Chapter 1

INTRODUCTION TO ELECTRONIC INSTRUMENTATION

- ***How are Electronic Instrumentation and related fields defined?***
- ***How has Instrumentation evolved into Electronic Instrumentation?***
- ***What is the internal structure of an instrument like?***
- ***How should the quality of an instrument be specified?***
- ***What is the smallest signal that can be measured?***
- ***What sources of error are acting on an instrument?***

1.1 Instruments for solving a measurement problem

Electronic Instrumentation is about the design, realisation and use of electronic systems for the **measurement** of electrical and non-electrical quantities. Strongly related fields are measurement science and data acquisition. Each of these disciplines has a specific function in solving a measurement problem. The activity that is the basis of electronic instrumentation is **measuring**.

Measuring (performing a measurement) is formally defined as the set of operations having the objective of determining a value of a quantity.

Measurement science provides the theoretical framework, which is used in an initial feasibility study prior to the actual measurement to assess whether the parameter of interest (the **measurand**) is **measurable**. Measurability can be impaired due to a fundamental limitation or a practical constraint. For instance, the mass and dimensions of a sculpture can be measured, but its artistic quality for fundamental reasons cannot. For practical reasons, a certain voltage level could prove non-measurable in cases where a high noise level is injected into the input circuit by Electro-Magnetic Interference (EMI) from, for instance, a nearby vacuum cleaner.

Measurability is analysed in terms of the requirements imposed by the measurement problem (such as accuracy), the practical constraints (for instance measuring at a high temperature or in water or in the case of a high noise level), and the theoretical considerations (can the parameter be measured when considering fundamental constraints, such as quantisation of electric charge?). The results of this study include, amongst others, a listing of the requirements (= **problem specifications**) that any instrument considered for use should satisfy.

Electronic instrumentation provides the knowledge that is needed to select the commercially available instruments if available with sufficient performance (= **instrument specifications**) to do the job. Such an instrument may be available for a routine measurement, but perhaps not in the case of a highly demanding measurement problem. In the latter case, Electronic instrumentation also provides the required knowledge at the sub-system and component levels to design and realise customised instruments to meet the requirements of demanding applications (= **measurement problems**).

Typical questions addressed in instrumentation are: what is the system structure like?; what bandwidth is required?; where in the system structure should an amplifier be implemented?; where are sampling and quantization implemented? and what is the system transfer function or **sensitivity**? Last but not least; what is the minimum level of the parameter of interest that can actually be measured when using a specific instrument (or, more generally, when using a specific instrumentation technique) when considering the practical constraints? (= what is the **detection limit**?). These are questions about the basic functioning of the system. However, it should be noted that the answers depend on issues such as the flexibility, costs and performance of the instrument.

Data acquisition refers to a field that covers a cross-section of the disciplines of measurement science and instrumentation with a focus on solving practical industrial and consumer problems in measurement and control, rather than aiming for the lowest possible detection limit using sophisticated instruments.

1.2 Impact of electronic instrumentation

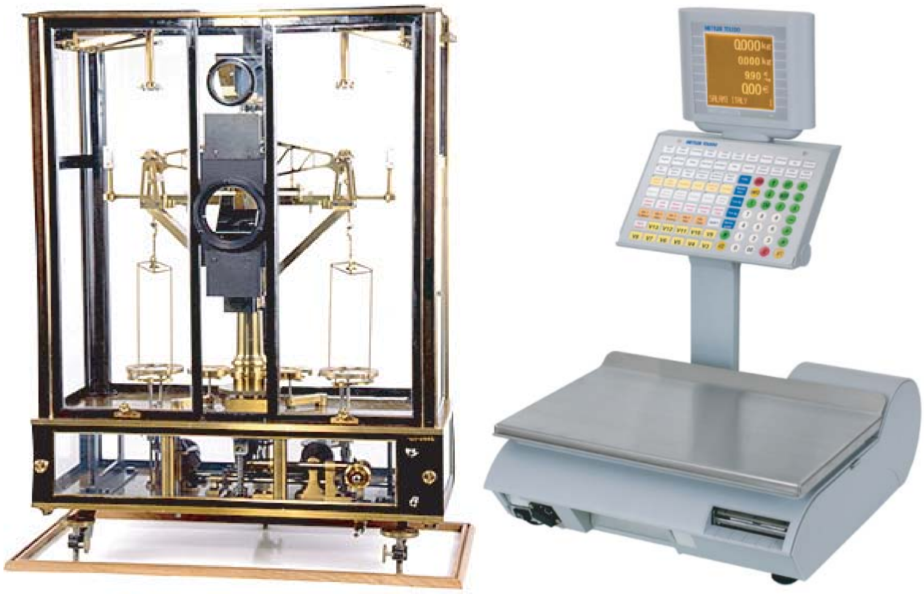
Non-electronic instrumentation would be unthinkable with the present state-of-the art. However, it should be noted that modern electronic instruments are the result of microelectronic technology, which has enabled the fabrication of the high-performance electronic components required in an instrument. The vast majority of early instruments were based on mechanical principles and was intended for the measurement of mechanical quantities (length and mass) and also time. The measurement of electrical parameters became relevant only after the discovery of electricity and electro-magnetism. Electronic components emerged in the 20th century, but their potential in instrumentation was quickly recognised. This explains the commonly used -but in historic context not obvious- combined term ‘electronic instrumentation’ for describing the field.

