Chapter 1
Basic Ideas of Measurement

To measure is to compare.

1.1 Prologue

As is generally known, only physical quantities of the same quality can be compared. Looked at in practical terms, measurements require a system of well-defined physical units. According to the recommendation of the International Bureau of Weights and Measures, those are given by the International System of Units, SI, for short. There is, however, a difference between definitions and realizations of physical units. A defined unit is a theoretical quantity. In contrast, a realized unit is implemented via a physical device and may therefore be affected by imperfections. Other errors come into effect when realized units are put into practice.

1.2 True Values

Natural scientists explore the laws of nature and express their findings in mathematically formulated physical laws, or, as the case may be, in principles.

Physical relationships complying with nature are considered true. Consequently, their variables and constants represent true values.—"True", self-evidently, in terms of the given set of physical units. Defective numerical data do not fulfill a physical law. Let us reflect on how to take in the idea of what a true physical relationship implies. To this end, we turn to a mathematical relationship, say, $f(x, y; a, b)$. Given the four quantities $x, y; a, b$, we find

$$z = f(x, y; a, b).$$

Vice versa, for any appropriately connected set of five data $x, y, z; a, b$, we expect, trivially,

$$f(x, y; a, b) - z = 0.$$

M. Grabe, *Measurement Uncertainties in Science and Technology*,
DOI 10.1007/978-3-319-04888-8_1, © Springer-Verlag Berlin Heidelberg 2014

Now assume $z = f(x, y; a, b)$ to correctly implement a law of nature and that there are true data, say, x_0, y_0, z_0 and a_0, b_0.—The index 0 will always be used to distinguish true values.

Obviously, we again have

$$f(x_0, y_0; a_0, b_0) - z_0 = 0\,.$$

The statement illustrates the togetherness of a true law of nature and the idea that the experimenter needs to have available, at least theoretically, the true values of the physical quantities in view. Now imagine he inserted defective values—he has to in the end, as he cannot proceed otherwise—then, by necessity, the relationship would break down. Let us cast this in an indeed elementary proposal:

Physically true relationships imply physically true values.

Hence, to every quantity to be measured, i.e. to every *measurand*, a true value has to be attributed. To deny this basic requirement would render physics per se meaningless. And in fact, it is precisely here where the drama of reasoning starts: Due to experimental imperfections or theoretical shortcomings, metrologists cannot find the true values of the physical quantities they are aiming at. Rather, they have to be content with more or less reasonable approximations. As the true values of the measurands are blurred, the answer to the question of whether or not the envisaged law of nature happens to be true remains also blurred—at least up to a certain degree. In other words, for one, the laws of nature require the existence of true values, for another, due to imperfections of whatever kind, true values remain unknown in principle. Indeed, these mutually exclusive findings present us with a dilemma.

1.3 Measurement Uncertainties

To localize the true value of a physical quantity, all experimenters can do is, firstly, to conceive an abstract mathematical construct, a so-called *estimator*, aiming at the true value of the measurand,

$$\text{estimator} \approx \text{true value}\,,$$

and, secondly, an appropriately estimated *measurement uncertainty*. As has been discussed, measurement uncertainties are induced experimentally, on the part of the operating mechanisms of the measuring devices. On formal grounds, the numerical values of measurement uncertainties turn out positive and have, afterwards, to be given a $\pm$ sign.

Naturally, it remains unknown whether the estimator is larger or smaller than the true value. If this were known, executable corrections would not yet have been implemented. Viewed in this light, experimenters have no choice but to resort to symmetrical uncertainty intervals.

Let x symbolize the quantity to be measured, $\bar{x}$ some conceived estimator and, finally, $u_{\bar{x}}$ the associated measurement uncertainty. Figure 1.1 discloses: The range $2 \times u_{\bar{x}}$ is expected to localize the true value x_0 of the measurand,

$$\bar{x} - u_{\bar{x}} \le x_0 \le \bar{x} + u_{\bar{x}}\,. \tag{1.1}$$

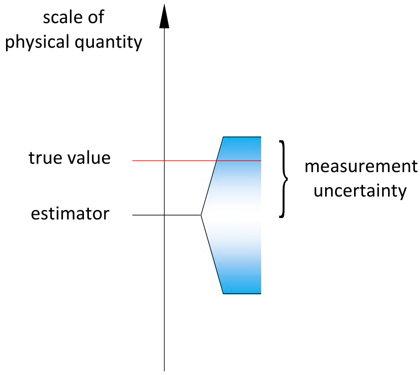

Fig. 1.1 The interval (estimator ± measurement uncertainty) is supposed to localize the true value of the measurand

