

INSTRUMENTATION AND COMPUTER CONTROL SYSTEMS

SENSORS AND SIGNAL CONDITIONING

Steve Collins

Michaelmas Term 2012

Introduction

An instrumentation system obtains data about a physical system either for the purpose of collecting information about that physical system or for the feedback control of the physical system

Any instrumentation system must include an **input transducer (sensor)**, such as a strain gauge, whose response to a particular stimulus can be measured electrically. **The other component that is generally present in modern instrumentation systems is a digital processor, such as a computer or a micro-controller.** These programmable components have the flexibility to be used for a variety of functions. The most important function that they perform is to **convert data into information.** In the simplest situation the processing required to extract information may only involve converting an input signal by a scale factor so that the final result is in conventional units. For example, the output voltage signal from a strain gauge may be converted to the corresponding actual strain. Alternatively, within a more sophisticated system the signal from a strain gauge placed on an engine mounting might be processed to extract the vibrational spectrum of an engine, which is then used to detect any unusual frequency that might be indicative of wear. This information can then be displayed to a

user, stored for later analysis, transmitted to a remote location or used by a controller.

The signal from a transducer is usually **analogue** in nature, ie. it is **continuously** varying and can take **any** value (**within an allowed range**). **This continuous analogue data has to be converted to a digital format prior to being transferred to the digital processor. Any instrumentation system must therefore include an analogue-to-digital (A/D) converter (ADC for short) to convert an analogue signal into a digital format, such as those discussed in the first year P2 course.**

A typical **ADC** will be an existing component that has been designed to **convert an analogue input voltage, typically with a range of a few volts, into a digital word, which usually contains 8 or more bits.** However, the output from a typical transducer, such as a **strain-gauge**, might have an amplitude of less than **10 mV**. This transducer output signal must therefore be **amplified** in an **analogue signal conditioning circuit** before it can be converted into a digital word.

Another aspect of the performance of the ADC that must also be taken into consideration when designing the signal conditioning circuit is that **the ADC samples the transducer output at specific time intervals. An unfortunate consequence of this is that several frequencies will become indistinguishable at the ADC output. This is referred to as aliasing, and the effect can only be avoided by using a low-pass, anti-alias filter** to ensure that only the low frequencies that can be represented accurately are present in the signal applied to the ADC input. Since the

requirement for the anti-alias filter arises from a fundamental property of the ADC, this type of filter should always be present.

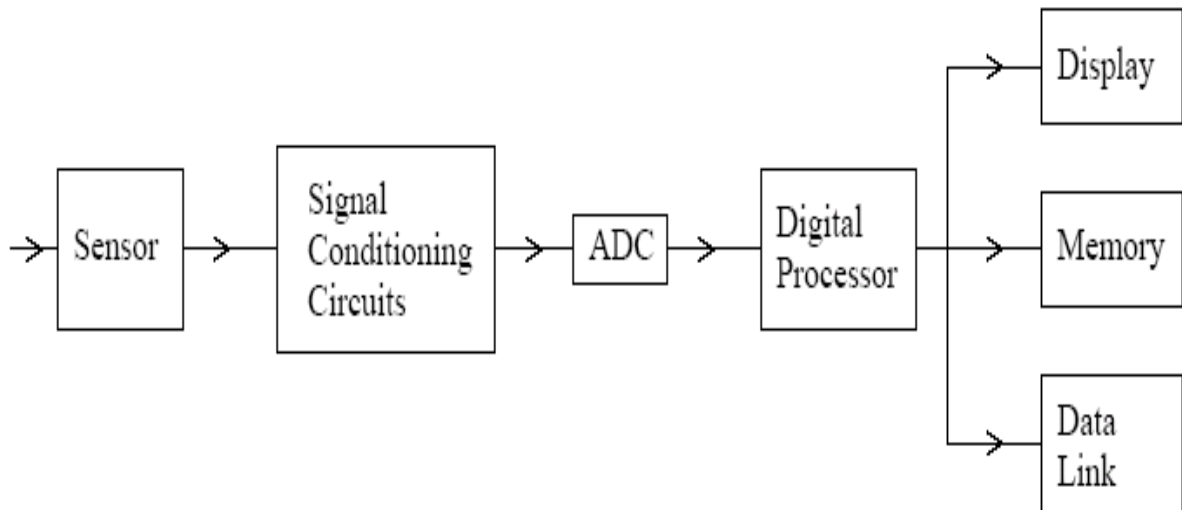


Figure 1: A block diagram of a typical instrumentation system with several different output devices

As shown in Figure 1 the characteristics of typical sensors and ADCs mean that the data collection (or acquisition) part of a typical modern instrumentation system can be split into the three functional blocks, a sensor, signal conditioning circuits and an ADC. The digital output from the ADC can then be processed in a programmable digital processor to **extract information that can be displayed to an operator, stored in a memory or transmitted via a data link or used in feedback control.**

The costs of all the components are continually falling. It is therefore becoming economically viable to gather an increasing amount of data, and hence hopefully information, from an every expanding range of host systems. One example of this trend is a 2.25 Km suspension bridge

constructed for the 2000 Olympics that had approximately 300 sensors embedded within the structure. These include:

- strain gauges to keep track of framework fatigue

- sensors to monitor motion in the stay cables caused by cross winds

- accelerometers in the roadway to measure the impact of earthquakes

The data from these sensors are gathered by four separate data-acquisition units (one in each pier of the bridge). These linked units are then connected to offices at the bridge site, in the headquarters of the bridge operating company in Athens and in the headquarters of the structural monitoring division of one of the bridge builders, which is in France.

This technology trend and its impact on every conceivable system means that **all engineers** should be familiar with instrumentation systems.

Aims and Organisation of the Course

The aim of the sensors and signal conditioning course is to develop an understanding of the function of the first key parts of a typical instrumentation system, such as the one in Figure 1.

The first parts of the instrumentation system that will be considered are the sensors. There are many different sensors that rely upon one of a range of different physical phenomena to create an output signal in response to different stimuli. A comprehensive survey of all sensors is time-consuming and beyond the scope of this course. However, we will aim to give a brief survey (or list) of sensor types and the types of signals which might be produced. Such signals are typically rather weak. These signals must therefore be amplified before they are converted into digital words.

One problem caused by the small amplitude of the output signals from sensors is that they can be easily confused with other small voltage changes within the instrumentation system. Techniques to reduce the **interference** caused by these other small voltage changes, including careful design of the circuit layout, shielding it from external electromagnetic fields and creating a signal represented by the voltage difference between two signals, will be briefly described. The resulting small **differential** signals could be amplified by a differential amplifier containing a single operational amplifier (op-amp). However, this circuit is not ideal and the more complex, but easier to use, **instrumentation**

amplifier, is therefore used. This was discussed in the first year P2 course, but it is so important we will cover it again here.

Once the analogue output signals have been amplified they need to be converted into a digital word. Any instrumentation system must therefore include an **analogue to digital converter (ADC)**. Two types of ADCs that are often used in instrumentation systems were discussed in the first year P2 course. The flash converter is conceptually simple and fast. However, it necessarily contains a large number of components and it is therefore relatively expensive. In some situations it is necessary to use an alternative type of converter. The alternative converter that was described in the P2 course, known as the successive approximation converter, contains a digital to analogue converter (DAC). These DACs are also useful in digital control systems, such as the one shown in Figure 2. In this course, the discussion of ADCs will concentrate on issues relating to **interfacing**.

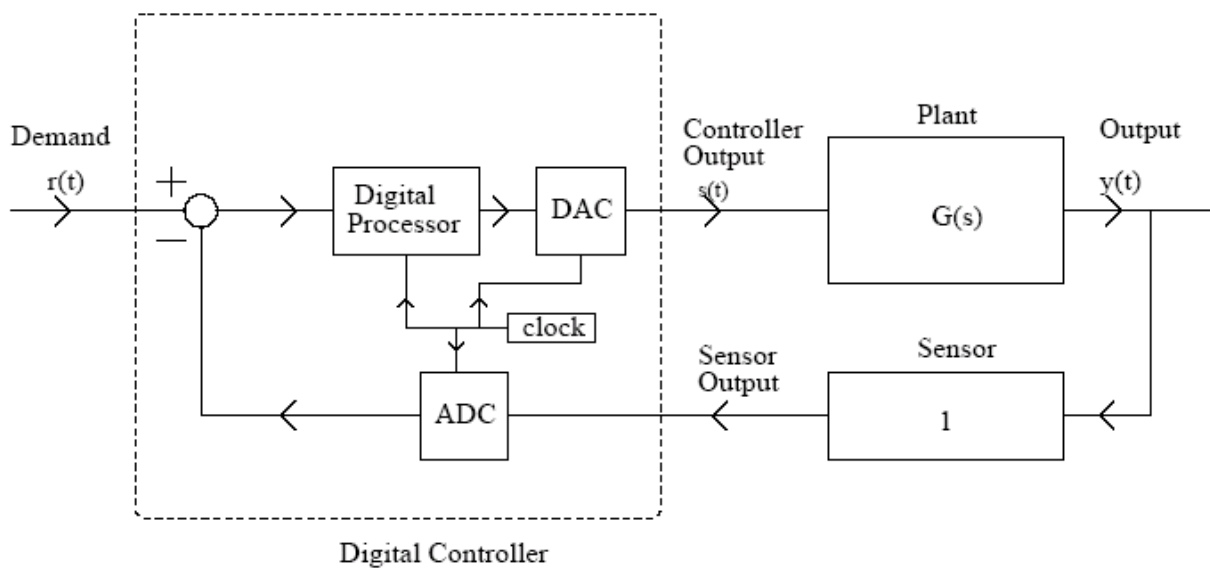


Figure 2: A block diagram of a typical digitally based control system.

Amplification of the sensor output signal is designed to match the maximum expected sensor output signal to the maximum input signal of the ADC. The sensitivity of the resulting data is then determined by the minimum change in the signal that can be reliably detected. The performance of ADCs means that this important aspect of the system performance is often limited by undesirable signals which are generated **within** the components of the electronic circuit. The origin of these **noise** signals, their effect on instrumentation systems and methods to limit their effects will therefore be described.

The discussion of **noise highlights the fact that one of the important system parameters that determines the amount of noise in a system is the system bandwidth. Filter circuits that can be used to control the bandwidth of the system therefore play a critical role in limiting the impact of noise on a system.** In fact there are four key different types of filters that are commonly used in instrumentation systems to fulfil a variety of functions. The circuits and characteristics of such filters were discussed in the first year P2 course. In this course the characteristics, applications and implementation of each of various filters will be further described.

Amplifiers and filters within an instrumentation system are typically based upon op-amps. The function of a particular circuit within one of these systems can be understood by analysing the circuit with the assumption that the op-amp is ideal. However, all real op-amps have a finite input impedance, gain and output impedance. **One of the key stages of designing any op-amp based circuit is the selection of the**

particular op-amp that should be used so that it 'appears' to be ideal in the particular circuit that is being designed. Even after an op-amp has been selected that appears to be ideal additional components may be required in a circuit to compensate for other non-ideal aspects of an op-amp's behaviour. Finally, **all op-amps have a gain that reduces at high frequencies. This means that any op-amp will only appear to be ideal for frequencies less than a maximum value.**

The output voltage from a strain gauge and several other sensors is a d.c. voltage. However, **there are sensors in which the stimulus of interest causes a change in either capacitance or inductance. Changes in these two parameters can only be sensed if an a.c. signal is applied to the sensor.** In addition, to avoid strong sources of interference, a.c. signals can also be applied to circuits containing sensors such as strain gauges. The use of this approach, with a lock-in amplifier, will be discussed.

Syllabus and Learning Outcomes

Sensors and signal conditioning. Interference avoidance and instrumentation amplifiers. Non-ideal op-amps. Sources of noise (including quantisation noise) and noise reduction by bandwidth limitation; Filters and their applications.

At the end of this course students should be able to:

1. An appreciation of the importance of signal conditioning for the interfacing of sensors.
2. An understanding of the key types of signal conditioning: amplification, filtering and isolation.
3. An understanding of interference, and the roles of differential and instrumentation amplifiers.
4. An appreciation of the impact of “real” operational amplifiers, and an understanding of how engineers can allow for real op-amp parameters in circuit analysis.
5. An appreciation of the origins of noise in signal conditioning circuits, and how its impact can be estimated.
6. An understanding and knowledge of basic filter types, together with their implementation.
7. An understanding of the importance of bandwidth limitation using filters (including anti-aliasing) and the lock-in amplifier.

Textbooks

A textbook that covers the majority of the topics in this course at an appropriate level is

Design with operational amplifiers and analog integrated circuits
by Sergio Franco, published by McGraw-Hill

Additional material on noise and the design of low-noise systems is contained in

Low-Noise Electronic System Design
by C.D. Motchenbacher and J.A. Connelly
published by John Wiley and Sons.

For further reading on the subject of sensors

Sensors and Transducer: Characteristics, Applications,
Instrumentation and Interfacing
by M.J. Usher and D.A. Keating,
published by MacMillan Press Ltd

or

'Instrumentation for Engineers and Scientists'
by John Turner and Martyn Hill,
published by the OUP.

AN INTRODUCTION TO SENSORS

The quantities measured by instrumentation systems are almost invariably **non-electrical**; for example, pressure, displacement, temperature, etc...

The first step in any electronic system that gathers data on this type of quantity must be to find a device that will transform a change in the physical quantity of interest into an electrical signal. This

transformation occurs in a device called a **transducer**; thus, within a platinum resistance thermometer a change in temperature is converted into a change in resistance using the temperature dependence of the resistance of the platinum wire.

A transducer may be described as an input transducer (now more usually known as a sensor) or an output transducer (now more usually known as an actuator), depending on the direction of information flow. Examples of input transducers are thermometers, microphones, pressure sensors and photodiodes; the corresponding output transducers are heaters, loudspeakers, pistons and light-emitting diodes.

There are many different types of sensor for some physical quantities (for example, temperature, strain, light flux etc...). In addition, there are other physical quantities, such as pressure and viscosity, which can only be measured if a mechanical transducer is used to convert the primary variable, such as a pressure, into a secondary mechanical variable, such as strain in a thin membrane, which can be measured. A description of the various sensors available to measure each physical quantity could be the

subject of an entire course of lectures. This type of review is not within the scope of this course, but we discuss some key issues in relation to sensors and some typical sensors.

Some key sensor parameters

A sensor's sensitivity indicates how much the sensor's output changes when the measured quantity changes. As a simple example, if a platinum resistance thermometer changes resistance by 0.4 ohm when the temperature changes by 1 °C, the sensitivity is 0.4 ohm/°C. Sensors that measure very small changes must have very high sensitivities. Sensors also have an impact on what they measure; for instance, a room temperature platinum resistance thermometer inserted into a hot liquid cools the liquid while the liquid heats the thermometer. Sensors need to be designed to have a small effect on what is measured. Making the sensor smaller often improves this and may introduce other advantages. Technological progress allows more and more sensors to be manufactured on a microscopic scale, such as microsensors using MEMS technology. In most cases, a microsensor reaches a significantly higher speed and sensitivity compared with macroscopic approaches.

A good sensor obeys the following rules:

- Is sensitive to the measured property
- Is insensitive to any other property
- Does not influence the measured property

Ideal sensors are designed to be linear. The output signal of such a sensor is linearly proportional to the value of the measured property. The sensitivity is then defined as the ratio between output signal and measured property. For example, if a sensor measures temperature and has a voltage output, the sensitivity is a constant with the unit [V/K]; this sensor is linear because the ratio is constant at all points of measurement. In general it is rather difficult to design sensors which are linear over wide ranges. For example, for the platinum resistance thermometer the resistance as a function of temperature can actually be expressed as:

$$R_T = R_0 \left[1 + AT + BT^2 + CT^3(T - 100) \right]$$

(see http://en.wikipedia.org/wiki/Resistance_thermometer)

where the quadratic and cubic terms are small, but not necessarily negligible. It might be possible to correct for this non-linearity in the signal conditioning stage. However the programmability of microcontrollers and microprocessors means that it is easier to perform this correction after analogue to digital conversion.

Sensor deviations

More generally, if the sensor is not ideal, several types of deviation can be considered:

- The sensitivity may in practice differ from the value specified. This is termed a sensitivity error, but the sensor may still be linear.
- Since the range of the output signal is always limited, either by the voltages powering any circuits or by the ADC input range, the output

signal will eventually reach a minimum or maximum when the measured property exceeds the limits – often referred to as saturation. The full-scale range defines the maximum and minimum values of the measured property.

- If the output signal is not zero when the measured property is zero, the sensor has an offset or bias. This is defined as the output of the sensor at zero input – it is very common.
- If the sensitivity is not constant over the range of the sensor, this is termed nonlinearity. Usually this is defined by the amount the output differs from ideal behaviour over the full range of the sensor.
- If deviation is caused by a rapid change of the measured property over time, there is a dynamic error. Often, this behaviour is described with a Bode plot showing sensitivity error and phase shift as function of the frequency of a periodic input signal.
- If the output signal slowly changes independent of the measured property, this is defined as drift.
- Long-term drift can indicate a slow degradation of sensor properties over a long period of time.
- Noise is a random deviation of the signal that varies in time.
- Hysteresis is an error caused when the measured property reverses direction, but there is some finite change required for the sensor to respond, creating a history-dependent offset error.
- The sensor may, to some extent, be sensitive to properties other than the property being measured. For example, most sensors are influenced by the temperature of their environment, even if that is not what they are designed to measure.

Such deviations can generally be classified as systematic errors or random errors:

- **Systematic errors** can sometimes be compensated for by means of some kind of calibration strategy. (e.g. the non-linearity of the platinum resistance thermometer mentioned above)
- Noise is a **random error** that can be reduced by signal processing, such as filtering, usually at the expense of the dynamic behaviour of the sensor. This is discussed further later.

Resolution

The resolution of a sensor is the smallest change it can detect in the quantity that it is measuring. The resolution is related to the precision with which the measurement can be made.

Sensors

The list of physical phenomena that can be measured is very long and includes:

Acoustic, sound, vibration

Chemical, Humidity

Electric current, electric potential, magnetic, radio

Flow, Pressure, force, density, level

Ionising radiation, subatomic particles

Position, angle, displacement, distance, speed, acceleration

Optical, light, imaging

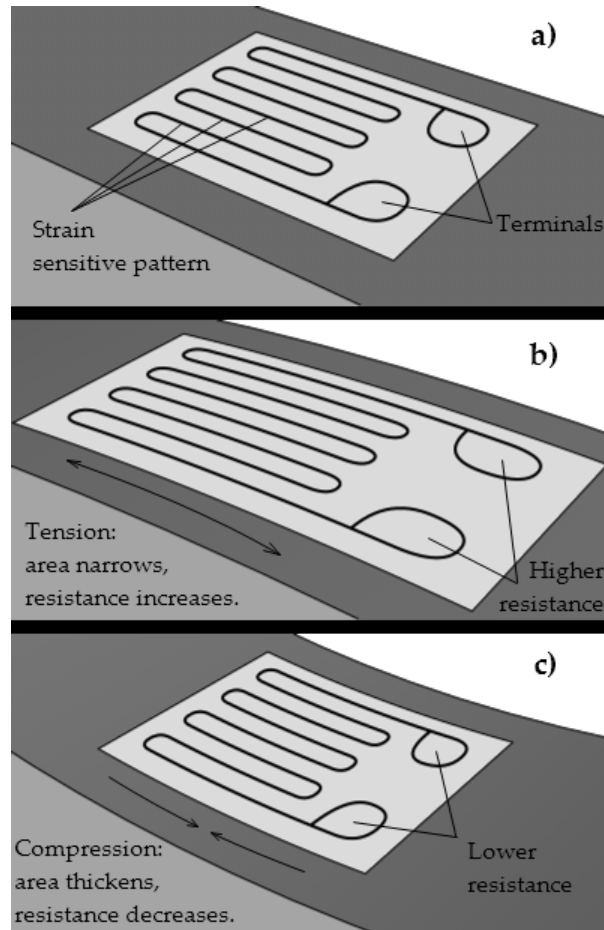
Thermal, heat, temperature

Proximity, presence

Since there are typically several different sensors per phenomenon then the list of sensors would be even longer. This is why a description of sensors would require a whole course. For this reason this course will be limited to using strain gauges to highlight the main problems that have to be considered when designing analogue signal processing circuits.

Strain gauge

Strain gauges were discussed in the P2 lectures as an example of a sensor. A strain gauge takes advantage of the physical property of electrical conductance's dependence on the conductor's geometry. When an electrical conductor is stretched (within the limits of its elasticity such that it does not break or permanently deform) it will become narrower and longer, changes that increase its electrical resistance end-to-end. Conversely, when a conductor is compressed such that it does not buckle, it will broaden and shorten, changes that decrease its electrical resistance end-to-end. From the measured electrical resistance of the strain gauge, the amount of applied stress may be inferred. A typical strain gauge arranges a long, thin conductive strip in a zig-zag pattern of parallel lines such that a small amount of stress in the direction of the orientation of the parallel lines results in a multiplicatively larger strain over the effective length of the conductor—and hence a multiplicatively larger change in resistance—than would be observed with a single straight-line conductive wire.



The so-called gauge factor is defined as:

$$G_F = \frac{\Delta R / R_G}{\varepsilon}$$

where R_G is the initial resistance of the strain gauge
 ΔR is the change in resistance when strain is applied
 ε is the applied strain.

$$\Delta R / R_G = G_F \varepsilon$$

Typically a gauge factors are ~2, which means that the fractional change in resistance is only twice as large as the strain. This means that the fractional changes in resistance are very small and they can only cause small changes in any electrical signals.

Signal Conditioning

The output from a transducer is generally a continuously varying or **analogue** signal. In contrast digital processors store and process signals sampled at particular times and represented as binary numbers. **Any modern instrumentation must therefore include a component** (described in the P2 course), **known as an Analogue-to-Digital converter (ADC), which converts the analogue input signal into a digital signal that can be read by the digital processor.**

The simplest and cheapest possible instrumentation system is one in which the output from the transducer is connected directly to the input of an ADC. However, both the transducer and the ADC are standard components that have not been designed for any particular application. More importantly, transducers rely upon physical processes that rarely, if ever, generate output signals that are compatible with the ADC input range. In particular **the maximum change in the output signal from a sensor is often smaller than the minimum change in signal that can be detected by the ADC.** This means that in this simplest system even the maximum change in the transducer output may be undetectable.

Instrumentation systems must therefore include a circuit before the ADC that amplifies the output from the transducer to make it detectable by the ADC. Such a circuit is referred to as **signal conditioning.**

In instrumentation, signal conditioning generally means manipulating an analogue signal (from a sensor) in such a way that it meets the requirements of the next stage of a system for further processing. In general the most common “next stage” will involve analogue-to-digital converters.

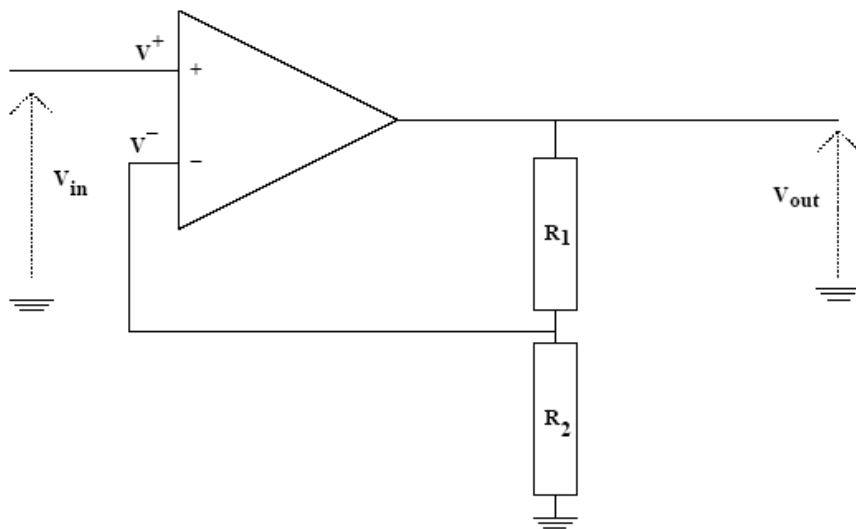
Signal inputs accepted by signal conditioning circuits include DC voltage and current, AC voltage and current (and possibly but rarely electric charge). The processes that are performed by these circuits will almost certainly include amplification and filtering. In addition to these functions the signal conditioning may also include a step to isolate the input circuits from the rest of the system (one area where this is important is in medical electronics where isolation protects the patient who is hosting the sensors) and/or a non-linear step (such as a log or an anti-log amplifier) to compensate for any non-linearity of the sensor.

Although isolation and non-linear stages are key to a few systems this course will focus on the amplification and filtering that must be included in almost all systems. Commonly used amplifiers for signal conditioning.

Simple Amplifier Circuits for Signal Conditioning

The output from many transducers is a voltage and there are two simple circuits (described in the P2 course), the non-inverting and the inverting amplifier, which can be used to amplify a voltage signal.

Non-inverting amplifier (P2 Revision):



A non-inverting amplifier circuit.

Don't forget the analysis of any op-amp circuit to understand its function is based on the simple ideal op-amp rules (taken from P2):

- **An ideal op-amp has an infinite input resistance, an infinite differential gain and an output resistance of zero.**

The infinite gain means that provided there is negative feedback:

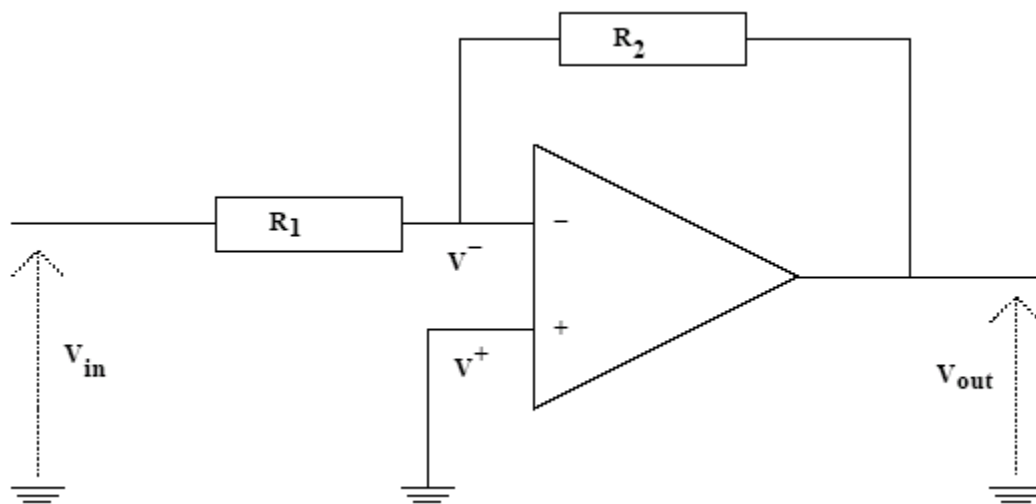
$$V^+ = V^- \quad (\text{be sure you understand WHY})$$

This equation and the characteristics of an ideal op-amp can be used to show that for the non-inverting amplifier

$$V_{out} = \left(1 + \frac{R_1}{R_2}\right)V_{in}$$

where the gain is given by the terms within the brackets.

Inverting amplifier (P2 Revision):



An inverting amplifier circuit.

Using the same rules as above, and noting that for the inverting amplifier V^+ is connected to ground (0V) we can show that:

$$V_{out} = \left\{ -\frac{R_2}{R_1} \right\} V_{in}$$

Mini-Summary

Instrumentation systems are widely used to control and monitor many different “host” systems.

Any physical variable that is being measured, has to be converted to an electrical signal, usually an analogue signal, by an input transducer or sensor. Useful information then has to be extracted from this signal, most often by a programme in a digital processor. Conversion from an analogue sensor output to a digital input for the processor is a two-stage process involving analogue signal processing (conditioning) and an analogue-to-digital converter.

Each instrumentation system therefore usually consists of four constituent parts, the sensor, analogue signal processing circuits, an analogue-to-digital converter and a digital processor.

Sensors rely upon physical processes that allow an electrical signal to be generated in response to a change in a physical variable. These physical processes usually result in small output signals. A key part of any analogue signal processing circuit is therefore a circuit that amplifies the changes in the output signal from a sensor.

