1. Instrumentation: characterising performance

1.1. *Measuring physical parameters: true value and error*

The *error* in a measurement is the difference between the value recorded by an instrument and the *true value*. The true value is a rather difficult concept, since it is in general both unknown and unknowable. Take for example the measurement of mass, for which the international standard was (until very recently!) a piece of platinum-iridium alloy held at the International Bureau of Weights and Measures in France. This individual weight was – for over a century – *the* kilogram.[1] There are various copies of this, which are held by national standards labs around the world, the accuracies of each are estimated by analysis and occasional comparison. The problem is that the actual mass of the standard kilogram can never be pinned-down to an arbitrarily high level of accuracy. For example, as soon as it is removed to be compared against something it must be touched, cleaned, exposed to air and so on, which inevitably results in some material being either lost or gained. Even if we could, at some instant in time, know exactly how many atoms it contains, what is the exact mass of each atom? As physicists, we understand that there is no such thing as an exact value for any physical parameter, which is why we always have an error in any good measurement. The important thing is to know how good our estimate of the error is. This is often harder to do than measuring the parameter itself.

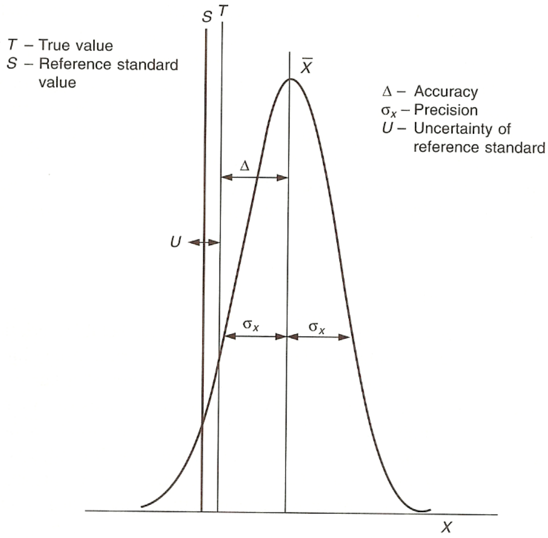

Figure 1.1: Precision, accuracy and related terms.

1.2. *Calibration*

So, until very recently the *primary standard* for the kilogram was the single mass held in France; nationally held copies are the *secondary standards*; and there are still a whole chain of copies of copies down to the weights used in laboratories in industry and academia, which can be traced back to the primary standard with some well established level of accuracy. *Calibration* is the process of comparing our instrument with some reference measurement which can be traced back to one of these standards, thereby allowing us to quantify the error. Whilst the true value of any measurement is inherently unknown, it is generally taken to be the value which we would measure if we used the best possible instrument available in the field, i.e. one which is best by both design and calibration. So, in a calibration laboratory, a set of scales could be calibrated by comparing the value of a test-mass on the 'device under test' with a reference instrument with a known and traceable calibration.

In general, the *reference measurement* is the best estimate we can make of the true value. We compare our instrument with the reference measurement, and can hence make some assessment of the *accuracy*, as illustrated in Figure 1.1. Usually, calibration also implies adjustment, whereby the output is tuned to minimise the error. This may be by some kind of electronic or mechanical compensation, with the aim being to release the instrument with a calibration certificate which states the estimated *accuracy* of the device.

1.3. *Precision is not accuracy*

The error described above is a *systematic error*. We also anticipate *random error* in a sequence of measurements due to natural stochastic variability (noise) in the underlying measurable, plus any additional noise added by the instrument or measurement process. Consider, for example, 10 trainee doctors taking a patient's temperature with a clinical thermometer which has as the smallest increment 0.1°C. This is the *resolution* of the sensor. We may expect to get a range of values with a mean of 37°C. We can plot a histogram of the results and, as we increase the number of measurements to, say, 1000, the histogram should tend to a normal distribution. If the thermometer has been calibrated (had its scale adjusted) then the best estimate of the reference measurement ("true value") is the mean of that histogram. The width of the distribution (its standard deviation) is the *precision* of the measurement. High-precision equates to high-repeatability. *Accuracy* tells us how far the mean is from the reference measurement. So, if we have an instrument with good resolution and repeatability, but poor calibration, then the result will be precise but not accurate. Figure 1.2 illustrates this point.

1.4. *Characterisation*

In order to fully characterise the behaviour of an instrument we need to consider how it responds to both a static

[1] In 2019 the kilogram was redefined by taking the fixed numerical value of the Planck constant h to be 6.62607015×10^{-34} kg m^2/s, where the metre and the second are defined in terms of the speed of light c and the transition between the two hyper-fine levels of the ground-state of the caesium-133 atom, $\Delta\nu_{Cs}$.

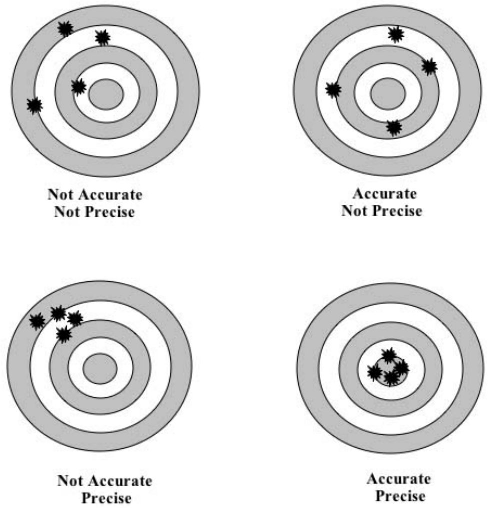

Figure 1.2: A precise measurement is not necessarily accurate.

input and under changing input conditions. How this is done and the relevance depends on the type of instrument being characterised but, in general, there are three types of measurement which can be made:

- Static response;
- Frequency response;
- Transient response.