1.2.1 Electronic instrumentation for a structured system

The most decisive advantage of electronic instruments is the opportunity that electronic components provide for structuring the various functions of the system in modules, each of which is designed to form a logical entity.

Figure 1.1, Measurement in ancient times (1420 B.C.). The papyrus of Ani shows a scene of the “Book of the Dead”: Anubis, god of the dead (with the jackal head) balances the weight of the heart of the deceased person standing on the left (Ani of Thebes) with “the feather of Maat”, which represents order, truth and justice. The god Thoth (with the bird’s head) writes the verdict.

Mechanical systems do not offer this flexibility. Generally the different functions are combined. A good example is the weighing balance, which is composed of a bar of well-defined length and attached in the middle to a supporting construction in such a way that it can rotate around the centre. The left-hand and right-hand parts of the bar are, in the basic concept, of equal length and are referred to as arms. The end of an arm can be loaded with an object. Loading one arm of the balance with an object and the other arm of equal length with another object causes the side of the balance loaded with the object of higher mass to drop. A series of measurements enables the sorting of objects on a scale from smallest mass to highest mass. This balance for mass measurements was already well known in ancient times, as is e.g. demonstrated on the so-called papyrus of Ani (Fig. 1.1).



(a) (b)

Figure 1.2, (a) Mechanical balance and (b) electronic weighing scale.

The weighing function (the imbalance) and the indicator (the angle of rotation of the arm with respect to the vertical suspension) are structurally combined. The instrument evolved in the 18th century to become the kind of fine mechanical structure shown in Fig. 1.2a. Instead of a crude comparison of the mass of two objects, the object to be weighted is attached to the tip of the measurement arm, while the opposite arm -the compensating arm- contains a reference mass that can be slid along the arm until balance is obtained. The value of the reference mass in combination with the arm lengths provides a direct measurement (i.e. measurement in an interval scale) of the mass of the object.

An electronic weighing scale has a much more modular structure. The system is composed of a separate component for mass measurement (the sensor), electronic circuits for read-out, a microprocessor for data processing and a digital display, as shown in Fig. 1.2b. The sensor is basically a structure which contains strain-sensitive resistors. These are connected in a Wheatstone bridge. As is shown in Chapter 2, this is basically the electrical equivalent of the mechanical balance. The structured design of the electronic weighing scale offers advantages such as: user-friendliness (reduced requirements on expertise and accuracy of the operator), flexibility (the price per unit mass can be used as additional input to give an indication in weight and price), objectivity of the measurement (no read-out error), and repair (simple replacement of modules).

1.2.2 Electronic instrumentation for miniaturisation

Traditional measurement systems are bulky, which complicates their use when measuring small objects or measuring objects with limited access. Integration technologies, such as Integrated Circuit (IC)-compatible Micro-Electro-Mechanical System (MEMS) technologies enable the fabrication of very small measurement systems on a single silicon chip, which can e.g. be implanted in the human eye to measure eye pressure, as shown in Fig. 1.3.

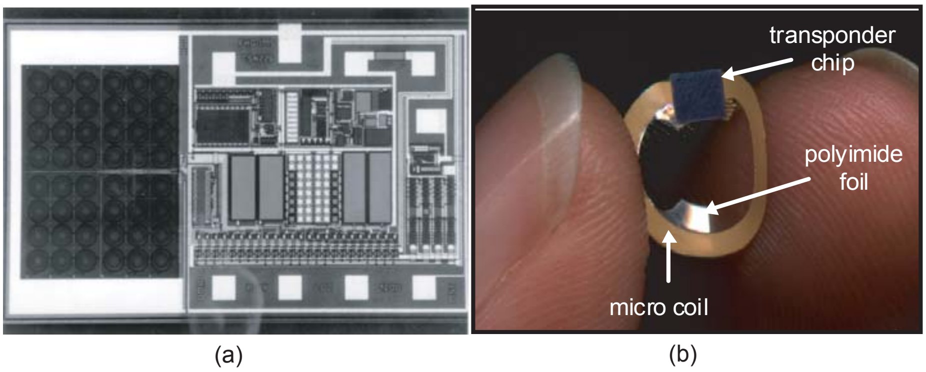
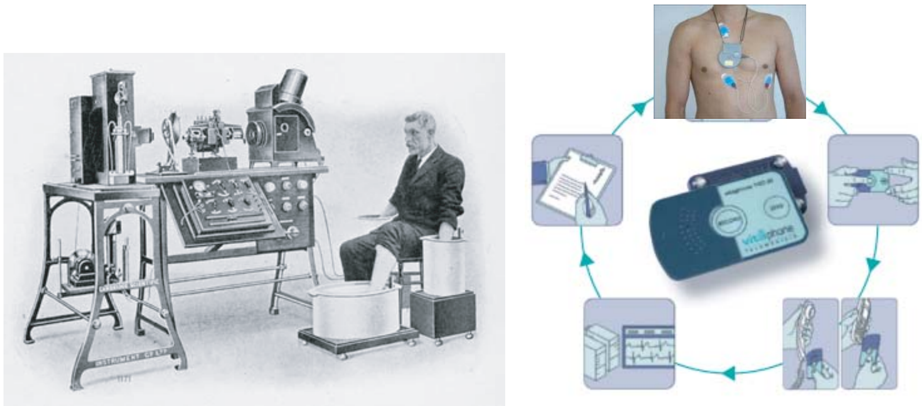


Figure 1.3, Eye pressure sensor (a) chip and (b) system containing sensor and circuits for communication ready for implantation in the human eye.