INTERFERENCE AND INSTRUMENTATION AMPLIFIERS

Introduction

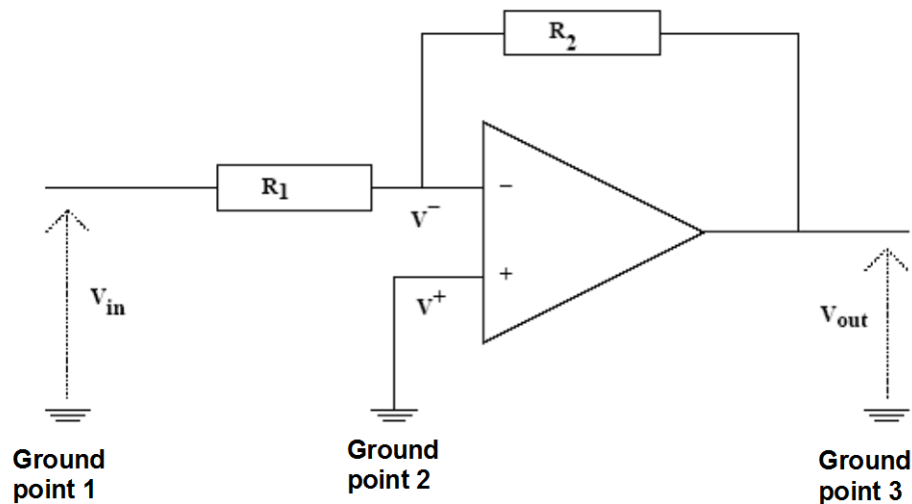
A lot of this material was also covered in the P2 course, but it is important, so it will be included here.

Inverting and the non-inverting amplifier circuits share a common problem - they both amplify the difference between the input voltage signal and the amplifier 'ground' connection. Any variations in the 'ground' voltage will be indistinguishable from changes in the sensor output voltage. The resulting interference can be minimised by carefully designing the analogue signal processing circuit to avoid shared ground connections and coupling to electromagnetic radiation. In addition, whenever possible the sensor should be included within a circuit that produces an output that is the difference between two voltages, a type of output known as a differential output, with the largest possible amplitude.

A differential output can be amplified using a simple differential amplifier circuit. However, variations between the actual and nominal values of the resistors in this circuit will create a response to changes in the average (common-mode) input signal. To avoid the problems that this causes the differential output from a transducer is usually amplified using a three op-amp instrumentation amplifier.

Illustration (see also P2 notes):

In the analysis of the inverting op-amp circuit presented above you will notice that there are three points which are shown as ground. Now, this can present a problem:



An inverting amplifier circuit (showing ground points)

How can we guarantee that these three “ground points” are in fact all at identical potential? In fact it was an assumption in the analysis, and if the assumption is not correct then the analysis changes. For example, if the V^+ input is not at zero-volts then the $V^+=0$ assumption must be dropped, and the output voltage becomes:

$$V_{out} = -\frac{R_2}{R_1}(V_{in} - V^+) + V^+ = -\frac{R_2}{R_1}V_{in} + \left(1 + \frac{R_2}{R_1}\right)V^+$$

It is clear from this that we would not be able to distinguish between sensor signal (V_{in}) and ground-point errors. This is a real problem!

There are at least two origins of such problems. In the P2 course the issue of careful grounding was discussed, illustrated by the following diagram:

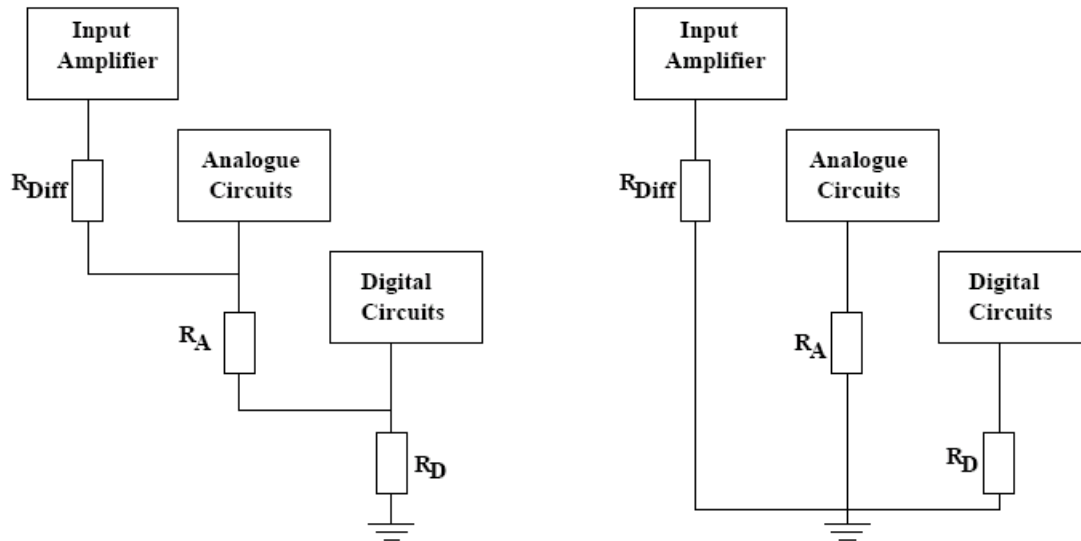


Figure (6) Schematic diagrams of a poor earthing scheme, on the left, and a good earthing scheme, on the right.

In Figure (6) the ground connections of three parts of the system on the left are connected to one another (“chained” together) within the overall circuit before a single connection is made to earth. **A hidden danger with this simple grounding scheme is that each connection has a small but finite resistance.** These resistances are represented by R_{Diff} , R_A and R_D in Figure (6), which shows that this grounding scheme creates common resistances to ground. **The problem caused by these common resistances is that the current flowing to ground through one circuit can change the ground potential of other circuits. As illustrated for the simple inverting op-amp circuit above, it is then impossible to tell the difference between real signals and such changes in ground potential.** Since the shared resistances created by this grounding scheme are formed by contacts, tracks and wires that have an ideal resistance of zero, it is understandable to think that this effect is negligible. However, the

total resistance that is shared could be a few Ohms. If we approximate this to say 10 Ohms then it would only require a current of $50\mu\text{A}$ to flow through this shared resistance to change the 'ground' potential of the input amplifier by 0.5mV . If the subsequent signal conditioning circuits have an overall gain of say 1000 this would cause an output voltage “error” of 0.5V . It is therefore quite possible for changes in the current drawn by one part of the circuit to create a fluctuation in the 'ground' potential of another part of the circuit that will be (incorrectly) interpreted as a significant signal.

In the P2 course it was suggested that this interference **can be avoided by using a single-point grounding scheme, often referred to as a star connection or star grounding scheme, shown on the right hand side of figure (6).**

However, there is a second problem...

Electromagnetic interference

All of the connecting wires act as aerials which can “receive” any electromagnetic fields in the environment. (This is especially so if there are any multiple ground paths, which can form beautiful loop aerials!) This can then result in significant time-varying currents flowing through the ground connections that could interfere with the circuit ground. Within small-scale circuits such effects can generally be avoided at low frequencies. However, in connecting sensors to signal conditioning circuits such electromagnetic interference on the cables can be problematic.

Bridge Circuits

Careful circuit layout reduces the impact of interference arising from undesirable signals on the earth connection of the input amplifier. However, it is better to use a design that prevents the problem by avoiding the need to rely upon a good ground connection. This can often be achieved by placing the sensor (such as a strain gauge) in a bridge arrangement, see figure (7), which generates both a signal voltage and a reference voltage, **giving a differential voltage**.

In figure (7), the transducer R_1 is subjected to the “influence” to be measured whilst R_2 , R_3 and R_4 are reference resistances subjected to the same conditions as R_1 , except for the influence.

Let ΔV be the bridge differential output voltage when the transducer “senses”, such that the sensor resistance changes from R_0 to $R_0 (1+\alpha)$, then in this particular configuration, the currents in the potential dividers on the left and right are:

$$I_{left} = \frac{V}{2R_0} \quad \text{and} \quad I_{right} = \frac{V}{2R_0 + \alpha R_0}$$

so that

$$\Delta V = (I_{left} - I_{right})R_0 = \frac{V}{2} - \frac{V}{2 + \alpha}$$

hence

$$\Delta V = \frac{\alpha V}{4(1 + \frac{\alpha}{2})}$$

For small values of α this expression shows that the output voltage is proportional to α and does not include a large constant term. This means that none of the limited input range of the ADC is wasted representing a large, constant voltage which contains no information.

Rather with this circuit the entire input range of subsequent circuits can be used to represent the useful signal.

(However, it should be noted that neither of the outputs are grounded, and this must be kept in mind in designing the next stage of the system.)

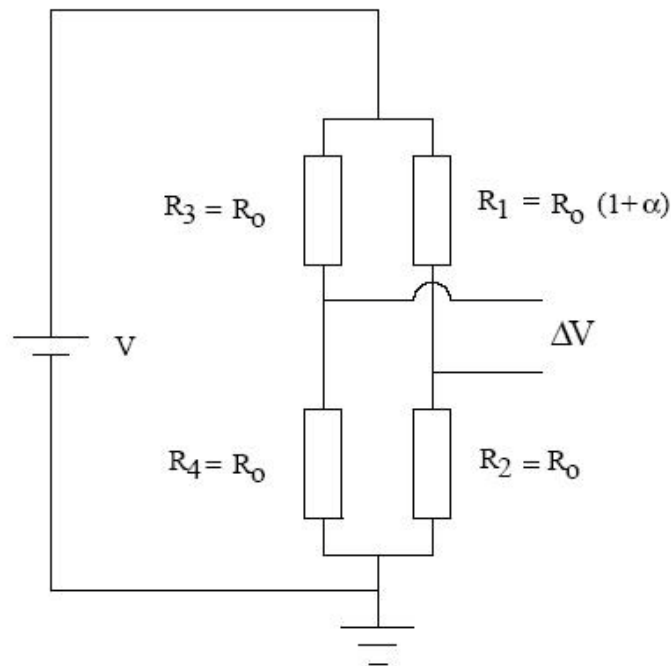


Figure 7: A sensor bridge circuit with a differential output.

Cables as a source of interference

The bridge circuit with a sensor avoids interference from signals on the ground connection of the instrumentation system. However, **the output signals are still small and are therefore vulnerable to electromagnetic interference. With a pair of wires connecting the bridge circuit to the instrumentation system it is possible for interference to be caused by fields which couple to the connecting wires.** There are particular effects with a connection formed by two wires.

- **The first is that the pair of wires can form a loop aerial that couples to any stray changing magnetic field. This can then generate an emf (i.e. voltage) around the loop.**
- **The second is that if the two wires run parallel, but, separately they will have different coupling capacitances to other conductors in the local environment, particularly the local mains power leads.**

To minimise the impact of both these phenomena the two wires are typically twisted around each other to form a twisted pair. This cheap, and neat, solution ensures that the cross-sectional area of any loop formed by the wire is minimal and that both wires have the same coupling capacitance to any other conductor. A twisted pair is therefore often used as a cheap but effective method of connecting two separate parts of an instrumentation system.

The full bridge circuit

The use of a twisted pair is effective, and represents a good solution to a problem. A better solution would be to avoid the problem altogether. **It is sometimes possible to minimise the impact of interference by adopting a good general design principle. This is to create the largest possible signal at the earliest opportunity.**

For example, the output signal from a bridge circuit can sometimes be increased by using the full bridge circuit with four sensors, as shown in **Figure (8)**. In this circuit, the sensors are arranged so that each branch of the circuit contains a pair of sensors acting in complement (for example one strain gauge in tension and the other in compression). In this way, the output signal from the bridge circuit is enhanced by a factor of 4.

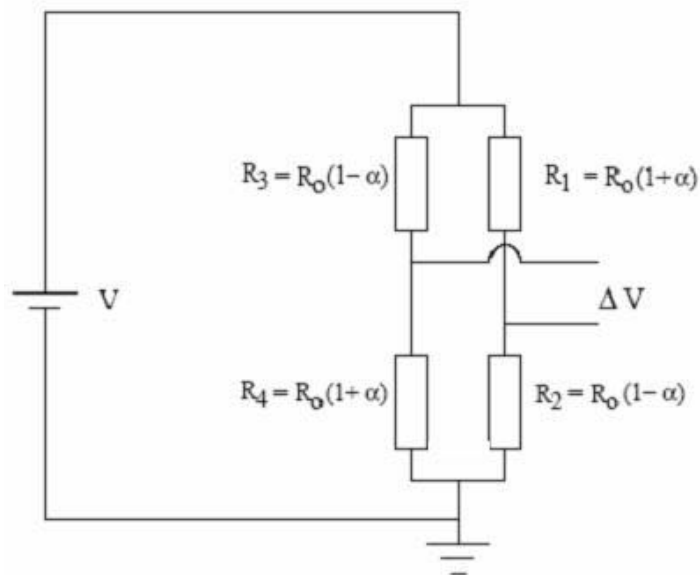


Figure 8: A full bridge circuit containing four strain gauges.

The differential amplifier

The output from the bridge circuit, and many other types of transducer, is the difference between two voltages. These differential voltages will be smaller than the input range of a typical ADC. A “signal conditioning” circuit is therefore required that amplifies a **differential** voltage, rather than a voltage relative to a ground point. i.e. given two wires coming from a sensor (or sensor in a bridge circuit) the voltage levels relative to a ground-point may be not well defined (due to electromagnetic interference, difficulty of “defining” the ground voltage, etc.). However, the voltage difference should be representative of the quantity being sensed.

A simple op-amp circuit that can amplify a differential voltage is shown in Figure (9). **This was analysed in the P2 course.**

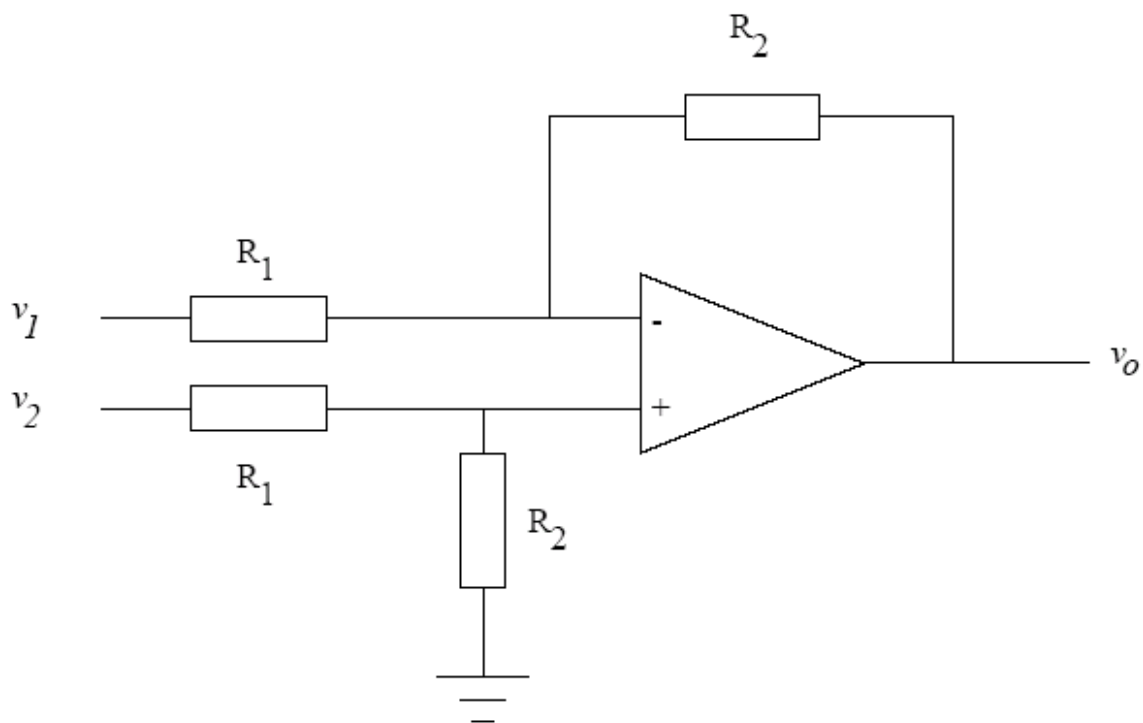


Figure 9: Single op-amp differential amplifier

Assume the op-amp is ideal and considering the connections to the op-amp inverting input:

$$\frac{v_o - v^-}{R_2} = \frac{v^- - v_1}{R_1}$$

The second input voltage is connected to the non-inverting input through a simple potential divider, so for this we can write:

$$v^+ = \frac{R_2}{R_1 + R_2} \cdot v_2$$

The “normal” ideal op-amp rules, together with negative feedback allow us, as usual, to equate the two op-amp inputs, writing $v^- = v^+$. Using this to eliminate v^+ and v^- in the above equations gives us:

$$\underline{v_o = \frac{R_2}{R_1} (v_2 - v_1)}$$

The magnitude of the gain is the same as we saw for the inverting amplifier earlier, but now it is the difference between the voltages $v_2 - v_1$ which is amplified, which is just what we need!

For signal conditioning applications a particularly useful aspect of this amplifier is that the output is dependent only on this difference and is independent of their absolute levels. This is important, so it is worth thinking about why this happens.

Consider the situation when the two inputs are the same, let's say $v_1=v_2=v$. Now the potential divider rule for the non-inverting input remains the same as before, and still we need $v^- = v^+$, so that means that

$$v^- = v^+ = \frac{R_2}{R_1 + R_2} v$$

But, the v^- input is “fed” by a potential divider between v_1 and v_o , and given that v_1 is also v the op-amp conditions can only be satisfied if v_o is also zero. This means that **the output depends only on the difference between the inputs.**

This type of differential behaviour is important, so we develop “standard” ways of expressing it. Firstly, we write the inputs in a different way. Rather than expressing the inputs in terms of the two voltages v_1 and v_2 we express the inputs in terms of the average and difference of these voltages. These are referred to as the “common mode” (cm) and “differential” (diff) voltages, expressed as:

$$v_{cm} = \frac{v_2 + v_1}{2}$$

$$v_{diff} = v_2 - v_1$$

REMEMBER: these are just a pair of simultaneous equations, so any pair of voltages can be expressed in terms of either their individual voltages or in terms of their common-mode and differential terms. We can re-arrange between the notations using the above expressions for the common-mode and differential voltages in terms of the input voltages, or their complement:

$$v_1 = v_{cm} - \frac{v_{diff}}{2}$$

$$v_2 = v_{cm} + \frac{v_{diff}}{2}$$

CAUTION: although there is no ambiguity in the definition of the common-mode term there is a sign ambiguity in the differential term, the sign of which will depend on the labelling of the two individual voltages – be careful!

Now, going back to the gain equation for the differential amplifier determined above:

$$v_o = \frac{R_2}{R_1} (v_2 - v_1)$$

and substituting for v_1 and v_2 in terms of the common-mode and differential voltages now makes it obvious that

$$v_o = \frac{R_2}{R_1} v_{diff}$$

which is independent of the common-mode term and has a gain for the differential term of R_2/R_1 . This differential gain is normally designated as A_{diff} .

More generally the output voltage from a circuit intended for differential amplification can then be written in the form:

$$v_o = A_{diff} v_{diff} + A_{cm} v_{cm}$$

or

$$v_o = A_{diff} \cdot (v_2 - v_1) + A_{cm} (v_2 + v_1)/2$$

where A_{diff} is the differential gain of the amplifier and A_{cm} is the so-called common mode gain. For the ideal differential amplifier

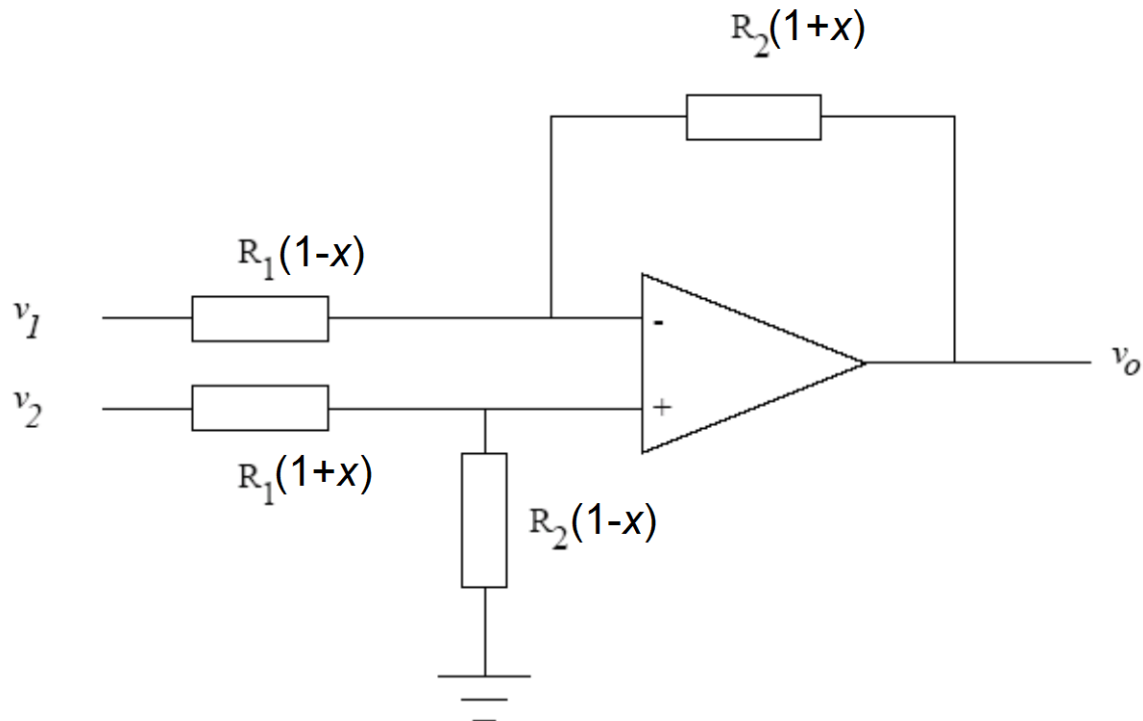
illustrated above the common mode gain is zero, $A_{cm} = 0$, so that the

output is independent of the common mode signal. The differential gain

A_{diff} is obviously (from the above analysis) just $A_{diff} = R_2 / R_1$

Why the circuit might not be non-ideal

The above analysis is an ideal situation, but of course in the real world the differential amplifier might not be so ideal. There are a number of reasons for this, but most commonly it is because the two resistors labelled R_1 will not be identical to one another and similarly the two resistors labelled R_2 will not be identical to one another. You may recall that you analysed this issue as a tutorial problem for P2! In this analysis you assumed a small error in the resistors, so that the circuit became something like:



where x is a small error ($\ll 1$) and is chosen to represent a “worst case”. Assuming the normal op-amp rules, and after a bit of algebra, in the P2 tutorial you came up with an equation for the output voltage of the form:

$$v_o = \left(1 + \frac{R_2(1+x)}{R_1(1-x)} \right) \frac{R_2(1-x)}{R_2(1-x) + R_1(1+x)} v_2 - \frac{R_2(1+x)}{R_1(1-x)} v_1$$

Now, expanding this for small x (i.e. ignoring terms in x^2 or higher order) this can be written:

$$v_o = \frac{R_2}{R_1} \left(1 + 2x \frac{R_2 - R_1}{R_1 + R_2} \right) v_2 - \frac{R_2}{R_1} (1 + 2x) v_1$$

then substituting for v_1 and v_2 in terms of v_{cm} and v_{diff} , and again assuming x is small we have:

$$v_o = \frac{R_2}{R_1} v_{diff} + \frac{R_2}{R_1} \left(\frac{-4xR_1}{R_1 + R_2} \right) v_{cm}$$

so now there is an extra term in addition to the differential gain found above. The differential gain remains the same ($A_{diff} = R_2 / R_1$), but now there is also a common mode gain:

$$A_{cm} = \frac{R_2}{R_1} \left(\frac{-4xR_1}{R_1 + R_2} \right)$$

(which, as expected, obviously vanishes in the limit of perfectly matched resistors, i.e. when $x=0$).

In general we would “like” a differential amplifier to have a small common mode gain in comparison with the differential gain. This allows it to “reject” common-mode signals and amplify the differences. **The ability of a differential amplifier to reject any common mode signal whilst amplifying the differential signal, is usually characterised by the ratio between the differential and common mode gains. It is called the common-mode rejection ratio (CMRR)**

$$CMRR = \frac{A_{diff}}{A_{cm}}$$

This means that the common mode rejection ratio of an ideal circuit is infinity. Despite non-ideal effects the CMRR of a good differential amplifier is usually large and it is therefore often quoted in decibels

$$CMRR = 20 \log(A_{diff}/A_{cm})$$

Input impedance

When using the basic differential amplifier illustrated above there is another problem, even if the circuit is carefully engineered with high tolerance resistors to maximise the CMRR. The two op-amp inputs are at the same voltage as one another (due to the negative feedback), so any differential input voltage “sees” the two resistors labelled as “ R_1 ”. The differential input impedance is therefore $2R_1$. You will recall from P2 that in general when connecting a signal source (in our case from a sensor/transducer) to an amplifier it is good to ensure that the input impedance of the amplifier is

large compared with the source output impedance. However, this cannot be guaranteed here. It would therefore be useful to arrange to INCREASE the input resistance of the differential amplifier.

As discussed in the P2 course the input impedance of the circuit can be increased by simply inserting an op-amp buffer circuit on each input of the differential amplifier as shown in figure (10). The source (sensor) output (i.e. the input to the circuit) is then directly connected to one input of an op-amp that has a very high, if not infinite, input impedance. Provided the op-amps are ideal then in the arrangement shown in figure 10 we have:

$$v_1 = v'_1 \text{ and } v_2 = v'_2,$$

which means that the signals from the sensor are applied to the differential amplifier as before. Adding op-amp buffers to the standard differential amplifier therefore simply increases its input impedance.

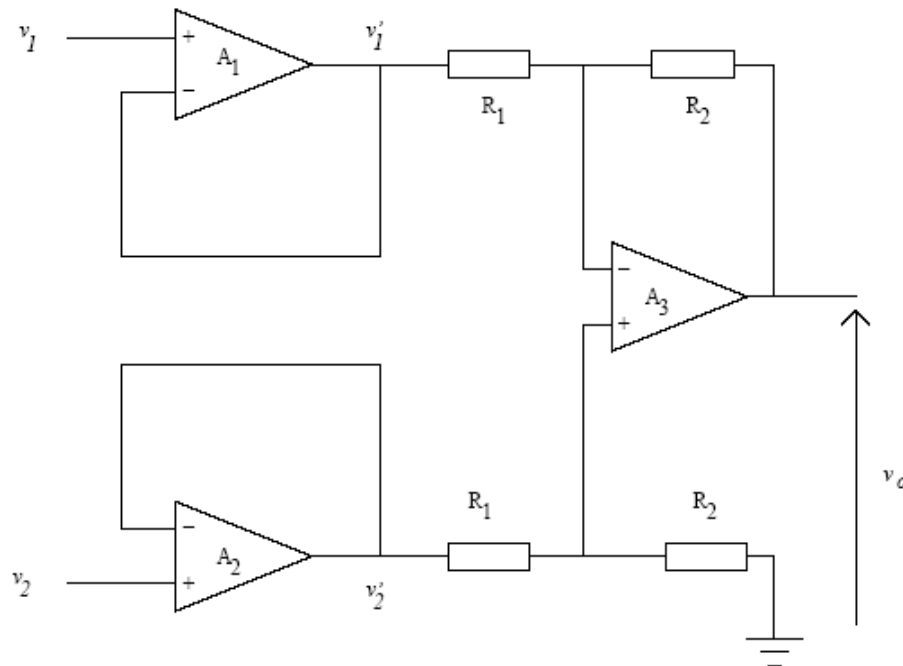


Figure 10: A differential amplifier with buffered inputs.

Instrumentation amplifiers

In principle we could add further gain to the input stages (buffers) by using the circuit for the basic non-inverting op-amp amplifier presented earlier. However, this is unwise because differences between the nominally identical resistors in the two non-inverting amplifiers will create different gains on the two input signals. It turns out we can do a bit better than the circuit in Figure (10) if we connect the non-inverting inputs through a common resistor, forming the **standard 3 op-amp instrumentation amplifier** shown in Figure (11).