The static response is what is generally considered to be the calibration of the instrument; the other two will give us further information about the time and frequency-domain behaviour of the device.

Static response

In this case the input is a set of static values which covers the *range* of the instrument, and the output is plotted for these – see Figure 1.3. We will usually want to be able to plot a straight line through the results which yields the *static sensitivity* (gradient) of the device and the *zero offset* (intercept). The input must be constant during the measurement and the output must be allowed to settle into the steady-state. If there is any non-linearity, then this can be expressed as the *non-linearity error*, usually taken to be the maximum deviation from the straight-line fit, divided by the full-range value, and expressed as a percentage. The input measurement is the reference value, so the static-response graph gives us our calibration; we can use it to correct the output of our instrument.

This approach is perfect for a set of scales or many other static parameters such as length. These are often referred to as *DC parameters*; however, many things we would like to measure are inherently variable *AC parameters* such as light intensity, electric/magnetic fields etc.

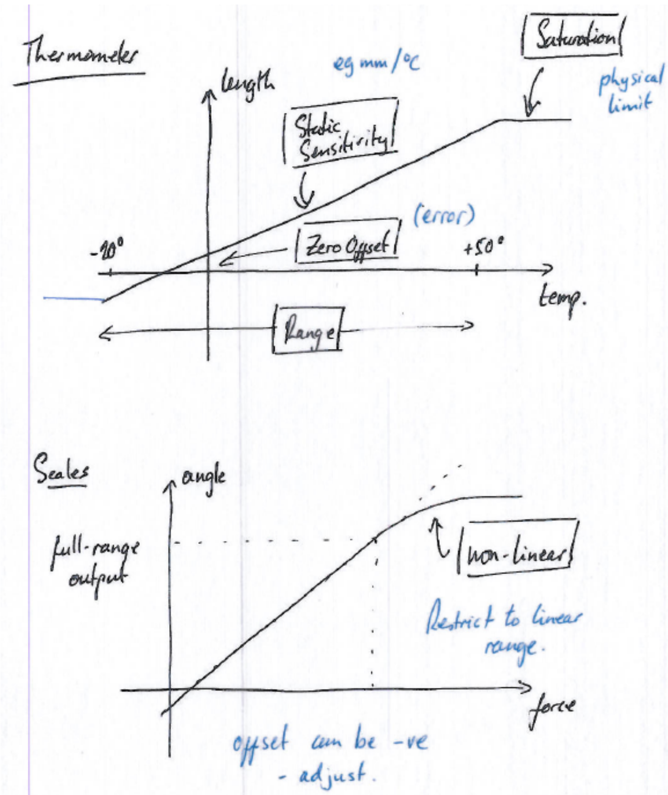

Figure 1.3: Examples of static response: thermometer and scales.

Frequency response

The *frequency response* is usually tested by applying a sinusoidal input signal which is stepped over the frequency range of the instrument. At each frequency step, the amplitude and phase of the output are measured with respect to the input. The result is graphed on a *Bode plot*, which is actually two plots: the first expresses the relative amplitude of output/input; the second gives the phase difference, with output leading input representing positive phase by convention. The frequency axis is usually logarithmic, and by convention the relative amplitude is given in dB, the phase in degrees. An example is shown in Figure 1.4.

The frequency response is the direct equivalent in the frequency domain of the static response. For any given input frequency we can predict both the amplitude and phase of the output. Since we know any arbitrary input signal can be decomposed into sinusoids, we can use the Bode plot to find the instrument's response to each, and hence predict what the output will be by summing the responses. This only applies if the instrument shows a nicely linear static response, which is why it is so important for instruments to be 'linear'. We will study this in much more detail later on; however, for now, it is important to understand that an instrument with *linear* characteristics can be analysed in the frequency-domain using Fourier techniques.

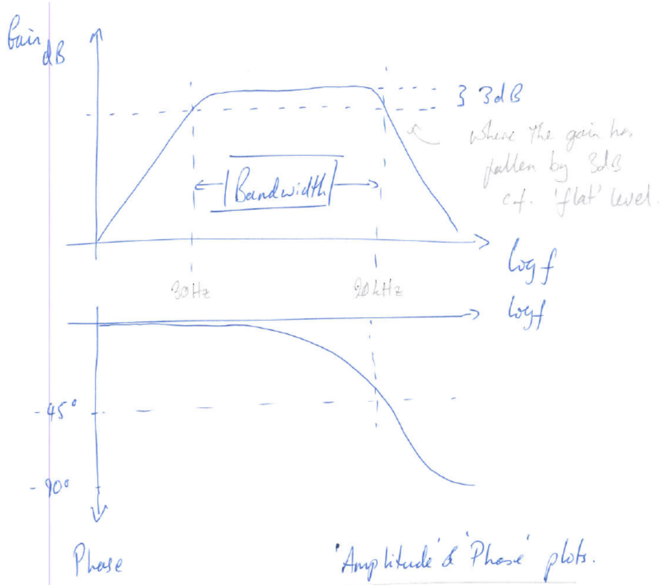

Figure 1.4: Bode plot for microphone.