Miniaturisation is not only pursued for the sake of small system dimensions, but also for achieving improved overall system performance. The instrument is often part of a much larger system. Embedding an instrument in such a system has been made possible by technological breakthroughs in microelectronics and communication technology. This aspect is very relevant in the medical application field, where the measurement is merely a part of the entire process leading to a diagnosis.

An example is the ElectroCardioGram (ECG), which is used for measuring heart activity using the electrical potential differences across the human body. A practical system for actually measuring these potentials was developed by Willem Einthoven in 1902 and was a major step forward towards the early detection of heart failure. However, as Fig. 1.4a indicates, the system was rather cumbersome to use. The patient had to keep his or her hands and one foot in a salt solution to ensure sufficient electrical contact.

One century of technological progress has resulted in the hand-held system shown in Fig. 1.4b. It can be operated by the patient at home and enables communication with a central database at a hospital. Whenever a patient feels insecure about his or her condition, the patient can attach the electrodes, an ECG



(a) (b)

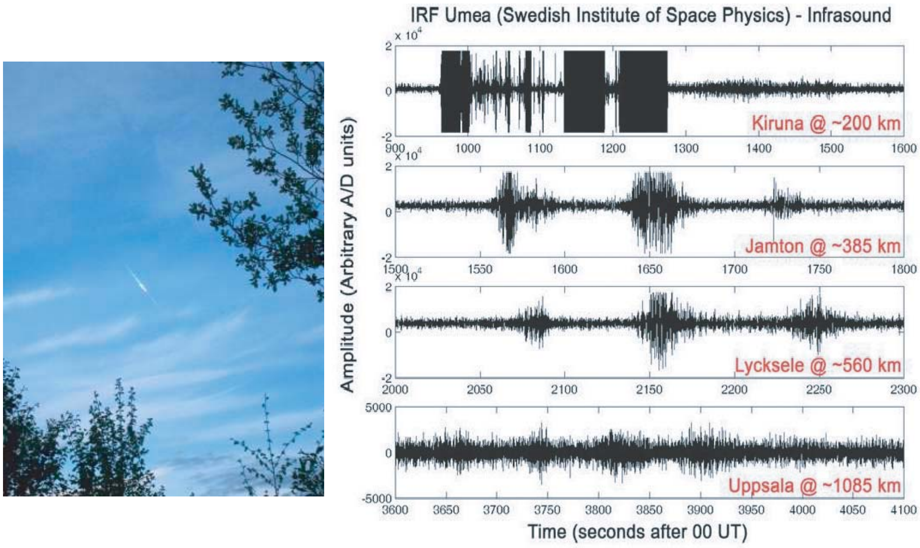
Figure 1.4, ECG (a) in 1902 and (b) incorporated in a complete health-care system (vitaphone-me.com).

scan is made and the result can be sent to a central database, which is coupled to an expert system. The ECG is evaluated at the medical centre connected to the system and direct action can be taken in case of an emergency. This approach strongly reduces the frequency of unnecessary visits to the hospital, reduces health care costs and adds to the quality of life of the patient. The increased benefits of the ECG are also brought about by the fact that electronic instrumentation is based on microelectronics, which allows a seamless integration of the instrument in computer and telecommunication systems.

1.2.3 Electronic instrumentation for data processing

Electronic instrumentation has also benefited from advances made in signal processing and data processing. In many measurement problems, the required information cannot easily be extracted from measurable signals. An example is shown in Fig. 1.5.

On June 7th 2006 a meteoroid impacted the earth's atmosphere over northern Norway. The effect was like a fireball followed by explosions, which were recorded at several locations, as shown in Fig. 1.5b. The position of the impact could be determined from the time delay experienced by the low-frequency components of the sound while these were traveling from the location of impact towards each of these different stations. Obviously, special signal processing techniques (i.e. correlation) have to be applied to obtain reliable information from a station more than 1000 km away from the location of impact.



(a) (b)

Figure 1.5, (a) Meteoroid entering the earth's atmosphere as a fireball over northern Norway and (b) recording of the sound produced at different stations (aquarid.physics.uwo.ca).

These are examples of electronic instrumentation as an enabling technology in many different applications by virtue of microelectronics technology. The obvious consequence of **electronic** instrumentation is the need for a component for taking the information contained in a non-electrical quantity or signal and transferring this information into an electrical quantity or signal and *vice versa*. This component is usually referred to as the **transducer** and the associated signal processing as **transduction**. The non-idealities in the transduction and the first stage of the read-out circuits (the analog front-end) largely determine the quality of the system and hence set the detection limit. This is a key issue in this book, which will be more extensively introduced in Section 1.5. The transduction determines the structure of the instrumentation system, as is discussed in the next section.

1.3 Structure of the instrument

The measurement instrument is basically an information conveyor between a source and a target. The source is an **object** (i.e. a phenomenon, process or system of physical, chemical or physiological nature) and the value of one or more quantities of the object is to be determined. The **target** defines the motive for performing the measurement. The target could be either a human observer reading the display in a measurement system (e.g. a thermometer) or an actuator that is part of a measurement and control system (e.g. a valve in a heating system). Three main parts of the measurement and control system can be distinguished:

- Data acquisition,
- Data processing and
- Data distribution, which are related as indicated in Fig. 1.6.

