1 Measurements, errors and estimates

1.1 Error and Uncertainty

In general we can say that we all agree more or less on the meaning of the word *error*. Error gives indications on experimental results and the quality of measurements. It thus helps to arrive at conclusions on experimental data and discern their reliability.

The experimentalist knows, starting from design phase, that his experiment will not be perfect despite all his efforts. He/she knows that the ideal *perfect* experiment will not be matched by the *real* one. Everyone knows that measurement is no more than an approximation of the ideal. To quantify this approximation one must employ common tools in describing experimental results, their quality and considerations on them as unambiguously and accessibly as possible. These tools, which are introduced here, are an essential part of the experimentalist's education.

Let us introduce a definition of error that best follows our *common sense* as 'doubt about a measurement' and let us make sure that others share this definition and, on this basis, draw the same conclusions.

The concept of error is associated, in the definition given in recent *I.S.O.* [1] rules, with the more general term of *uncertainty*, with which the experimental result is evaluated and with which it is associated.

As concerns the scientific disciplines, we take for granted that the best evaluation of an experimental result is the numerical one [2], which simplifies synthesis and facilitates the communication of results. We therefore consider this also as a criterion in evaluating the *uncertainty* we wish to communicate. In what follows we will try to give a numerical evaluation of *uncertainty* associated with experimental results. However, it is often stated that one wishes to determine the *accuracy* of

[2]This is not necessarily true but, in a modern teaching approach common to the great majority of teachers and researchers, is taken for granted.

the experimental result [3]. Contrary to what one may believe, it is not true that the only information one has is the result itself. The experimentalist also applies his/her own judgement in evaluating the validity of a measurement. This is the reason why the concept of uncertainty, which is a general one, must be kept separate, including in the numerical expression the different origins that lead to a certain error.

In what follows we will schematically divide the possible causes of errors and present an evaluation of the corresponding uncertainties which can be encountered, at least in most experimental situations. We will do this more for the sake of a common terminology than to limit the experimentalist's judgement.

1.1.1 Resolution and reading error

We can distinguish a type of error which we make, for example, <u>once</u> on measuring a single time interval with a single measurement. We suppose that this measurement has been taken using one of our sensorial organs: the eye. As an example, we can use the reading of an instrument in a certain interval of time.

This definition of measurement may seem arbitrary, but actually it is quite rigorous and points up arbitrariness in the measurement, both in the choice of the sensorial organs (the kind of experimental detection) and of the interval of time.

Let us try to understand through an example. Let us suppose we want to read temperature by means of a thermometer. We do this by observing marks associated with numerical values corresponding to the height of the mercury column. In reality, we are making a measurement

[3] "*Accuracy*" and "*precision*" are often confused, but the concepts are quite different depending upon the contexts in which they are used. Several definitions of *precision* can be found in the references. This term will not be used here. By *accuracy* we mean the qualitative evaluation, as presented here, of the validity of the measurement. Therefore, great accuracy signifies small uncertainty and vice versa. It is not, however, its opposite, since the latter will be given a precise numerical value.

with the eyes and we are choosing a reasonable interval of time. We commonly use the blink of the eye, which is more or less one second. If we chose a longer interval, for example 1000 seconds, we know very well that the measurement would be different. On the other hand, if we choose to determine the height of a person with a yardstick, we know very well that the measurement will give us the same result whether it is made during an interval of one second or during an interval of 1000 seconds. This example shows that we cannot define experimental measurement in a general way, but must relate the definition to the single experimental situation. To arrive at a general criterion we can base our reading error on the specifications of the instrument. The manufacturer provides the definition of the instrument's characteristics by giving the measurement resolution [4].

If the measurements were just as *accurate*, the resolution or uncertainty would correspond exactly with the *reading error*.

Let us go back to the example. The height of the mercury column has to correspond to a certain mark on the ruler. The manufacturer's mark corresponds, for example, to one degree. From that we cannot infer that the resolution of the instrument, namely the reading error, is 1 °C. This would be wrong, or in any case arbitrary. Very often, in fact, the resolution is not indicated in the technical specifications. Even when this is indicated, what we are about to say is still valid.