Unfortunately, the centers as well as the lengths of suchlike uncertainty intervals evade deterministic fixations. Rather, the whole system of physical constants and physical quantities at large floats on a slightly blurry level. In a sense, nature has inserted something like a gray area in between physical reasoning and experimental findings.

Not surprisingly, proposition (1.1) is a "hot topic" in metrology and has in no way found a generally accepted judgment.

This monograph's formalism strictly sticks to objective statements. In particular, no postulates outside physical experience will be put forward. Let us note:

Laws of nature can only be proven within the limits of measurement uncertainties.

1.4 Breakdown of the Gaussian Approach

In his tracts on the Method of Least Squares, Carl Friedrich Gauß[1] distinguished *irregular or random* errors from *regular or constant* errors. Oddly enough, Gauss himself dismissed the latter arguing that it were up to experimenters to get rid of them. As a consequence, he based his formalism exclusively on irregular or random errors. Conceivably due to his authority, for a long time regular or constant errors stayed very much in the background, covered by a veil of oblivion.

Regular or constant errors re-entered the consciousness of experimenters—maybe even without remembering Gauss—from the moment their influence became experimentally overwhelming. The issue was then, of course, of how to proceed. The common practice focused and still focuses on the measure to submit systematic perturbations to a verbal transformation in character, certainly for the simple reason that the classical formalism would need no amendment if only regular or constant

[1]Though this is the proper spelling, in what follows, we shall refer to the international notation "Gauss".

errors entered the evaluation procedures featuring the properties of random quantities. To my knowledge, wherever regular or constant errors were taken into account, they were subjected to a notional conversion in character.

The consequences thereof became visible in the wake of the internationalization of metrology. More and more often, experimenters were confronted with contradicting measuring results. It appeared that the discrepancies were not entirely attributable to experimental causes. Rather, the error calculus itself might have become insufficient.

In the late 1970s, a seminar entitled *On the Statement of the Measurement Uncertainty* was held at the Physikalisch-Technische Bundesanstalt, Germany. At this seminar the proposal was made to treat regular or constant errors for what they really are, namely as quantities evoking biased estimators, Grabe, [20]. This observation initiated the break-down of the Gaussian error calculus.

1.5 Non-Gaussian Prospect

Nowadays Gauss's "irregular or random errors" are shortened to *random errors* while "regular or constant errors" are referred to as *unknown systematic errors*.[2]

In general, we may assume random errors and unknown systematic errors to be of comparable orders of magnitude. Beyond that, at least in principle, both errors are random in character. Yet there is a difference: with respect to random errors we have available, say indirectly, quantitative access via the scattering of repeated measurements. With respect to unknown systematic errors all we know is that they are time-constant perturbations—as long as the experimental device is not subjected to what is colloquially called "a drift"—and, that they are, in particular, unknown with respect to magnitude and sign. Nonetheless, unknown systematic errors prove assessable via intervals of fixed lengths. Some unknown systematic errors happen to occur during the assemblage of the measuring device, others take effect at the start of the operation. Let f denote an unknown systematic error. Given the confining interval proves symmetric to zero, we have

$$-f_s \leq f \leq f_s\,; \quad f = \text{const.}\,, \quad f_s \geq 0\,. \tag{1.2}$$

At this the boundaries $\pm f_s$ should be safe. From there a complemental statement of the kind "f should not lie outside the quoted limits" appears needless. The assessment of the interval thresholds proves intricate. Not without good reason the procedure points at what metrologists perceive "the ultimate art of measurement".

[2]Errors are always unknown, hence it might well be adequate to contract the term "unknown systematic error" to "systematic error". Conceivably, the term unknown systematic error is due to the period when metrologists also considered "known systematic errors", meanwhile, however, labeled corrections.