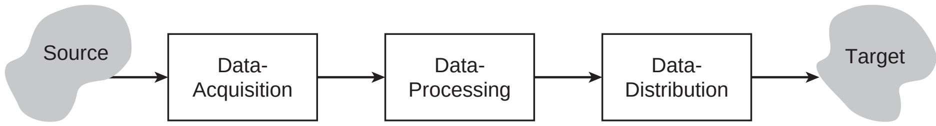


Figure 1.6, General structure of a measurement and control system.

In the **data acquisition** sub-system, information is obtained (= **acquired**) on the measurand and transferred into a measurement result (= **data**). The signals at the interface between data acquisition and electronic data processing, as well as those between the data processing and data distribution sub-systems, are by necessity electrical and digital. This defines the data type and format directly at the output of the data acquisition unit, hence the inclusion of a major part of the functionality therein. Data processing involves manipulating the input data and forwarding the result to the data distribution sub-system.

The information of interest at the source is contained in a signal domain. The signal domains are generally energetically defined; the following six signal domains are distinguished:

- Optical (radiant),
- Thermal,
- Mechanical,
- Magnetic,
- Chemical and
- Electrical.

Transferring information from one signal domain into another requires a transducer. In electronic instrumentation one of these domains is electrical. The data acquisition sub-system has to include an **input transducer** or **sensor** that enables the change from non-electrical carrier into the electrical domain. Similarly, the data distribution sub-system includes an **output transducer** that enables the change from the electrical domain into the actual energy domain used for display or actuation, as shown in Fig. 1.7.

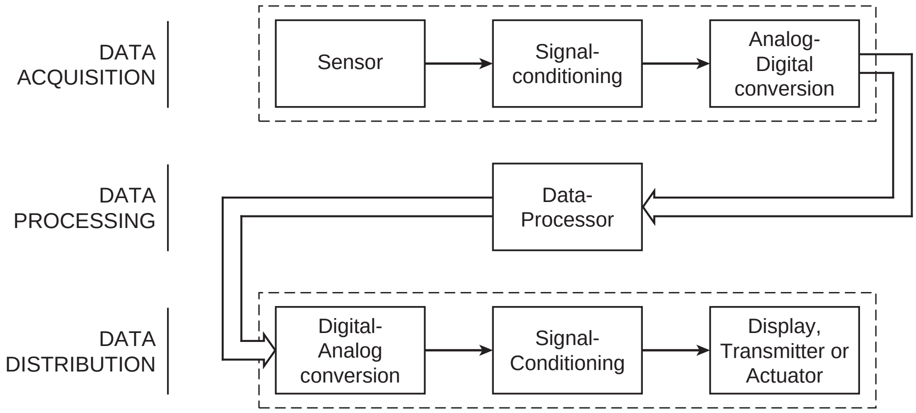


Figure 1.7, Structure of an electronic measurement and control system.

Proper coupling of the signal from the sensor output into the input of the **Analog-to-Digital Converter (ADC)** requires **signal conditioning**. Similarly, the data distribution unit usually includes signal conditioning between the **Digital-to-Analog Converter (DAC)** and an **actuator**. In the case of an instrument with only a digital display, a direct drive from the data-processor is possible.

The signal conditioning required in the data acquisition sub-system strongly depends on the specifications of the measurement problem and the details of the sensor used, and may contain the following functions:

- **Amplification**, which matches the range of the sensor output signal to that of the input of the AD converter,
- **Impedance transformation**, which enables read-out of a voltage source with a relatively high source impedance without a scaling error (see Section 2.4.1),
- **Frequency or time filtering**, which optimises the **Signal-to-Noise Ratio (SNR= power of the intended signal / power of the noise signal)**, and
- **Signal conversion within the electrical domain**. Some sensor types directly supply a voltage or current level proportional to the measurand. However, other sensors are designed to change the value of an electrical component, such as a resistor, capacitor or inductor in a manner that is proportional to the non-electrical input. Read-out of these types of sensors requires circuits for the conversion of the modulated value of the electrical component into an electrical signal (e.g. a fixed current source to generate an output voltage proportional to the change in resistance).

Example 1.1

The temperature-dependent resistor in the thermostat shown in Fig. 1.8 is specified by a resistance at 0 °C equal to R(0 °C)= 100 Ω and a sensitivity α= 0.4 Ω/K ($R(T)= R(0°C)+\alpha(T- 0°C)$). A fixed-value current source is used to convert resistance value into output voltage. A second resistor, R_o, is used to set the zero value. A differential amplifier is used to match the full-scale temperature to the maximum input voltage of the ADC. The electrical current supplied by the current source is limited to avoid self-heating of the sensor by dissipation. From the figure, the following results can be derived: $U_{i,ADC}= I\times[R(T)-R_o]G$, while the dissipation can be expressed as: $P_{diss}= I^2\times R(T)$. Heat dissipation would introduce a rise in temperature equal to $\Delta T= g_{th}\times P_{diss}$, with g_{th} [K/W] as the thermal conductivity of the resistor. This effect adds to the measurement uncertainty, which should be lower than the specification.

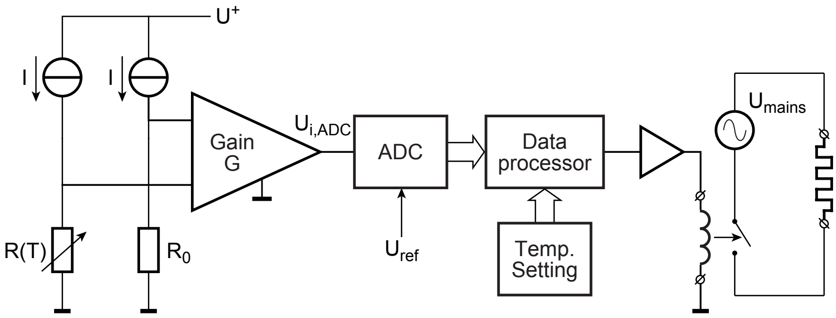


Figure 1.8, Structure of a simple thermostat.