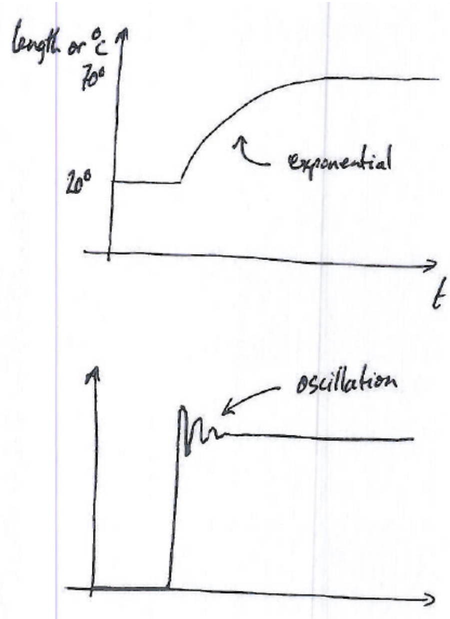

Figure 1.5: Transient responses for thermometer (upper panel) and scales (lower panel).

Transient response

We often wish to know the response to a non-sinusoidal transient change at the input such as when we place a mass on the scales (mathematically, a *step function*) – see Figure 1.5. Another useful test-input is the *delta function*. These can provide information on the settling-time for the output, as well as being used to test for system stability. Again, we will cover this in much more detail later on.

1.5. *The decibel*

The dB is a way of expressing ratios on a logarithmic scale. It is ubiquitous in engineering and takes some practice to get used to. By definition the decibel expresses the ratios of two power levels

$$n\,\mathrm{dB} = 10\,\log_{10}\frac{P_2}{P_1}\,.$$

If we *double* the power output of an amplifier then we say the power has "gone up by 3 dB". Four times the power is 6 dB, which is the same as doubling the voltage (amplitude) output of the amp, since power is proportional to amplitude squared:

$$n\,\mathrm{dB} = 20\,\log_{10}\frac{A_2}{A_1}\,.$$

Further confusion can arise from the fact that dB are sometimes used to express absolute amplitude quantities, such as a voltage relative to 1 μV: 1 volt in dBμV is expressed as 20 $\log_{10} 10^6/1$ dB. Another common quantity is the 'dBm' (decibel-milliwatt), used to indicate a power level with reference to 1 mW. This can be confusing and it is worth spending some time to become familiar with the usage.

1.6. *Linearisation*

As mentioned above, a linear instrument is important to allow Fourier analysis. In a future lecture we shall define mathematically what is meant by the term 'linear system', and then we will return to this later in the course to see how the Fourier and Laplace transforms can be brought to bear to understand the behaviour of really quite complicated systems involving feedback, amplification and other linear operations. For now, it is important to understand that this is the reason that instrument designers pay so much attention to ensuring that the static response of the instrument is a straight line through the origin. If there is an offset then this can be compensated for electronically. If the output deviates from a straight line at high inputs then we can artificially limit the input. Both are examples of *linearisation*. Consider also the thermistor (Section 2), which has a roughly exponential temperature response. This can be coupled with a logarithmic amplifier to yield a linear response.

1.7. *Loading*

As a final note on the subject of measurement errors, we must always be aware that the very act of measurement can disturb the measurement we are trying to make. Two simple examples. First, when we put a thermometer in a cup of tea, we cool the tea slightly. Secondly, imagine a pressure sensor which works by moving a diaphragm against a spring in response to pressure. In a constrained space, the volume of the sensor will change slightly and could influence the pressure measurement. These are both examples of the sensor *loading* the measurement. We will encounter many more examples of of loading in electronic systems.

2. Sensors and transducers

Generally, a *transducer* is any device which converts some physical parameter from one form into another, usually in a proportional and predictable way. All sensors are transducers, specifically they are *input transducers*, in that they convert some physical observable into a more useful form, usually – but not always – an electrical signal. The mercury thermometer, for example, is an input transducer which converts thermodynamic temperature into length. Typically though, we want the output of the transducer to be an electrical signal so that we can do further manipulation and processing using electronic circuits, and ultimately obtain a representation of the signal into a computer-readable form. Also, we may want to go in the other direction, i.e. converting an electrical signal into some kind of physical observable (such as light, heat, sound, movement etc.) for which we need an *output transducer*. Figure 2.1 illustrates the idea with a simple "public address" system. The microphone is an input transducer which converts speech (sound pressure variations) into a voltage signal. Then we have an amplifier, where we can also do more fancy things such as adjusting the tone (frequency content) of the signal. The amplifier drives a speaker which is the output transducer, whose job is to convert the amplified electrical signal back into a mechanical movement (of the speaker diaphragm) which in turn gives us sound. Note that we will be studying a more complicated system of sound/electrical transducers in the lab sessions.