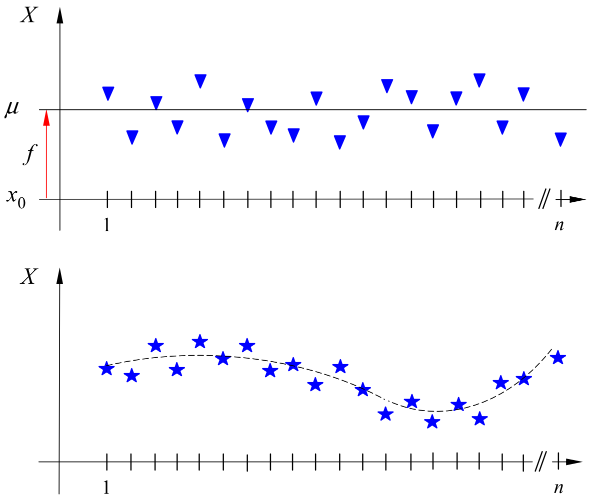

Fig. 1.2 Non-drifting (*above*) and drifting (*below*) sequences of repeated measurements

To reduce the influence of random errors, it seems reasonable to repeat the measurements, say, n times under the same conditions and to subsequently condense the set of values[3]

$$x_1, x_2, \ldots, x_n$$

to an appropriately designed estimator targeting at the unknown true value. The questions of how the estimator might look and of how large n should be, will be discussed soon. Experimenters aspire to non-drifting measuring devices. Figure 1.2 illustrates a non-drifting device (upper drawing) and a drifting device (lower drawing). In what follows, the treatise will confine itself to non-drifting experimental set-ups.

A non-drifting device displays repeated measurements scattering randomly about a "horizontal" line. Given that the width of the scattering remains unaltered during the period the measurements are taken up, the experimenter may consider the underlying statistical process stationary. Non-drifting measuring devices and statistical stationarity mark the most favourable experimental situation. For this, C. Eisenhart [9] once coined the vivid term *state of statistical control*.

Indeed, this is the best situation the experimenter can hope for. In terms of physical reality, the device is said to behave "perfectly". Over longer periods, however, even the best measuring devices may change their properties, notably, unknown systematic errors might start to drift or undergo abrupt changes.

Whether or not there is an unknown systematic error, the experimenter cannot observe directly. However, experience tells him that he has to anticipate systematic perturbations as indicated in Fig. 1.2, (upper drawing). Here, the experimentally

[3]As a zeroth measurement does not exist, the index 0 in x_0 unambiguously distinguishes the true value from the measured data $x_1, x_2, x_3, \ldots$.

invisible unknown systematic error f has shifted the bulk of measured data as a whole "upwards", away from the true value x_0. At the same time, the experimenter himself cannot perceive any shift at all. In particular, he cannot know of whether f has shifted the bulk of data "upwards" or "downwards" which might equally well have happened. As a result, a naive observer might contest the existence of a shift or preclude it altogether. But in fact, this is the long-term damage of Gauss's postulate that regular or constant errors are eliminable on the part of the experimenter.—Hence, the classical Gaussian error calculus needs to be improved.

In what follows, we shall assume the random scattering of the measured data to be normally distributed—at least in the sense of a reasonable approximation. Distributions other than normal, of interest e.g. in nuclear physics, are treated elsewhere [52]. Given normality is conceded, the parameter μ marks the center of the normal distribution density.

Let us draw a distinction between theoretically defined parameters and their empirical counterparts called *estimators*. The parameter μ is of theoretical origin and hence unknown by its very nature. All experimentalists can do to localize a theoretical parameter is to resort to an appropriate empirical estimator inclusive its uncertainty. This is metrology's strategy and the actual purpose of the book. Figure 1.2, (upper drawing) suggests the fundamental proposition

$$\mu = x_0 + f\,. \tag{1.3}$$

From a physical perspective the true value x_0 is fixed by nature and the parameter μ has got shifted away from x_0 by some unknown systematic error f. As neither the magnitude nor the sign of the systematic error f is known, the experimenter has no idea where to find x_0. Actually, the unknown true value may lie either "above" or "below" μ.