The control strategy and several parameters, such as power of the heating element and the heat capacity of the room to be heated, are entered into the data processor. User input, such as the desired 24-hour temperature profile, can be entered. The data distribution sub-system is simple and is composed of a relay to switch a heating element. Signal conditioning involves a power amplifier to drive the relay. DA conversion is superfluous, since only the on/off function is required. Basically the relay serves as a 1-bit DAC.

The task is to dimension the thermostat for operation between 10 °C and 25 °C, with 10 °C as the zero level and 25 °C as the full scale. The following additional information is available:

- The ADC operates on signals between U_i= 0 V and $U_i= U_{ref}$= 1.024 V.
- The resistors are biased using I= 10 mA.
- R_o and I are temperature-independent.

Solution:
R_o should be dimensioned to result in a zero differential amplifier input voltage at 10 ^oC:
$R_o = R(10\ ^oC) = R(0\ ^oC) + 0.4[10\ ^oC - 0\ ^oC] = 100 + 0.4\times10 = 104\ \Omega$.

Subsequently, the gain should be set to give full scale at 25 ^oC. Full scale implies code $11...11_2$. Assume an 8-bit ADC with $U_{ref} = 1.024$.

The ADC has a quantisation interval $q = 1.024/2^8 = 4$ mV.
Thus full scale is associated with $U_{i,ADC}$ between 1.020 V and 1.024 V.
The gain setting required results from:
$U_{i,ADC} = 1.024 = I[R(25\ ^oC) - R_o]G = 10^{-2}[15\times0.4]G$. Hence $G = 17$.

1.4 Error and uncertainty in an instrument

An instrument basically performs a mapping of the input quantity to be measured over its full measurement range onto the available range of an output quantity (e.g. the scale of a display), as shown schematically in Fig. 1.9.

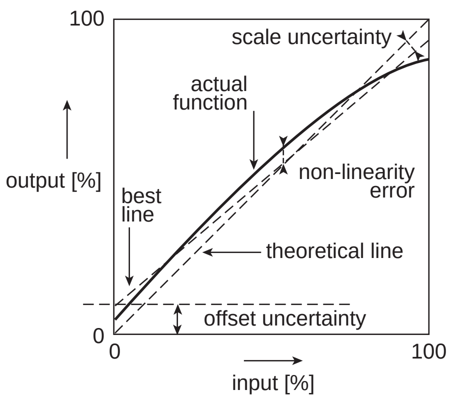


Figure 1.9, Errors and uncertainties in a measurement.

A practical instrument introduces a certain amount of uncertainty (error) that causes ambiguity about the actual value of the measurand being measured and ultimately results in a detection limit.

Errors can be classified, according to their impact on the measurement result, as:

- Scale errors
- Additive errors
- Non-linearity errors
- Cross-sensitivity for well-defined, although undesirable, quantities

While all these error types need to be considered to determine the measurement uncertainty, additive errors and cross-sensitivities are the most important in the calculation of the detection limit. For this reason the emphasis is on additive errors and cross-sensitivities, with multiplicative and non-linearity errors only discussed when relevant to the detection limit in a specific application.
The sources of error enable a classification into:

- Systematic errors (= deterministic errors),
- Random errors (= stochastic errors) and
- Human errors,

which can further be divided into six categories:

1. Systematic errors that occur when the measurement activity affects the measurand (source loading by the instrument). This type of error usually results from an incorrect measurement technique or improper use of an instrument. An example is voltage measurement across a relatively high impedance using an instrument with a relatively low input impedance. Another example is temperature measurement using a thermometer with a large heat capacitance.

2. Systematic errors due to the use of an improperly calibrated instrument. An example is an amplifier that is set at the incorrect gain. If one is unaware of this, then the deviation from the presumed gain is specified as the scale error.

3. Systematic errors due to a parasitic sensitivity to an undesired parameter (e.g. common-mode voltage in a differential voltage measurement, or a parasitic temperature sensitivity in a pressure sensor when operated at a temperature not equal to the calibration temperature).

4. Random errors due to instrument noise, which could be either generated inside the instrument or of external origin and injected into the measurement setup (e.g. capacitive coupling).

5. Random errors due to the remaining uncertainty in the internal reference after proper calibration. This is included in the instrument specification and is not discussed here.

6. Human errors due to incorrect reading of a display. This type of error can be largely avoided by an ergonomically optimised design of the instrument or by using a fully automated measurement setup and is not considered here.

Basically, the first two classes of systematic error mentioned above are also due to human error and are also referred to as '**improper use of instruments**'. Nevertheless, these need to be considered. Loading of the source by the measurement is undesirable, yet often unavoidable, since physical contact is often required in order to get access to the information. In this sense it is a prerequisite for the measurement and has to be included in the detection limit analysis. Especially in temperature measurement, source loading can be an important factor that limits the detectivity.

The errors of the 2nd, 3rd and 4th categories are additive errors that ultimately determine the detection limit. Depending on the application, one of these four sources of uncertainty is usually dominant. The signal level associated with the measurand should exceed the magnitude of these parasitics to enable a reproducible detection.