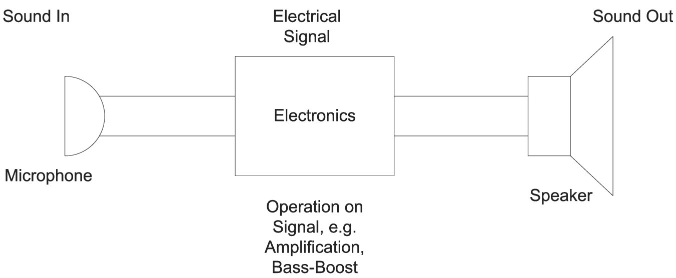

Figure 2.1: Input/output transducers.

There are thousands of different types of transducers and we will only look at a few. This is not a course on detector physics, and we do not have the time to study the details of many individual transducers. However, we will see that there are many similarities in their behaviour, so we will study them in a more general (that is to say device-independent) way. In this handout we will look briefly at two different types of input transducer: temperature and photon sensors. Note that, according to general convention, we will usually use 'sensor' and 'input transducer' interchangeably during this course.

2.1. Temperature transducers

Thermocouple. This is the simplest temperature transducer/sensor since it directly generates a voltage. A thermoelectric voltage is generated across any metal when the two ends are at different temperatures. Using two junctions between dissimilar metals – one held at a known temperature, the other at the point we wish to measure – the voltage generated is proportional to the temperature difference between junctions. The voltage is small – of the order a few μV/°C – so precision voltage measurement is required. The technique requires some care; however, the thermocouple is robust, small, and operates over a very wide temperature range (-200°C to $+1500$°C).

Platinum Resistance Temperature Device. This is the standard for high-accuracy measurement over the temperature range -200°C to $+600$°C. The PRTD shows a reasonably linear variation in resistance as a function of temperature. Consequently, we just need to measure the resistance (e.g. with a multimeter) then apply a conversion factor to get the temperature. They typically have a rather slow response time, and can be reasonably expensive.

Thermistor. This is a semiconductor device where resistance decreases rapidly with temperature. The temperature coefficient of resistance is of the order 4%/°C, which is an order of magnitude larger than that of the PRTD. The thermistor is also small, cheap and robust, but not as accurate as the PRTD. It is a good choice for temperature measurement over the range -75°C to $+150$°C. They are available in all sorts of packages, for example the glass-encapsulated ones are good for harsh or chemical environments (though since glass is an insulator this slows down the response time to temperature changes). One drawback is that the temperature-resistance relationship looks like:

$$\frac{1}{T} = A + B\ln R + C(\ln R)^3 \,.$$

To go from R to T we need either a calculation or a look-up table, both of which are possible with electronics and/or a microprocessor, but this does mean that we cannot simply hook the thermistor up to a meter in the lab and take a quick reading of the temperature.

The point of this discussion is that none of the above are perfect and we have to make a choice according to: *i*) the temperature range we need to cover; *ii*) the accuracy required; *iii*) the harshness of the environment and *iv*) the ease of conversion into a temperature value. This is a general rule for transducers: the perfect device does not exist and we often need to compromise.

AD590. Where such a situation exists, the semiconductor industry is always ready to sell us a solution. The AD590 from Analog Devices is an integrated circuit device which produces a highly linear current output of 1μA/K. It works over a reasonable temperature range of -55°C to $+150$°C. It comes in a range of packages (Figure 2.2) and can be bought for under \$10 (more expensive than a thermistor but cheaper than a good PRTD). A simple circuit of resistors converts the current output to a voltage, then a multimeter gives us a very quick and easy direct reading of the temperature. The complexity of linearising the output as a function of temperature is all handled within the circuitry of the device (this is what we are paying for). We do not generally need to be concerned with what goes on inside the device as long as we have a good datasheet to describe its behaviour. The front-page of the AD590 is reproduced in Figure 2.3. It is useful to browse the complete datasheet and you are recommended to download a copy from the manufacturer's website at http://www.analog.com. Look through it to see the level of detail that the manufacturer thinks is important for the user to know. You also might want to identify the ways in which the performance of this transducer departs from 'ideal' behaviour.

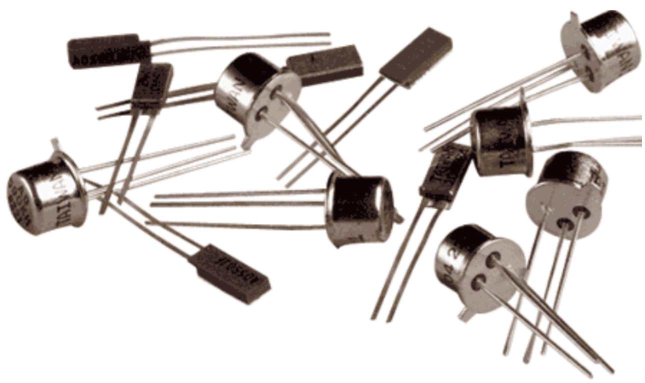

Figure 2.2: IC temperature transducer: Analog Devices AD590.

2.2. Photon sensors

For the temperature transducers we are not, generally, expecting to operate at the limits of physics (unless we are trying to measure mK temperatures, but then we would use a different technique anyway). We often need to measure the temperature of our experiment or equipment and the above techniques will usually suffice. Next we will look at something more sophisticated: photon detection.

Photomultipliers tubes (PMTs) have been around for many decades and are still the sensor of choice for many photon detection applications. With the modern trend of implementing sensors in silicon, benefiting from mass production, robustness, etc., silicon photomultipliers (SiPMs) have become very popular in recent times. Also in this case sensor selection involves trade-offs across *many* parameters, and we shall discuss some examples here along with their operating principles.