Let us assume the interval confining f to be unsymmetric to zero, say,

$$-f_{s,1} \le f \le f_{s,2}\,; \quad f = \text{const.}\,,\; f_{s,1},\, f_{s,2} \ge 0\,. \tag{1.4}$$

With a view to Fig. 1.3 the systematic error and the measured data may be submitted to an elementary transformation reestablishing symmetry. The idea is as follows: If there are experimental findings *definitely* speaking against a symmetric interval $-f_s \le f \le f_s$, this information can serve to subtract the term

$$\frac{(-f_{s,1} + f_{s,2})}{2} \tag{1.5}$$

from both the asymmetric boundaries and the measured data, thus producing a re-symmetrized interval

$$-f_s \le f \le f_s \quad \text{where} \quad f_s = \frac{(f_{s,1} + f_{s,2})}{2}\,; \quad f_s \ge 0 \tag{1.6}$$

and a shifted data set

$$x_l' = x_l - (-f_{s,1} + f_{s,2})/2, \quad l = 1, \ldots, n. \tag{1.7}$$

For convenience, we shall abstain from explicitly priming the transformed data. Rather, we shall tacitly assume the corrections to have already been made. Figure 1.4

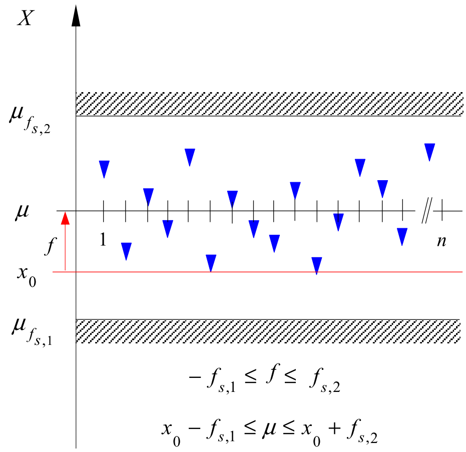

Fig. 1.3 Unsymmetric interval confining f

summarizes the basic metrological operations resting on an error model assuming normally distributed random errors and time-constant unknown systematic errors of the property (1.2).

While random errors occur instantaneously, i.e. in the course of the current measurements, systematic errors become fixed before any measurement has been carried out or immediately at the start thereof. Random errors are to be attributed to microscopic fluctuations of any events affecting the measuring process and prove to be uncontrollable on a macroscopic scale. Though we might not be in a position to understand these phenomena in detail, we find guidance in the central limit theorem of statistics, see e.g. [50]. According to this, loosely stated, the superposition of sufficiently many similar random perturbations tends to a normal distribution. Actually, for practical reasons, the statistical properties of measured data are rarely scrutinized. From there we observe that if the assumption of normally distributed data fails the procedures estimating the random part of measurement uncertainties, as outlined in the following step by step, will not apply.

Compared with random errors, the status of unknown systematic errors is very different. In experimental terms, systematic errors may be due to the finite exactness of adjustments and pre-settings, to the limited knowledge of environmental and boundary conditions. As an example, let us address air buoyancy in mass calibrations. Though the density of air is predictable with remarkable accuracy, calibration procedures are still burdened with non-negligible systematic uncertainty components [28]. Other systematic errors may stem from instabilities of power supplies, thermal changes of mechanical components or detector off-sets. Even the switching-on of a measuring device may not lead to a unique stationary state.

Concerning theoretical conceptions, systematic errors may enter when physical models are put into practice. In particular, inherently different procedures for measuring one and the same physical quantity are apt to bring about varying systematic errors [14, 31]. Last but not least, hypotheses entering the test of a physical model are to be considered to give rise to systematic effects.