A systematic error (= deterministic error) can be avoided during the measurement or corrected for at any time provided that the source of deterministic error is completely described.

The key issue in this definition is the **complete description** of the source of error. It is quite possible in a complex system that an experimenter is unaware of a particular source of error at the time of a measurement. However, it can be corrected for, even after completion of a measurement program, provided that:

1. this unexpected and undesirable effect is predictable in time and with temperature (remains constant= repeatable, or with a known trend) and
2. the measurement procedure, the equipment used and the results obtained are described in sufficient detail.

Having the option to correct data after gaining new insight is especially important in the case of an expensive or dangerous experiment. This is the reason why professionals record their findings in detail in a logbook. Moreover, in industry each page is dated in case of a legal dispute about an infringement of a patent that is based on a particular finding described in the logbook.

Example 1.2
A Voltmeter with an input resistance R_i= 100 kΩ is used to measure a voltage source, U_{sr}, with unknown source resistance, R_{sr}. Measurement at 20 °C gives: U_{mo}= 1.961 V. The operator becomes aware of the poor quality of the instrument at a later stage. This is a systematic error of category 1. Fortunately, the voltage source and the instrument are still available. The source resistance, although of unknown value, is known to be very stable. However, it is uncertain whether the voltage has drifted in time. Two measurements are performed for an *ex post* correction of this mistake:
(1) performing the same measurement again under identical conditions and
(2) measuring the voltage with a shunt resistor R_p= 10 kΩ connected in parallel to R_i. The measurement results are: U_{m1}= 1.902 V and U_{m2}= 1.590 V respectively.

The tasks are to:
(1) determine the voltage U_{sr} and the source resistance R_{sr} at the time of the new measurements (at $t = t_1$) and
(2) use this information to correct the original measurement for loading of the source by the instrument, based on these additional measurements (at $t = t_o$).

Solution:
Note that the effect of any temperature dependence can be disregarded, because the repeated measurement is at the same temperature.
(1) $U_{m1} = R_i \times U_{sr}(t_1)/(R_i+R_{sr}) = 100\ \text{k}\Omega/(100\ \text{k}\Omega+R_{sr})U_{br}(t_1) = 1.902$ V.
$U_{m2} = R_i//R_p \times U_{sr}(t_1)/(R_i//R_p+R_{sr}) = 10^5/11/(10^5/11+R_{sr})U_{sr}(t_1) = 1.590$ V.

Solving this set of two independent equations with two unknowns yields: $U_{sr} = 1.938$ V and $R_{sr} = 2.002$ kΩ.

(2) $U_{sr}(t_o) = U_{mo}(R_i+R_{sr})/R_i = 1.961 \times 102/100 = 2.000$ V.
From these results it can be concluded that the voltage source has drifted -62 mV in the time interval between the two measurements.

An essential condition for *ex post* correction of a systematic error (as in Example 1.2) is the reproducibility of the error: i.e. the error affects the result in a repeated measurement in the same way. Obviously, a great deal of extra work can be avoided in the case of an *ex ante* correction. However, this requires one to be aware of the limitations of the instruments prior to the measurement.

The amplitude of a random error (= stochastic error) is described by a statistical distribution and the value cannot be predicted at every moment in time. Therefore, this error cannot be reduced to zero by compensation or correction.

The most prominent type of random error is noise, which is introduced in the next sections. Although noise cannot be reduced to zero by compensation or correction, instrumentation techniques are available to minimise this type of error.

A deterministic error can in principle be reduced to zero using a perfect compensation or correction. As a consequence, deterministic errors should not be a source of uncertainty and, thus, should not be a significant factor for limiting the detectivity. Unfortunately, in a practical instrument there is a residual deterministic error signal left after correction or compensation. Since this is not expected in the sense that measures have been implemented with the purpose of eliminating the error, this residual should be considered a random signal.

This overlap between deterministic error and random error is also the reason for the somewhat subtle difference between error and uncertainty. An **error** can in principle be avoided (by eliminating human error and deterministic error), whereas **uncertainty** is due to the combined effect of all the unknowns and random properties of the setup or instrument.

As is discussed in the next section, error or uncertainty does limit the performance of an instrument, which is usually specified as the detection limit.

1.5 Performance of an instrument

The most fundamental question in Instrumentation and Measurement is:
'**What is the smallest value of the quantity of interest that can still be measured reproducibly?**'

This question needs further refinement, as the fundamental (physical) limitations need to be distinguished from the practical (instrumental) constraints. Fundamental limitations are due to inescapable laws of nature, whereas practical limitations are imposed by shortcomings of the instrument. The fundamental type of limitation does not change in time, whereas those of the instrument can be relieved by an improved instrument design or by breakthroughs in technology. Progress will bring the instrumental detection limit closer to the fundamental (physical) limitation.

Example 1.3
Assume one would like to visualise a 10 kHz sine-wave AC-current with: i_{rms}= 1 fA. This is an extremely small current. In fact a fundamental physical limit is exceeded. Electrical current is defined as the amount of charge passing through a cross-section of a conductor per unit of time:
$i = \partial Q / \partial t$.

The amount of charge stored in a half-period is equal to:
$Q = \hat{\imath} \int_0^{\pi} \sin(\omega t)\, dt = 2\hat{\imath}/\omega = 2\times10^{-15}/2\pi\times10^{4} = 3\times10^{-20}$ C,
which is smaller than the unit charge. This AC-current, therefore, cannot be displayed (basically this is a shot noise related detection limit).