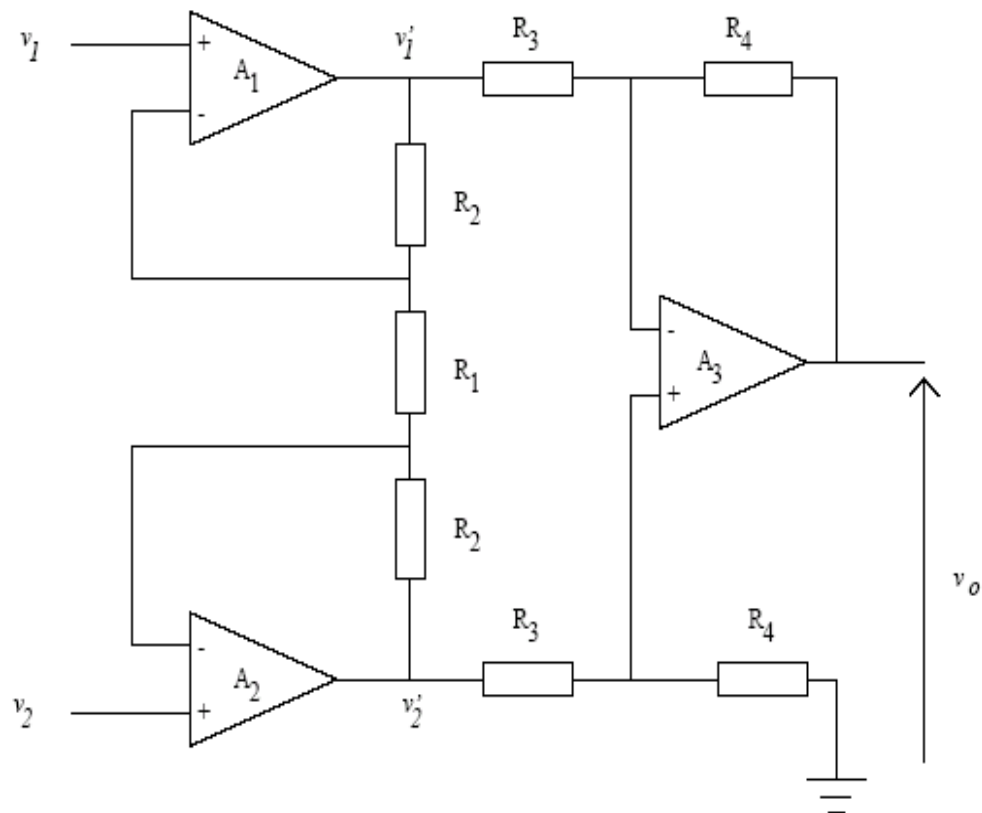


Figure 11: The standard 3 op-amp instrumentation amplifier.

The analysis of this circuit was presented in the P2 lectures, and is repeated here: To understand the circuit first consider the current flowing vertically through the resistors connecting v'_1 to v'_2

$$i = \frac{v'_1 - v_1}{R_2} = \frac{v_1 - v_2}{R_1} = \frac{v_2 - v'_2}{R_2}$$

the middle two equations can be re-arranged to give

$$v'_1 = \left(1 + \frac{R_2}{R_1}\right)v_1 - \frac{R_2}{R_1}v_2$$

and the right hand pair give

$$v'_2 = \left(1 + \frac{R_2}{R_1}\right)v_2 - \frac{R_2}{R_1}v_1$$

then subtracting the first of these two equations from the second gives

$$v'_2 - v'_1 = (v_2 - v_1)\left(1 + \frac{2R_2}{R_1}\right)$$

So, the differential output from the first pair of op-amps (referred to as the first stage) is equal to the differential input multiplied by a differential gain factor, let's say A_{diff_1}

$$A_{diff_1} = 1 + \frac{2R_2}{R_1}$$

Adding the equations for v'_1 **and** v'_2 together (and dividing by two) we also have:

$$\frac{v'_1 + v'_2}{2} = \frac{v_1 + v_2}{2}$$

i.e. the common mode term in the output from the first stage is identical to the common mode term in the input to the first stage, so the common-mode

gain is unity. **Thus the first stage can provide a differential gain and unity common-mode gain, thereby increasing the CMRR as well as the input resistance.**

Additionally, analysis shows that the CMRR of the first stage is not substantially damaged by small errors in the resistor values.

Normally the R_2 resistors are part of the package and the single R_1 resistor is chosen by the user.

The output from this first stage is fed to a differential amplifier circuit which is effectively identical to that considered above. This last stage is often given a fixed low differential gain (unity, for example) by the manufacturer. The advantage of this is that all the resistors within this part of the circuit can be integrated within the instrumentation amplifier package. Matching to the required accuracy is then achieved as part of the manufacturing process. With a low differential gain the main function of the differential amplifier is to provide a single output whilst rejecting any common-mode input signal.

The overall CMRR of the instrumentation amplifier can be written in a number of ways. Given the way that we have expressed the differential and common-mode gains we can now gather things together. For the first stage:

$$v'_2 - v'_1 = (v_2 - v_1) \left(1 + \frac{2R_2}{R_1} \right)$$

and

$$\frac{v'_1 + v'_2}{2} = \frac{v_1 + v_2}{2}$$

or

$$v'_{diff} = v_{diff} A_{diff_1}$$

and

$$v'_{cm} = v_{cm}$$

For the last stage:

$$v_o = \frac{R_4}{R_3} (v'_2 - v'_1) + \frac{R_4}{R_3} \left(\frac{-4xR_3}{R_3 + R_4} \right) \left(\frac{v'_2 + v'_1}{2} \right)$$

or

$$v_o = A_{diff_laststage} v'_{diff} + A_{cm_laststage} v'_{cm}$$

Hence:

$$v_o = A_{diff_laststage} A_{diff_1} v_{diff} + A_{cm_laststage} v_{cm}$$

So the overall differential gain is:

$$A_{diff_laststage} A_{diff_1}$$

and the overall common-mode gain is:

$$A_{cm_laststage}$$

Generally the CMRR is therefore:

$$CMRR = \frac{A_{diff_laststage}}{A_{cm_laststage}} \times A_{diff_1}$$

If the first stage does also have common mode gain then this generalises to:

$$CMRR = \frac{A_{diff_laststage}}{A_{cm_laststage}} \times \frac{A_{diff_1}}{A_{cm_1}}$$

In both cases this can also be expressed as:

$$CMRR = CMRR_{firststage} \times CMRR_{laststage}$$

Mini Summary

Amplification is required in the vast majority of systems to match the maximum output signal from the sensor to the maximum input signal of an ADC.

Simple op-amp amplifiers with a single input amplify the difference between their input signal and the local ground voltage. Since this might be different from the ground voltage at the sensor this type of amplifier is **vulnerable to interference caused by fluctuations in the ground potential.** When signals are small careful design is therefore required to prevent interference from sources, including electromagnetic radiation and current flowing through connections shared by different circuits. **Techniques including shielding, good grounding and twisted pairs can significantly reduce interference.**

One approach to creating a system that is robust to interference is to use a sensitive transducer to create a large signal as early as possible. One problem that may arise with sensitive transducers is that they can be sensitive to more than one physical effect, for example a strain gauge may be sensitive to temperature. **In some situations it is possible to design a circuit that distinguishes between changes in two physical quantities that both affect the output of a sensor.** In other situations it may become necessary to measure both quantities and then to use the digital processor to correct for the undesirable dependence.

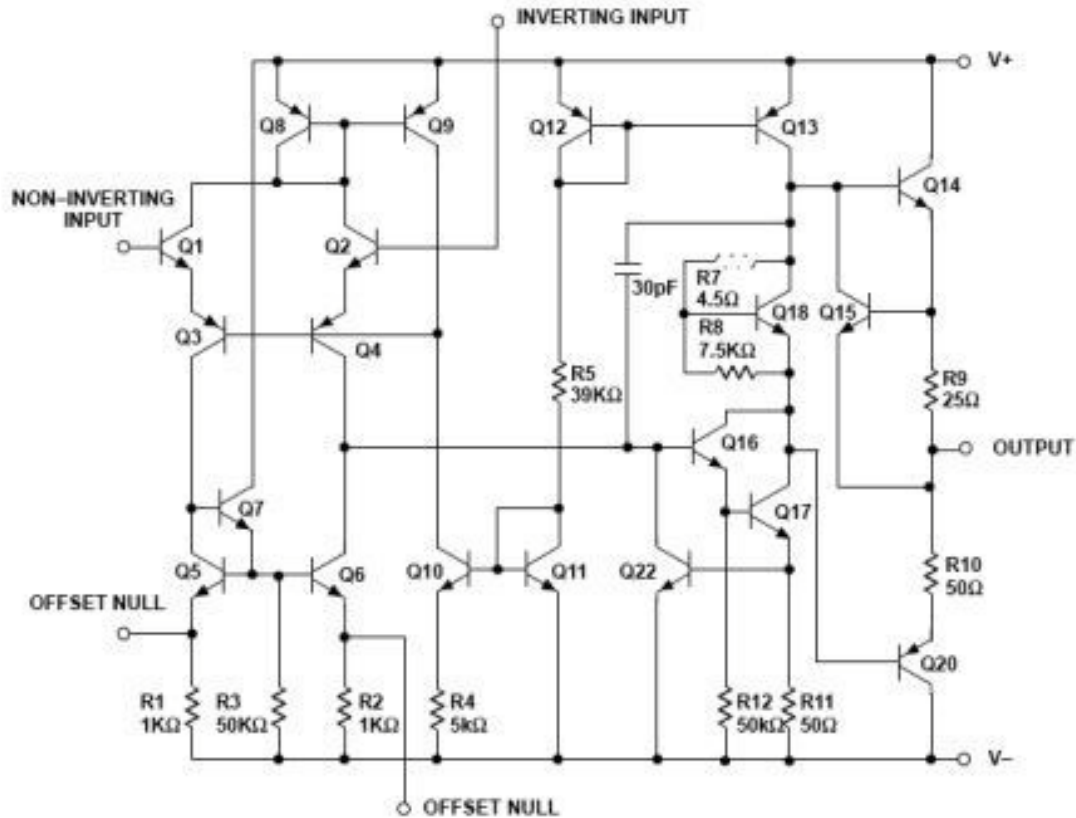
Another approach to creating a system that is robust to interference, that has the added advantage of creating a signal that only represents the change induced by the physical effect of interest, is to use **a sensor with a differential output. This type of sensor must then be connected to a differential amplifier.** A simple differential amplifier can be created from a single op-amp and two pairs of identical resistors. However, variations between the resistances of nominally identical components leads to circuits which respond to the average input (common mode) signal.

The problems caused by a finite common-mode gain are usually reduced by using a three op-amp instrumentation amplifier. This circuit has a high input impedance, a high common-mode rejection ratio and an easily set differential gain. An additional attractive feature is that it is very easy to use!

REAL OP-AMPS

Introduction (– for information only)

The inside of a real op-amp (for example, the 741) is rather complicated:



This makes the full analysis of a signal conditioning circuit built using op-amps potentially rather difficult, as the op-amp may not be as “ideal” as initially assumed!

There are two stages in analysing any real op-amp circuit. Initially, the function of the circuit can be determined assuming that the op-amp is ideal. Subsequently, it may be necessary to perform a more detailed analysis of any circuit, including the non-ideal behaviour of the op-amp, in order to quantify the performance of the circuit more precisely. Alternatively, this type of detailed analysis is needed to ensure that an op-amp is chosen which appears to be ideal in the context of a particular circuit.

Behaviour of Real Op-amps

For a real op-amp the gain and input impedance are large and the output impedance is small, however, they are all finite. A model of the op-amp that includes these effects is shown in figure (44). This model of the op-amp can be included in the analysis of a particular circuit to ensure that the non-ideal behaviour of the op-amp selected for a particular application has a negligible effect on the circuit performance.

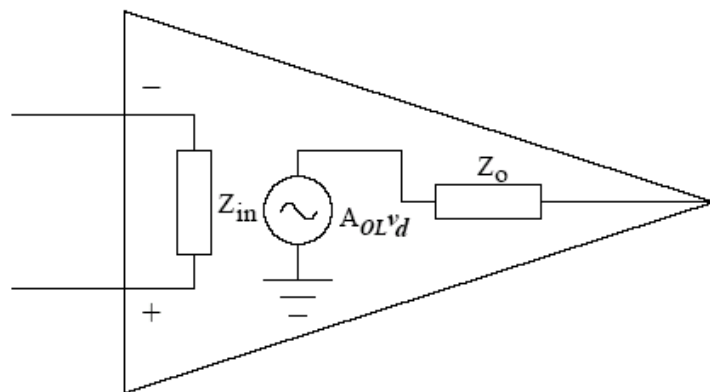


Figure 44: A model for an op-amp that includes its finite input impedance, output impedance and gain.

In addition, to the finite gain and impedances of the op-amp there are a few other non-ideal aspects of the behaviour of the op-amp which must be taken into account when either designing/analysing a circuit or selecting an op-amp.

Limited Output Voltage Range – Power is supplied to the op-amp via two power supply connections, not shown on most circuit diagrams. **The output voltage of the op-amp is limited by the supply voltages applied to these pins.** Thus if an op-amp has a positive supply voltage of +5V and a negative supply voltage of -5V the output voltage will be limited to the range between +5V and -5V. In fact the actual output voltage range is likely to be less than the supply voltage range. If the input conditions require an output voltage outside the allowed range then the output voltage will saturate to a maximum positive or minimum negative value. If this range is too small for a particular circuit then an alternative op-amp should be used which has a larger maximum supply voltage and hence output voltage range.

Input Offset Voltages – The two input signals are connected to two different input transistors within the op-amp. Ideally, these two transistors are identical. However, variations in the manufacturing process mean that this ideal condition is rarely achievable, which means that for **a real op-amp a small differential input voltage is required to create an output of 0 V.** This input offset voltage can be accommodated by one of two **alternative techniques.** In some op-amps extra connections are provided to allow the user to add a variable resistance that can then be used by the user to zero the input offset voltage. This **trimming** is the least expensive technique of compensating for the offset voltage, however, it

is inconvenient. For example the resistance may need to be adjusted if the operating temperature changes. **The more usual solution to this problem is therefore for the op-amp manufacturer to include an equivalent resistance within the op-amp package that is adjusted as part of the manufacturing process.** The additional manufacturing processes required to trim a circuit means that these components are more expensive. However, they are more ideal when received by the user and they are therefore very popular.

Common Mode Rejection

(In the material above the issue of differential and common mode gains for an op-amp based circuit were considered. Here the situation for the op-amp itself which is considered.)

The two input signals to the op-amp can be considered to consist of an average, or common-mode, component v_{cm} and a differential component v_{diff} so that

$$v^+ = v_{cm} + v_{diff}/2$$

$$v^- = v_{cm} - v_{diff}/2$$

The ideal op-amp will only respond to the differential part of these two input signals. However, a **real op-amp will respond to the common-mode signal with a common-mode gain A_{cm} .** By convention this aspect of the performance of an op-amp, and other circuits, is characterised by the

logarithm of the ratio of the differential gain A_{diff} to the common mode gain A_{cm} . This **common-mode rejection ratio** CMRR is

$$CMRR = 20 \log_{10} [A_{diff}/A_{cm}]$$

Many important op-amp circuits are based upon a constant bias, usually 0V, applied to v^+ and a feedback loop connected to v^- . The operation of these circuits is based upon a high differential gain this ensures that $v^+ = v^-$, and since v^+ is a constant this means that v_{cm} is constant. For these circuits the effect of a finite common-mode gain is therefore negligible. However, in instrumentation amplifiers the input voltages to the op-amps will vary with the common-mode signal. For these circuits the CMRR of the op-amp can have a significant impact on performance.

Input Bias Current - Each input to some op-amps is connected to the base of one of a pair of bipolar transistors. The base current for each device flows through the op-amp inputs. These two dc currents are referred to as the **input bias currents**. A technique to compensate for the presence of these currents will be described later.

Noise - Finally, op-amps contain several devices, each of which will generate **noise**. This aspect of the behaviour of real op-amps will be described later.

Real Op-amps in an Inverting Amplifier

In order to understand how the finite gain and impedances of an op-amp can be included in the analysis of a circuit we consider an example, the inverting amplifier shown in figure (45).

Some definitions before starting the analysis:

open-loop voltage gain $A_{OL} = v_o/v_d$ (measured without feedback)

closed-loop voltage gain $A_{CL} = v_o/v_{in}$ (gain with feedback)

the feedback fraction $\beta = v_f/v_o = R_1/(R_1 + R_2)$

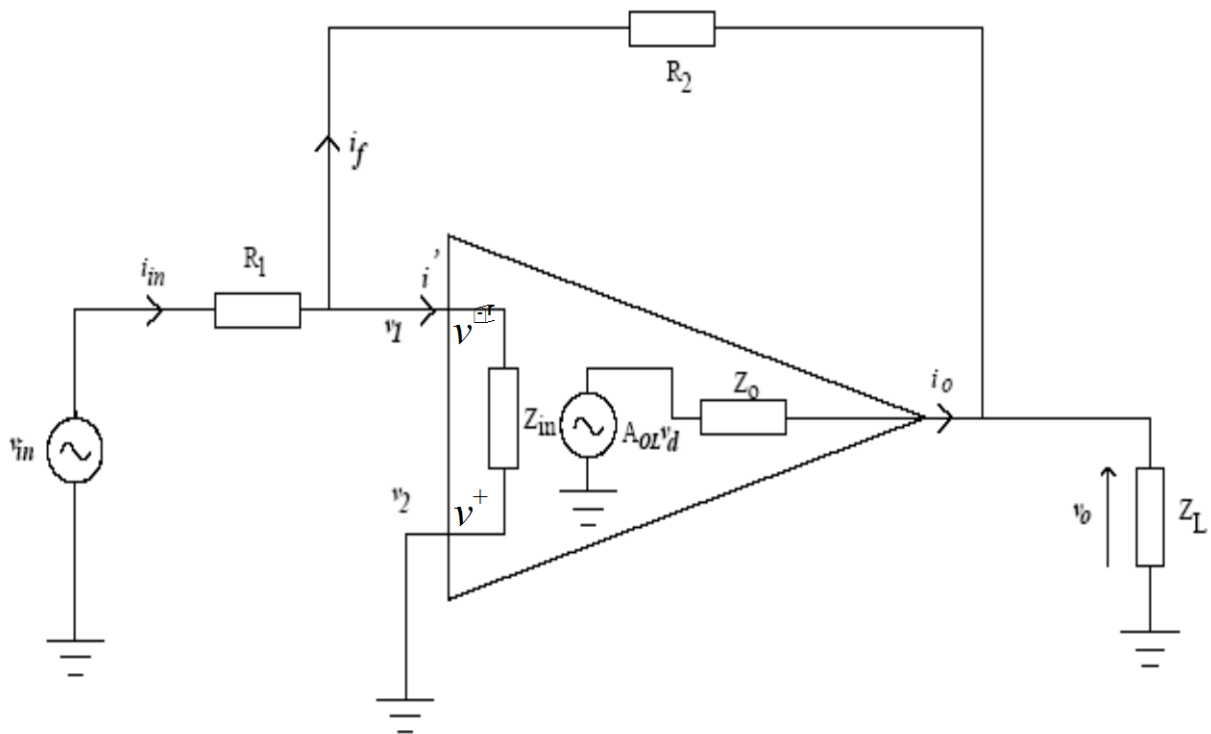


Figure 45: Inverting amplifier circuit including non-ideal op-amp model.

The voltage at the inverting input to the op-amp can be determined by linear superposition of the voltages created by the two voltage sources within the circuit. To simplify the analysis assume that the op-amp is almost ideal so that the input impedance is large and the output impedance is small. In particular, for this circuit assume that $Z_{in} \gg R_1$ and that $Z_o \ll R_2$.

First to determine the effect of the op-amp output voltage, assume that the input voltage is zero. Then the voltage at the inverting input arising from the output of the op-amp is equal to v_f and is given by:

$$v^- = \frac{R_1}{R_1 + R_2} \cdot v_o$$

Then to determine the effect of the input voltage assume that

$A_{OL}v_d = 0$ and that for simplicity also assume $Z_o \ll R_2$ and $Z_{in} \gg R_1$.

$$v^- = \frac{R_2}{R_1 + R_2} \cdot v_{in}$$

Linear superposition means that the voltage at the inverting input can then be determined by adding the effects of the input voltage and the op-amp output voltage

$$v_d = v^+ - v^- = 0 - \left\{ \frac{R_2}{R_1 + R_2} v_{in} + \frac{R_1}{R_1 + R_2} v_o \right\}$$

This expression for v_d can now be included in the equation for the voltage at the output from the op-amp. Taking into account that the input is connected to the inverting input to the op-amp and the direction of the output current

$$v_o = A_{OL}v_d - i_o Z_o$$

and then re-arrangement (eliminating v_d) gives

$$v_o = -\frac{R_2}{R_1 + R_2} \cdot \frac{A_{OL}}{1 + \beta A_{OL}} \cdot v_{in} - i_o \frac{Z_o}{1 + \beta A_{OL}}$$

The closed-loop gain is therefore given by:

$$A_{CL} = -\frac{R_2}{R_1 + R_2} \cdot \frac{1}{\beta} \cdot \frac{1}{1 + \frac{1}{\beta A_{OL}}}$$

but since

$$\frac{R_2}{R_1 + R_2} = 1 - \beta$$

this expression becomes

$$A_{CL} = \left(1 - \frac{1}{\beta}\right) \cdot \frac{1}{1 + \frac{1}{\beta A_{OL}}}$$

Finally, if $\beta A_{OL} \gg 1$ then

$$A_{CL} = \left(1 - \frac{1}{\beta}\right)$$

The conclusion is that:

- (i) the output impedance has decreased by a factor of $1 + \beta A_{OL}$
- (ii) as long as the op-amps open loop gain is greater than the closed loop gain the closed-loop gain is independent of the exact value of the op-amp's open-loop gain.

Real op-amps in a non-inverting amplifier

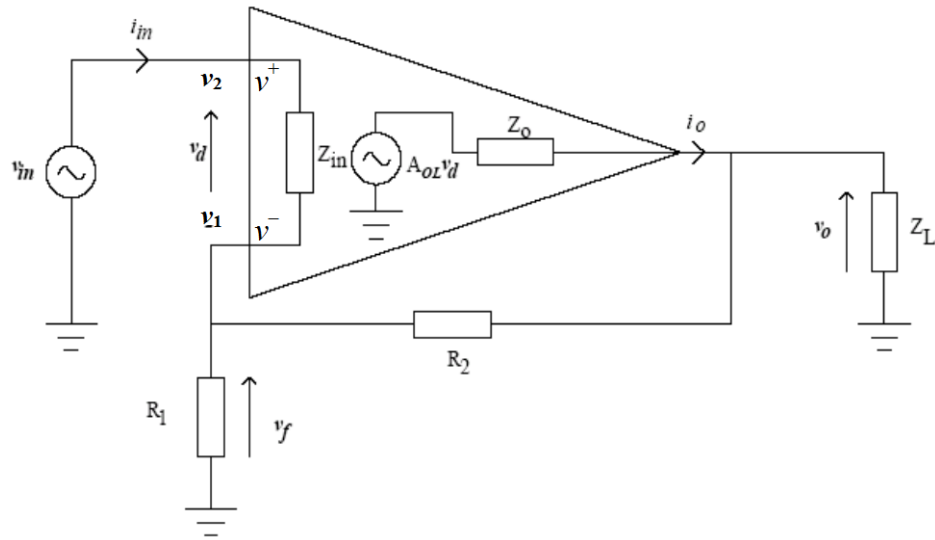


Figure 46: A non-inverting amplifier circuit including non-ideal op-amp model.

The circuit shown in figure (46) can be analysed to include the effects of a non-ideal op-amp, see the tutorial problem, to show that;

$$v_o = \frac{A_{OL}}{1 + \beta A_{OL}} \cdot v_{in} - i_o \cdot \frac{Z_o}{1 + \beta A_{OL}}$$

This means that the circuit behaves as an amplifier with closed-loop gain:

$$A_{CL} = \frac{A_{OL}}{1 + \beta A_{OL}}$$

and the output impedance is

$$Z_o / (1 + \beta A_{OL})$$

The expression for the closed-loop gain can also be re-written as follows:

$$A_{CL} = \frac{1}{\beta} \cdot \frac{1}{\left(1 + \frac{1}{\beta A_{OL}}\right)}$$

If $\beta A_{OL} \gg 1$, then $A_{CL} \simeq 1/\beta$ and the closed-loop gain once again only depends upon the components in the feedback circuits and is independent of the exact value of A_{OL} as required.

Since $A_{CL} \simeq 1/\beta$, the condition $\beta A_{OL} \gg 1$ is equivalent to $A_{OL} \gg A_{CL}$. Thus one of the conditions for selecting an op-amp is that the open-loop gain is much larger than the closed-loop gain of the final circuit.

Input impedance- Assume that the circuit design is power efficient, so that any currents through the feedback resistors are small compared to the current delivered to the load impedance Z_L , then Z_o and Z_L form a potential divider and the output voltage becomes:

$$v_o = A_{OL} v_d \frac{Z_L}{Z_L + Z_o} \quad (a)$$

Now the voltage v_2 can be determined from v_o and the R_1 and R_2 potential divider, so:

$$v_2 = v_o \frac{R_1}{R_1 + R_2} = \beta v_o$$

where β is $R_1/(R_1+R_2)$. Substituting in for v_o from eq. (a) we have:

$$v_2 = \beta A_{OL} \frac{Z_L}{Z_L + Z_o} v_d \quad (b)$$

Now, considering figure 46, the input voltage v_{in} can be written as the sum of v_2 and v_d , i.e.:

$$v_{in} = v_2 + v_d$$

Substituting in for v_2 from eq. (b) we then have:

$$v_{in} = \beta A_{OL} \frac{Z_L}{Z_L + Z_o} v_d + v_d = v_d \left(1 + \beta A_{OL} \frac{Z_L}{Z_L + Z_o} \right)$$

We can now determine Z_{inCL} by dividing by the input current i_{in}

$$Z_{inCL} = \frac{v_{in}}{i_{in}} = \frac{v_d}{i_{in}} \left(1 + \beta A_{OL} \frac{Z_L}{Z_L + Z_o} \right)$$

Thus

$$Z_{inCL} = Z_{in} \left(1 + \beta A_{OL} \frac{Z_L}{Z_L + Z_o} \right)$$

Finally, assuming that $Z_L \gg Z_o$, so that most of the power from the op-amp output is delivered to the load and not dissipated internally.

$$Z_{inCL} = Z_{in} (1 + \beta A_{OL})$$

so the already large input impedance is increased by the feedback loop.