2.2.1. Photomultiplier tubes

The photomultiplier tube (PMT) is effectively a type of light intensity transducer. It is based on vacuum-tube (i.e. 'valve') technology. It is capable of efficient photon detection at extremely low-light levels (down to the single photon). Manufacturers such as Hamamatsu Photonics or ET Enterprises Ltd (ETEL) have developed devices specifically for the physics community. At the extremes of measurement science we need to understand the physics of how the detector works in order to get the best out of it, which is why we will look into the principle of the PMT. Some models are shown in Figure 2.4.

The operation of the device is illustrated in Figure 2.5. A photon striking the photocathode liberates a photoelectron. The first "dynode" is held more positive than the cathode so the liberated electron is accelerated towards it and, on striking the dynode, several electrons are released – assume only 2 in this example. Each subsequent dynode is progressively more positive so, in this case, we get an electron multiplication of 2^n where n is the number of dynodes. Now, in real devices we can get a gain in excess of 10 at each dynode. At the end of the tube the electrons are collected at a final anode and, flowing back to the power supply through a resistor, generate a voltage signal, the amplitude of which is proportional to the number of photons detected.

For a single photon detected (one photoelectron emitted from the photocathode) we can easily achieve a *gain* of 1–10 million (electrons at the anode per photoelectron emitted from the cathode), producing a few mV across a load resistor – which is an easily detectable voltage.

The PMT electrodes (photocathode, dynodes, anode) are typically biased using an external voltage divider circuit (see Figure 2.6) connected to a high voltage (HV) power supply. It is important to understand that the source of all the electrons which flow inside the tube is ultimately the power supply. In "dark" operation, a current flows around the loop which consists of the power supply and the resistors in the voltage divider alone. In the presence of light, some of these electrons are diverted into the tube via the dynodes, such that part of this current flows out of the tube via the anode. Ultimately, the tube will be saturated when all of the electrons from the HV supply are diverted into the tube. The entire current from the supply then flows through the anode and the voltage signal readout from the tube will be the maximum possible. This understanding is quite subtle and requires a working knowledge of basic DC circuits. We will cover this ground again in more detail in future lectures.

PMTs are not without their difficulties. A large (and somewhat dangerous) HV supply is needed (1–2 kV is usually applied across the tube). Even when there is no incident light, thermal emission of electrons from the photocathode means that a current always flows in the tube. This called the *dark current* and we want to minimise it since it adds noise and reduces the dynamic range of our measurement (these are both topics which will be discussed in more detail later on). One way to do this is to cool down the PMT to reduce the thermal emission. Note that the dark current appears as an *offset*, since the sensor shows an output even with zero light input.

ANALOG DEVICES

2-Terminal IC Temperature Transducer

AD590

FEATURES

Linear current output: 1 μA/K
Wide temperature range: −55°C to +150°C
Probe-compatible ceramic sensor package
2-terminal device: voltage in/current out
Laser trimmed to ±0.5°C calibration accuracy (AD590M)
Excellent linearity: ±0.3°C over full range (AD590M)
Wide power supply range: 4 V to 30 V
Sensor isolation from case
Low cost

GENERAL DESCRIPTION

The AD590 is a 2-terminal integrated circuit temperature transducer that produces an output current proportional to absolute temperature. For supply voltages between 4 V and 30 V, the device acts as a high impedance, constant current regulator passing 1 μA/K. Laser trimming of the chip's thin-film resistors is used to calibrate the device to 298.2 μA output at 298.2 K (25°C).

The AD590 should be used in any temperature-sensing application below 150°C in which conventional electrical temperature sensors are currently employed. The inherent low cost of a monolithic integrated circuit combined with the elimination of support circuitry makes the AD590 an attractive alternative for many temperature measurement situations. Linearization circuitry, precision voltage amplifiers, resistance measuring circuitry, and cold junction compensation are not needed in applying the AD590.

In addition to temperature measurement, applications include temperature compensation or correction of discrete components, biasing proportional to absolute temperature, flow rate measurement, level detection of fluids and anemometry. The AD590 is available in chip form, making it suitable for hybrid circuits and fast temperature measurements in protected environments.

The AD590 is particularly useful in remote sensing applications. The device is insensitive to voltage drops over long lines due to its high impedance current output. Any well-insulated twisted pair is sufficient for operation at hundreds of feet from the receiving circuitry. The output characteristics also make the AD590 easy to multiplex: the current can be switched by a CMOS multiplexer, or the supply voltage can be switched by a logic gate output.

PIN CONFIGURATIONS

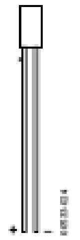

Figure 1. 2-Lead CQFP

Figure 2. 8-Lead SOIC

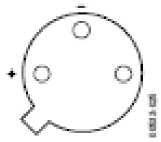

Figure 3. 3-Pin TO-52

PRODUCT HIGHLIGHTS

1. The AD590 is a calibrated, 2-terminal temperature sensor requiring only a dc voltage supply (4 V to 30 V). Costly transmitters, filters, lead wire compensation, and linearization circuits are all unnecessary in applying the device.
2. State-of-the-art laser trimming at the wafer level in conjunction with extensive final testing ensures that AD590 units are easily interchangeable.
3. Superior interface rejection occurs because the output is a current rather than a voltage. In addition, power requirements are low (1.5 mW @ 5 V @ 25°C). These features make the AD590 easy to apply as a remote sensor.
4. The high output impedance (>10 MΩ) provides excellent rejection of supply voltage drift and ripple. For instance, changing the power supply from 5 V to 10 V results in only a 1 μA maximum current change, or 1°C equivalent error.
5. The AD590 is electrically durable: it withstands a forward voltage of up to 44 V and a reverse voltage of 20 V. Therefore, supply irregularities or pin reversal does not damage the device.