After a closer look at the requirements the user realized that a mistake had been made. The measurement of a 10 kHz square-wave AC current with $\hat{\imath}$= i_{rms}= 1 nA had been intended and not 1 fA. Therefore, the above-mentioned fundamental limitation does not apply and the specifications of the instrument can be analysed. The oscilloscope available has an extremely good noise specification with an equivalent input current noise source over a bandwidth of 100 kHz (to maintain shape information) equal to 300 pA.

This specification implies that the signal is nevertheless barely measurable using state-of-the-art technology.

The answer to the big question raised in this section is that a practical instrument is designed to measure the actual value of a well-defined property with given instrument shortcomings (specifications). The lower limit that can be measured at a specified inaccuracy level is generally referred to as the **detection limit**. This is a key specification of the instrument and is the red line through this book.

The detection limit of an instrument is the minimum level of the desired input signal that can be measured, considering parasitic signals and accuracy specification.

The performance of an instrument used for measuring a non-electrical quantity is usually limited by the data acquisition unit. Any information lost or corrupted can in principle not be retrieved by data processing, unless additional information is included, such as redundant coding, or when prior information is available about the signal properties.

Since the practical detection limit is due to instrument constraints, technological advances in component manufacturing and system design may lead to a shift in the detection limit. This progress could result in two types of progress:
(a) an instrument could become available that is able to measure a phenomenon which has, so far, been undetectable or
(b) the costs, weight or dimensions of an instrument are reduced to such values that a known measurement technique becomes suitable for a new application.

The first effect is the most spectacular, and the consequences of a new measurement result for science have sometimes been revolutionary. One single measurement that does not comply with prevailing theory could force a re-thinking of the concepts. One example is the photo-electric effect in a vacuum tube, which was expressed in terms of quantisation of energy and, consequently, provided another source of experimental validation of the concept of quantum physics introduced earlier by Planck. This concept had completely changed the understanding of physics by the early 20th century.

The second effect is decisive for the introduction of a particular measurement system in a consumer product. Market acceptance requires a reasonable ratio between the costs of the added electronic feature and those of the consumer product (e.g. the costs of a system for use in an automotive application, such as an airbag, should at introduction not exceed €100). Obviously, the size and weight of additional systems are also restricted. This economic constraint could be bypassed by government regulations on safety or exhaust emissions. Nevertheless, the availability of cheaper, lighter and smaller systems makes new applications possible and thus does cause a shift in the detection limit, albeit related to a particular application and not in the absolute sense.

Very powerful in this respect are approaches that combine fabrication technologies that are suitable for mass production with transduction effects in the materials used. Combining a micro-electronic process with silicon sensor fabrication has demonstrated to be successful, as it enables the high-volume production of a combination of sensor and read-out circuits in a single chip.

1.6 Sensitivity and detection limit in an instrument

The sensitivity of an instrument can in principle be increased without constraints by increasing the system gain. In a practical system, however, the useful gain is limited by the detection limit.

When using or designing an instrument, it is crucial to be aware of the fact that the smallest measurable value of the measurand is set by the detection limit of the instrument, which is determined by system uncertainties, parasitic signals and interfering effects, and not primarily by the system gain.

The word ‘primarily’ refers to the fact that the system gain needs to be maximised within bounds for a low detection limit value, as is discussed in this section.

Non-perfect component performance, uncertainties in the system transfer function and interference effects can be expressed in terms of **equivalent input error sources** (the term **input-referred error sources** is also used). The input signal measured can be directly compared with this equivalent input error source and conclusions with respect to measurability can be made. This concept is used in the next chapters.

Equivalent input error sources are theoretical (calculated) quantities. These are a set of signal sources connected to the input of an ideal system (i.e. a system without error). The magnitudes of the signal levels of these input sources are such, that their effect on the output of the system is identical to that of the combined effect of all the distributed error sources of the same kind in the actual (i.e. non-ideal) system without input signal applied.

As discussed in Chapters 3-5, the error signals could be generated inside the measurement system (such as offset and internal noise), or could be due to a system limitation (such as the finite rejection of a common-mode level in a differential measurement), or due to external noise injected into the system by its non-zero susceptibility to that particular source of noise.

Figure 1.10a shows a simple voltage amplifier with gain G and with noise assumed as the dominant source of error. All the distributed noise contributions acting on the circuit are represented by the equivalent input noise source u_n(t), in series with the input signal u_i(t) (note that the use of only one noise source is a simplification, which is allowed because the source voltage is assumed without source impedance, Z_g= 0 Ω -see Chapter 5). The range of the possible voltage levels at the output of the amplifier is limited to the supply voltages ($U_{o,peak}$ is in between V_+ and V_-).

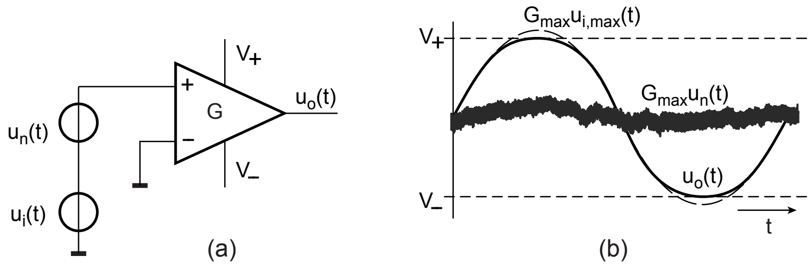


Figure 1.10, Dynamic range of a voltage amplifier.