To summarize:

- (i) The output impedance has been decreased by a factor $(1 + \beta A_{OL})$
- (ii) The closed-loop gain equal to $1/\beta$ which is independent of the exact value of the op-amp's open-loop gain.
- (iii) The input impedance has been increased by the factor $1 + \beta A_{OL}$

and the conditions for selecting an op-amp to ensure that an op-amp appears to be ideal for this application are

$$Z_{in} \gg R_1$$

$$A_{OL} \gg A_{CL}$$

$$Z_o \ll Z_L$$

Correcting for finite input bias currents

When designing an op-amp circuit it is often necessary to consider the finite input bias current which many op-amps require. The input impedance of the op-amp that has been considered so far represents the changes in this current that occur when the two input voltages change. It is also necessary to **consider the average current that must be supplied to these two inputs**. Each of the inputs pins may require a dc bias current to ensure that a bipolar transistor within the op-amp operates correctly. **These bias currents can be represented by two current sources between the relevant input and ground, shown in figure (47).**

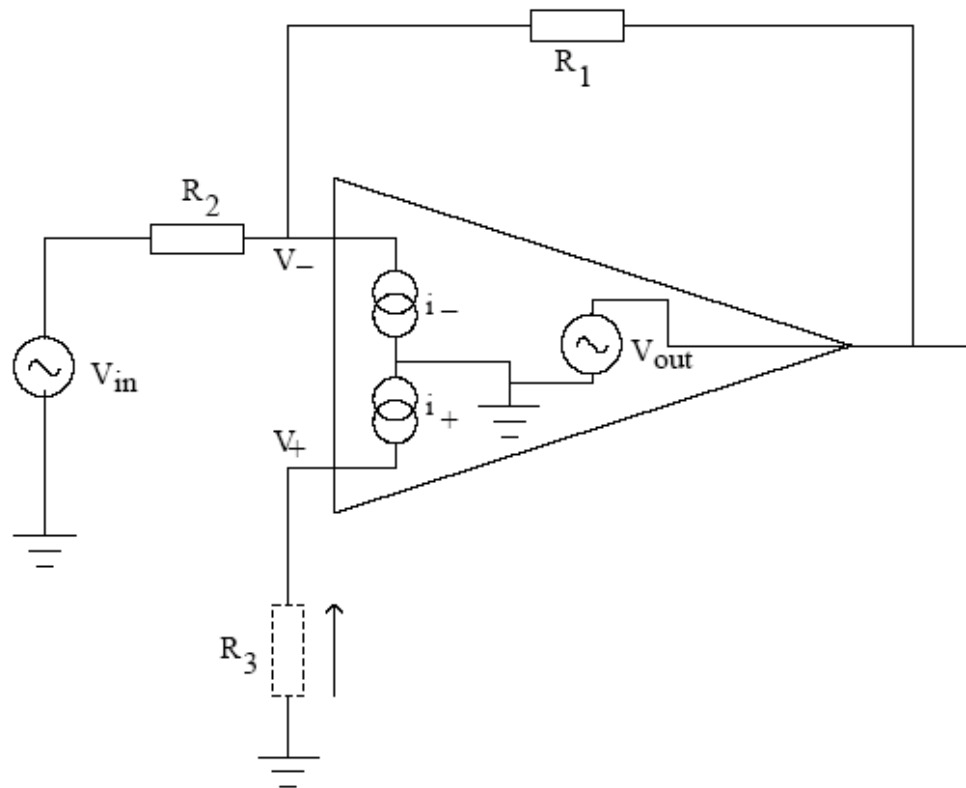


Figure 47: An inverting amplifier circuit including input bias currents. For an ideal op-amp only two resistors are required, however, a third resistor is required to compensate for the input bias currents.

In the example circuit shown, an inverting amplifier, there are three signal sources, V_{in} , V_{out} and i_- that determine the voltage at the inverting input to the op-amp. Superposition means that the total voltage at the inverting input, can be calculated by adding together the effect of each source. During the calculation of the effect of each signal source, any other voltage source is short-circuited and any other current source is assumed to be open circuit. This means that whilst calculating the effects of the two voltage sources, the bias current source is open circuit, i.e. the op-amp is ideal. **The error in voltage at the inverting input caused by the presence of i_- is therefore the voltage generated at this circuit node when the two voltage sources are short-circuited, i.e. both voltage sources are set to zero.** With the output voltages from both the voltage sources set to zero then both R_1 and R_2 connect the inverting input to ground. The bias current flowing through this parallel combination of resistors will then cause a voltage signal

$$\delta V_- = i_- \cdot R_1 // R_2$$

At first it would appear that an input bias of an op-amp would be insignificant. However, this expression shows that with typical resistance values of 10K or more, an input bias current of 1 μ A can generate an error of 10mV. This is significant compared to the signal from the typical sensor.

There are **two possible solutions to this problem:**

- (i) select an op-amp with a small bias current or
- (ii) use small value resistances in the feedback circuit.

Unfortunately **in some situations neither of these two solutions are possible. In this case the circuit can be modified to significantly reduce the impact of the undesirable signal.**

To understand how the circuit must be modified remember that the op-amp has a high, if not quite infinite, differential gain. It will therefore amplify any differential input voltage, such as the one created by the bias current flowing in resistors R_1 and R_2 , even in the absence of an input signal. The effects of this bias current will become negligible if the inverting input is at the same voltage as the non-inverting input so that the differential voltage is zero. Then there will be no differential input voltage when the input to the circuit is ground. The effects of the input bias currents will then be negligible.

In the ideal circuit the voltage at the non-inverting input is zero, even when there is a finite input bias current. **To allow the input bias current to create a finite input voltage a resistor is required between the non-inverting input and ground, see the dashed box in figure (47).** Assuming that the resistance of this extra resistor is R_3 , then $V_+ = i_+R_3$. However, in order to obtain an output voltage of zero when the input is zero the value of R_3 should be chosen so that under these conditions $V_+ = V_-$. Thus

$$i_+R_3 = i_-R_1//R_2$$

and since in op-amps the two inputs to the op-amp will be connected to nominally 'identical' devices we can assume that $i_+ = i_-$ and hence

$$\underline{R_3 = R_1//R_2}$$

that is the new resistor is equivalent to the other two resistors acting in parallel. The addition of this extra resistor, which would have no effect on an ideal op-amp, is therefore to cancel the effects of the finite input bias current.

Frequency Response

So far it has been assumed that the gain of an op-amp is independent of frequency. However, because of unavoidable parasitic capacitances that form part of each transistor within the op-amp the gain of a real op-amp is frequency dependent. The result is a complex frequency dependent gain (and phase shift) that is both difficult for the manufacturer to control and/or specify accurately and for the designer to use confidently.

The solution to this problem that has been adopted by manufacturers is to include a large capacitance within each op-amp. This capacitance is then designed to ensure that over the frequency range of interest the op-amp behaves as if it had a single dominant RC response. Thus **compensated op-amps** (such as the 741 and all the op-amps that you are likely to use) have the frequency response characteristics of a low-pass RC filter

$$A_{OL}(f) = \frac{|A_{OL}|_{dc}}{1 + j\frac{f}{f_c}}$$

where $|A_{OL}|_{dc}$ is the dc gain of the op-amp and f_c is the 3dB break-point frequency created by the frequency compensating capacitor.

The op-amp is then designed to ensure that its gain falls to unity at a frequency that is low enough so that any other frequency dependant responses within the op-amp are negligible. The consequence of this is that the 3dB break-point frequency f_c for compensated op-amps is extremely low: 5 Hz for the 741, for example.

The convention in an op-amp data sheet is to give the values of the dc gain, $|A_{OL}|_{dc}$, and the unity gain frequency f_1 (which is the frequency at which the open-loop gain has fallen to 0 dB). This frequency represents the maximum frequency at which the op-amp could be used to create a unity gain buffer. It is therefore a measure of the maximum useful frequency of the op-amp.

The closed-loop frequency response $A_{CL}(f)$ of any circuit can now be determined by replacing the frequency independent open-loop gain A_{OL} by the frequency dependant gain. Thus for the non-inverting amplifier

$$A_{CL} = \frac{1}{\beta} \cdot \left(\frac{1}{1 + \frac{1}{\beta A_{OL}}} \right)$$

becomes

$$A_{CL}(f) = \frac{1}{\beta} \cdot \left(\frac{1}{1 + \frac{1}{\beta A_{OL}(f)}} \right)$$

Because for this particular circuit the feedback network only contains resistors β is independent of frequency and $1/\beta$ is a straight line parallel to the frequency axis, as shown in figure (48).

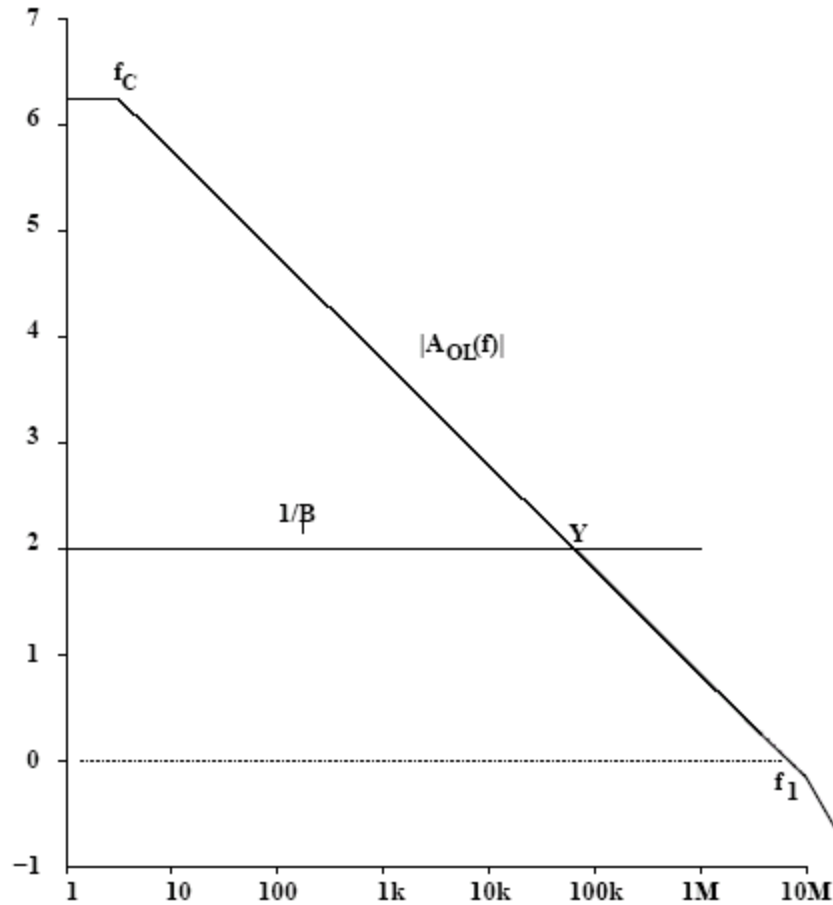


Figure 48: Closed-loop frequency response of a non-inverting amplifier.

At the point Y in figure (48) $|A_{OL}(f)| = 1/\beta$ and hence at this point:

$$|A_{OL}(f)| = \frac{|A_{OL}|_{dc}}{\sqrt{1 + \left(\frac{f}{f_c}\right)^2}} = \frac{1}{\beta}$$

The bandwidth of the circuit with the feedback loop, the **closed-loop bandwidth**, f_1' is therefore given by:

$$\sqrt{1 + \left(\frac{f_1'}{f_c}\right)^2} = \beta |A_{OL}|_{dc}$$

Thus although the open-loop bandwidth of the op-amp alone is only f_c , the bandwidth for the circuit with feedback is f_1' where, assuming $f_1' \gg f_c$,

$$f_1' = f_c \cdot \beta | A_{OL} |_{dc}$$

The greater the amount of feedback (i.e. the greater the value of β and the lower the gain $1/\beta$), the greater the bandwidth f_1' . In fact this equation shows that **the gain bandwidth product of the closed loop circuit**

$$\frac{1}{\beta} \cdot f_1'$$

is a constant and the value of the constant is determined by the op-amp.

Frequency compensation and slew rate

One effect of the inclusion of a compensating capacitor within the op-amp is that it limits the rate at which the op-amp can respond to any sudden changes of input. In the data sheets the parameter used to characterise this effect is the slew rate (usually expressed in $V/\mu s$) which is the maximum rate of change of the output voltage of the op-amp.

Op-amp data sheets usually give the slew rate for unity gain. Assume, therefore, that the input to a voltage follower is a large-amplitude high-frequency sinewave where:

$$v_{in} = V_{max} \sin \omega t = v_o \quad (\text{unity gain})$$

The maximum rate of change of the output is given by:

$$\left| \frac{dv_o}{dt} \right|_{max} = \omega V_{max}$$

For an output **free of distortion**, the slew rate determines the maximum frequency of operation f_{max} for a desired output swing. Thus:

$$f_{max} = \frac{SR \times 10^6}{2\pi V_{max}}$$

where SR is the slew rate in $V/\mu s$ for a unity gain circuit.

Mini-Summary

An ideal op-amp has an infinite differential gain, an infinite input impedance and an output impedance of zero. The characteristics of an ideal op-amp are used to select a circuit which performs the required function.

Real op-amps have a large differential gain, a large input impedance and a small output impedance. These are represented in a model of the op-amp that can be used to derive conditions for an op-amp to appear to be ideal in a particular circuit.

Other aspects of the behaviour of real op-amps also affect the performance of circuits. In particular the output voltage is limited by the voltages used to power the op-amp. There is also a finite dc input bias current whose effects can be cancelled by including an additional resistance in the circuit.

The gain of a real op-amp is frequency dependant. To simplify the design of circuits the op-amps are designed to have a single dominant pole and therefore the gain-bandwidth product of an amplifier including an op-amp is constant.

The slew-rate of the op-amp limits the maximum rate of change of the output voltage of an op-amp.

NOISE

Introduction

So far the emphasis has been on how to amplify an analogue signal so that it matches the maximum input voltage of an ADC. This will determine the maximum signal that can be detected without saturation. It is at least as important to know the minimum reliably detectable signal.

Usually, the minimum detectable input signal is approximated to the maximum error that is made during the conversion process. This is half of the change in input voltage that corresponds to a change in output of one least significant bit (LSB). For a 12-bit converter with a maximum input voltage of 5V this error is

$$5/(2 \times 2^{12}) = 610\mu V.$$

In itself this is a small voltage that shows that the system is vulnerable to interference. However, the original source of this ADC input signal is a sensor output signal that is considerably smaller. Thus for example if the maximum transducer signal was itself 10 mV, i.e. an amplifier with a gain of 500 has been used between the transducer and the ADC, then 1/2 LSB is equivalent to a change in transducer output of 1.2 μ V. Careful design is then required to ensure that this small signal can be reliably detected despite interference and other signals generated by various unavoidable physical processes within components. Assuming that interference

can be avoided it is the noise signals, caused by the physical processes underlying charge flow in components, that determine the sensitivity of modern instrumentation systems.

Fundamental noise sources

There are three noise mechanisms that are found in many components,

Thermal or Johnson noise

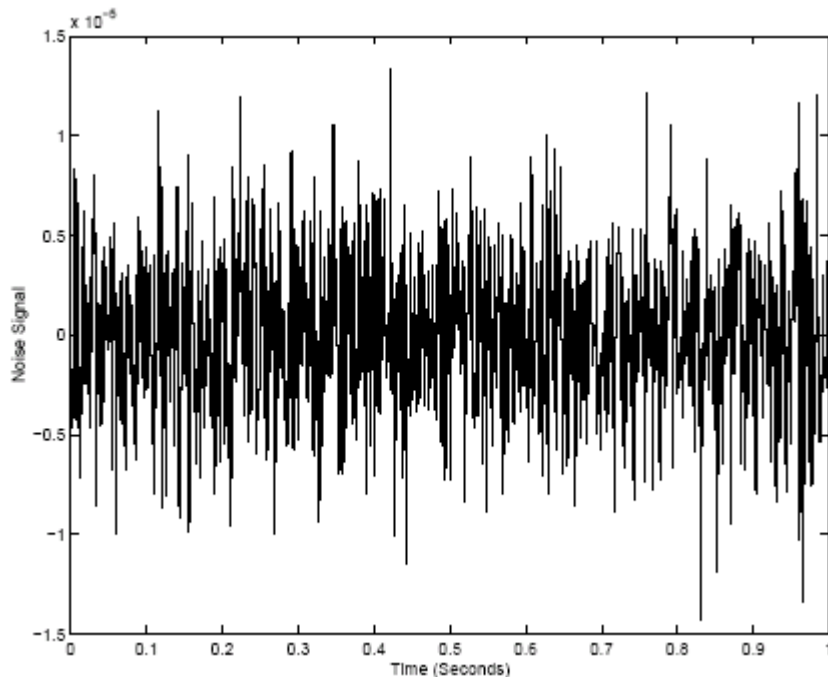


Figure 24: White noise from a 100 k Ω resistor in a bandwidth of 10 kHz.

Any resistor generates a noise voltage across its terminals, known as **thermal** or **Johnson** noise. This type of noise arises because the measured macroscopic current actually represents a flow of charge carriers through a component. **The average velocity of these carriers,**

and hence the average current, is determined by the electric field, however, interactions between carriers mean that there are variations in the velocity of individual carriers. These fluctuations in the motion of charge carriers in the resistor then lead to fluctuations in the current flowing through the resistor, which translate into momentary random changes in the voltage across the resistor such as those shown in figure (24). In the frequency domain these random fluctuations correspond to equal amounts of power at all frequencies and it is this flat frequency spectrum which leads to this type of noise also being known as white noise.

The flat frequency power spectrum means that the total amount of noise power will be proportional to the noise bandwidth of the system B . It is therefore not surprising that the root-mean-square output voltage generated by a resistor is proportional to \sqrt{B} . In fact for a resistance R

$$\underline{V_{\text{noise}}(\text{rms}) = (4kTBR)^{\frac{1}{2}}}$$

where k is Boltzmann's constant and T the absolute temperature.

To estimate the voltage change that this can represent, consider a 100 k Ω resistor at room temperature, over a modest bandwidth of 10 kHz:

$$V_n(\text{rms}) = \{4 \times 1.38 \times 10^{-23} \times 300 \times 10^4 \times 10^5\}^{\frac{1}{2}} = 4\mu\text{V}$$

This is a small voltage, however, it is not negligible compared to the signal voltages that may need to be reliably detected in any application which requires a sensitive measurement. The impact of

this type of noise should therefore be considered in any carefully designed instrumentation system.

The equation for thermal noise shows that two general principles should be followed in order to reduce the amount of Johnson noise.

(i) **Any resistances should be kept as small as possible.**

However, the ability of the designer to follow this strategy will always be limited by the necessity to limit power consumption and to ensure that any parasitic resistances in the circuit, such as contact and track resistances, are negligible. This usually limits the smallest resistance used to be greater than 1 k Ω .

(ii) **The frequency bandwidth of the system should be limited to the frequency range of interest.** Filters should therefore be used to limit the frequency range of the ADC input signal to the frequency range required to capture any necessary information. Thus for example if vibrations in the range 0.1-1kHz from a motor are indicative of wear which requires maintenance then the system should be designed to limit the signal to frequencies in the range 0.1-1kHz. The filter circuits used for this will be discussed later.

In addition to these general precautions it is sometimes necessary to **calculate the effects of Johnson noise in a particular circuit. The equivalent circuit for Johnson noise that is used in these calculations is an ideal noise-free resistor in series with a voltage**

noise source or in parallel with a current noise source as shown in figure (25).

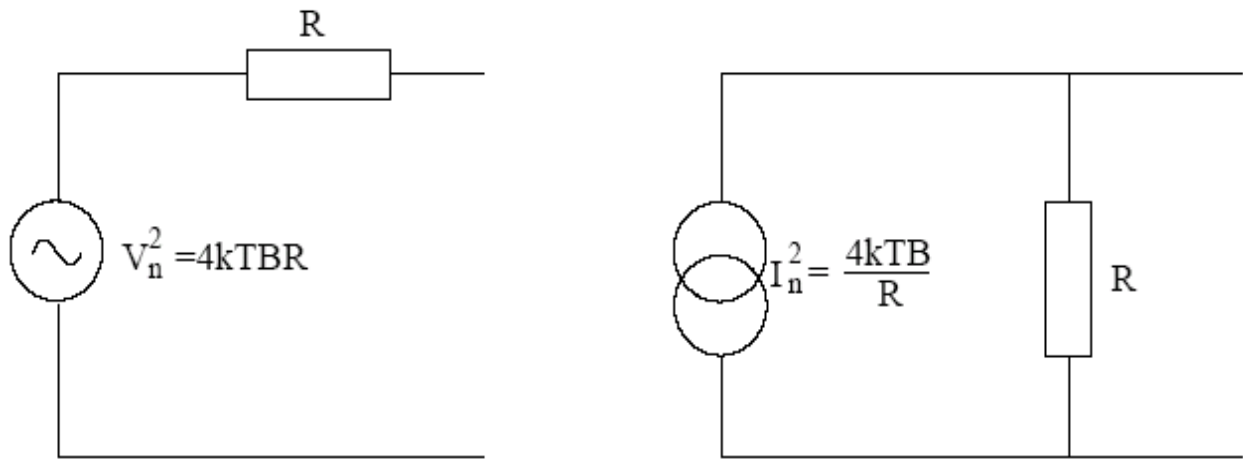


Figure 25: The Thevenin and Norton equivalent circuits to represent noise in a resistor.

Shot noise

Like thermal noise, shot noise also arises because the measured current represents a flow of charge carriers through a device. **Thermal noise arises from instantaneous variations in the average velocity of carriers. Shot noise arises from fluctuations in the number of carriers.** The result of these microscopic fluctuations in carrier flow can be represented by a macroscopic current noise source I_n . If a current I is flowing then for a noise bandwidth B

$$\underline{I_n^2 = 2 \cdot q \cdot I \cdot B}$$

where $q = 1.602 \times 10^{-19}$

Shot noise occurs in diodes, but, most importantly it arises as a consequence of the small, but, finite base current of a bipolar transistor.

This is an important contribution to the noise in some op-amps that will be considered later.

1/f or Flicker noise.

Flicker noise is present in many different physical systems. In electronic components it is associated with fluctuations in current flow caused by temporary trapping of charge carriers.

Unlike the other two noise sources that have been mentioned the flicker noise power per unit frequency depends upon frequency. **Thus when a dc current I is flowing the noise in a frequency range of 1Hz centred on frequency f , is**

$$\underline{i_n^2 = K.I^\alpha / f}$$

K and α are device dependant constants.

To calculate the total noise between two frequencies f_h and f_l the expression for i_n^2 is integrated to give

$$I_n^2 = K.I^\alpha \text{Ln}(f_h/f_l)$$

this shows that each decade of frequency adds an equal amount to the total noise of the device.

The most important aspect of 1/f noise is that it shows that the significance of 1/f noise decreases as the frequency of interest increases. In all devices it will therefore always be less than the thermal noise at all

frequencies above a critical frequency. The value of this critical frequency depends upon the particular device. However, **there are some important general trends.**

(i) **1/f noise is most significant in MOSFET transistors.** These form the basis of almost all digital circuits and are increasingly used to design analogue circuits on the same substrate as digital circuits. For this type of device 1/f noise can be the dominant noise source for frequencies up to 10kHz or above. 1/f noise therefore usually has to be taken into account in noise calculations.

(ii) **Compared to thermal noise, 1/f noise is only important in most bipolar devices at frequencies below 10-100Hz.** Its effects are therefore usually insignificant, particularly when low-frequencies are removed by filtering.

(iii) **1/f noise is usually negligible in resistors, however, normal carbon composite resistors can have an order of magnitude more 1/f noise than a wire-wound resistor.** In critical parts of a circuit it is therefore sometimes necessary to use expensive wire-wound resistors rather than the more usual carbon composite resistors.

Noise in Ideal Capacitors and Inductors

An ideal inductor has no resistance and therefore it will not generate noise. Noise arises from fluctuations in either the average velocity or number of carriers. **An ideal capacitor contains a perfect insulator that blocks the flow of carriers, which means that an ideal capacitor will not generate any noise.**

Addition of noise sources

Noise arises from random processes within devices, which cause voltages and currents to fluctuate around mean values. **Since it is impossible to ascribe a meaningful frequency or phase to this type of random signal, the only parameter that can be used to characterise noise is its power or root mean square amplitude.**

$$V_n = \left(\frac{1}{T} \int_0^T v_n^2 dt \right)^{1/2}$$

The quality of a signal in the presence of noise is most often specified by the signal-to-noise ratio (SNR)

$$SNR = 10 \log_{10}(V_s^2/V_n^2)$$

where V_s is the rms value of the signal and V_n is the rms value of the noise.

A critical step in determining the SNR at the output of a circuit is to calculate the total rms noise arising from the different noise sources within the circuit.

For two noise sources the total mean square output voltage V_n^2 is

$$V_n^2 = \frac{1}{T} \int_0^T [v_{n1} + v_{n2}]^2 dt$$

which can be expanded to give

$$V_n^2 = V_{n1}^2 + V_{n2}^2 + \frac{2}{T} \int_0^T v_{n1}v_{n2} dt$$

If the two noise sources are independent, they will be uncorrelated and

$$\int_0^T v_{n1}v_{n2} dt = 0$$

so that

$$\underline{V_n = (V_{n1}^2 + V_{n2}^2)^{1/2}}$$

Since the fluctuations in different devices will be independent this equation can be used to add the noise from different devices. The first important consequence of this can be highlighted by a simple example. Consider a situation in which the rms output voltages of two noise sources are $10\mu V$ and $5\mu V$, then the total rms output will be

$$(10^2 + 5^2)^{1/2} = 11.2\mu V.$$

This clearly demonstrates that **if there are two unequal noise sources then the larger noise source will dominate. In order to reduce the amount of noise in this situation the designer should concentrate upon reducing the noise from the larger noise source. A good strategy when designing a system is therefore to ensure that all noise sources are approximately equal.**

The equation

$$V_n = (V_{n1}^2 + V_{n2}^2)^{1/2}$$

can also be used to add the effect of two noise sources in the same circuit. However, before it can be used the effect of each noise source within the circuit must be calculated. This can be done by simply assuming that each source acts alone in a circuit in which all other voltage sources are short-circuited and other current sources are open-circuited. This first stage of these calculations is therefore similar to using superposition to calculate the response of a circuit to several signal sources. This principle states that in a linear circuit the response of two or more sources acting simultaneously is the sum of the responses for each source acting alone with the other voltage sources short-circuited and other current sources open-circuited. However, in the principle of superposition it is implicitly assumed that there is a fixed relationship between the phase of the various sources. The overall response is therefore calculated by simply adding the individual responses. This assumption is not valid for the signal sources representing noise that contains many frequency components. In this case the equation

$$V_n = (V_{n1}^2 + V_{n2}^2)^{1/2}$$

must be used to calculate the total amount of noise.

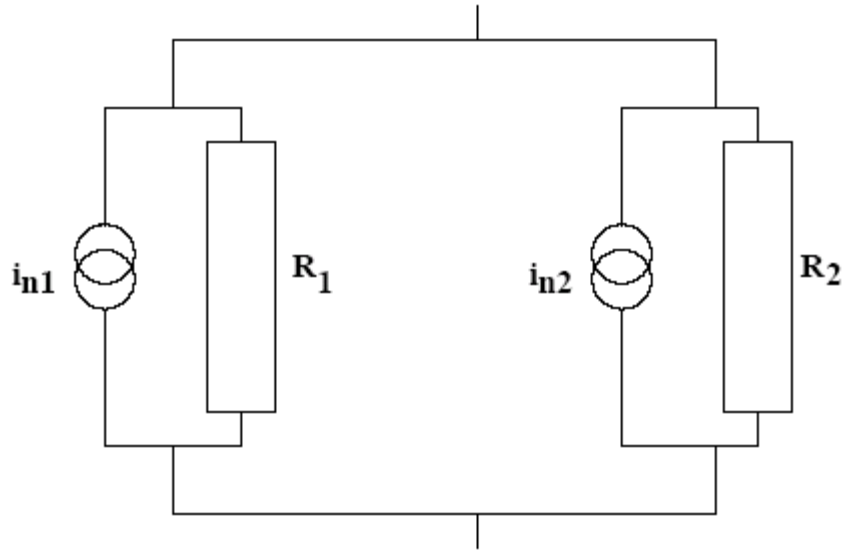


Figure 26: Two resistors in parallel, including the current sources that represent noise generated within each resistor.