One Technology Way, P.O. Box 9106, Norwood, MA 02062-9106, U.S.A.
Tel: 781.329.4700 www.analog.com
Fax: 781.461.3113

Figure 2.3: AD590 datasheet (link).

Figure 2.4: Various PMT models from Hamamatsu Photonics.

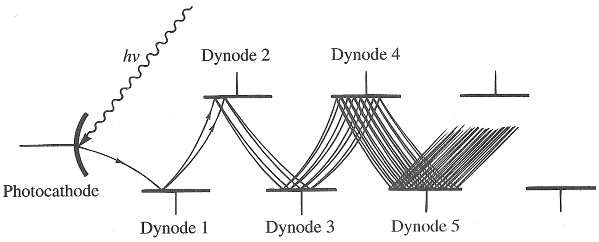

Figure 2.5: PMT operating principle (from Diefenderfer & Holton). This example shows a side-window PMT; most devices are front-illuminated PMTs using a semi-transparent photocathode.

The *quantum efficiency* (QE) of the photocathode is a very important parameter. Not all photons incident on the photosensitive layer lead to the mission of a photoelectron – since photons can simply be transmitted through, or a photoelectron may be created but never be emitted into the PMT vacuum. The QE is the fraction of incident photons which generate detectable electron emission. A QE of 25% is a typical minimum number, often as high as 40% in some PMT models.

Further Reading. Diefenderfer & Holton chapter 7 is good on the general topic of transducers and has more details of temperature transducers and PMTs. Note, however, that their figure 7.18 is wrong! (it is a useful exercise to

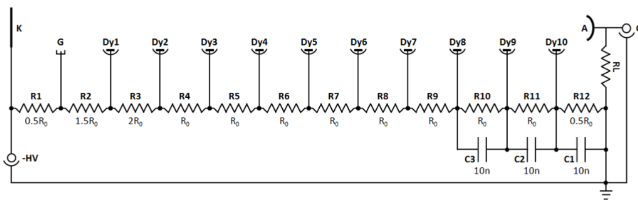

Figure 2.6: PMT voltage divider circuit. A single HV (here −1.0 kV) is applied to the cathode, and the circuit generates the voltages needed to bias the dynodes. The signal is collected at the anode and the current develops a voltage across the load resistor, R_L. The three capacitors ensure that the dynode electrodes can provide the required number of electrons without a significant voltage drop.

work out why). Horowitz & Hill chapter 15 is also very good, though it will make more sense once we have covered some of the electronics involved. For PMTs very good background and technical information can be found on the Hamamatsu and ETEL websites at `www.hamamatsu.com` and `www.et-enterprises.com`.

2.2.2. Silicon photomultipliers

SiPMs are silicon-based photon sensors consisting of an array of very small single-photon avalanche devices (SPADs) operating in Geiger mode, i.e. a few volts above breakdown. A small sensor 5 mm across may contain some 10,000 SPAD micro-cells (each typically 10–100 μm across). SiPMs are robust devices, requiring low bias voltages (tens of V), and offer extremely good characteristics for low-light (photon counting) applications. Some examples are shown in Figure 2.7.

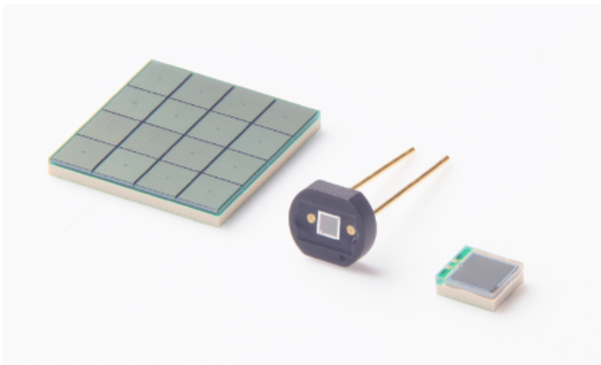

Figure 2.7: Hamamatsu SiPM devices and small arrays.

The silicon implementation of a SiPM is complex and varies depending on the wavelength of light and other factors – but the basic operating principle is relatively simple. A photon absorbed in a SPAD pixel produces a photoelectron which gets 'multiplied' in an avalanche process in the silicon. Each SPAD has a quenching resistor in series that limits ('quenches') the response of the pixel, i.e. the output signal of a cell is independent of the number of electrons which initiated the avalanche. Every signal in a single pixel produces the same-size response, typically ~1M electrons – i.e. a gain comparable to a PMT. The outputs from all SPADs in the SiPM are connected in parallel, as illustrated in Figure 2.8, and summed across a common load resistor. At low light levels the output is highly linear and is proportional to the number of SPADs that 'fired' in response to the optical stimulus.