The maximum useful level of the gain is set by distortion due to saturation at the output at given supply voltage levels. Figure 1.10b shows that for gain $G > G_{max}$ the amplified input level $Gu_{i,max}$ would result in an output voltage peak value larger than V_+. Since the amplifier cannot be driven beyond the power supply voltage, the output voltage u_o(t) is flattened at the peak values and distortion is generated. Consequently, at a given maximum input signal level, the maximum useful gain, G_{max}, is set by the onset to saturation and thus by the maximum acceptable output signal that satisfies the accuracy specifications due to distortion.

On the other hand, the minimum measurable input signal is limited by the equivalent input noise source u_n(t). If the level of the desired input signal, u_i(t), is lower than that of u_n(t), then this desirable signal can in principle not be reproducibly measured. Consequently, the detection limit, $u_{i,min}$(t), is defined. Although the subsequent chapters demonstrate that it makes sense to maximise system gain to reduce the effect of noise generated at the output, the minimum input signal that can be detected is primarily determined by the value of the equivalent input noise source and not by the system gain.

1.7 Procedure for determining the detection limit

The analysis of the detection limit generally involves the following steps:

1. Identification of the dominating detection-limiting effect.
2. Calculation of the equivalent input error signal sources.
3. Calculation of the minimum measurable signal, while also considering the inaccuracy or SNR specification.

An introduction to errors and uncertainties and how these limit the detection is provided in the next section. In electrical measurements the equivalent input error sources are voltage and current sources, whereas in, for instance, an accelerometer this is an equivalent acceleration. This analysis should provide the tools for lowering the detection limit when redesigning a sub-optimum system.

In Chapter 2 the effect of a non-perfect transducer is discussed, which includes the errors introduced by the non-selective transduction and source loading. Offset is the most likely detection-limiting factor when measuring the amplitude of quasi-static signals (Chapter 3), whereas CMRR is crucial in differential systems (Chapter 4). Noise is the limiting factor in wide band systems (Chapter 5). Circuits and signal processing techniques for improving detectivity are presented in Chapter 6. The technique for mixed-signal domain modelling presented in Chapter 7 enables the calculation of the overall system detection limit.

1.8 Exercises

1.1 Refer to Fig. 1.8, with *R*(T) at 0 °C, *R*(0°C) is equal to 100 Ω, and sensitivity (TCR)= $\partial R/\partial T$= 0.4 Ω/K. The input range of the ADC extends from 0 V to 5.12 V. The bias current sources are set to I= 10 mA. Give the values for R_o and gain G to have the input range of the ADC covered in the case of T between 5 °C and 30 °C. An 8-bit AD converter is used with U_{ref}= 5.12 V.

Solution:

R_o should be dimensioned to result in a zero differential amplifier input voltage at 5 °C: R(5 °C)= R(0 °C)+ 0.4[5 °C-0 °C]= R_o= 102 Ω.
Subsequently, the gain should be set to give full scale at 30 °C. Full scale implies code $11...11_2$. The ADC with U_{ref}= 5.12 V has a quantisation interval: q= 5.12/2^8= 20 mV, thus full scale is associated with $U_{i,ADC}$ between 5.10 V and 5.12 V. The gain setting required results from: $U_{i,ADC}$= 5.12= $I\times[R(30\ ^oC)-R_o]G$= $10^{-2}\times[25\times0.4]G$. Hence G= 51.2.

1.2 What is the maximum bias current, I, if the uncertainty in temperature due to self-heating is limited to ΔT= 0.1 °C. The temperature of $R(T)$ increases 20 °C in the case of a dissipation of 1 W. Assume T= 0 °C.

Solution:
ΔT_{max}=0.1 °C→ P_{max}= 0.1/20= 5 mW= $I^2 \times R_o$.
Hence, I= $(5\times10^{-3}/100)^{1/2}$= 7.07 mA.

A voltage measurement on a signal source with negligible source impedance results in a reading of 1.01400 V at 7 °C (=280 K). Long after the measurement, the experimenter realises that this source has a parasitic temperature dependence. Although unknown, a linear temperature dependence can be assumed. However, the source voltage may have changed over time. To attempt an *ex post* reconstruction of the temperature coefficient, two voltage measurements are performed: The first is at 20 °C (= 293 K), and results in 1.0965 V. The second at 0 °C (= 273 K) yields 1.0865 V.

1.3 Calculate the temperature coefficient of the voltage source from this measurement information.

1.4 How much has the voltage source drifted in time between the moment of the initial measurement and that of the two additional measurements?

Solution:
(1) U_{mo} (t_o, 280K)= 1.0140 V.
(2) U_{m1} (t_1, 293K)= 1.0965 V= $U_{mo}(t_o$, 280K)+ $\Delta U(t)$+ (293–280)$(\partial U/\partial T)$→
$\Delta U(t)$+ 13$(\partial U/\partial T)$= 82.5 mV.
(3) U_{m2} (t_1, 273K)= 1.0865 V= $U_{mo}(t_o$, 280K)+ $\Delta U(t)$+ (273–280)$(\partial U/\partial T)$→
$\Delta U(t)$- 7$(\partial U/\partial T)$= 72.5 mV.

(2)-(3)→ 20$(\partial U/\partial T)$= 10 mV. Hence, $(\partial U/\partial T)$= 0.5 mV/K .
(7/13)×(2)+ 1×(3)→ $\Delta U(t)$= 76 mV.