An important simple example of a circuit containing two noise sources is a parallel combination of two resistors, R_1 and R_2 . First calculate the noise voltage caused by noise source, i_{n1} . For this calculation i_{n2} is assumed to be open-circuit, i.e. disconnected or removed from the circuit. The current i_{n1} therefore flows through the parallel combination of R_1 and R_2 . The corresponding output voltage is therefore

$$v_{n1} = i_{n1} \cdot (R_1 // R_2)$$

and similarly

$$v_{n2} = i_{n2} \cdot (R_1 // R_2)$$

The total noise can then be calculated

$$v_n^2 = v_{n1}^2 + v_{n2}^2 = (i_{n1}^2 + i_{n2}^2) \cdot (R_1 // R_2)^2$$

However, for a resistance R

$$i_{nR}^2 = 4kTB/R$$

which means that

$$V_n^2 = 4kTB(1/R_1 + 1/R_2).(R_1//R_2)^2$$

$$V_n^2 = 4kTB(R_1//R_2)^2/(R_1//R_2)$$

$$V_n^2 = 4kTB(R_1//R_2)$$

This is the noise voltage arising from a resistor with the same resistance as the parallel combination. In this example it would therefore be easier to combine the resistors to determine the effective total resistance and then calculate the corresponding noise. In fact this is an example of a general principle. **A simple approach to calculating noise in networks of resistors is to calculate the effective resistance and then add a single noise source to represent the noise of the whole resistor network.**

A circuit model for op-amp noise

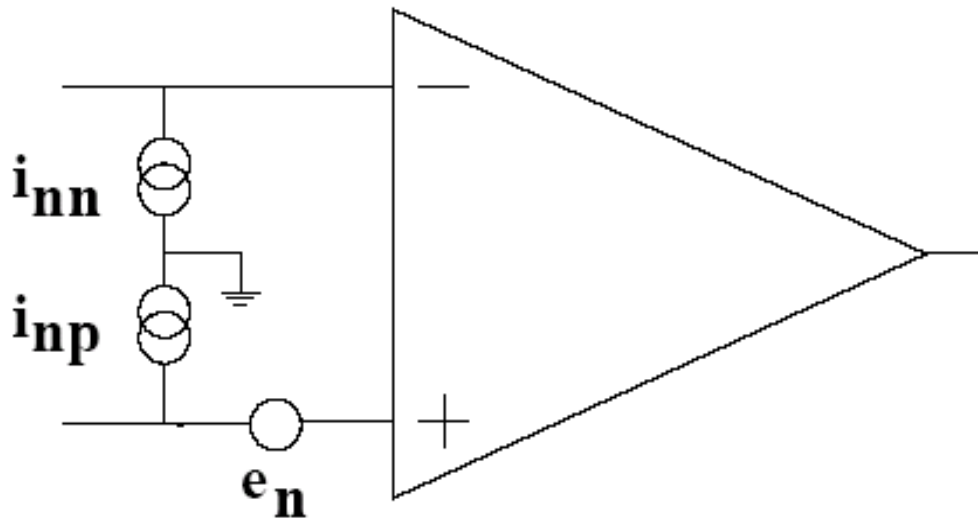


Figure 27: Circuit model to represent noise generated within an op-amp

Calculation of noise in a circuit that includes an op-amp requires a model for noise generated within the op-amp. A noisy resistance can be represented by an ideal resistance and a voltage or current source that generates the noise. Similarly, **the noise produced internally by an op-amp is modelled, by a noiseless op-amp and three noise generators. As shown in figure (27) these are a noise voltage generator and two noise current generators.**

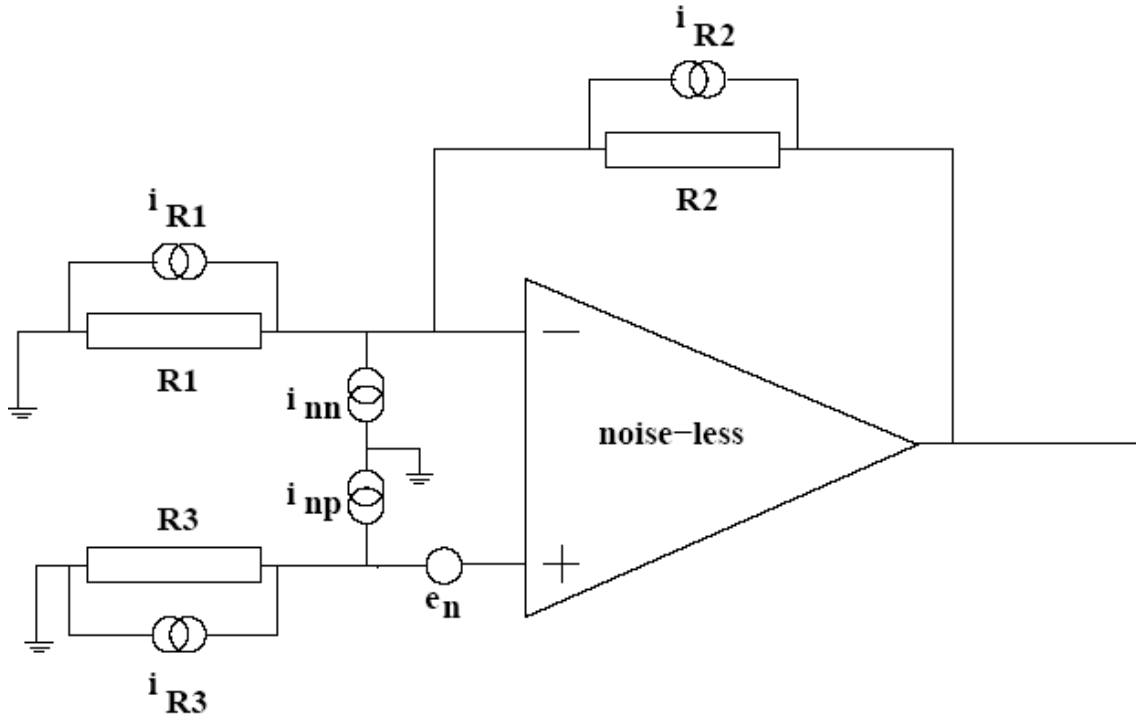


Figure 28 Circuit diagram showing the various noise sources in an inverting amplifier circuit.

This model can be used within any circuit to calculate the total noise. For example consider the inverting amplifier circuit, which includes a resistor connecting the non-inverting input to ground to compensate for the input current. Figure (28) then shows the circuit including the model of a noisy op-amp.

The first stage in calculating the total equivalent noise at the input is to calculate the effect of each of the individual signal sources.

Start by calculating the noise at the non-inverting input. Both the noise current sources, i_{R3} and i_{np} , create a noise voltage when they flow through resistor $R3$. Adding these two voltages to the

noise voltage from the op-amp itself leads to an expression for the total voltage noise at the non-inverting input

$$e_+^2 = R_3^2 \cdot (i_{R_3}^2 + i_{np}^2) + e_n^2$$

Now in the circuit attached to the inverting input there are two resistors. These can be combined into an equivalent circuit before proceeding. Since the voltage source at the output of the op-amp is assumed to be short-circuited during the calculation of the signal from all the noise sources the resistors R_1 and R_2 appear in parallel to create an effective resistance

$$R_{eff} = \frac{R_1 \cdot R_2}{R_1 + R_2}$$

This effective resistance both generates its own thermal noise and acts as the effective resistance to ground for the current noise i_{nn} . The noise voltage at the non-inverting input is therefore

$$e_-^2 = R_{eff}^2 \cdot (i_{R_{eff}}^2 + i_{nn}^2)$$

The final step of the calculation is to combine the noise at the two input pins to give the total noise at the non-inverting amplifier input

$$e_{total}^2 = R_3^2 \cdot (i_{R_3}^2 + i_{np}^2) + e_n^2 + R_{eff}^2 \cdot (i_{R_{eff}}^2 + i_{nn}^2)$$

Assuming that $i_{np}^2 = i_{nn}^2 = i_n^2$ this reduces to

$$e_{total}^2 = e_n^2 + i_n^2 (R_{eff}^2 + R_3^2) + R_3^2 \cdot i_{R_3}^2 + R_{eff}^2 \cdot i_{R_{eff}}^2$$

To simplify this expression further assume that the amplifier circuit has been well designed so that $Reff = R1//R2 = R3$ ¹. Hence

$$e_{total}^2 = e_n^2 + 2 i_n^2 \cdot Reff^2 + 2 Reff^2 \cdot i_{Reff}^2$$

this can be further simplified using the expression

$$i_{Reff}^2 = 4kTB/Reff$$

$$e_{total}^2 = e_n^2 + 2 \cdot i_n^2 \cdot Reff^2 + 2 \times 4kTB \cdot Reff$$

This last expression shows the three contributions to the total noise at the input to the amplifier;

- (i) the voltage noise generated by the amplifier**
- (ii) the voltage noise generated by the current noise in the amplifier flowing through an effective resistance, $Reff$, which arises from a combination of the resistors associated with the amplifier**
- (iii) the thermal noise generated by the effective resistance of the amplifier circuit**

The importance of the existence of the current noise, which primarily arises from shot noise associated with the input bias current, is now clear. **If $Reff$ is small then the total noise will be dominated by the voltage noise of the amplifier.** However, as $Reff$ increases the contribution from the other two terms will also increase and in particular the

¹The resistance R3 has been included to compensate for the DC bias current that flows into the inputs of some non-ideal op-amps (see earlier section.)

contribution from the current noise of the amplifier, which is proportional to R_{eff}^2 , will begin to dominate. **This result demonstrates the importance of minimising R_{eff} .** Unfortunately, the minimum value of R_{eff} will be restricted by the output resistance of the source of the input signal, R_S . This arises because the values of R_1 and R_2 in the circuit will be designed to be large compared to R_S . This is necessary to both reduce signal loss arising from the current drawn through R_S and to ensure that the amplifier gain is determined by R_1 and R_2 . For an amplifier, for which $R_2 > R_1$ these requirements mean that $R_2 > R_1 \gg R_S$. Then, since R_3 is equivalent to the parallel combination of R_1 and R_2 , $R_{eff} > R_S$ and there is a limit on the minimum value of R_{eff} . **This is one reason why it is important to select a sensor with a small output resistance. If this is not possible it may be necessary to select an op-amp that has a small input bias current and therefore a small noise current.**

The noise at the input to the amplifier circuit is indistinguishable from the input signal. Like the signal it will be amplified by the gain of the amplifier circuit. Unfortunately the presence of noise in both the op-amp and the resistors in the feedback circuit reduces the signal to noise ratio at the output of the amplifier compared to that at the signal source. **Since the SNR will only be reduced by any analogue circuits it is important to use a sensitive sensor so that the input SNR is as large as possible.** This is another advantage of using a full-bridge circuit that can amplify the signal without adding any extra noise.

Quantisation Noise

A further mechanism that degrades the signal-to-noise ratio in an instrumentation system is the errors introduced by quantisation in the analogue-to-digital converter.

In an n-bit ADC every analogue input voltage is represented by a digital output that corresponds to one of 2^n analogue reference levels. **The maximum error that will be made during each conversion is half the difference between two of these levels.** If the full-scale input voltage of the ADC is V_{FS} , then the difference between two reference levels is $V_{FS}/2^n$ and **the corresponding maximum error is $V_{FS}/(2 \cdot 2^n)$.** **With the ever improving performance, and reducing cost, of ADCs it is now quite common for systems to be designed so that this error is smaller than the input noise. The output from the ADC will then be accurate enough to represent the noise as well as the signal.**

For the situations in which cost means that it is not possible to use an ADC with a large number of bits, it is possible to calculate the variance of the error introduced by the quantisation process. This error is often referred to as **quantisation noise. It can be treated as another voltage noise source acting at the input to the ADC, however, the effects of this can only be quantified using an expression for the variance of this error.**

To calculate this variance it is conceptually easiest to consider a flash converter, however, the result will apply to any converter. Let the voltage difference between reference levels in the ADC be δ . Then the error, e , for each conversion will be between $\delta/2$ and $-\delta/2$. Assuming that each error is equally probable, then the probability density of error e , $p(e)$, is

$$p(e) = 1/\delta$$

Since the mean value of the error is zero, its variance is

$$\sigma_e^2 = \int_{-\delta/2}^{+\delta/2} p(e) e^2 de$$

and hence

$$\sigma_e^2 = (1/\delta) \cdot \int_{-\delta/2}^{+\delta/2} e^2 de = \delta^2/12$$

This suggests that the ADC is more accurate than the simple estimate based upon the largest error, $\delta/2$, suggests. However, the difference is relatively small. In many situations it is therefore appropriate to use the estimated error $\delta/2$, but, to remember that this is a conservative figure so that if an n-bit ADC appears to be just adequate based upon the simple estimate, it will be adequate because the quantisation noise has been overestimated by a factor of 1.7, which is almost the equivalent of an extra bit.

Mini Summary

Noise signals arise from unavoidable fluctuations in the flow of charge carriers through a conductor. The small signals that this generates can be comparable with the signals that can be represented at the output of modern ADCs. **Noise signals can therefore be critical in determining the accuracy of modern instrumentation systems that are no longer limited by ADC performance.**

Thermal noise arises from fluctuations in the average velocity of carriers in a conductor. This is the dominant unavoidable noise source in any resistor in a circuit. It can only be limited by either reducing the value of resistance used or limiting the bandwidth of the signal to the frequency range required to capture any necessary information.

Shot noise arises from fluctuations in the number of carriers passing a particular point in a device This mechanism is important in bipolar devices and hence in any op-amp which incorporates these transistors. As with thermal noise shot noise can be reduced by limiting the signal bandwidth.

The power per unit bandwidth of flicker ($1/f$) noise reduces as frequency increases. In most devices this noise source is only important at low frequencies. However, it is important in MOSFET devices that are increasingly used to design analogue circuits in the same package as digital circuits.

Noise is a randomly fluctuating signal which is characterised by its root means square amplitude.

Ideal capacitors and inductors do not generate noise themselves.

The quality of a signal in the presence of noise is most often specified by the signal-to-noise ratio (SNR) which is

$$SNR = 10 \log_{10}(V_s^2/V_n^2) = 20 \log_{10}(V_s/V_n)$$

where V_s is the rms value of the signal and V_n is the rms value of the noise.

If the two noise sources are independent, they will be uncorrelated which means that

$$\int_0^T v_{n1}v_{n2}dt = 0$$

and the root mean square output of two noise sources is

$$V_n = (V_{n1}^2 + V_{n2}^2)^{1/2}$$

This means that a good strategy when designing a system is to ensure that

all noise sources are approximately equal.

A simple approach to calculating noise in networks of resistors is to calculate the effective resistance and then add a single noise source to represent the noise of the whole resistor network.

If a circuit with a large output resistance is connected to the input of an

op-amp amplifier the total noise of the circuit can be dominated by the current noise of the op-amp. It is therefore important to minimise output resistances. When this is not possible it may be necessary to select a op-amp with a zero input bias current, and hence a negligible current noise.

The standard deviation of quantisation noise, which represents fluctuating errors caused by the process of digitising a signal, is

$$\sigma_e = \delta/\sqrt{12}$$

In order to achieve a requirement for both high-gain and low-noise it is often necessary to have a low-gain, low-noise circuit, usually called a pre-amplifier, followed by a high-gain circuit.

NON-IDEAL DIGITAL-TO-ANALOGUE AND ANALOGUE-TO-DIGITAL CONVERSION

Introduction

Specification of D/A converters

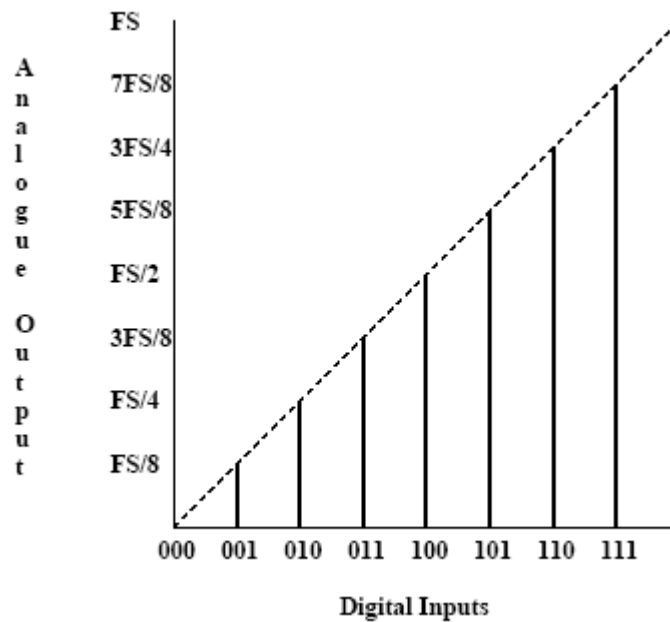


Figure 12: The ideal response of a 3-bit DAC, showing the analogue output voltage as a fraction of the full scale output FS. Each bar represents the output for a particular input and the dashed line shows the line connecting the ideal outputs.

As you learnt in the P2 course last year, a digital to analogue converter (DAC) converts a digital input represented as a binary number to an analogue voltage (or current) that is proportional to the

value of this input. The **ideal** relationship between the analogue output and digital input for a 3-bit converter is shown in figure (12).²

Various physical processes occur when circuits, including DACs, are manufactured which mean that **it is very difficult, if not impossible, to manufacture a circuit which achieves the ideal performance specification. In the case of DACs the resulting error is characterised as the maximum deviation between the actual and ideal outputs. This absolute accuracy is expressed as a fraction of the output change caused by a change in the digital input of one least significant bit (LSB).**

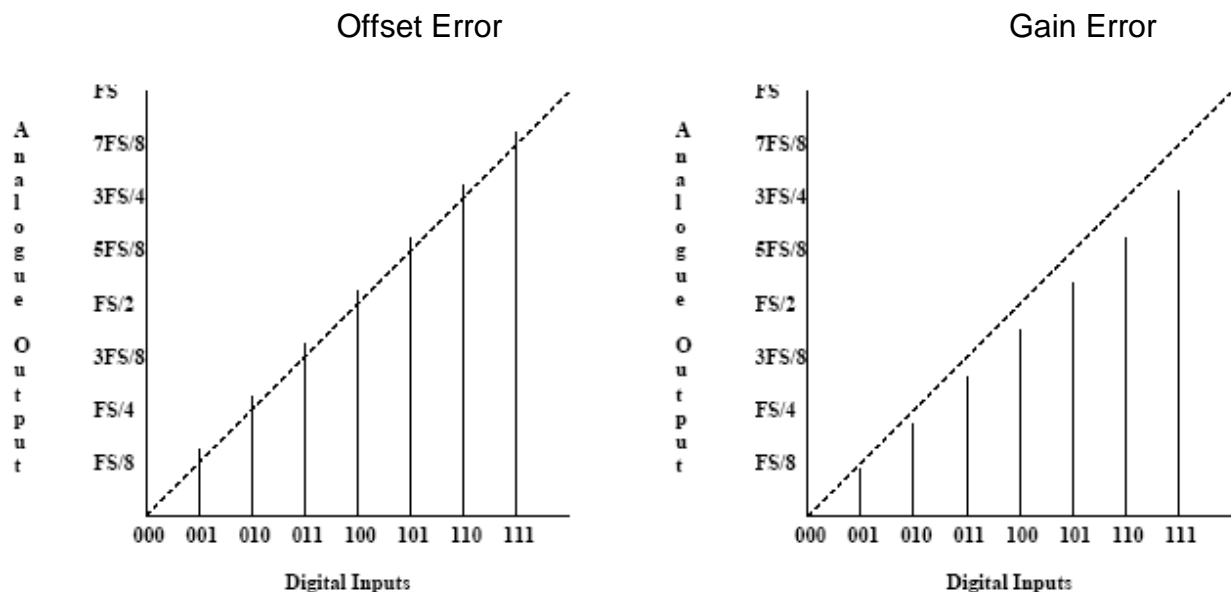


Figure 13: Exaggerated examples of DAC offset error, on the left, and gain error, on the right.

² In figure 13 only 3 bits have been shown for clarity. However, in real instrumentation systems DACs with 6, 8, 10, 12 and 14 bits are often used.

The maximum deviation of the actual output from the ideal output is the **absolute accuracy** of the DAC. There are several different types of error that might occur in the output of a DAC. An **offset error** means that the error between the actual output and the ideal output is the same for all binary inputs. In contrast, a **gain error** means that the slope of the ideal and actual outputs are different, see figure (13).

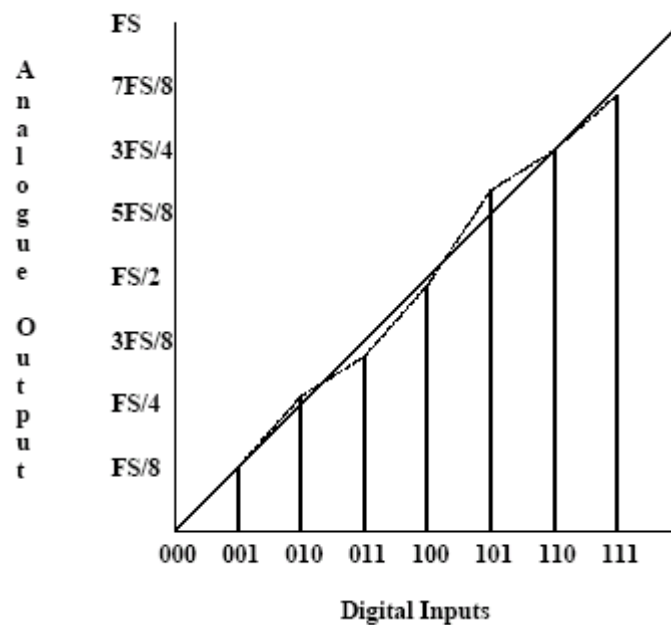


Figure 14: The response of a DAC showing the two lines used to define the non-linear response of the DAC.

Even for a system with no offset error and an ideal gain the individual outputs may still deviate from their ideal values. These errors are characterised as the **integral non-linearity (INL)** and the **differential nonlinearity (DNL)**. The **INL** is the maximum difference between the actual and ideal output calculated for each digital input. As with all errors the INL should be less than 1 LSB. The DNL is a measure of

the changes in output between successive inputs expressed as a fraction of the LSB. Again this should be less than an LSB. However, this may not be achieved and in an extreme situation the DNL may be more than -1LSB . In this case for a particular digital input an **increase** by 1 LSB in the input will cause a **decrease** in the output. This might be acceptable, but, it has to be avoided in many applications, particularly when the DAC output forms part of a control system or an ADC.

Clearly, **for an n-bit DAC to be credible all errors should be less than $\text{LSB}/2$.** This will almost certainly be correct for low-frequency operation. However, there are other sources of error that can degrade the DAC performance at high frequencies. The most common cause of high frequency errors are **glitches**. These are spikes in the output which occur at transitions between different inputs because of imperfections in the DAC circuitry, for example during a transition from 0111 to 1000 the most significant bit (the MSB) may change fractionally faster than the other bits so that there is an instant at which the output corresponds to 1111. These glitches only become important when the difference between the switching times of different bits become a significant fraction of the time for which each different digital input is applied. These glitches are critical to an increasing number of systems that rely upon a digital circuit and a DAC to create a well controlled, programmable output signal (An example of this type of system is the signal generators which you use in labs.). Despite these applications many data sheets for DACs still specify the maximum operating frequency as the highest frequency at which the DAC output attempts to change in response to a new input. This may be dramatically higher than the frequency at which the absolute error is less

than an LSB. All data sheets therefore have to be read carefully when selecting a component to ensure that the required performance will be achieved at the required frequency.

Specification of A/D converters

As with DACs, the ADC characteristics are rarely ideal. Once again the ADC performance is specified in terms of offset error and gain error together with integral and differential non-linearities. The offset and gain error are each defined in terms of the input voltage at which the code transitions occur. Ideally, as shown in figure (19), for an ADC with a maximum input of V_{FS} these transitions occur when the ratio V_{in}/V_{FS} is an odd multiple of $\frac{1}{2}$ LSB.

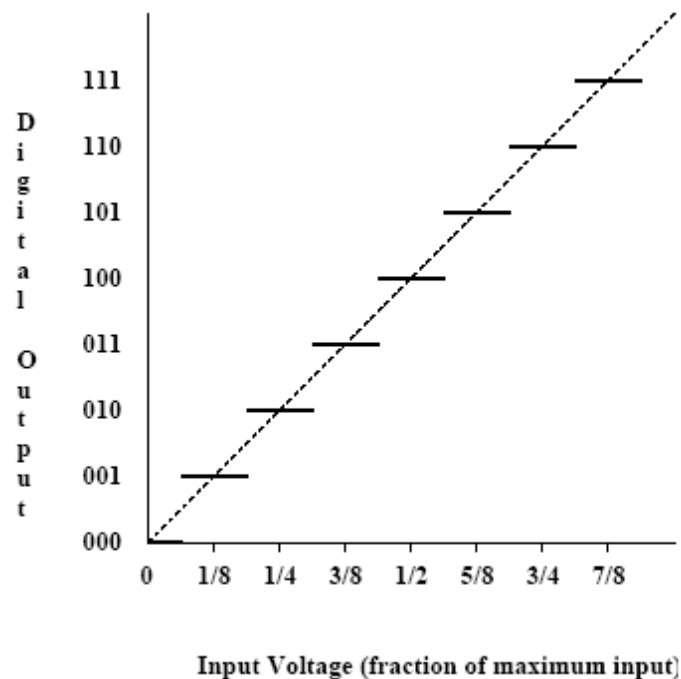


Figure 19: The ideal response of a 3-bit ADC.

The offset error is the difference between the input corresponding to the first code transition and ideal value. The gain error is the difference between the first and last transitions and the ideal value for this parameter.

Non-linearities of an ADC are defined with respect to the code centre line which is the line joining each of the mid-points of the measured code ranges shown in figure (20). The integral non-linearity is then the maximum difference between the code centre line and its ideal location. Similarly, the differential non-linearity is the maximum difference between neighbouring code transitions.

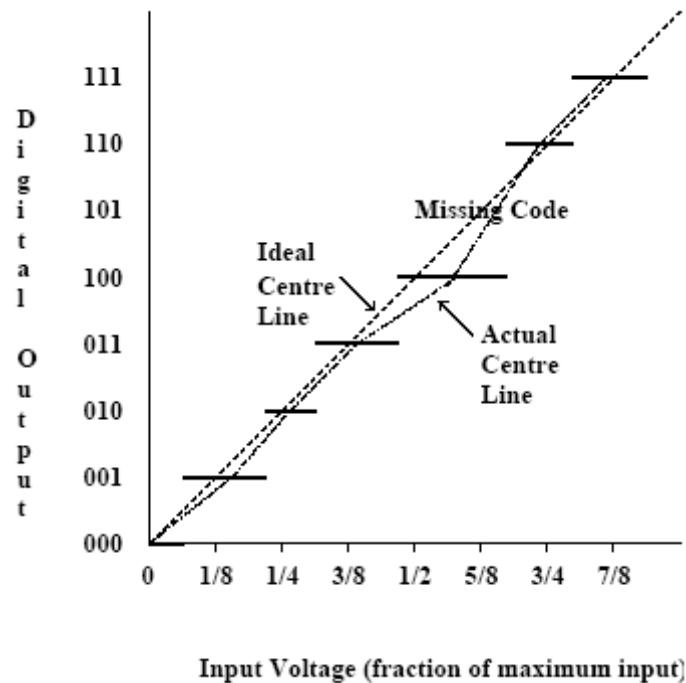


Figure 20: The response of an ADC showing the differential and integral non-linearities and a missing code.

Finally, figure (20) also shows one last possible error. In some situations it is possible for the ADC to have a missing output code that is never generated.

Oversampling Converters

Over the past few years a new type of conversion architecture has emerged for low and medium speed applications, for example high-quality digital audio. These new architectures exploit the increasingly low cost and high clock frequencies available in digital circuits to sample an input signal at many times the rate required to represent its maximum frequency. This oversampled digital signal is then processed by algorithms, known as digital filters, that filter the signal to reduce its bandwidth whilst increasing the number of bits representing the signal. The advantages of this approach are:

- (i) The requirements placed upon the accuracy of the analogue components of the converter are reduced.
- (ii) The digital filtering after sampling allows the requirements on the anti-alias filter to be relaxed.
- (iii) In many applications a sample-and-hold circuit is no longer required.

Mini Summary

DAC's and ADC's are important components of instrumentation and control systems. They convert analogue input signals into a digital format and enable the generation of analogue output signals to control actuators.

Like all electronic components **the actual performance of real DACs and ADCs are different from their ideal performance.** Several different measures of these performance errors are used to characterise each type of DAC or ADC. This information is contained in a component data sheet. However, these data sheets need to be read carefully when selecting a particular component to ensure that it will perform to the required accuracy.

FILTER CIRCUITS

Introduction

A filter is a circuit whose transfer function, that is the ratio of its output to its input, depends upon the frequency of the input signal. The resulting frequency selectivity of filters means that they are used to fulfil a variety of functions in instrumentation and signal conditioning systems. In particular they are a vital part of any well-designed analogue signal processing circuit between the sensor and the ADC.

Filters were part of the first year P2 course, so much of the material should be familiar.

Basic filter ideas

There are a number of ways to classify filters, but the simplest way is to classify them in terms of their frequency-dependent transfer function. We can write this as:

$$V_{out}(j\omega) = G(j\omega)V_{in}(j\omega)$$

where “G(j ω)” is the transfer function.