The probability that an incident photon is detected is not controlled by a quantum efficiency alone, as in the case of a traditional PMT, but rather by a *photon detection efficiency* (PDE). In addition to the quantum efficiency of the silicon, this parameter must take into account the probability that a photoelectron actually initiates an avalanche and, significantly, the 'fill factor' of the device, i.e. the ratio of active to total area of a pixel as a result of the dead space between SPADs. Nevertheless, SiPM PDEs

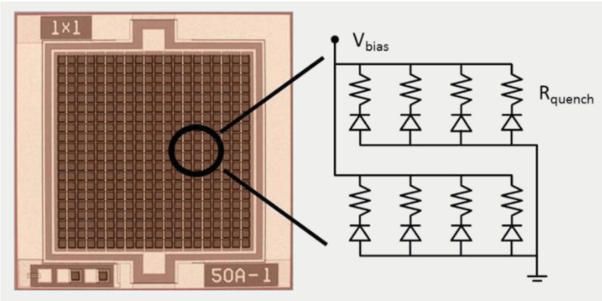

Figure 2.8: SiPM principle of operation (from Ketek GmbH).

can match PMT QEs in most wavelength regions, and we saw above that the gain is also comparable.

SiPMs have a number of advantages over traditional vacuum PMTs, as well as some disadvantages. Their ability to 'count photons' with extremely high resolution is perhaps their main advantage. On the other hand, they have extremely high *dark count rates* at room temperature, as well as *correlated noise* (e.g. an avalanche in one SPAD can trigger one in an adjacent cell).

In conclusion, photon detection is a good example of where a long and very successful technology is being slowly replaced by a more modern, silicon-based one, but even in this case no device is perfect: in any application it is critical to understand the detailed technical requirements and how this maps to the specification offered by each technology or device. The solution is often not obvious.

3. Input and output impedance

Thevenin and Norton equivalent circuits. We can consider any device such as a sensor to be modelled as a Thevenin or Norton source. Usually, we model a device which is inherently voltage-producing (e.g. LM35 temperature sensor) as a Thevenin source, and a device which is inherently current-producing (e.g. PMT) as a Norton source – but in practice the two models are interchangeable (you can convert a Thevenin model to a Norton model and vice-versa if you wish). Both these models give us the idea that the source is always associated with some kind of impedance known as the source or *output impedance.*

Output impedance. This is the impedance which we would measure 'looking into' the terminals of a source. This we can do by varying the load resistance which we connect it to, and seeing how the output signal (voltage or current) varies as a function of load, then using this information to calculate the effective impedance of the device. This is how it is done in first-year lab, in the experiment to calculate the output impedance of a signal generator.

We hope that the output impedance of a voltage source is low, and for a current source that it is high, and in both cases that it is real (resistive). This is generally true; however, any length of wire will have an inductance, and consequently there will always be some small reactive component of impedance too. This is a function of the design of the sensor, so we can try to keep the inductance low by minimising cable length and (especially) avoiding loops and coils of wire. In practice, sensor inductances are usually small and, since the reactance $X_L = \omega L$, the contribution to total impedance is negligible for all but the highest frequencies. There is also some capacitance in any sensor, and this causes more problems, typically. Any wire running next to another wire or the case of the sensor will result in there being some capacitance between the two conductors. This is called *stray capacitance.* There may also be some capacitance due to the physics of the sensor, e.g. the PMT where the anode and dynodes are metal plates. Overall, these capacitances can add up to some tens of pF, resulting in reactances of some kΩ at MHz frequencies. Stray capacitance is usually modelled as being in parallel to the sensor terminals, while stray inductance (should it exist) will appear in series.

Input impedance. Any device to which we connect a sensor, such as an amplifier, a meter or an oscilloscope, will have an *input impedance* which is the apparent ratio of the voltage applied to the device to the current it draws. Generally, we want this to be high and resistive for voltage measurements, and low for current measurements.

The same arguments as above apply for stray capacitance and inductance. A typical oscilloscope input is equivalent to a resistance of 1 MΩ in parallel with a capacitance of about 20 pF. Due to the fact that the capacitance is in parallel, it has no effect at low frequencies, but at MHz frequencies it is clear that the total input impedance $\mathbf{Z_T} = (1/R + 1/\mathbf{Z_C})^{-1}$ will be significantly reduced. Most scopes come with special probes designed to boost the impedance and hence prevent the probe from loading the circuit being measured.

3.1. Impedance matching

When the source and load impedance are equal then we find that the maximum power is transferred from source to load. We will also find that this is the condition for minimising any reflections of signals. Note that, in the case of the voltage source, the source and load form a voltage-divider in such a way that the voltage measured in the load is only half the source voltage – see Figure 3.1. This is due to the significant current which the load (same low impedance as the source) is pulling from the source. This is an example of *loading* a source; we might say that by loading the sensor with a low-impedance we have "pulled down" the voltage. This is OK, as long as our intention is to transfer maximum power.