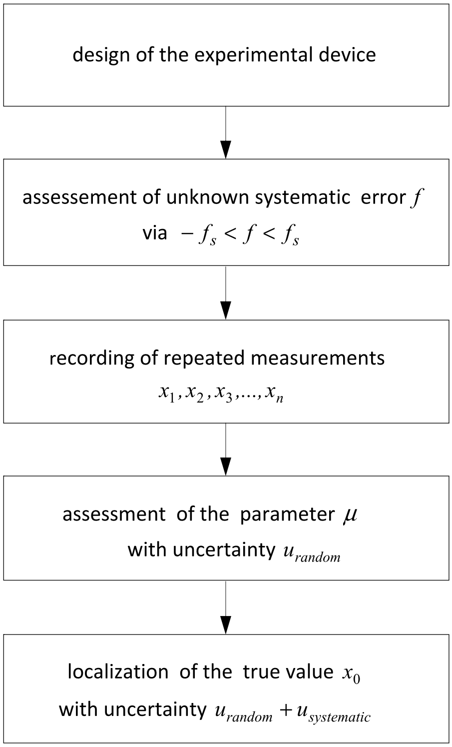

Fig. 1.4 Inference of the true value, schematic

As there is no chance of ignoring or eliminating time-constant systematic errors, experimenters have to localize them within intervals, the respective lengths of which stemming from the properties of the measuring device, the conditions under which it is operated and, as the case may be, from theoretical conceptions.

Let us revert to the question of how often a measurement should be repeated. At the first, a large number of repeated measurements turns out meaningful only if the resolution of the measuring device is sufficiently high. If the resolution is low, it should hardly make sense to record a substantial number of repeated measurements as a coarse device responds to one and the same input, applied repeatedly, with a limited number of different outcoming values. But even if the resolution of the measuring device is high, large numbers of repeated measurements will in no way reduce the influence of systematic errors.

Another practical experience points to potential drifts of measuring devices. Even though we excluded suchlike drifts explicitly, in the long run measuring devices may nevertheless tend to drift more or less. Hence, experimenters should control the measuring conditions so as to keep unknown systematic errors constant during the time needed to register a series of repeated measurements, say, $x_1, x_2, \ldots, x_n$. Given the measurements are repeated at a later instant in time or in another place, there may be other systematic errors. From there, the extensions of the intervals confining systematic errors should be such that their effective values steadily remain

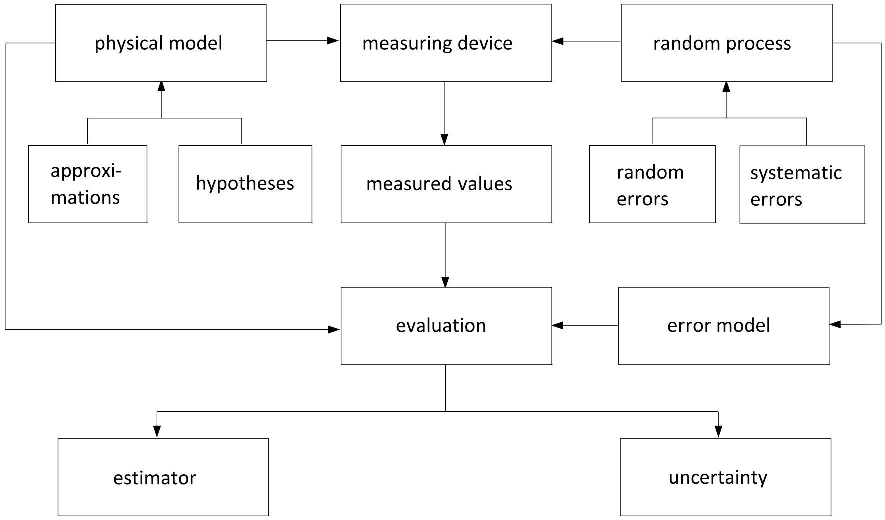

Fig. 1.5 Measuring process, diagrammatic

enclosed within the allotted intervals. Indeed, this basic requirement should guide the assessment of systematic errors. After all, independent of the time and the place in which a measuring device is used, the intervals limiting systematic errors should be reliable.

From there, smaller samples of repeated measurements may appear more meaningful than larger ones. At any rate, as stated before, averaging arbitrarily many repeated measurements, burdened by time-constant unknown systematic errors, does in no way reduce the influence of the latter.

Ultimately Fig. 1.5 particularizes the measuring process depicted in Fig. 1.4. As stated above, we shall address unknown systematic errors more briefly than systematic errors as, by their very nature, they are just as unknown as random errors.

1.6 Appraising Measurement Uncertainties

The non-Gaussian prospect reminds us to handle terms such as *accuracy*, *repeatability*, *reproducibility*, *precision*, etc. with some detachment. Here, we shall limit ourselves to the terms accuracy and repeatability as they appear serviceable as well as unambiguous.