Note: Although this is very simple, there are variations in the way this transfer function is presented. For example, sometimes it is written without

the “j” in front of the ω , i.e. $V_{out}(\omega) = G(\omega)V_{in}(\omega)$. The advantage of including the “j” is that the same form can then be used with Laplace transforms, i.e.

$$G(j\omega) \rightarrow G(s)$$

The transfer function contains “information” about how the amplitude and phase of the signal is influenced when passing through the filter. It is common to “classify” filters in terms of how the amplitude is influenced. Many types of filter can then be classified, but the most important are:

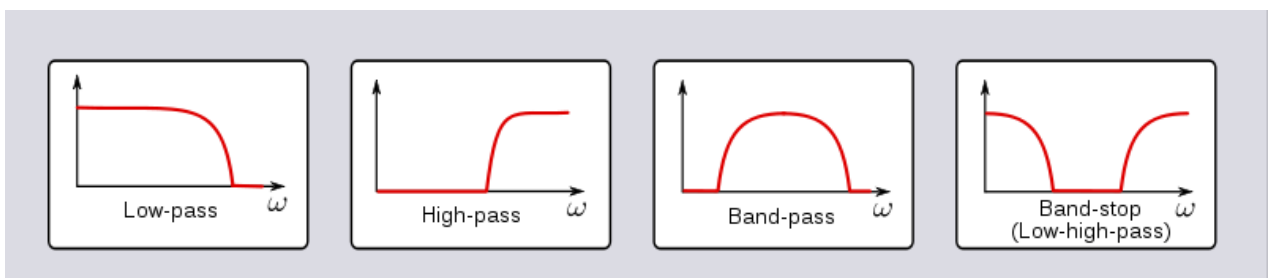
Low-pass filters allow any signal at a frequency below a characteristic frequency to pass (ideally unattenuated).

High-pass filters allow signals above a characteristic frequency to pass (ideally unattenuated).

Band-pass filters allow frequencies in a particular range to pass (ideally unattenuated).

Band-stop filters block frequencies in a particular range (ideally fully attenuated).

These are illustrated in the figure below:

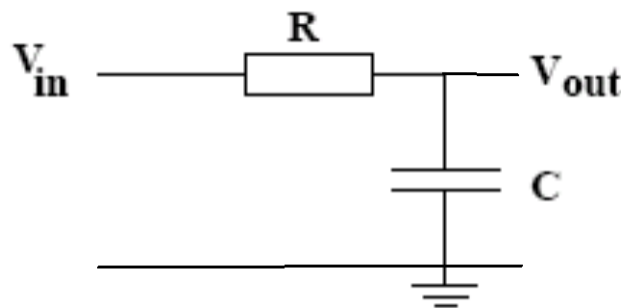


Of course, practical filters do not have ideal pass/block properties – therefore it is important to consider the behaviour of various basic filter elements.

First order filters

The most basic filter elements are first-order filters. This can be either low-pass or high-pass – other functions are not available with first-order filters.

The basic first order low pass filter consists of a resistor and capacitor:

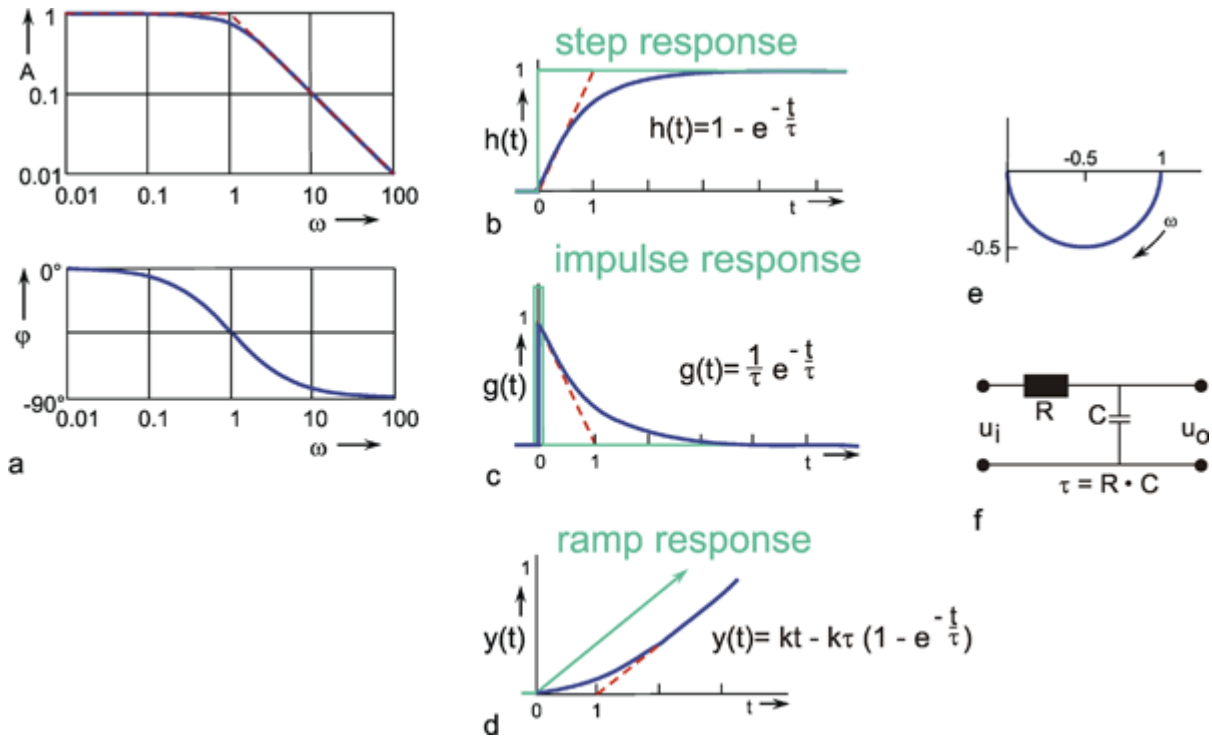


You should be very familiar with the analysis of this circuit (remind yourselves if not!), leading to a transfer function between the input and output of the form:

$$G(j\omega) = \frac{1}{1 + j\omega RC} \equiv \frac{1}{1 + j\omega T} \equiv \frac{1}{1 + j\omega / \omega_0}$$

where RC is equal to the time constant T (sometimes given the symbol τ) and the reciprocal of the cross-over (or 3dB) frequency ω_0 .

The behaviour of this basic filter element can be represented in various ways:



Here, (a) is a Bode-plot of the frequency response of the amplitude (A) and phase angle (ϕ) of the transfer function, where the time-constant is set to be equal to unity. Logarithmic axes are used for the amplitude and frequency, showing that above the cross-over frequency the amplitude of the transfer function drops by a factor of ten for each factor of ten increase in frequency. This is equivalent to 20dB per decade (remember dB is a logarithmic power scale). We can determine the magnitude (or amplitude) of the transfer function as an equation very simply:

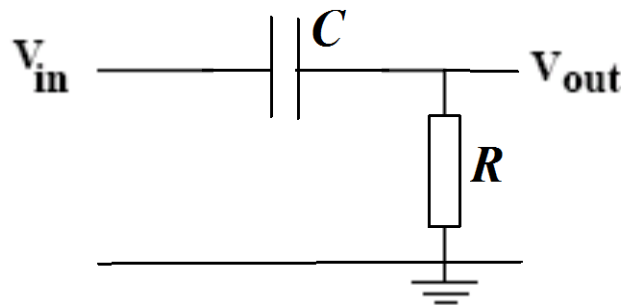
$$|G(j\omega)| = \left| \frac{1}{1 + j\omega/\omega_0} \right| = \sqrt{\frac{1}{1 + \omega^2/\omega_0^2}}$$

which at high frequency is equivalent to ω_0/ω , i.e. inversely proportional to ω , hence the 20dB per decade drop-off.

The plots (b), (c) and (d) show the amplitude responses as a function of time to a step, impulse and ramp input – these will be important when you consider control systems.

Plot (e) is a polar plot of the transfer function. To produce this plot the real and imaginary components of the transfer function are plotted on an argand diagram as the frequency is varied. For the basic first order low-pass filter we can see that when $\omega=0$ we have $G(j\omega)=1$, and as ω tends to infinity G tends to zero. Another “easy” point to see is that when $\omega=\omega_0$ we have $G = 1/\sqrt{2} - j/\sqrt{2}$. As ω varies from zero to infinite frequency we follow the semi-circle from unity around to zero. This polar notation for transfer functions will also be important when considering control systems.

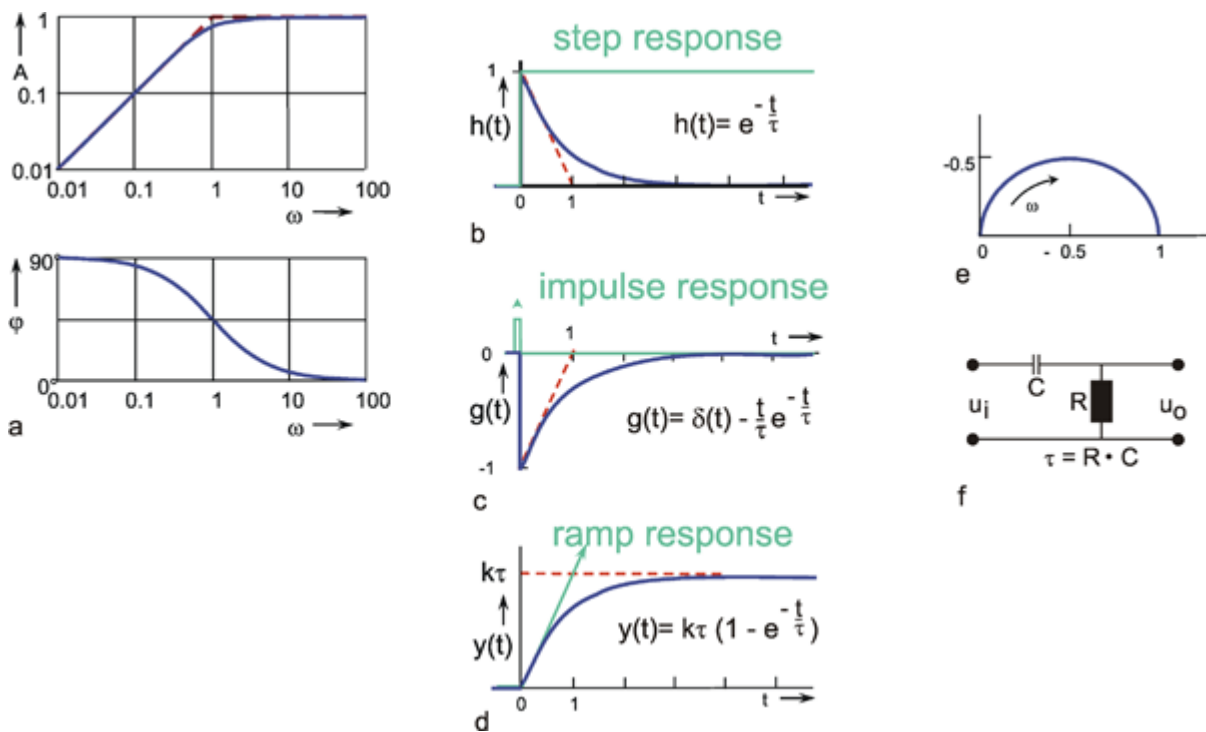
The basic first-order high-pass filter can also be constructed with a resistor and capacitor:



Again, you should be very familiar with the analysis of this (remind yourselves if not!), leading to a transfer function between the input and output of the form:

$$G(j\omega) = \frac{j\omega RC}{1 + j\omega RC} \equiv \frac{j\omega T}{1 + j\omega T} \equiv \frac{j\omega / \omega_0}{1 + j\omega / \omega_0}$$

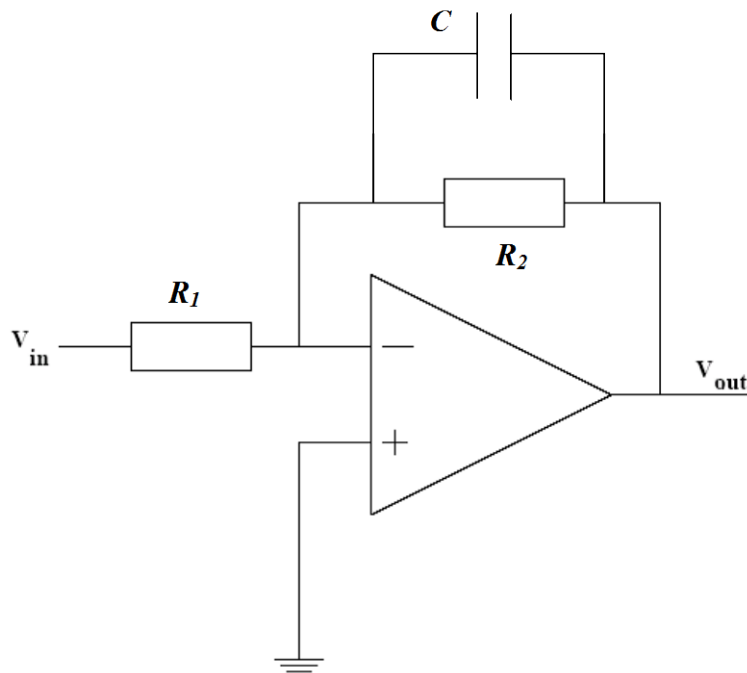
where RC is again equal to the time constant T (sometimes given the symbol τ) and the reciprocal of the cross-over (or 3dB) frequency ω_0 . Similar to the diagrams of the low-pass filter transfer function and responses, we have:



Now (a) shows a 20dB per decade cut-off BELOW the cross-over frequency (again chosen to be ω_0 is unity), and the polar plot (e) is “inverted”.

Active first order filters

The maximum “gain” of the passive filters shown above is unity. However, we can see how to engineer basic filter circuits with gain if we consider the inverting op-amp circuit in two variations.

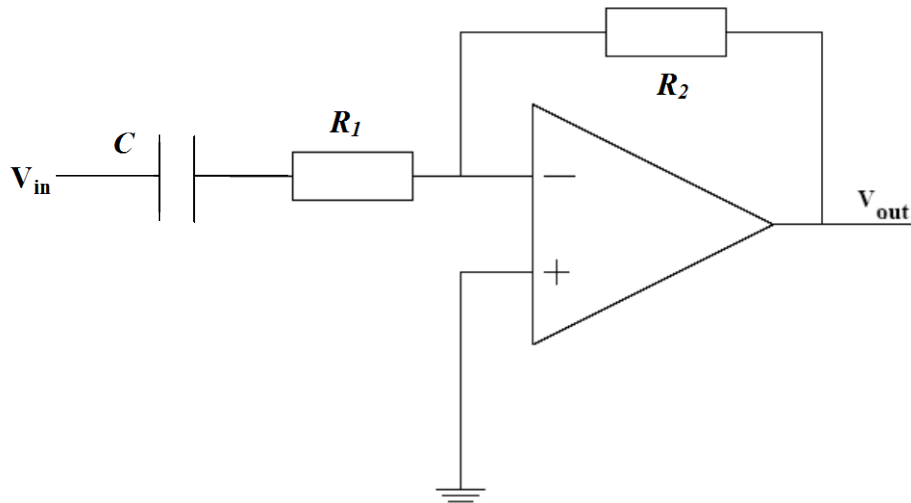


which has a transfer function:

$$G(j\omega) = -\frac{\left(\frac{R_2}{1 + j\omega CR_2}\right)}{R_1} = -\frac{R_2}{R_1} \left(\frac{1}{1 + j\omega CR_2}\right)$$

which is identical to the low-pass filter form shown above, but now multiplied by a “gain” of $-R_2/R_1$.

and



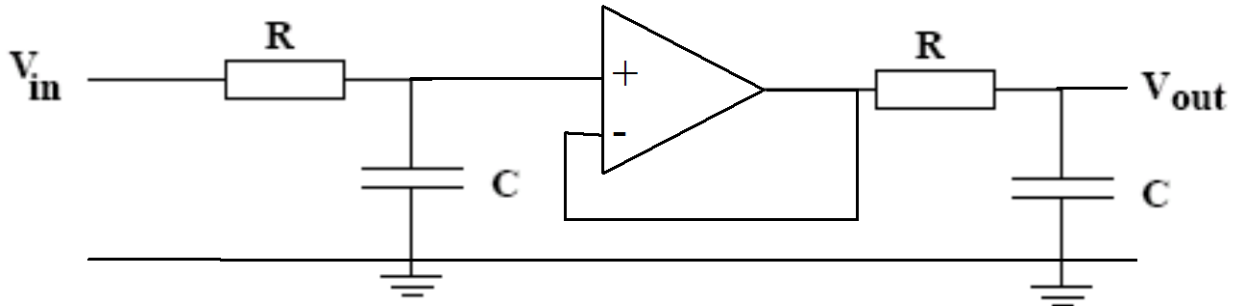
which has a transfer function:

$$G(j\omega) = -\frac{R_2}{R_1 + 1/(j\omega C)} = -\frac{R_2 j\omega C}{R_1 j\omega C + 1} = -\frac{R_2}{R_1} \left(\frac{j\omega R_2 C}{1 + j\omega R_2 C} \right)$$

which is identical to the high-pass filter form shown above, but now multiplied by a “gain” of $-R_2/R_1$.

Second order filters

Whilst the filter ideas introduced above are useful, they also have some limitations. Only low-pass and high-pass functions can be implemented and the cut-off is limited to 20dB per decade. What if we would like steeper cut-offs etc.? One basic answer is simple, we can chain filters together. For example if we connect two low-pass filters together (putting a buffer between them) we have:



This gives a transfer function equivalent to the first order low pass filter multiplied by itself, i.e:

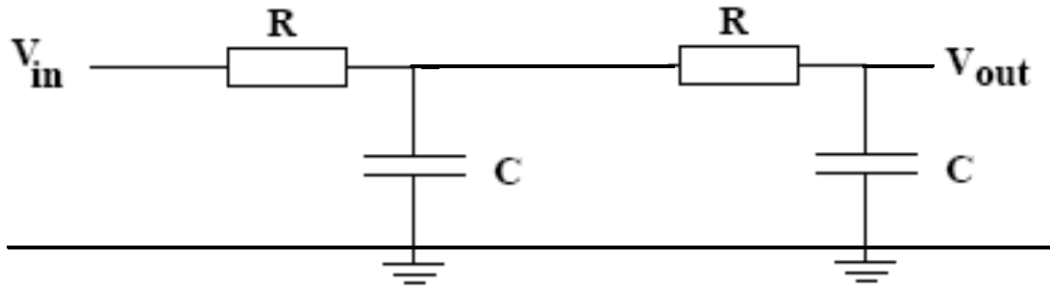
$$G_2(j\omega) = \frac{1}{1 + j\omega RC} \times \frac{1}{1 + j\omega RC} \equiv \frac{1}{(1 + j\omega T)^2} \equiv \frac{1}{(1 + j\omega / \omega_0)^2}$$

which for high frequencies becomes ω_0^2 / ω^2 . Therefore at high frequencies the magnitude is inversely proportional to ω^2 , whereas for the first order filter it was inversely proportional to ω . Hence, for this second-order filter the drop-off at high frequencies is 40dB per decade.

For the first order low-pass filter the transfer function is always of the form:

$$G(j\omega) = \frac{1}{1 + j\omega T} \equiv \frac{1}{1 + j\omega / \omega_0}$$

However, does the second order low-pass filter always have a transfer function of the form determined above? The answer is yes and no! This can be illustrated by determining the response if we had not included the buffer. The circuit is now:



To determine the transfer function of this we could use mesh/loop or node analysis – a very similar problem was on a P2 tutorial sheet. The resulting transfer function is:

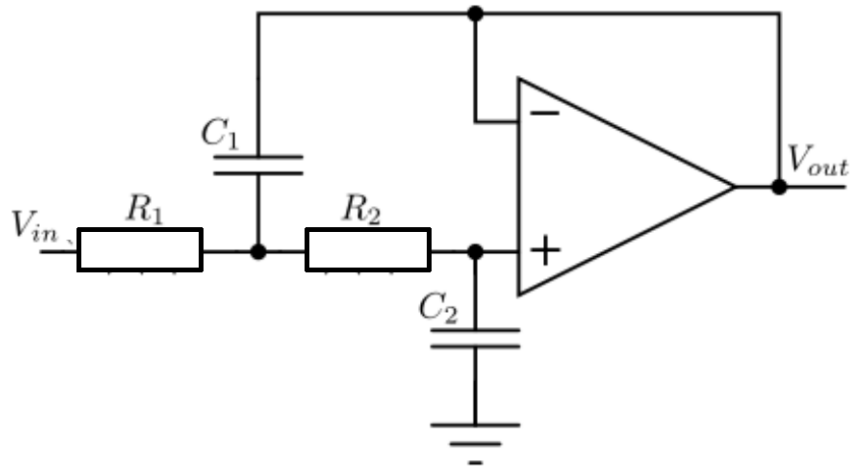
$$G_2(j\omega) = \frac{1}{1 + 3j\omega RC - \omega^2 R^2 C^2} \equiv \frac{1}{1 + 3j\omega T - \omega^2 T^2} \equiv \frac{1}{1 + 3j\omega / \omega_0 - \omega^2 / \omega_0^2}$$

which is not identical in form to the transfer function determined above, but has a similar overall structure. Note that it is the middle term which has changed, and in general we can write an expression for a second-order low-pass filter which is of the form:

$$G_2(j\omega) = \frac{1}{1 + 2\zeta j\omega / \omega_0 - \omega^2 / \omega_0^2}$$

where ζ is referred to as the damping. In the buffered case above $\zeta=1$ and in the un-buffered case $\zeta=1.5$.

In general, we can conveniently design a second-order low-pass filter where the damping ratio ζ and characteristic frequency ω_0 can be controlled, by using a Sallen-Key topology. Such a circuit is illustrated below:



For this circuit the transfer function is:

$$G_2(j\omega) = \frac{1}{1 + j\omega C_2(R_1 + R_2) - \omega^2 C_1 C_2 R_1 R_2}$$

so,

$$\omega_0 = \frac{1}{\sqrt{C_1 C_2 R_1 R_2}} \quad \text{and} \quad \zeta = \frac{C_2(R_1 + R_2)}{2\sqrt{C_1 C_2 R_1 R_2}}$$

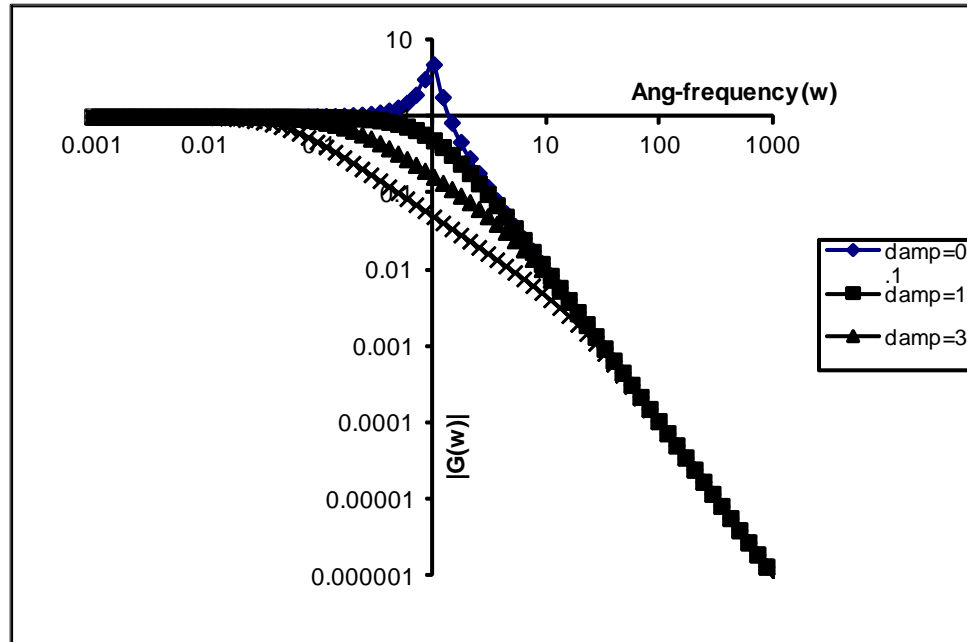
and given that there are four components and only two parameters to set we have enough freedom to independently control the damping ratio ζ and characteristic frequency ω_0 .

To illustrate the effect of varying the damping parameter a series of plots of $|G|$ for a second order low pass filter, where we have fixed $\omega_0=1$, but changed the damping. Taking the form used above:

$$G_2(j\omega) = \frac{1}{1 + 2\zeta j\omega / \omega_0 - \omega^2 / \omega_0^2}$$

⇒

$$|G_2(j\omega)| = \frac{1}{\sqrt{(1 - \omega^2 / \omega_0^2)^2 + 4\zeta^2 \omega^2 / \omega_0^2}}$$



For any damping we can see that when $\omega \ll \omega_0$ (in this case $\omega \ll 1$) then $G=1$ and when $\omega \gg \omega_0$ (in this case $\omega \gg 1$) $G = \omega_0^2 / \omega^2$, i.e. the roll-off is 40dB per decade. However, the behaviour around $\omega = \omega_0$ (in this case $\omega = 1$) varies substantially.

For large values of damping (greater than 1) the transfer function for the second order low pass filter can be factorised. For example, when the damping is 10 we have:

$$G_2(j\omega) = \frac{1}{1 + 20j\omega / \omega_0 - \omega^2 / \omega_0^2} \cong \frac{1}{1 + 19.95 \times j\omega / \omega_0} \times \frac{1}{1 + 0.05 \times j\omega / \omega_0}$$

So the second order filter behaves as the “product” of two first order filters, with cross-over frequencies of $\omega=\omega_0/19.95$ and $\omega=\omega_0/0.05$. This is evident from the diagram above, where when the damping is 10 we can see two cross over frequencies, the first at $1/19.95$ and the second at $1/0.05$, where between them we see a first-order roll-off of 20dB per decade.

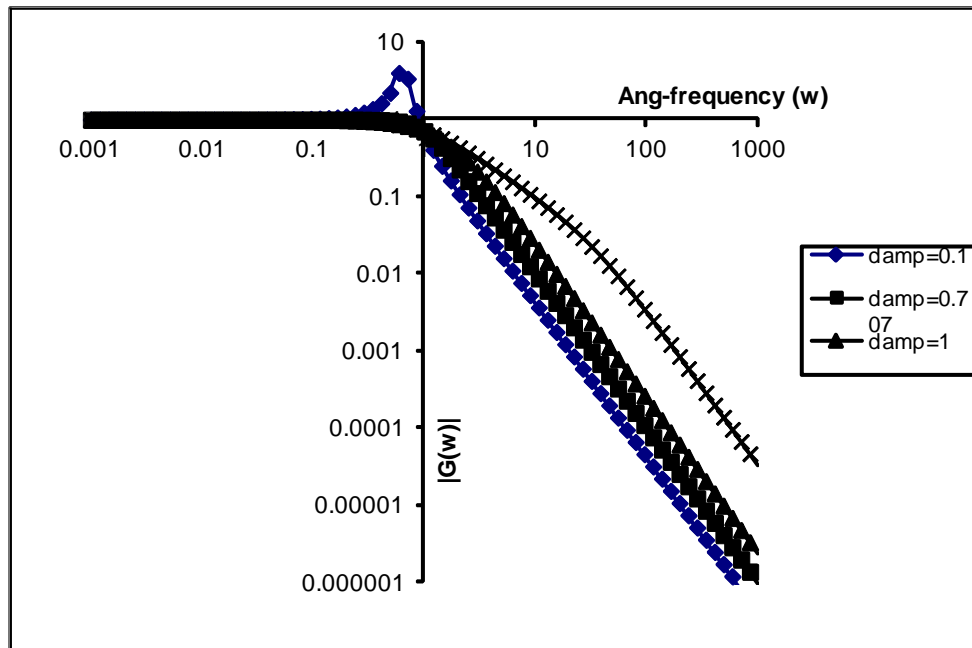
As the damping ratio is reduced the two cross-over frequencies come together, until when the damping is unity we reach the point where the behaviour is equivalent to two “cascaded” (and buffered) identical first-order filters as shown earlier. That is, when damping is unity:

$$G_2(j\omega) = \frac{1}{1 + 2j\omega/\omega_0 - \omega^2/\omega_0^2} = \frac{1}{1 + j\omega/\omega_0} \times \frac{1}{1 + j\omega/\omega_0}$$

If the damping is further reduced (less than unity) then it is less convenient to factorise into two first order terms. The factors in front of the $j\omega$ terms become complex! The resulting behaviour can be seen in the diagram above. For example, when the damping is 0.1, we see a resonance around the cross-over frequency.