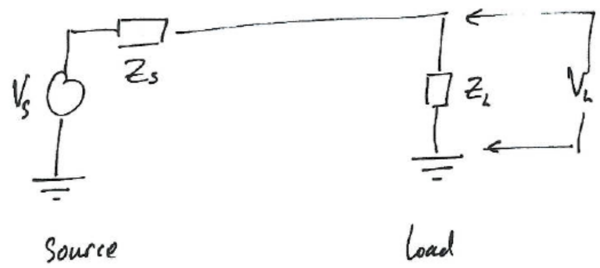

Figure 3.1: Impedance matching: the power transfer to the load is maximised when $\mathbf{Z_S} = \mathbf{Z_L^*}$ (or $R_S = R_L$ if both are purely resistive). Note that when the impedance is matched the source has been loaded, i.e. we now measure only a fraction of the voltage: $V_L = V_S/2$ – though $P_L = P_S$ in this instance.

The *maximum power transfer theorem* states that for source and load resistances R_S and R_L the maximum power is transferred for $R_S = R_L$. In the general case of complex impedances, we find that we need $\mathbf{Z_S} = \mathbf{Z_L^*}$. This results in equal power being dissipated in both source and load, since the voltage-divider splits the source voltage equally across source and load impedances. This is not very efficient, but we do achieve maximum power, which we sometimes wish to do. The theorem is easily proved; this will be left as a problem sheet question.

If source and load impedance are fixed and we want to achieve maximum power we can use an *impedance transformer*, as shown in Figure 3.2. This works for AC signals only, as it requires use of a transformer to magnetically couple the power from the primary to the secondary circuits. A transformer consists of two coils of wire (solenoids) wound together often on a core of high magnetic permeability. The inductance of each solenoid is proportional to the square of the number of turns. We have $L \propto n^2$ and $Z \propto L$ for both the primary and secondary sides of the transformer. Since the transformer is entirely inductive (reactive) it couples power from primary to secondary sides *with no loss* (in practice there is some small resistance in the wire which results in ohmic loss, of course; however, transformers can be very efficient). By choosing the number of turns on the primary side to give the same impedance as the source we couple max power from source into the transformer. By choosing the number of turns on the secondary winding to give the same impedance as the load we couple all the power out of the transformer into the load. As if by magic, the transformer has 'transformed' the impedance of the load into something which matches the source. In practice, this works quite well for frequencies in the kHz to MHz range.

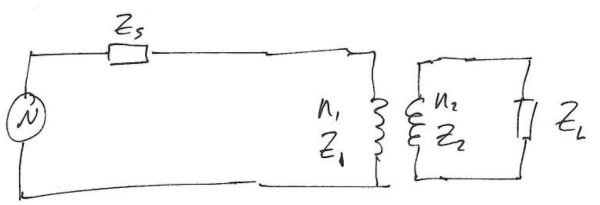

Figure 3.2: Impedance transformer: stepping up or stepping down impedances.

Efficiency. We can define *efficiency* as the ratio of load power to total power: $\epsilon = P_L/P_T \times 100\%$. Maximum power transfer results in only 50% efficiency! This seems counter-intuitive, but consider a battery with $R_S = 1\ \Omega$: if we connect a 1 Ω load we will get the maximum power we can out of the battery. Both battery and load will get very hot, and the battery will be exhausted very quickly. Half of the chemical energy has been wasted heating the battery. We can make better use of the stored energy by releasing it more slowly into a higher-resistance load. It can easily be shown that:

$$\epsilon = \frac{R_L}{R_L + R_S}\,.$$

Hence, for maximum efficiency we want either a low R_S or a high R_L. Either will deliver improved efficiency, at the expense of reduced overall power. In other words, it is important to understand that maximum power transfer and maximum efficiency are different things. We can increase the efficiency by increasing the load resistance; however, this decreases the overall power delivered by the source. If we want instead to maximise the absolute amount of power transferred to the load, we should match the two impedances.

3.2. Impedance bridging

In many situations we do not wish to transfer maximum power from source to load. For example, for many voltage-source sensors we just wish to measure the voltage which the sensor produces. Indeed, drawing any current from the sensor is a problem, since the loading effect will drop voltage across the sensor's source impedance. In this case, we will use *impedance bridging* where a low-impedance source is connected to a high-impedance load. For example, measuring the voltage of a battery we use a digital multimeter (DMM) with an input of 1 MΩ or so. The efficiency here is very high and, in practice, we transfer negligible power from the battery into the meter. The meter is a voltage-sensing device; we do not wish it to draw any more current than is strictly necessary for accurate operation. *Impedance bridging* is the norm for instrumentation applications where we simply wish to move a voltage signal around without transferring significant amounts of power – a typical example is illustrated in Figure 3.3. The exception to this is where we need to *impedance match* specifically to avoid signals being reflected (e.g. if the signal out of the sensor is very fast, such that reflections from a mismatched amplifier would distort the signal shape or even appear as new, delayed signals).

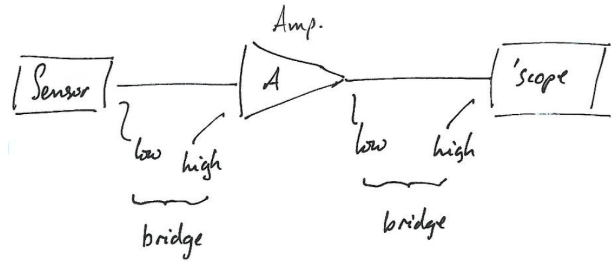

Figure 3.3: Impedance bridging to preserve the amplitude of a voltage signal. In this regime the source impedance is low and the load impedance is high, so that the efficiency approaches 100%.