Let us consider two different thermometers, one small enough to hold in our hand, the other two meters high. Let us suppose they have two identical scales from one end to the other. By identical scales we really mean that they have the same value in degrees°C from the beginning to the end of the scale (for example from 30°C to 40°C) and that they have the same number of marks (eleven). They will have the same mercury bulb at the base, of course of different size. If we follow the same line of reasoning as before, they should give measurements with

[4] In reality, what is given is an *intrinsic error* or a *fiducial error*, determined with respect to standard reference values.

the same accuracy. However, on the larger thermometer two consecutive marks are much further apart (without even quantifying, it is enough to consider the fact that there are several centimeters between them) than the marks on the smaller one. The eye can distinguish positions of the mercury column between two marks on the larger thermometer that it cannot evaluate on the smaller one. Experimentalists may thus be induced to determine the resolution of the instrument as being greater than the distance between two marks. The reading error, as well as the measurement, would not have a general definition. On the contrary, it would be arbitrarily determined by each experimentalist in a different manner. In fact, this is the case. Here we can only give commonsense indications, which are discussed at length in chapter ??.

When reading instruments, the marker (for example the arrow of a scale or, as in this case, the top of the mercury column) should be much smaller than the interval between two consecutive marks. If this is not the case, evaluation of uncertainty, which is to say estimating the reading error, must be performed by taking the inter-distance between two consecutive marks. Depending on to what extent this is verified, we can estimate the resolution as equal to the interval between two marks divided by two, three, four, and so on. It is clear, however, that the manufacturer will put appropriate, and not arbitrary, marks on the instrument. We do not consider reasonable, and experience supports us in this, a presumed resolution larger than the inter-distance between the marks divided by four.

These are indications of a wide consensus in the evaluation of the reading error of an instrument in the case of experimental measurements made by the sensorial organ of the eyes [5].

In what follows we will find also a more rigorous justification of the estimation of one half or of one quarter of the interval between two

---

[5] In the case of measurements taken by other sensorial organs, like the ear, it is more complicated. Suffice it here to indicate that measurements taken by the sensorial organ of the ear are far more arbitrary [2].

consecutive marks.

For the moment let us go on with the discussion of error now that we have a definition of the estimate of uncertainty due to the reading error, which will be used for the definition of casual error.

1.1.2 The illegitimate error

While measuring, we may make wrong assumptions, mistakes in calculating or use the wrong units. We can call this kind of error *illegitimate.* Illegitimate errors can later be corrected by more careful analysis. Here we can also include all possible operations that allow us to make better measurements by improving environmental conditions or using better instruments. In the following we shall give an example of the difficulty in defining illegitimate error.

Example: the problem of meter positioning in the experiment with coaxial cables

In the case of the experiment for the determination of the speed of light in coaxial cables (described in the appendix), we need to measure an electrical cable with a meter bar that is slightly shorter than the cable. One tries to place one end of the cable exactly at the end of the metre bar. But then, since the cable is longer, one has to move the meter bar. During this operation the cable often bends owing to its normal elastic properties. One therefore has to put one finger on the cable at the one-metre mark and at the same time keep the two parallel - a difficult operation, especially if one is using a carpenter's metre. To complete the measurement one then has to position the end of the metre once again and measure the remaining part (uncertainties must be added up, as discussed in the following chapters).

In the first reading it is legitimate to consider just the reading error of the meter, which corresponds to the smallest mark on the scale, let us say, 0.1 centimetres. But in the second measurement, if we take the same reading error, which is later to be added to the first to have the full uncertainty of the measurement over the entire length of the cable, we could argue that the smallest mark, or better the smallest interval between two marks, underestimates the error. This is due to the fact that positioning for the second measurement is not as açcurate as it was for the first. One student pointed out that the error could be estimated as equal to the width of the thumb with which he marked the end of the first measurement of the cable and positioned it for the second.

This case is typical of an illegitimate error which experimentalists have to keep to a minimum. It is a question of the experimentalist's properly positioning the cable and not allowing it to bend, and the meter. One could also mark the end of the first measurement on the cable with a pen, for instance, which would certainly be more accurate than the thumb.

Similar to this are reading errors on a scale due merely to incorrect positioning of the reading plane with respect to the scale plane [6]

1.1.3 Uncertainty and systematic error

By this we mean something with less defined contours, which are common to many different experimental situations. This is due to the very nature of this kind of error. The systematic error must be distinguished from the illegitimate error because it is quite difficult to know in advance if we can avoid it. For example, one can call systematic error the one due to the uncalibrated scale of a thermometer or meter bar. If the manufacturer of a meter makes a mistake in numbering a scale, e.g. going directly from 10 to 30 omitting 20, one then all measurements

[6]This is very often called *parallax error*, because the axis orthogonal to the scale plane is not orthogonal to the plane of our eyes. But in this case the definition refers to the orthogonal axes and not to the planes themselves. This is very often classified among the systematic errors, which have a completely different meaning in this book.