It is easier to understand the relevance of the above to practical filter usage if we introduce an alternative way of characterising the low-pass filter. The cross over frequency is a key characteristic, but if we observe the behaviour in the responses shown above then we can see that the roll-off “appears” to begin at different points for different damping values. This can be allowed for by adjusting the cross-over frequencies so that the roll-off occurs at a similar point in each case. To do this we must select a point to characterise the

“start” of the roll-off. The point normally chosen is the -3dB point. That is where the gain of the low-pass filter has dropped by 3dB from its low frequency value. This is also where the amplitude of the transfer function has dropped by a factor of $\sqrt{2}$ (because $20\log_{10}(1/\sqrt{2}) = -3dB$). Adjusting the values of the cross-over frequencies in this way for a range of damping we get the plots shown below:



We can now see that for a fixed 3dB point, reducing the damping in a second order low pass filter causes the transfer function to roll-off earlier (qualitatively, this is sometimes referred to as “sharpening” the response). The cost of this is the potential introduction of a peak around the cross-over frequency. (Note: ω_0 now varies from curve-to-curve, and is not always =1!)

It turns out that the “best” which can be done without introducing any peak is when the damping is equal to $1/\sqrt{2} \cong 0.707$. (This has the advantage that ω_0 is the 3dB frequency – which avoids any possible confusion when

describing the filter) In this case the second-order low pass transfer function becomes:

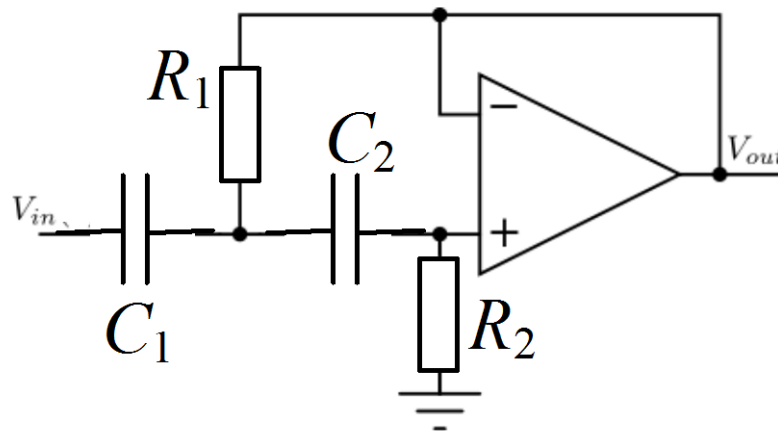
$$G_2(j\omega) = \frac{1}{1 + \sqrt{2}j\omega/\omega_0 - \omega^2/\omega_0^2}$$

and the amplitude of the transfer function is given by:

$$|G_2(j\omega)| = \sqrt{\frac{1}{1 + \omega^4/\omega_0^4}}$$

This is referred to as a second-order low-pass Butterworth filter, named after the British engineer Stephen Butterworth (~1930).

A high-pass second order filter can be constructed in a similar way to that outlined above, but swapping the resistors and capacitors, to give:



For this circuit the transfer function is:

$$G_2(j\omega) = \frac{-\omega^2 C_1 C_2 R_1 R_2}{1 + j\omega R_1 (C_1 + C_2) - \omega^2 C_1 C_2 R_1 R_2}$$

so with,

$$\omega_0 = \frac{1}{\sqrt{C_1 C_2 R_1 R_2}} \quad \text{and} \quad \zeta = \frac{R_1(C_1 + C_2)}{2\sqrt{C_1 C_2 R_1 R_2}}$$

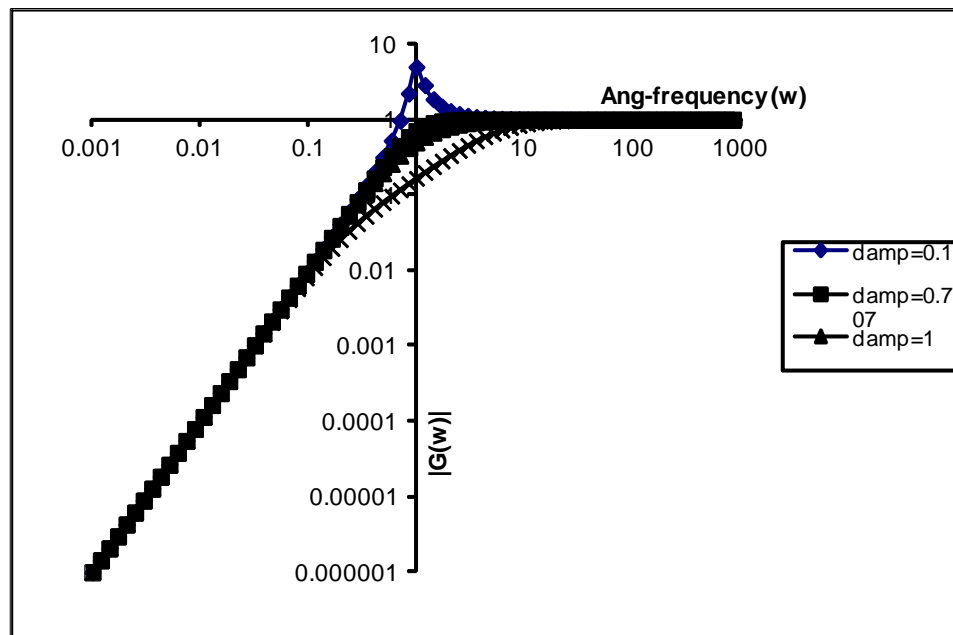
we have:

$$G_2(j\omega) = \frac{-\omega^2 / \omega_0^2}{1 + 2\zeta j\omega / \omega_0 - \omega^2 / \omega_0^2}$$

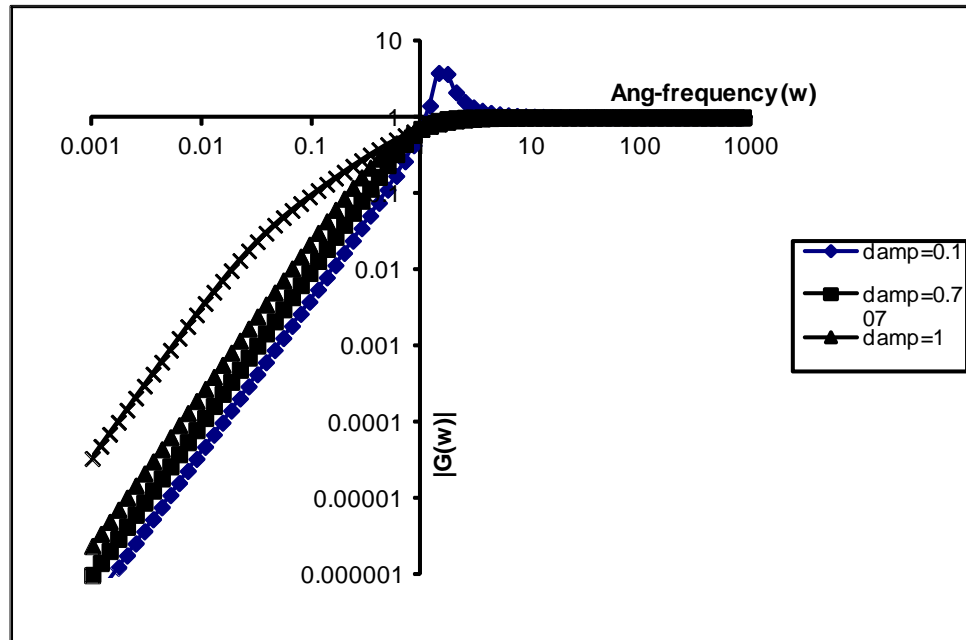
and the amplitude of the transfer function is then:

$$|G_2(j\omega)| = \frac{\omega^2 / \omega_0^2}{\sqrt{(1 - \omega^2 / \omega_0^2)^2 + 4\zeta^2 \omega^2 / \omega_0^2}}$$

This leads to amplitude Bode plots of the form:

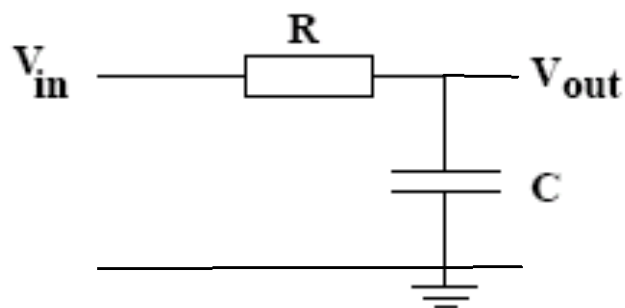


Or, if the cross-over frequencies are adjusted to make the -3dB points identical for each case we have:



Band pass

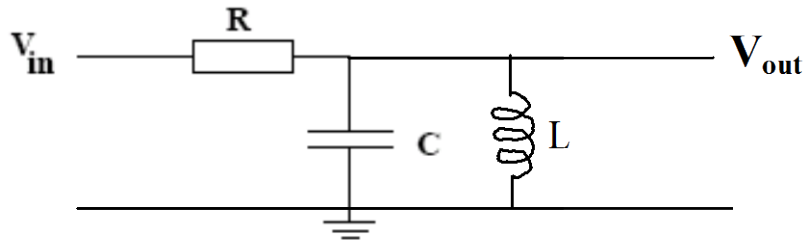
A further form of second-order filter is possible. A good way to see this concept is to re-consider the first order low-pass filter introduced earlier. Remember the basic first order low pass filter consists of a resistor and capacitor:



which resulted in a transfer function between the input and output of the form:

$$G(j\omega) = \frac{1}{1 + j\omega RC} \equiv \frac{1}{1 + j\omega T} \equiv \frac{1}{1 + j\omega / \omega_0}$$

Now, consider what happens if an inductor is placed in parallel with the capacitor, to give:



This changes the impedance of the reactive part from

$$\frac{1}{j\omega C} \quad \text{to} \quad \left(j\omega C + \frac{1}{j\omega L} \right)^{-1}$$

or, if we consider admittances it changes

$$j\omega C \quad \text{to} \quad j\omega C + \frac{1}{j\omega L}$$

If we substitute this into the low pass filter transfer function we generate a new transfer function:

$$G_{low}(j\omega) = \frac{1}{1 + j\omega / \omega_{0_low}} \rightarrow$$

$$G_{bp}(j\omega) = \frac{1}{1 + j\omega / \omega_A + \omega_B / (j\omega)}$$

$$= \frac{j\omega}{j\omega - \omega^2 / \omega_A + \omega_B}$$

$$= \frac{j\omega / \omega_B}{1 + j\omega / \omega_B - \omega^2 / (\omega_A \omega_B)}$$

which is the transfer function for a band-pass filter, conventionally expressed in the form:

$$G_{bp}(j\omega) = \frac{2\zeta j\omega / \omega_0}{1 + 2\zeta j\omega / \omega_0 - \omega^2 / \omega_0^2}$$

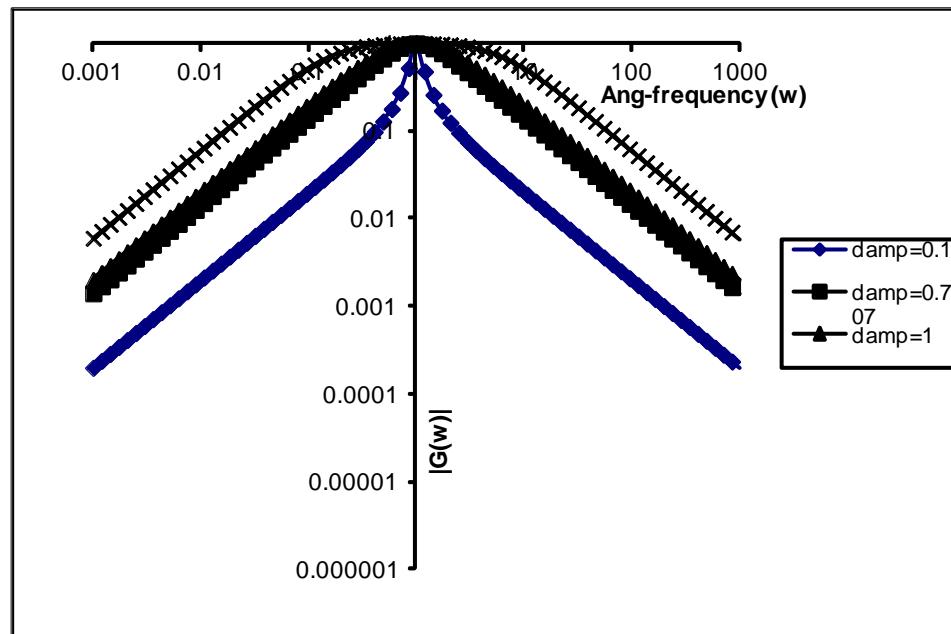
or, introducing the “quality factor” Q this can be written:

$$G_{bp}(j\omega) = \frac{\frac{1}{Q} j\omega / \omega_0}{1 + \frac{1}{Q} j\omega / \omega_0 - \omega^2 / \omega_0^2}$$

and the amplitude of the transfer function is:

$$|G_{bp}(j\omega)| = \frac{2\zeta\omega / \omega_0}{\sqrt{(1 - \omega^2 / \omega_0^2)^2 + 4\zeta^2 \omega^2 / \omega_0^2}}$$

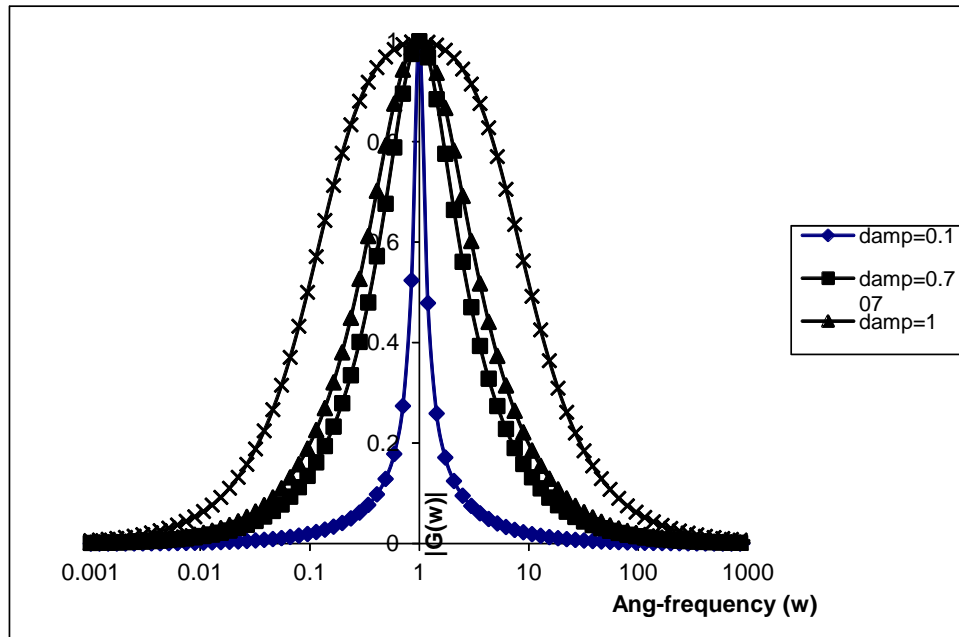
We can again plot the amplitude Bode diagram for various damping, giving:



In this case the curves are symmetric, and the cross-over (natural) frequency is the useful characteristic. It is also clear that at $\omega = \omega_0$ we have:

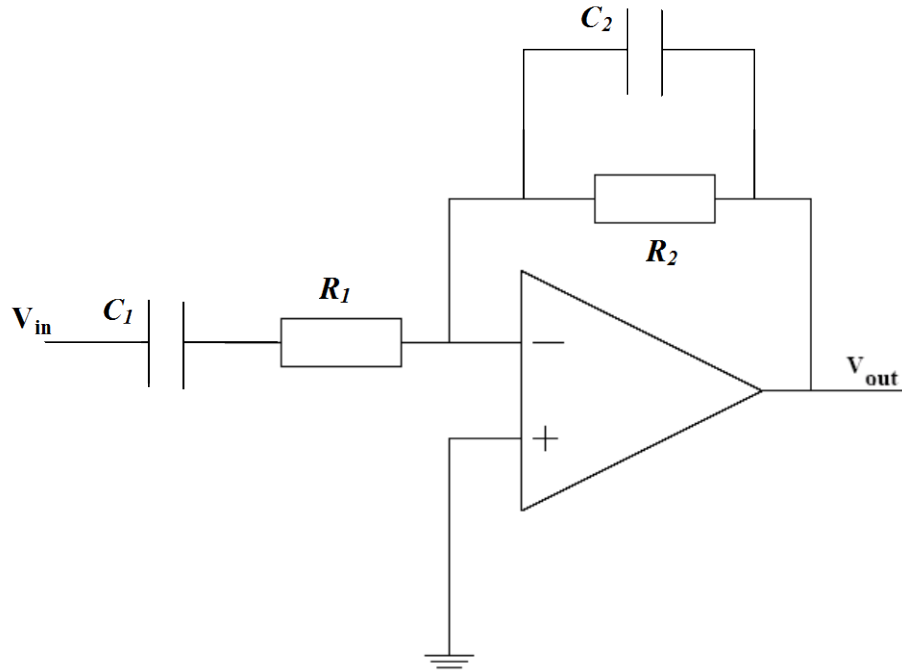
$$|G_{bp}(j\omega_0)| = \frac{2\zeta\omega_0 / \omega_0}{\sqrt{\left(1 - \omega_0^2 / \omega_0^2\right)^2 + 4\zeta^2\omega_0^2 / \omega_0^2}} = 1$$

The behaviour of this filter is also nicely illustrated if we plot $|G|$ on a linear scale, giving:



Clearly, the smaller the damping the narrower the range of frequencies that are passed (centred around the natural frequency).

This is exactly what happens with the resistor-capacitor-inductor circuit shown above. However, inductors can be inconvenient (they are both large and imperfect), and it is often more useful to use active circuits based on resistors and capacitors only (together with an op-amp). There are many active implementations of the band pass filter. A simple implementation is:



Recalling the well-known transfer function when only the resistors are present (i.e. $G = -R_2 / R_1$) we can easily see that we now have:

$$\begin{aligned}
 G(j\omega) &= -\frac{\left(\frac{R_2}{1 + j\omega C_2 R_2}\right)}{R_1 + \frac{1}{j\omega C_1}} \\
 &= -\frac{j\omega C_1 R_2}{1 + j\omega(C_1 R_1 + C_2 R_2) - \omega^2 C_1 R_1 C_2 R_2} \\
 &= -\left(\frac{C_1 R_2}{C_1 R_1 + C_2 R_2}\right) \frac{j\omega(C_1 R_1 + C_2 R_2)}{1 + j\omega(C_1 R_1 + C_2 R_2) - \omega^2 C_1 R_1 C_2 R_2}
 \end{aligned}$$

which is of the form introduced above for a second-order band-pass filter, but now with a gain (or loss) at the cross-over (natural) frequency given by the term in front.

General second order filter implementation.

In the P2 course you were introduced to a multiple feedback circuit design which could be used to implement second-order low-pass, band-pass and high-pass filters. The circuit is reproduced below:

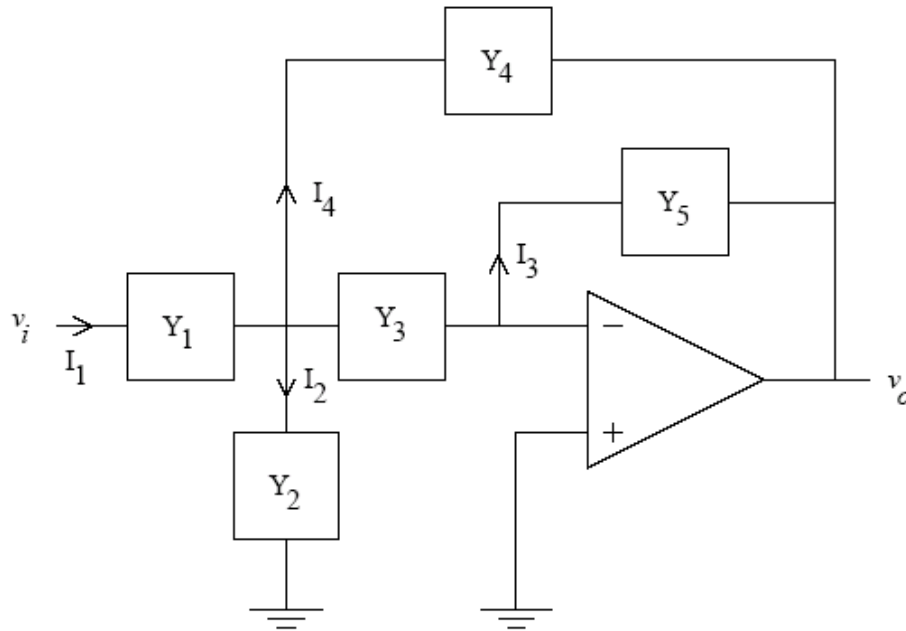


Figure 40: Generalised multiple feedback circuit.

Where the “Y”s represent the admittances of the various feedback components, i.e. $Y_n=1/Z_n$, so for a resistor $Y_R=1/R$ and for a capacitor $Y_C=j\omega C$. The analysis of the circuit was presented in the P2 course (if you are not confident you could analyse it if asked to do so then remind yourself of this because it is important!). The resulting transfer function is:

$$G = \frac{v_o}{v_i} = \frac{-Y_1 Y_3}{Y_5(Y_1 + Y_2 + Y_3 + Y_4) + Y_3 Y_4}$$

Now, by comparing this transfer function with the general transfer functions for the various filter types we can see that by selecting appropriate components they can be implemented. For example, if we use:

$$Y_1 = 1/R_1$$

$$Y_2 = 1/R_2$$

$$Y_3 = j\omega C_3$$

$$Y_4 = j\omega C_4$$

$$Y_5 = 1/R_5$$

we get the transfer function:

$$G_{bp}(j\omega) = \frac{-j\omega C_3 \frac{R_5}{R_1} \left(\frac{R_1 R_2}{R_1 + R_2} \right)}{1 + j\omega \left(\frac{R_1 R_2}{R_1 + R_2} \right) (C_3 + C_4) - \omega^2 C_3 C_4 R_5 \left(\frac{R_1 R_2}{R_1 + R_2} \right)}$$

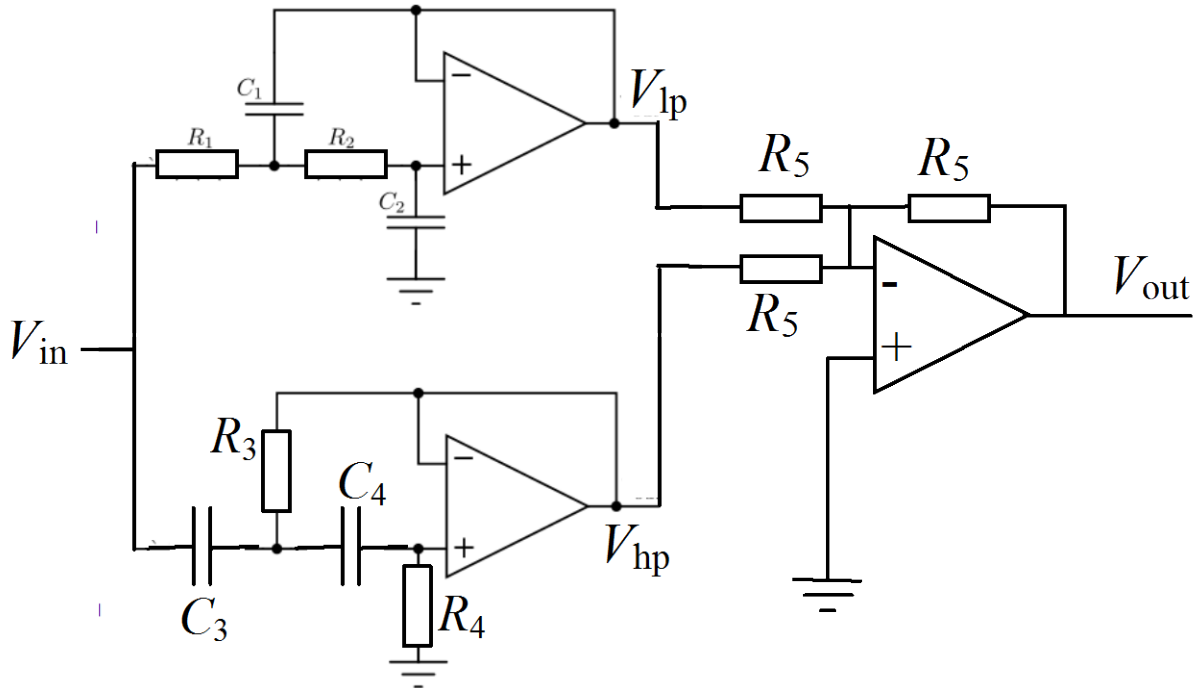
$$= - \left(\frac{R_5}{R_1} \frac{C_3}{(C_3 + C_4)} \right) \frac{j\omega \left(\frac{R_1 R_2}{R_1 + R_2} \right) (C_3 + C_4)}{1 + j\omega \left(\frac{R_1 R_2}{R_1 + R_2} \right) (C_3 + C_4) - \omega^2 C_3 C_4 R_5 \left(\frac{R_1 R_2}{R_1 + R_2} \right)}$$

which is identical in form to the general second-order band-pass filter transfer function introduced above.

Alternative component selection allows low and high pass to be implemented.

Band-stop filter

The band-stop filter (sometimes called a notch filter) is intended to remove signals around a narrow frequency range. For example we mentioned earlier that 50Hz mains interference is a common problem, so a 50Hz “notch” filter might be useful. A simple implementation is to feed a signal through both a second order low-pass filter and a second order high-pass filter simultaneously and then to combine the signals with a further op-amp. Consider the following circuit:



The transfer function for the circuit around the upper left op amp, between V_{in} and V_{lp} is given by:

$$G_{lp}(j\omega) = \frac{1}{1 + 2\zeta j\omega / \omega_0 - \omega^2 / \omega_0^2}$$

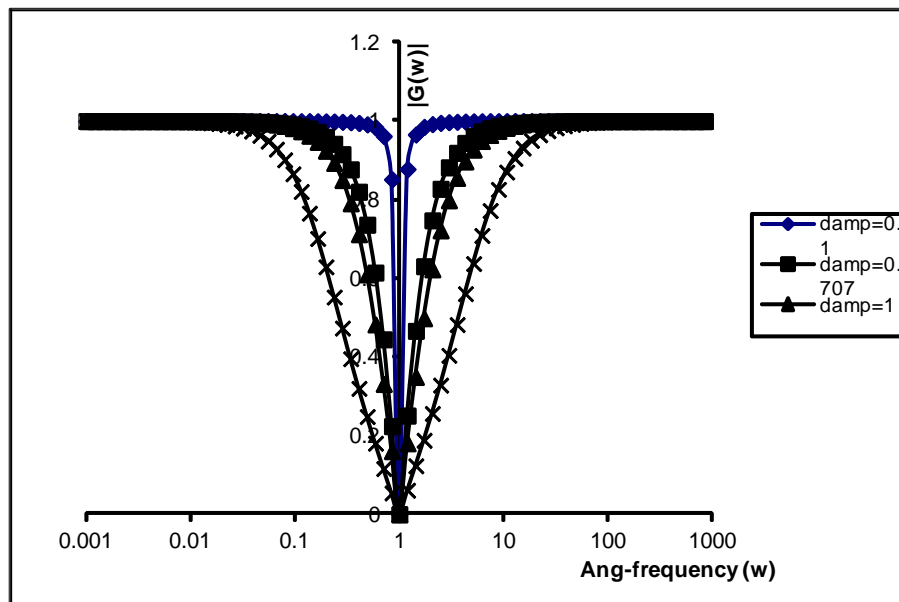
The transfer function for the circuit around the lower left op amp, between V_{in} and V_{hp} is given by:

$$G_{hp}(j\omega) = \frac{-\omega^2 / \omega_0^2}{1 + 2\zeta j\omega / \omega_0 - \omega^2 / \omega_0^2}$$

The circuit around the right hand op amp is then an inverting summing amplifier, so the overall transfer function between V_{in} and V_{out} is:

$$G_{bs}(j\omega) = \frac{\omega^2 / \omega_0^2 - 1}{1 + 2\zeta j\omega / \omega_0 - \omega^2 / \omega_0^2}$$

If we plot the modulus of this for various damping ratios (where we have fixed the cross-over (natural) frequency to be unity) we have:



Higher order filters

In the above material we have discussed first-order filters (low-pass and high-pass) and second order filters (low-pass, band-pass, high-pass and stop-band). Clearly it is possible to engineer “higher” order filters, which generally have steeper characteristics. For example, the first-order low-pass filter had a roll-off of 20dB per decade and the second order low-pass filter

had a roll-off of 40dB per decade. Clearly third, fourth, etc. order low-pass filters would have roll-offs of 60, 80, etc. dB per decade. These are created by connecting the correct number of first and second order filters in series.