Accuracy refers to the localization of the true value of the measurand: the smaller the uncertainty the higher the accuracy. Virtually, accuracy is an illustrative transcription of that what measurement uncertainty numerically quantifies.

Repeatability, in the proper sense, refers to one and the same measuring device operating under invariable conditions. The idea is to quantify the statistical scattering of the measured data at it requiring strictly time-constant systematic errors: the closer the scattering the higher the repeatability. Thus, repeatability aims at nothing else but the observable effects of random errors alone, disregarding the invisible influence of the hidden unknown systematic errors.

For the terms reproducibility and precision potential definitions will not be discussed here. At any rate, the user should be mindful of the conditions and the meaning of such specifications.

A highly controversial, intricate question refers to the reliability of measurement uncertainties. On the one hand the true values of the measurands are inaccessible. On the other, uncertainties are due to formal appraisals. Hence, at first glance, there seems to be no way to verify as to whether a given uncertainty localizes the quested true value. Though this applies in principle, there nevertheless is an escape in so far as we may resort to simulated data. Indeed, modeling data so as to implement the properties of the considered experimental device, we may verify of whether or not the related uncertainty, as issued on the part of the simulated data, localizes the pre-defined, fictional true value. This is how we shall proceed.

It is neither necessary nor advisable to always strive for highest accuracies. It is, however, imperative that we know the (quantified) uncertainty of the measurement result and that we can be sure that the result localizes the true value of the measurand.

The national metrology institutes are committed to ensure the realization of units and measures by means of hierarchies of standards. The uncertainties are considered to be clearly and reliably documented so that experimenters can count on them at any time and any place.

In the natural sciences, measurement uncertainties play a crucial role. Hypotheses concerning principles of nature are either to be endorsed or to be rejected by comparing theoretical predictions with experimental findings. Though there will always be something like a "gray area", eventually measurement uncertainties may well decide whether new ideas should be considered true or better be left in a waiting position.

In engineering sciences measurement uncertainties provide an almost unlimited diversity of intertwined dependencies which, in the end, are crucial for the functioning, the quality, the service life and the competitiveness of industrial products.

Trade and industry categorically call for mutual fairness: both, the buyer and the seller expect the measuring processes to subdivide merchandise and goods so that neither side will be overcharged.[4]

To have a basic example for the statement (1.1), let us consider the result of the weighing of some mass m,

$$(10.31 - 0.05)\ \mathrm{g} \le m_0 \le (10.31 + 0.05)\ \mathrm{g}\,.$$

Hence, the true value m_0 is expected to lie somewhere within the interval 10.26 g ... 10.36 g.

Notably, we recall:

[4]One might wish to multiply the relative uncertainties of the counters for gas, oil, petrol and electricity by their annual throughputs and the respective unit prices.

Rating measurement uncertainties, we should abstain from using terms like optimistic or pessimistic assessments. If uncertainties are kept too tight they might pretend effects which do not exist; if kept too large they might hide valuable findings. Rather, uncertainties should be kept objective thus mediating the contest between theory and experiment on the basis of equitable criteria.

The uncertainty $u_{\bar{x}}$ is related to other specifications. For instance, the quotient

$$\frac{u_{\bar{x}}}{\bar{x}} \tag{1.8}$$

denotes the relative uncertainty. We might wish to substitute x_0 for $\bar{x}$. This, clearly, is prohibitive, as true values are unknown. Relative uncertainties vividly illustrate the accuracy of measurements independent of the order of magnitude of the estimators. Let U_1 and U_2 designate any physical units and let

$$\bar{x}_1 = 10^{-6}\,\mathrm{U}_1\,, \quad \bar{x}_2 = 10^{6}\,\mathrm{U}_2$$

be estimators with uncertainties

$$u_{\bar{x}_1} = 10^{-8}\,\mathrm{U}_1\,, \quad u_{\bar{x}_2} = 10^{4}\,\mathrm{U}_2\,.$$

Obviously, the two estimators have been assessed with equal relative uncertainties, though their absolute uncertainties differ by no less than twelve orders of magnitude.