above 10 will contain a systematic error of 10 cm. *Systematically* or *repeatedly* each reading will contain an error that can be evaluated as a *systematic uncertainty* equal to 10 cm. If we notice the defect after taking all the measurements, we can evaluate this defect numerically and make allowances for it in the experimental results. In this case we call it *illegitimate error.*

Experimentalists may also fail to realise that there are illegitimate errors or may have introduced more subtle measurement defects or ones generated by themselves. In this case they can only compare their measurements with those made by others *with different instruments*. In this way, on noting a discrepancy they can attribute to the measurement an evaluation of the systematic error, *in all compared measurements and for all measurements of all other experimentalists.*

1.2 Uncertainty and casual error

The conceptual distinction between the errors presented above is that *the casual error* [7] cannot be cancelled, neglected or evaluated a posteriori. This is true even if we reduce it so much as to be far smaller than the experimental result, which may occur in one of the two previous cases of errors. This is precisely the error that we have to evaluate and associate with the measurement since we are assuming that we can-

[7]As defined in most textbooks, it would be more appropriate to term the casual error the *probable*. This is also the name with an analogous definition given in reference [1]. Here it would be more suitable not to fault the definition but the name to ascribe to it for teaching purposes. Students are often confused by the word *casual*; they ascribe to it a meaning that is not synonymous with *accidental* and include in it the concept of systematic error and illegitimate error. By *probable* error one often means a quantity corresponding to an interval of values. In this textbook we discuss this quantity again in the last chapter. But we do not adopt its definition in common use (also not adopted in *I.S.O.*) regulation) nor its use.

l_1	$8.0cm$
l_2	$8.1cm$
l_3	$7.9cm$
l_4	$8.1cm$
l_5	$8.2cm$

Table 1.1: Values of five measurements of length of a pencil.

not have *perfect* measurements in any case [8]. By definition we can have only an estimate of it. If we knew the exact error, we would know the result itself with absolute accuracy and therefore it would not be necessary to evaluate the error. This would contradict what we said earlier on. It is the estimate which can be quantified in a rigorously mathematical way. Here we call the estimate of this error *casual uncertainty*. The methods used to obtain this estimate are the basis of the statistical analysis of experimental data. To obtain it, in fact, concepts and methods of mathematical analysis and probability calculations are required.

A first, easily intuitive and commonly used method can be deduced by a simple estimate of the numerical result itself by determining the arithmetical mean.

Let us consider the measurements in table 1.1, of lengths l_i of a common object such as a partly used pencil.

With $n = 5$ measurements, one can obtain the arithmetical mean of the measurements of the length, which is indicated as $\bar{l}$:

$$\bar{l} = \frac{\sum_{i=1}^{n} li}{n} = \frac{(8.0 + 8.1 + 7.9 + 8.1 + 8.2)cm}{5} = \qquad (1.1)$$
$$= \frac{40.3cm}{5} \cong 8.06cm \cong 8.1$$

[8]This is taken for granted here, but in reality there is experimental evidence that this is really the case. It will be verified later on.

l_1	$8.0cm$	l_{11}	$8.0cm$	l_{21}	$8.0cm$
l_2	$8.1cm$	l_{12}	$8.0cm$	l_{22}	$8.0cm$
l_3	$7.9cm$	l_{13}	$8.1cm$	l_{23}	$7.8cm$
l_4	$8.1cm$	l_{14}	$8.0cm$	l_{24}	$8.0cm$
l_5	$8.2cm$	l_{15}	$7.8cm$	l_{25}	$8.0cm$
l_6	$8.2cm$	l_{16}	$7.7cm$		
l_7	$8.0cm$	l_{17}	$8.0cm$		
l_8	$8.2cm$	l_{18}	$8.0cm$		
l_9	$8.1cm$	l_{19}	$8.1cm$		
l_{10}	$8.0cm$	l_{20}	$8.0cm$		

Table 1.2: Values of 25 measurements of length of a pencil.

By making a much larger number of measurements, for example by having 24 other colleagues measure it, we can expect to obtain data of the kind in table 1.2.

We therefore have

$$\bar{l} = \frac{\sum_{i=1}^{n} li}{n} = \frac{\sum_{i=1}^{25} li}{25} = \frac{200.4cm}{25} \cong 8.0cm \tag{1.2}$$

This is an *estimate* of the value of length of the pencil.

We will come back(chapter 3.2) to this problem of significant digits. The digits of equations 1.1 and 1.2 are correctly evaluated, but they are not justified.

