) if they meet such goals. By suitably formalizing them, the measurer is able to express the available information of both the measurand value and its estimated quality. Quoting the GUM again, no method for evaluating the measurement uncertainty can be a “substitute for critical thinking, intellectual honesty, and professional skill”: indeed “the quality and utility of the uncertainty quoted for the result of a measurement ultimately depends on the understanding, critical analysis, and integrity of those who contribute to the assignment of its value”.

5: THE EXPRESSION OF MEASUREMENT RESULTS AND THEIR UNCERTAINTY

For both type A and type B evaluation methods, the GUM assumes that measurands (but the same holds for all the quantities involved in the measurement system: influence quantities, correction factors, properties of reference materials, manufacturer or reference data, …) can be formalized as {random variables}, and as such characterized by statistical parameters:

- the measurand value is estimated as the {mean value} of the random variable; in the case of type A evaluations, for which an experimental population X of n repeated reading data $\{x_i\}$ is available, it is computed as:

$$m(X) = \frac{1}{n}\sum_{i=1}^{n} x_i \tag{1}$$

- the uncertainty of the measurand value is estimated as the {standard deviation} of the measurand value, being itself a random variable; this parameter is termed by the GUM {*standard uncertainty*} and denoted $u(m(X))$; in the case of type A evaluations it is computed as:

$$u(m(X)) = \sqrt{\frac{1}{n}\frac{1}{n-1}\sum_{i=1}^{n}(x_i - m(X))^2}$$

(2)

Measurement results can be then reported for example as m_S = 100,021 47(35) g, meaning that the evaluated mass m of the system S (whose specification should include the indication of the operative condition in which the measurement has been performed) is 100,021 g with a standard uncertainty of 0,35 mg.

The same couple of values (measurand value, standard uncertainty) is adopted to express measurement results as intervals. To this goal a *coverage factor k* (typically in the range 2 to 3) is introduced, such that $U(X)=ku(m(X))$, termed *expanded uncertainty*, is adopted as the half-width of the interval representing the measurement result: $[m(X)-U(X), m(X)+U(X)]$, commonly written as $m(X) \pm U(X)$ (if the probability distribution of the random variable is known this interval can be thought of as a *confidence interval*, whose confidence level is depends on k).

In the case of *derived measurement*, i.e., when the measurand Y is a quantity depending on N input quantities X_i, i=1,…,N:

$Y = f(X_1,\ldots,X_N)$ (3)

and for each quantity X_i, the estimated value $m(X_i)$ and uncertainty $u(m(X_i))$ are given, the issue arises of how to obtain the corresponding values $m(Y)$ and $u(m(Y))$ for Y.

The measurand value $m(Y)$ is simply obtained by introducing the estimates $m(X_i)$ in the model function f:

$m(Y) = f(m(X_1),\ldots, m(X_N))$ (4)

The uncertainty $u(m(Y))$ is instead evaluated by means of the so-called {*law of propagation of uncertainty*}, that for statistically non-correlated quantities is:

$$u^2(m(Y)) = \sum_{i=1}^{N} c_i^2 u^2(m(X_i)) \tag{5}$$

where the "sensitivity coefficients" c_i that define the extent to which Y is influenced by variations of the input quantities X_i are computed as:

$c_i = \dfrac{\partial f}{\partial X_i}$ evaluated at $X_i = m(X_i)$

In the general case of correlated input quantities (i.e., their covariance $u(m(X_i), m(X_j)) \neq 0$), the equation (5) becomes:

$$u^2(m(Y)) = \sum_{i=1}^{N} c_i^2 u^2(m(X_i)) + 2\sum_{i=1}^{N-1} \sum_{j=i+1}^{N} c_i c_j u(m(X_i), m(X_j)) \tag{6}$$

in which the combined standard uncertainty of the measurement result $m(Y)$ is computed on the basis of a first-order Taylor series approximation of equation (3) (see also mm_160, mm_161).

6: THE PROCEDURE FOR ASSIGNING THE MEASUREMENT UNCERTAINTY: AN EXAMPLE

Measurement uncertainty is a *pragmatic* parameter: its value is not intrinsic to the measurand but is to be established in reference to the specific goals according to which the measurement is performed. No "true uncertainty" exists, and the preliminary step of a procedure aimed at assigning a value to the uncertainty of a measurand value is therefore to decide a {*target uncertainty*}, the maximum value of uncertainty compatible with the given goals. In any step of the procedure, if the estimated value is reliably considered less than such a target uncertainty, then the procedure should be stopped with a positive result: the measurand can be evaluated with a satisfying uncertainty, and no further resources are required to refine the procedure. On the other hand, whenever the estimated uncertainty becomes greater than the target uncertainty the procedure must be definitely stopped with a negative outcome, conveying the information that better measurements are required to meet the specified target uncertainty.

According to the GUM viewpoint, *any* measurand Y should be actually evaluated by derived measurement, i.e., by firstly identifying its dependence on a set of "input quantities" X_i, such as influence quantities, calibration parameters, correction factors, … For each X_i, the values $m(X_i)$ should be obtained by statistical or other methods (e.g., as part of instrument specifications), and for each of such $m(X_i)$ the corresponding standard uncertainties $u(m(X_i))$ and covariances $u(m(X_i),m(X_j))$ should be evaluated, again by either type A or type B procedures.

In the case the functional relation f is known in its analytical form, the sensitivity coefficients c_i can be then computed; if, on the other hand, the complexity of the measurement system prevents the explicit formalization of the equation (3), each coefficient c_j can be experimentally estimated by a suitable setup of the system in which Y is repeatedly measured while all the quantities X_i but X_j are kept constant.

When at least some $u(m(X_i))$, $u(m(X_i),m(X_j))$, and c_i are available, the equation (6) can be computed to obtain an estimation of the measurand uncertainty $u(m(Y))$, to be compared to the specified target uncertainty.

REFERENCES

CIPM (1981) Proc.-Verb. Com. Int. Poids et Mesures 49, 8-9, 26, 1981 (in French).

ISO (1993) *Guide to the Expression of Uncertainty in Measurement*, International Organization for Standardization: Geneva, Switzerland.

ISO (1993) *International Vocabulary of Basic and General Terms in Metrology*, second edition, International Organization for Standardization, Geneva, Switzerland.

ISO (1999) *Geometrical Product Specification – Part.2: Guide to the Estimation of Uncertainty in Measurement in Calibration of Measuring Equipment and Product Verification*, International Organization for Standardization, Geneva, Switzerland.

Mari, L.; Zingales, G. (2000) *Uncertainty in Measurement Science*, in: Measurement Science – A discussion, Karija, K.; Finkelstein, L. (eds.), Ohmsha, IOS Press.

Taylor, B.N.; Kuyatt, C.E. (1994) *Guidelines for Evaluating and Expressing the Uncertainty of NIST Measurement Results*, NIST Technical Note 1297 (available on the NIST web site).

List of figure captions

Figure 1 – The role of measurement in the relations between empirical and symbolic realms. [mm_155_fig_1]

Figure 2 – The (possible) equivalence of empirical and informational procedures. [mm_155_fig_2]

Figure 3 – Abstract schematization of a measurement. [mm_155_fig_3]

Figure 4 – Abstract schematization of the empirical component of a measurement. [mm_155_fig_4]