There are uncertainty statements being still more specific. The frequently used ppm uncertainty, ppm meaning “parts per million”, takes reference to the decimal power of 6. The quotient

$$\frac{u_{\bar{x}}}{\bar{x}} = \frac{u_{\mathrm{ppm}}(\bar{x})}{10^6}$$

leads to

$$u_{\mathrm{ppm}}(\bar{x}) = \frac{u_{\bar{x}}}{\bar{x}}\,10^6\,. \tag{1.9}$$

The idea is to achieve a more comfortable notation of numbers by getting rid of decimal powers which are known a priori. As an example, let $u_{\mathrm{ppm}} = 1$. This implies $u_{\bar{x}}/\bar{x} = 10^{-6}$.

Example 1.1: Josephson effect The frequency ν of the inverse ac Josephson effect involves dc voltage steps $U = (h/2e)\nu$; e designates the elementary charge and h the Planck constant. Let the quotient $2e/h$ be quantified as

$$4.8359767(14)\,10^{14}\ \mathrm{Hz\,V^{-1}}\,.$$

How is this statement to be understood? It is an abbreviation and means

$$2\overline{e/h} \pm u_{2\overline{e/h}} = (4.8359767 \pm 0.0000014)10^{14}\ \mathrm{Hz\,V^{-1}}\,.$$

Converting to ppm yields

$$u_{\mathrm{ppm}} = \frac{0.0000014 \cdot 10^{14}\ \mathrm{Hz\,V^{-1}}}{4.8359767 \cdot 10^{14}\ \mathrm{Hz\,V^{-1}}} \cdot 10^6 = 0.289\ldots$$

or, rounding up,

$$u_{\mathrm{ppm}}(2\overline{e/h}) = 0.30\,.$$

1.7 Quotation of Numerical Values

Evaluation procedures nearly always endow estimators with dispensable, insignificant decimal digits. To get rid of them, the measurement uncertainty should tell us where and how we have to round and, in particular, how to quote the final result [44].

Rounding of Estimators

First of all, we determine the decimal place in which the estimator should be rounded. Given that the first decimal of the measurement uncertainty differing from 0 turns out to be one of the digits

$$\left.\begin{array}{r} 1 \text{ or } 2 \\[1.5em] 3 \text{ up to } 9 \end{array}\right\} \quad \text{the estimator should be rounded} \quad \left\{\begin{array}{c} \text{in the place to the right} \\ \text{of this place} \\[0.5em] \text{in just this place} \end{array}\right.$$

Having determined the decimal place of the estimator in which it should get rounded, we round it

$$\left.\begin{array}{r} \text{down,} \\ \text{up,} \end{array}\right\} \quad \begin{array}{l} \text{if the decimal place to the right} \\ \text{of this place is one of the digits} \end{array} \quad \left\{\begin{array}{l} 0, \text{ up to } 4 \\ 5, \text{ up to } 9 \end{array}\right.$$

Rounding of Uncertainties

The uncertainty should be rounded up in the decimal place in which the estimator has been rounded.

Sometimes the numerical value of a physical constant is fixed in view of some quoted condition. It goes without saying that fixing a physical constant does not produce a more accurate quantity. A possibility to emphasize a fixed digit would be to print it boldly, e.g. 273.1**6** K.

Example 1.2: Rounding of estimators and uncertainties Let

$$(7.238143 \pm 0.000185)\,\mathrm{g}$$

be the result of a weighing. The notation

$$\begin{array}{l} 7.2381\,4\,3 \\ \uparrow \\ 0.0001\,8\,5 \end{array}$$

facilitates the rounding procedure. The arrow indicates the decimal place in which the estimator should be rounded down. In just this decimal place the uncertainty should be rounded up, i.e.

$$(7.23814 \pm 0.00019)\ \text{g}.$$

For simplicity in Table 1.1 the units have been suppressed.

Table 1.1 Rounding of estimators and uncertainties

Raw results			Rounded results		
5.755018...	±	0.000194...	5.75502	±	0.00020
1.9134...	±	0.0048...	1.913	±	0.005
119748.8...	±	123.7...	119750	±	130
81191.21...	±	51.7...	81190	±	60

http://www.springer.com/978-3-319-04887-1

Measurement Uncertainties in Science and Technology
Grabe, M.
2014, XIV, 401 p. 80 illus., 41 illus. in color., Hardcover
ISBN: 978-3-319-04887-1