It is clear that the experimentalist knows that the pencil has only one length [9]. This is what is called *true value* or *exact value* obtained by a conceptual extrapolation and mathematically defined. We can then

[9] It is incorrect to write in this way, but to give full explanation it is necessary to introduce concepts of quantum mechanics, which the students may not be prepared for.

state that the arithmetical mean of equation 1.2 is the estimate of the true value. Actually, this is not exactly the case, and we will see later on how we can improve on this definition.

We started our discussion on the determination of uncertainty by assuming that we can make an estimate of it, as well as of the real value. We will therefore have to make an idealisation both of the measurement and of the error.

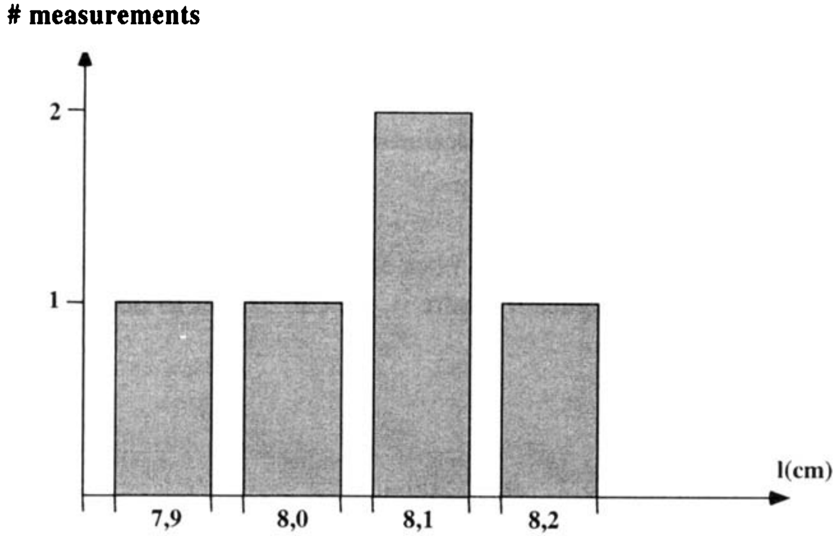

Figure 1.1: Histogram for five measurements of length of a pencil.

Let us suppose we have made an even larger number of measurements. For example, by having 24 other colleagues measure the pencil, just *"to be more certain"* of the result. In this case we are assuming that by increasing the number of measurements used in the determination of the arithmetic mean, the result should be more accurate. To quantify this statement, we are therefore stating that the error we are making should be smaller and therefore the uncertainty associated with the measurement should also be smaller.

Before we continue improving the definition, we see that by increasing the number of measurements, the arithmetical mean becomes a progressively better estimate of the true value of a quantity [10].

To show that, it is useful to write the values in a simple plot. In the abscissa we will put the values of the measurements and in the ordinates the values of the number of times this value has been obtained by measurement [11]. This plot is called a *histogram*. Figure 1.1 illustrates the distribution of data in table 1.1.

By comparing this figure with figure 1.2, for table 1.2 it is clear that measurements tend to assume a value that approaches the value obtained with the arithmetic mean.

We can now implement an idealised reasoning process. By making an ever-increasing number of measurements, we will obtain a progressively more accurate measurement, which is to say a better estimate of the true value, from the arithmetic mean. For an infinite number of measurements, we obtain a true value, which is the best estimate of itself, and which is the true value itself. By translating this reasoning into mathematical terms, we can write that the true value is :

$$\mu = \lim_{n\to\infty} \bar{l} = \lim_{n\to\infty} \left(\sum_{i=1}^{n} \frac{l_i}{n}\right) \tag{1.3}$$

In figure 1.3 the curve we see the curve of the *parent distribution* of the *experimental distribution* of figure 1.2.

The curve will not necessarily have the same shape as figure 1.3, it may also be asymmetric. A parent distribution can be associated with

[10]In reality we will see some cases in which this is not true. This is due to the fact that the whole reasoning has been built on the assumption that we now have to demonstrate. Thus, for the moment we do not need further indications but we are simply introducing intuitive concepts and defining them in a *more rigorous* manner.

[11]This is very often called *frequency*, but in statistical analysis this name usually indicates a different quantity. Here we therefore avoid this use so as to avoid ambiguities.