Applications of filters

Noise rejection

Filters are often used to limit the noise bandwidth (see definition below) of the input signal in order to reduce the amount of noise in the input signal.

The quality of a signal in the presence of noise is most often specified by the signal-to-noise ratio (SNR) which is

$$SNR = 10 \log_{10} \left(\frac{V_s^2}{V_n^2} \right) \equiv 20 \log_{10} \left(\frac{V_s}{V_n} \right)$$

where V_s is the rms value of the signal and V_n is the rms value of the noise.

This definition shows that **a signal-to-noise ratio of one means that the noise power equals the signal power.** Initially, it would appear that **under these conditions it would be difficult to distinguish a signal from noise to obtain an accurate measurement.** The signal in figure (31) shows that **as expected with a SNR of one it is difficult to make an estimate of the signal component over a short interval,** in this case less than 0.1s. However, this figure also shows **that over a longer timescale it can be possible to distinguish a periodic signal from the random signal caused by noise.** In effect you can recognise the underlying 10 Hz signal

using its predictable periodic behaviour. In an instrumentation system the same effect can be achieved by filtering this voltage to reduce the amount of noise.

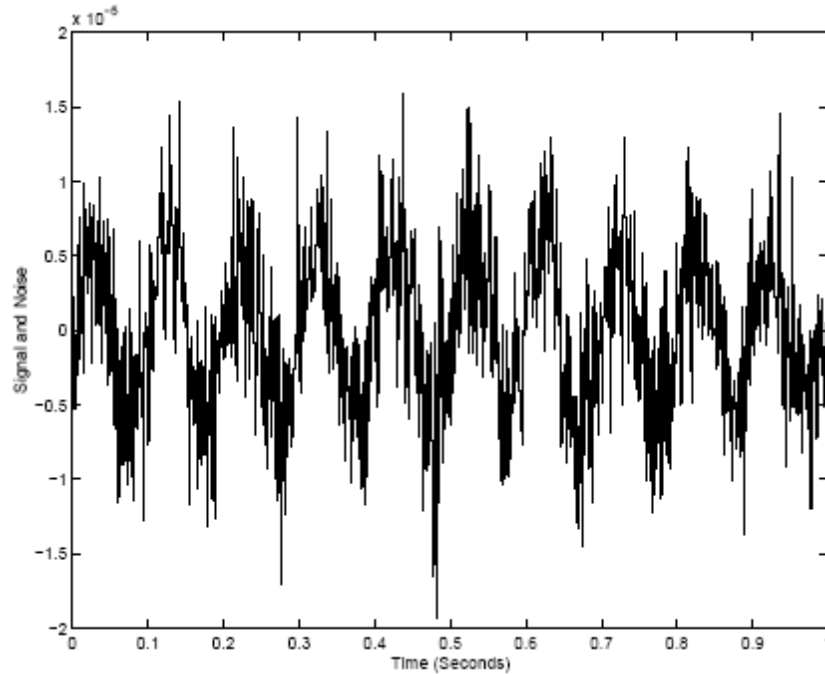


Figure 31: A signal at 10Hz with noise when the signal-to-noise ratio is one.

Signal Bandwidth and Noise Bandwidth

One of the useful functions of a low-pass filter is that it limits the noise bandwidth of the signal.

For example, the transfer function of a simple RC low-pass filter, with a characteristic frequency, f_c , can be written in the form

$$\frac{1}{1 + j \cdot f / f_c}$$

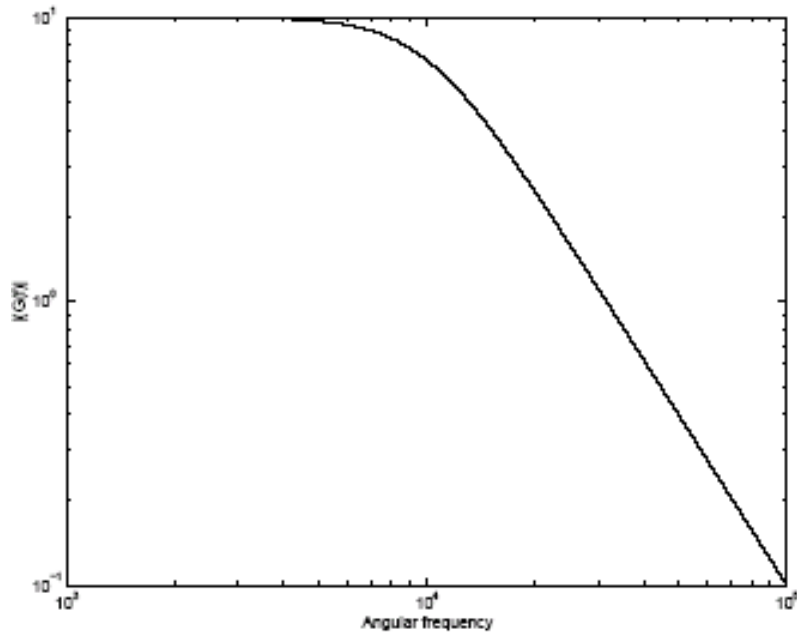


Figure 37: The frequency response of a simple RC low-pass filter.

This is a smooth transfer function and it has to be assumed that some attenuation of the frequencies of interest is acceptable. By convention it is usually assumed that this filter will be used for signal frequencies less than f_c . This means that it is assumed that it is acceptable to attenuate some signal frequencies by as much as 3dB and by convention the signal bandwidth of this filter is f_c .

The signal bandwidth of the low-pass filter is only determined by convention, and, in practice it is determined by the amount of acceptable signal attenuation in a particular application. **In contrast, the noise bandwidth unambiguously arises from a calculation of the noise at the output of the filter.**

For a white noise source the noise power in a small frequency range df is independent of the actual centre frequency. As with independent noise sources the total noise arising from contributions in different frequency ranges is determined by adding the power in each frequency band. Hence for a white noise source with a noise power of v_{wn}^2 per Hertz, the output power from the filter will be

$$V_{wn}^2 = \int_0^{\text{inf}} \frac{v_{wn}^2 df}{1 + (f/f_c)^2}$$

which can be integrated to give

$$V_{wn}^2 = v_{wn}^2 1.57 f_c$$

This is equivalent to the output of an ideal low-pass filter with a bandwidth of $1.57 f_c$. **This is the noise equivalent band-width of the filter. Using the characteristic frequency of the filter to determine the noise bandwidth of the filter would therefore significantly under-estimate the amount of noise that would be present at the output of the filter.** However, not unexpectedly, for filters with faster roll-offs at higher frequencies the noise equivalent bandwidth quickly approaches the characteristic frequency of the filter.

Anti-aliasing

Aliasing is a problem that arises when a signal that is continuously varying in time is sampled in time. If a high frequency signal is sampled too infrequently then once it is sampled it appears to be a signal at a different (lower) frequency. To avoid aliasing problems the

sampling frequency must be twice the maximum frequency of the input signal. This is the **Nyquist (or Nyquist–Shannon) sampling theorem**.

Increasing the sampling rate of a signal will prevent aliasing problems for signal frequencies, but, it can never prevent aliasing of high frequency noise. Analogue anti-aliasing filters are therefore included in all well designed instrumentation systems. However, good analogue filters, with a sharp roll-off, requires a large number of components and must be carefully designed. Even then they are not ideal. Once they have been sampled signal are therefore sometimes filtered by the digital processor.

AC SIGNALS IN INSTRUMENTATION – the lock-in amplifier

Introduction

In the earlier description of a bridge circuit it was assumed that a d.c. bias voltage was applied to the elements of the bridge. However, **there are some situations in which an a.c. bias voltage is either necessary or beneficial. One example of a situation in which an a.c. bias voltage is necessary is when taking measurements from a bridge circuit formed by capacitors.** The problem with using a d.c. bias voltage with this type of bridge circuit is that the d.c. impedance of all capacitors is infinite and it is impossible to detect a change in capacitance. In this case to reduce the impedance of the capacitors an a.c. bias voltage should be applied to the bridge circuit.

An example of a situation in which an a.c. bias voltage is beneficial is when there is a small signal expected near 50 Hz. Since this is the same frequency as the mains power supply, it is particularly vulnerable to interference. An a.c. bias voltage can then be used to shift the sensor output signal to a different frequency to separate the signal frequency of interest from the frequency of the interfering signal. Filters can then be used to reject the interference whilst amplifying the signal. To support these two critical functions an important system component is required – the lock-in amplifier. The lock-in amplifier removes the effects of the a.c. bias from the filtered and amplified output of the bridge circuit. It also acts as a narrow pass-band filter around the operating frequency.

Bridge Circuits and Capacitance Based Sensors

A change of resistance is only one of the means that can be used to create a circuit element that is sensitive to a physical variable. The other two properties that could be used are capacitance and inductance. An example of the use of inductance to measure a variable, is the use of a moving core within an inductor to detect displacement. In addition, changes in capacitance arising from variations in dielectric constant caused by absorbed gases can also be detected. This mechanism then forms the basis of gas detectors, including detectors for humidity and explosive or poisonous gases.

One approach to converting changes in inductance or capacitance to a voltage change is to simply form a bridge circuit containing either inductors or capacitors. However, the low-frequency impedance of an inductance is very small and its dc impedance is zero. A real dc power supply, with a small but finite output impedance, would be unable to sustain a voltage across this type of bridge. In contrast the impedance of a capacitor at low frequencies is very large. It would therefore be easy to sustain a dc voltage across a bridge circuit containing capacitors. However, this large impedance will make it impossible to sense the voltage across the bridge circuit without altering its value.

The problems with dc bridges containing inductors or capacitors can only be overcome if the impedance of the inductors can be increased and the impedance of capacitances can be decreased. Both these

objectives can be achieved using the same approach: applying an ac voltage of a known frequency to the bridge circuit.

Output Signal Processing

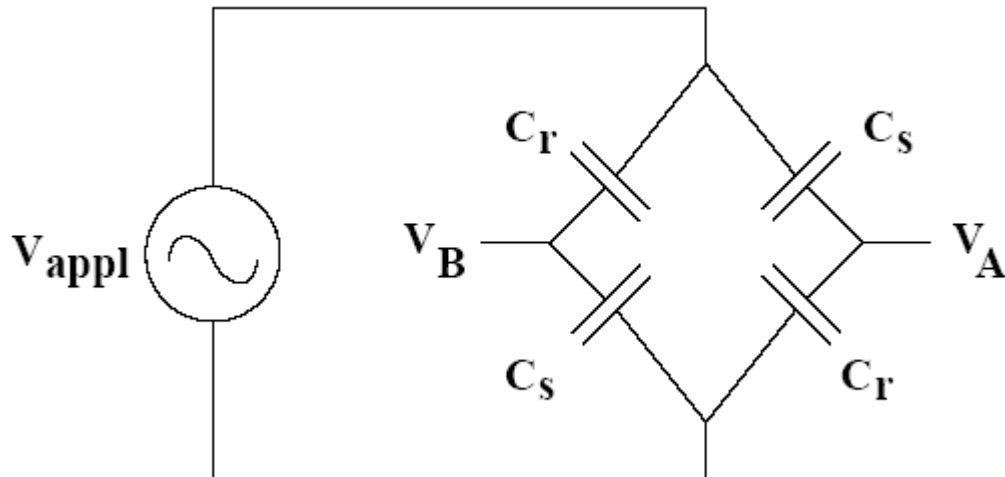


Figure 54: Bridge circuit formed with capacitors biased by an ac signal generator.

To understand the effects of applying an a.c. bias voltage to a bridge circuit consider a bridge formed from two sensor capacitors and two reference capacitors as shown in figure (54). If the voltage applied to this bridge circuit is V_{appl} , then the differential output voltage is

$$V_A - V_B = \frac{1/C_r - 1/C_s}{1/C_r + 1/C_s} V_{appl}$$

Now assume that the value of the sensing capacitor has been chosen so that

$$C_s = C_r + \Delta C$$

and the modulation of the capacitance value caused by the physical variable, ΔC , is only a small fraction of its average value so that $\Delta C \ll C_r$, then

$$V_A - V_B = \frac{\Delta C}{2 C_r} V_{appl}$$

The final stage of the analysis is to assume that the physical stimulus has a single dominant frequency so that

$$\frac{\Delta C}{2 C_r} = A_m \cos(\omega_m t)$$

where ω_m is the frequency of the physical effect itself, and that the bias voltage has the form

$$V_{appl} = V_a \cos(\omega_r t)$$

where ω_r is the frequency of the applied bias voltage.

Then the output signal from the bridge circuit is:

$$\begin{aligned} V_A - V_B &= A_m V_a \cos(\omega_m t) \cos(\omega_r t) \\ &= \frac{A_m V_a}{2} (\cos(\omega_r + \omega_m)t + \cos(\omega_r - \omega_m)t) \end{aligned}$$

Assuming that $\omega_r > \omega_m$ (reference/bias frequency is greater than the physical-effect/signal frequency) this means that the output signal contains two frequencies centred on ω_r separated by $2 \omega_m$.

With recent and continuing improvements in ADC performance it may be feasible to digitise this output signal before extracting the required information in a digital processor. However, in order to obtain a reasonable impedance in the bridge circuit it may be necessary to operate at frequencies of more than 100KHz. This could preclude the use of the cheaper successive approximation ADCs. Even in situations in which it may be possible to directly digitise the ac signal, it might therefore be cheaper to process the signal to reduce its frequency without losing relevant information. The sampling rate of the ADC will then be determined by the rate of change of the physical variable, ω_m , rather than the higher frequency of the applied ac stimulus, ω_r . **An instrumentation system that employs an ac bridge may therefore require a technique to remove the effects of the a.c. signal that has been applied to a bridge circuit.**

One approach to removing the applied ac signal is to use the circuit shown in figure (55). This circuit relies upon the existence of a reference signal, $r(t)$, which can be used to switch the input signal, $s(t)$, through one of two alternate paths. These two paths are designed so that one path acts as an amplifier with a gain of +1, whilst the other path acts as an inverting amplifier with a gain of -1. In effect this switching between two paths multiplies the input signal by a square wave with an amplitude of ± 1 .

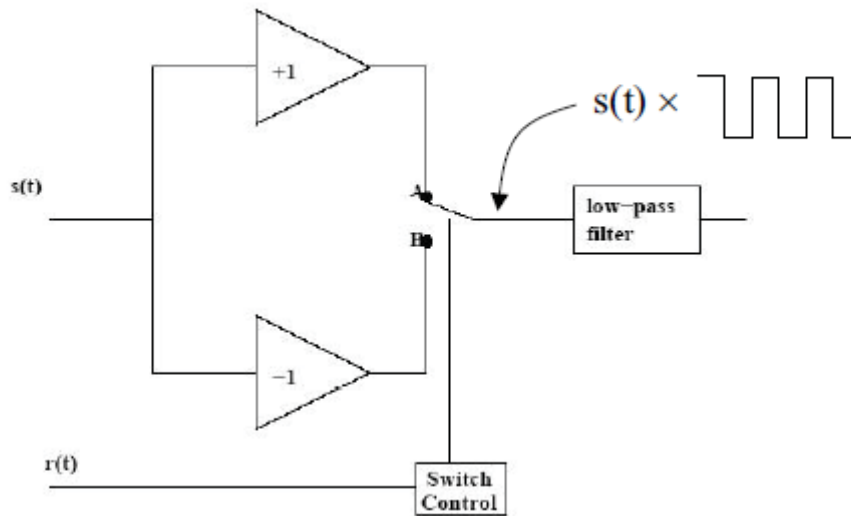


Figure 55: Block diagram of a lock-in amplifier.

To analyse the system assume that the reference signal $r(t)$ is a square wave of frequency f_r (where f_r is the frequency of the a.c. signal generator). By determining the position of an electronic switch between A and B this signal effectively multiplies the signal $s(t)$ by a square wave. The Fourier series of this type of square wave is (see HLT):

$$r(t) = \frac{4}{\pi} \left\{ \cos \omega_r t - \frac{1}{3} \cos 3\omega_r t + \frac{1}{5} \cos 5\omega_r t - \dots \right\}$$

This is then multiplied by the input signal

$$s(t) = A_m V_a \cos(\omega_m t) \cos(\omega_r t)$$

and hence the output from the switch multiplier is

$$v_p = \frac{A_m V_a}{\pi} \left\{ \cos(2\omega_r + \omega_m)t + \cos(\omega_m t) + \cos(2\omega_r - \omega_m)t + \cos(-\omega_m t) + \dots \right\}$$

The low-pass filter which follows the switch is then designed so that its cut-off frequency is significantly less than $2\omega_r - \omega_m$. Hence the output from the low-pass filter is:

$$v_{out} = \frac{A_m V_a}{\pi} |G(j\omega_m)| (\cos(\omega_m t) + \cos(-\omega_m t))$$

where $|G(j\omega_m)|$ is the magnitude response of the LPF at frequency ω_m and since $\cos(\omega_m t) = \cos(-\omega_m t)$

$$v_{out} = \frac{2 A_m V_a}{\pi} |G(j\omega_m)| \cos(\omega_m t)$$

The combination of a multiplier and a low-pass filter that respond to a narrow range of frequencies is a commonly used component of an instrumentation system. It is often referred to as a **lock-in amplifier**.

Interference and Noise Reduction

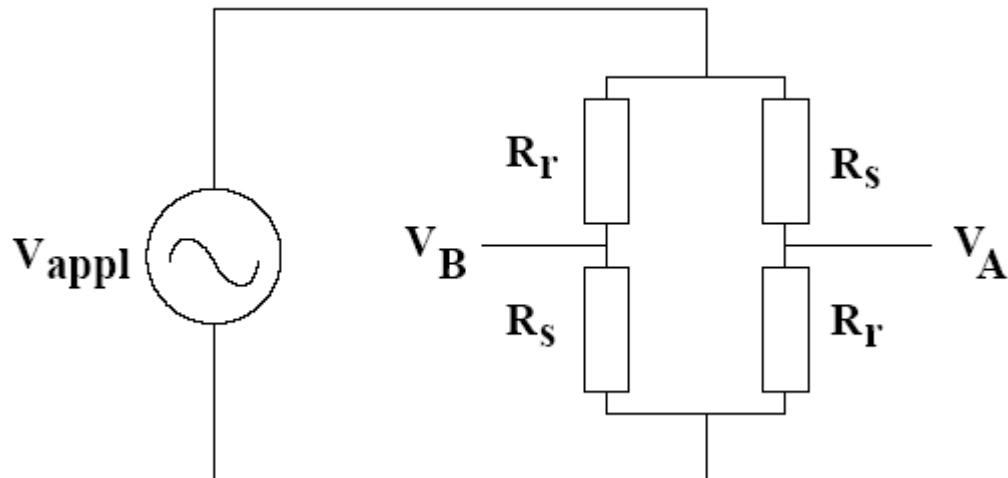


Figure 56: A bridge circuit of resistors with an ac input voltage.

The use of a lock-in amplifier has been introduced in the context of either all-capacitance or all-inductance bridge circuits. In these

situations the bridge circuit is biased with an a.c. signal in order to control the impedance of the elements within the bridge. The lock-in amplifier is then used to reduce the frequency at which the output signal has to be sampled without losing any information about the physical process being monitored. However, **lock-in amplifiers can also be used in other systems in order to avoid strong sources of interference or noise.**

To understand how a lock-in amplifier can be used to avoid interference or noise consider a bridge formed from two sensor resistors and two reference resistors. If the voltage applied to this bridge circuit is V_{appl} , then the differential output voltage is

$$V_A - V_B = \frac{R_r - R_s}{R_r + R_s} V_{appl}$$

Now assume that the value of the sensitive resistor has been chosen so that

$$R_s = R_r - \Delta R$$

and the modulation of the resistance value caused by the physical variable, ΔR , is only a small fraction of its average value so that $\Delta R \ll R_r$, then

$$V_A - V_B = \frac{\Delta R}{2 R_r} V_{appl}$$

Problems arise when the system has to be designed to detect small changes in resistance. In this situation small changes in the output signal will be vulnerable to either external interference or noise arising within subsequent analogue circuits. **Some protection from interference can be obtained by filtering the signal.** However, any filters must be designed to pass frequencies that contain information about the physical variable that is

being measured. It can then be very difficult to obtain a filter with a sharp enough cut-off to reject a strong source of interference and in the worst-case situation the interfering signal may occur within the interesting frequency range. Analogue filtering is then impossible and there is a risk that the interfering signal could sometimes be large enough to cause saturation of the input signal to the ADC.

The solution in these situations is to apply an a.c. bias voltage to the bridge circuit

$$V_{appl} = V_a \cos(\omega_r t)$$

To understand the effect of this consider a situation in which the physical stimulus has a single dominant frequency, ω_m , so that

$$\Delta R/2R = A_m \cos(\omega_m t)$$

In this situation the output signal from the bridge circuit is:

$$V_A - V_B = A_m V_a \cos(\omega_m t) \cos(\omega_r t)$$

This can be rewritten in the form

$$V_A - V_B = \frac{A_m V_a}{2} \cos((\omega_r + \omega_m)t) + \cos((\omega_r - \omega_m)t)$$

to show that the output now contains two signals at frequencies $\omega_r + \omega_m$ and $\omega_r - \omega_m$.

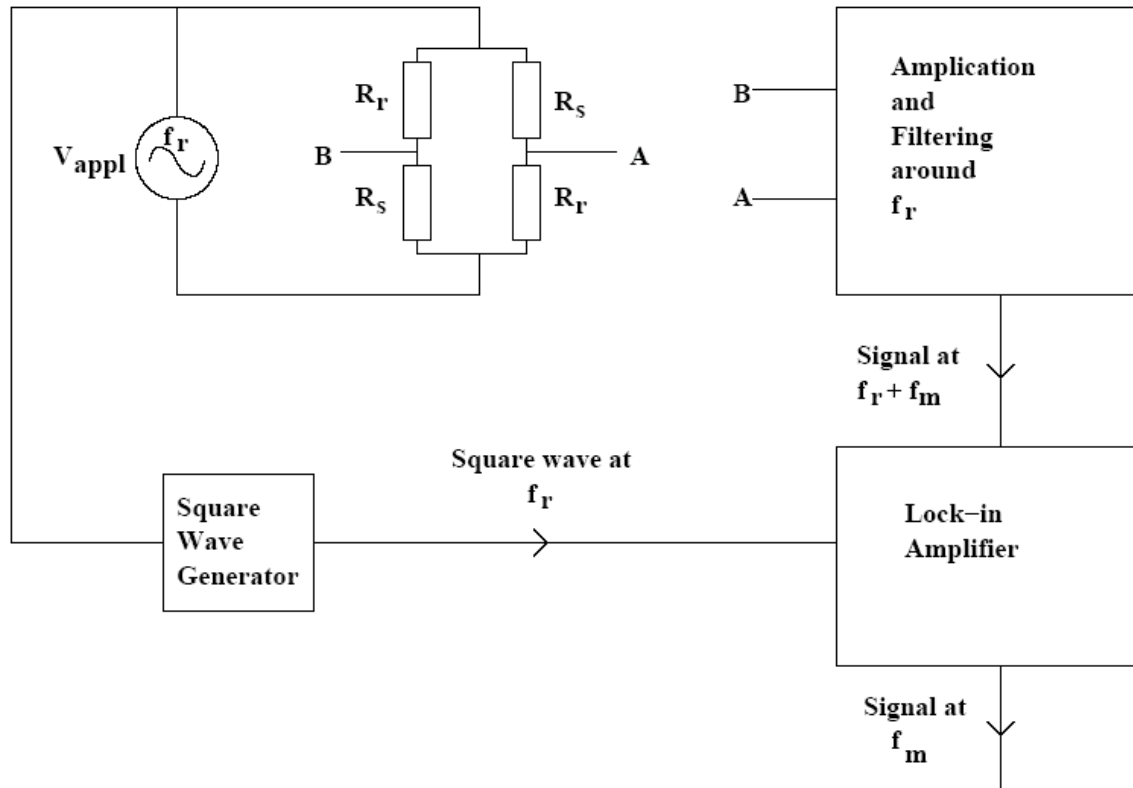


Figure 57: A schematic diagram of the component parts required to shift the intermediate signal frequency in order to avoid noise and or interference. Once all the analogue amplification and filtering has been performed in a frequency range chosen by the designer a lock-in amplifier can be used to reduce the output frequency (Note that the two points labelled A are connected together as are the points labelled B)

Applying an a.c. signal of known frequency to the bridge circuit therefore enables the designer to create an output signal with two well controlled frequency components. Although the amplitude of these two components is small the designer can select the value to ω_r in order to avoid frequencies at which interference is expected (usually the frequency of the local mains and/or its harmonics). Furthermore, **the value of ω_r can be chosen to shift the output frequencies so far from the frequency of**

any interference that even a relatively simple band-pass filter can amplify the signal at frequencies around ω_r whilst rejecting interference. Once the signal has been amplified without amplifying any interference its frequency can be reduced using a switching multiplier and low-pass filter prior to digitisation without fear of degradation.

Phase-Difference Detection or Phase-Sensitive Detection (not really part of this course, but rather interesting....)

Another approach to detecting a change in capacitance is to place a reference resistor in series with a variable sensor capacitance to create a low-pass filter. If the input voltage to this circuit is close to the 3dB frequency of the filter then variations in the capacitance values will cause a measurable change in the phase difference between the input signal to the filter and its output. This phase difference can then be detected using a combination of a switching multiplier and low-pass filter

To understand how this can occur consider a difference between the phase of two signals consider the situation in which the transducer modulates the **phase** of a signal with a known frequency. In this situation the input signal can be expressed as

$$s(t) = V_s \cos(\omega_r t + \phi)$$

Once again assume that the reference signal $r(t)$ is a square wave of frequency f_r and amplitude ± 1 , equivalent to a Fourier series of:

$$r(t) = \frac{4}{\pi} \left\{ \cos \omega_r t - \frac{1}{3} \cos 3\omega_r t + \frac{1}{5} \cos 5\omega_r t - \dots \right\}$$

The output v_p of the switching multiplier is then given by the product $r(t)s(t)$.

Thus:

$$v_p = \frac{2V_s}{\pi} \left\{ \cos(2\omega_r t + \phi) + \cos \phi - \frac{1}{3} \cos(4\omega_r t + \phi) - \frac{1}{3} \cos(2\omega_r t + \phi) + \frac{1}{5} \dots \right\}$$

The 3-dB cut-off frequency of the low-pass filter ω_c is chosen to be well below ω_r . Hence the multiplier products at frequencies $2\omega_r$, $4\omega_r$, $6\omega_r$, etc are eliminated and the output of the Phase-Sensitive Detector contains only the phase-sensitive d.c component:

$$v_{out} = \frac{2V_s}{\pi} \cos \phi$$

(assuming that the d.c. gain of the LPF is 1) **Hence using this circuit it is possible to detect any phase difference between the reference signal and input signal. The circuit is therefore often referred to as a phase sensitive detector.**

Mini summary

Alternating current (a.c.) biasing signals can be used in conjunction with bridge circuits containing inductors or capacitors to control the impedance of the elements of the bridge circuit so that a measurable output signal can be created.

Alternating current (a.c.) biasing signals can be used with bridge circuits of resistors in order to shift the frequency of the output signal to a pre-determined frequency range. Filter circuits can then be designed in the analogue signal processing stage which can amplify the signal but reject interference whose frequency would otherwise be too close to the signal frequency of interest.

A lock-in amplifier can be used after amplification and filtering to remove the effects of the a.c. bias. The maximum required ADC sampling rate is then determined by the rate of change of the physical variable being measured rather than the much higher a.c. frequency chosen by the designer.

The key component of the lock-in amplifier is a circuit that can also be used to detect the phase between two signals of the same frequency. It is therefore often known as a phase sensitive detector.