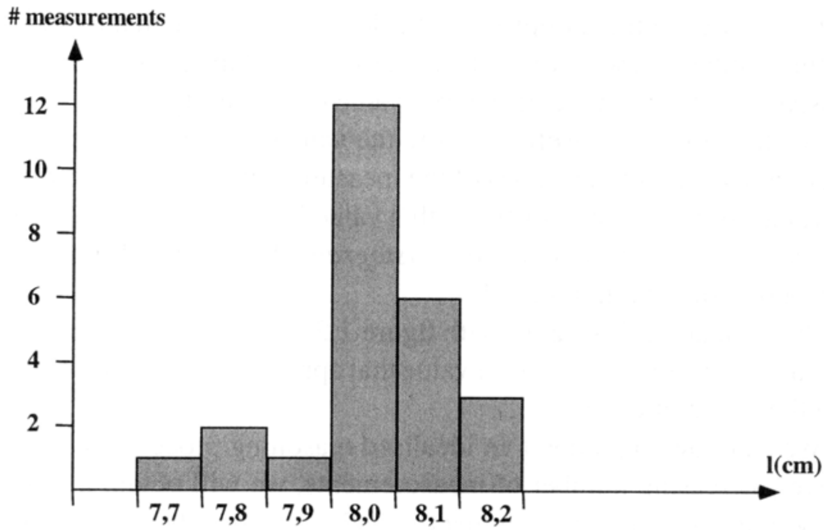

Figure 1.2: Histogram for 25 measurements of length of the same pencil used in the previous figure. This is an *experimental distribution.*

any experimental distribution. *Mean value* is the value of the abscissa obtained from equation 1.3, which ideally corresponds to the arithmetic mean. A warning note is that not only will the curve not always have a single and symmetric bell shape, but in general, the parent distribution will be difficult to identify from the experimental distribution. The limit will not have a possible numerical equivalence in practical terms. Therefore it is difficult to obtain the true value. For this reason we can use methods based on the concept of *maximum likelihood.* This will be the subject of chapter 4. Here we wish only to limit the discussion to arrive at a better understanding of the definition.

For example, in the curve in figure 1.4, the mean value is neither the one to which the maximum value of the ordinates corresponds, nor the one around which (to the right and to the left) the curve is symmetric in the values of the integral. The latter points are called *most probable*

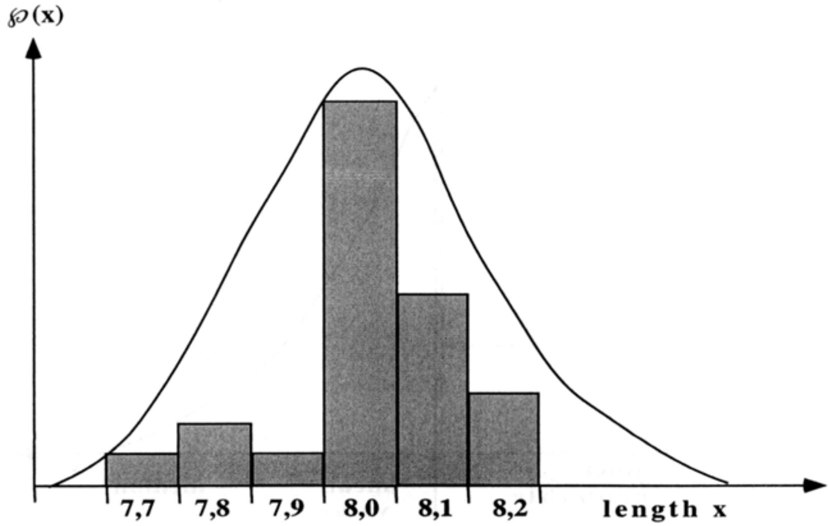

Figure 1.3: Parent disribution curve from the limit procedure, corresponding to the experimental distribution of the previous figure.

value or *mode* and *median* respectively. In the following we will justify this terminology. Let us go back to uncertainty.

We are following a line of reasoning to obtain a quantitative evaluation of the casual uncertainty.

In the curve in figure 1.4 we indicate the *deviations* of each value in abscissa, which means the differences between the value of the abscissa and the mean value:

$$d_i = x_i - \mu \tag{1.4}$$

It is clear that the only deviation equal to zero is that of the mean value. This statement may be obvious but it is best to keep in mind in the following discussion.

For the deviations, we can implement a limit procedure as well. Even though the d_i are as many as the l_i with $i = 1, \dots, \infty$ we can

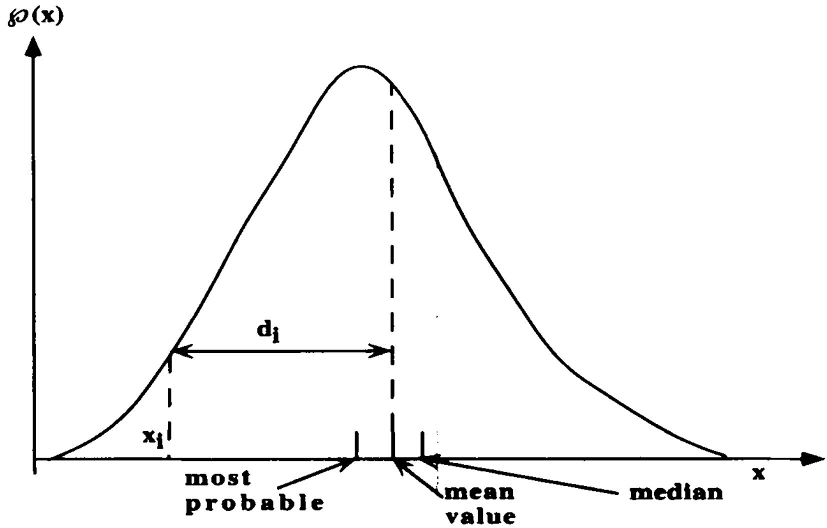

Figure 1.4: Curve of the parent distribution corresponding to the previous experimental distribution. The shape is asymmetric; it is possible to define the characteristic parameters, *mean value*, *median*, *most probable value* , and the *deviations*.

calculate the arithmetical mean of a subsample of n measurements and calculate the deviations d_i with $i = 1, \dots, n$, therefore the arithmetical mean of the deviations themselves, obtaining the mean deviation $\overline{d}$:

$$\overline{d} = \frac{\sum_{i=1}^{n} d_i}{n} \tag{1.5}$$

Going to the limit for the deviations similar to what we have done for the l_i:

$$\lim_{n\to\infty} \bar{d} = \lim_{n\to\infty} (\sum_{i=1}^{n} \frac{d_i}{n}) = \lim_{n\to\infty} (\sum_{i=1}^{n} (\frac{l_i - \mu}{n})) = \tag{1.6}$$
$$= \lim_{n\to\infty} (\frac{1}{n}(\sum_{i=1}^{n} (l_i - \mu))) = \lim_{n\to\infty} (\frac{1}{n}(\sum_{i=1}^{n} (l_i) - \frac{1}{n}(\sum_{i=1}^{n} (\mu))) =$$
$$= \lim_{n\to\infty} (\frac{1}{n}) \sum_{i=1}^{n} (l_i) - \lim_{n\to\infty} (\frac{1}{n})(\sum_{i=1}^{n} \mu) =$$
$$= \lim_{n\to\infty} (\frac{1}{n}) \sum_{i=1}^{n} (l_i) - (\mu) = \mu - \mu = 0$$

With what we said before the deviations defined in this way are equal to zero for the mean value. The next steps in arriving at equation 1.2 are needed to state that the d_i are *the best estimate of the error.* In fact, their mean value is equal to zero at ∞, namely when the arithmetical mean of the measurements reaches the true value. This is exactly what we wished to state. We have, in fact, found a correspondence between the idealisation of being able to find the true value and the fact that in this case we do not make *any error.*

Thus the uncertainty for each measurement is obtained with a limit procedure on the deviations.

To maintain a definition of an always positive quantity, the common practice is to consider the modulus of the deviations, in the following way [11]:

[11]To continue the discussion we introduce here the definitions of modulus and squares of the deviations, which may seem arbitrary. In reality, they are introduced simply because they keep the meaning of deviation from the true value and they do not equal zero at ∞.

$$\eta = \lim_{n\to\infty} (\sum_{i=1}^{n} (\frac{|l_i - \mu|}{n})) > 0 \tag{1.7}$$

It is common use to define the square of same quantity in the summation.

We can now reconsider the limit of equation 1.5 which is called *variance* [12]:

$$\xi = \lim_{n\to\infty} (\sum_{i=1}^{n} (\frac{|l_i - \mu|^2}{n})) \tag{1.8}$$

To have dimensions comparable with the measured quantity, it is advisable to consider the square root of this quantity:

$$\sigma = \sqrt{\xi} = \sqrt{\lim_{n\to\infty} (\sum_{i=1}^{n} (\frac{|l_i - \mu|^2}{n}))} \tag{1.9}$$

which is called *standard deviation of the parent distribution.*

We judge the standard deviation as the best estimate to associate with uncertainty due to the casual error. We usually tend to *identify standard deviation with casual uncertainty.*

It is still necessary to find a quantity that we can use in a correct quantification of the uncertainty to associate with the measurement.

In practice, it is not easy to calculate the true value. Therefore in the definition of equation 1.9 we replace the mean value with the arithmetical mean with approximation:

$$\xi \cong \omega = \sqrt{\sum_{i=1}^{n} (\frac{|l_i - \bar{l}|^2}{n})} \tag{1.10}$$

[12]This terminology is of more general use in statistics. Variance must be specified with respect to the intended quantity. In our case it will always mean variance with respect to the mean value

Going back to the previous observation, that the only deviation equal to zero is that of the mean value, it is obvious that we cannot calculate the deviation of the mean if we include the mean value itself in the division. In fact this contributes to the deviation as zero and would therefore lead to underestimation of the result of the mean deviation. To avoid this contradiction [13] in the redefinition of the standard deviation, we will introduce the $n - 1$ factor in the calculation instead of n. It is to be noted that, in any case, in the approximation for $n \to \infty$, the difference is negligible. Therefore the equation 1.10 becomes:

$$\xi \cong s = \sqrt{\sum_{i=1}^{n} \left(\frac{|l_i - \bar{l}|^2}{n - 1}\right)} \tag{1.11}$$

Thus this is the quantity considered as being the *standard deviation of the experimental distribution* or simply *standard deviation*. We identify the estimate of the casual uncertainty with the experimental standard deviation which will be used from now on.

Going back to the example of table 1.2, and considering that all measurements were performed with the same instrument having resolution 0.1cm.

In the case of the table the standard deviation, is :

$$s = \sqrt{\left(\sum_{i=1}^{n} \left(\frac{|d_i|^2}{n - 1}\right)\right)} = \sqrt{\frac{0.36}{24}} \cong \sqrt{0.02} \cong 0.12 \cong 0.1cm \tag{1.12}$$

By repeating the same exercise for table 1.1 we obtain $s \cong 0.2cm$. This result justifies what was said above.

---

[13] Although this reasoning could have been introduced in the definition of the parent standard deviation, the author feels that it is easier to understand at this point

d_1	$0.0cm$	d_{11}	$0.0cm$	d_{21}	$0.0cm$
d_2	$0.1cm$	d_{12}	$0.0cm$	d_{22}	$0.0cm$
d_3	$0.1cm$	d_{13}	$0.1cm$	d_{23}	$0.2cm$
d_4	$0.2cm$	d_{14}	$0.0cm$	d_{24}	$0.0cm$
d_5	$0.2cm$	d_{15}	$0.2cm$	d_{25}	$0.0cm$
d_6	$0.2cm$	d_{16}	$0.3cm$		
d_7	$0.0cm$	d_{17}	$0.0cm$		
d_8	$0.2cm$	d_{18}	$0.0cm$		
d_9	$0.1cm$	d_{19}	$0.1cm$		
d_{10}	$0.1cm$	d_{20}	$0.0cm$		

Table 1.3: Values of the deviations for the 25 measurements of length of a pencil.

1.2.1 The expression of expanded uncertainty

Here we wish to give a *global* or *extended*, expression of uncertainty that includes the various types of error introduced and that allows their numerical evaluation in a brief and common notation, with which we can accompany the experimental result.

To obtain this we must discuss the combination of uncertainties and thus determination of combined uncertainty. It is also essential to discuss *confidence levels* based on concepts of probability and on distribution functions.

Before arriving at this, it is necessary to mention at least the notation commonly used to indicate the global uncertainty to associate with a measurement. This will be justified later on in several places in the text.

To indicate uncertainty and the measurement, we write:

$$x \pm \Delta x \tag{1.13}$$

meaning that it is highly plausible that quantity x, measured in a certain experiment or in several experiments performed one after another or independently in combined experiments, belongs to the interval of values

$-\Delta x, +\Delta x$.

To exemplify: if the global uncertainty were only one, obtained simply by estimation of casual uncertainty, $s = 0.2cm$, then, going back to the measurements in the previous chapter, we would have:

$$l \pm \Delta l = (8.1 \pm 0.2)cm \tag{1.14}$$

Or we could simply express only the reading error (equal to instrument resolution), for each measurement made with the meter. For example:

$$l_i \pm \Delta l_i = (7.7 \pm 0.1)cm \tag{1.15}$$

We will return to this concept of the expression of expanded uncertainty. We underline here that in the previous equation 1.14, the associated uncertainty should be equal to the expanded uncertainty determined by the casual uncertainty and evaluated as described in discussions to follow.

Example: the problem of the reading of zero with the polarimeter

We present here the case of the reading of zero or *offset*, when using the polarimeter, a common laboratory instrument.

The experiment is described in detail in appendix A.2. Here it is given simply as an example in which the declared accuracy of the instrument, and thus the reading error, as read on the scales, is much smaller than the systematic error introduced when taking measurements, not only due to the inadequacy of the instrument in its manufacture, but to the very principle of measurement. This is much less accurate than the scale appears to indicate.

Let us consider a normal polarimeter used in chemistry, biology and physics. In figure 1.5 we see the picture obtained in the so-called *equal-shadow* position in the eyepiece of the instrument. In this position the

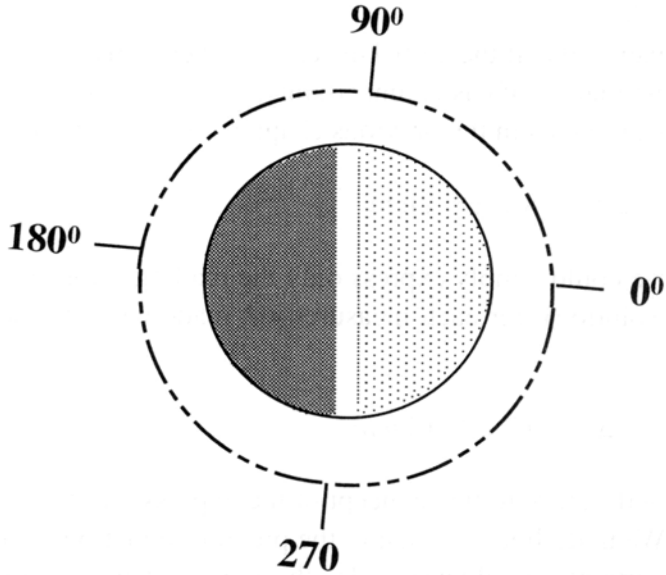

Figure 1.5: View through the eyepiece of *equishadow* in a position close to 0 on the graduated scale.

two half-moons should have the same intensity. At least the manufacturer refers to this position as the 0 position of the instrument if, for example, there is some de-ionised water inside the sample holder (see description of the instruments and experiments in Appendix A).

We could argue, however, that an *illegitimate* error caused by the difficulty of accurately determining light intensity in the two half-moons is quite plausible. For the rest of the discussion we will suppose that we are able to neglect this error, which in any case will make a contribution.

Let us consider the reading scale. The vernier should give a nominal resolution for the global reading of the scale, equal to $0°, 1', 30''$. Therefore, by repeating the measurement several times, the value of $0°, 0', 0''$ should be at the centre of the distribution, and the greater the number of

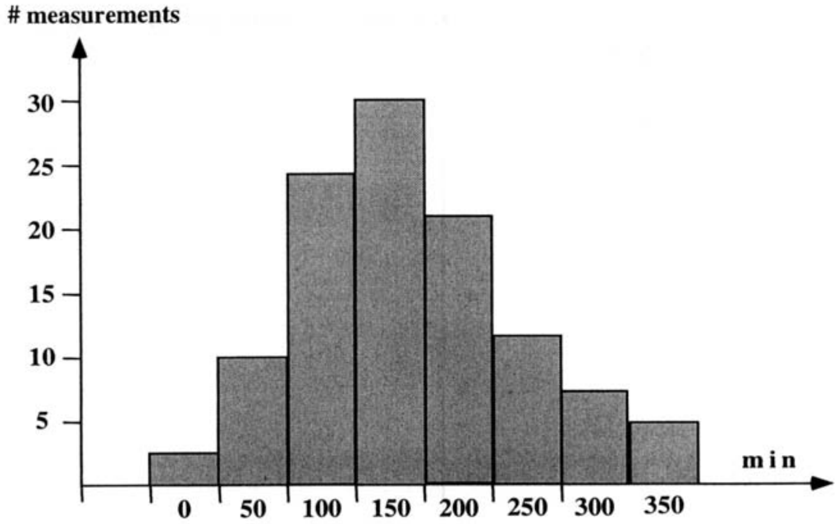

Figure 1.6: Distribution of values of the angles around the equishadow position. The scale is 1/60 of a degree.

measurements the closer it will be to the centre. The data are given in figure 1.6. It is clear that the distribution is not symmetric and its centre is not a $0°, 0', 0''$. Since the number of measurements is large (more than a hundred), we can state that there is a systematic error which does not make deviations symmetric around the mean. This indicates that the scale of the instrument is not centred around $0°, 0', 0''$, and the position of equi-intensity does not correspond to it (figure 1.5).

Exercise: *tossing of dice*

Take two dice, with faces numbered from l to 6; toss them 20 times.

Write down the number of times you obtain a number corresponding to a number in the table then draw a histogram.

Determine the mean, the median, the most probable value, the standard deviation.

Repeat for 100 tosses.