Analogue Electronics 9: Improving on RC filters

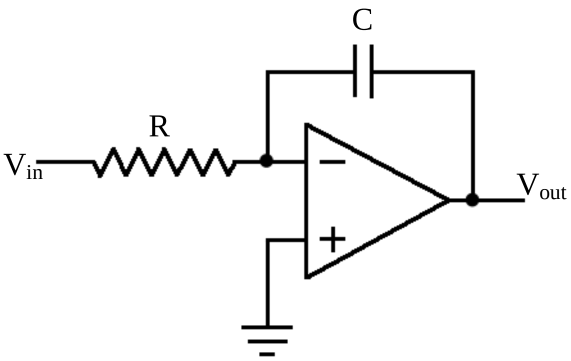
In the last lecture we introduced the Op Amp and basic ideas about feedback. In checkpoint A2 you are going to use an Op Amp first as an *amplifier* and then as a *filter*. A while back we started looking at filters built using only resistors and capacitors (RC filters). Using Op Amps it is possible to obtain filter performance which is *more ideal*.

Contents:

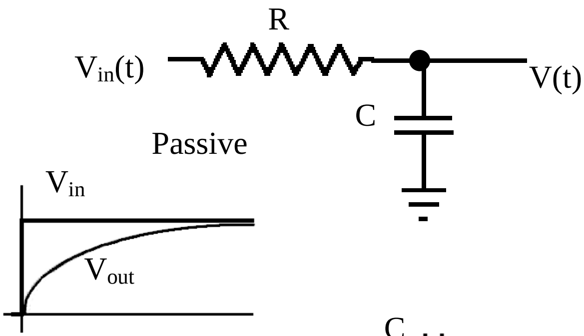
- Integrators & differentiators
- Active filters
- Frequency response of RC filters

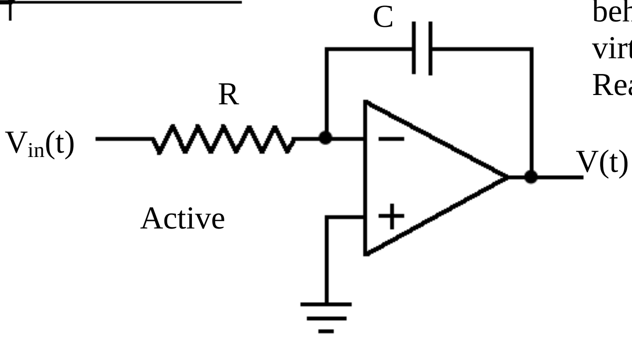
Integrators: (H&H, 4.19, p. 222)

Our initial op-amp circuits just involved resistors. These could be used to amplify, add and subtract. Here we are going to look at op-amp circuits involving *capacitors* as well. We begin by replacing the *feedback* resistor with a *capacitor*.



Intuitively one might expect that this would mean that only rapidly varying parts of the output are fed back for comparison with the input. Using the golden rules we will show that this circuit actually performs an *integration of the input voltage*. We encountered an integrator previously in circuits only involving resistors and capacitors. The op-amp version of this **integrator** proves to be a significant improvement on the passive version.





With the op-amp circuit **all** the variation in time can be across the capacitor.
The feedback loop mirrors the behaviour (because of the virtual ground).
Readout remains possible.

From the golden rules we deduct:

- the inverting input must be a virtual ground
- in addition the op-amp draws no current. Hence whatever current is feeding back through the capacitor must be cancel current through the resistor.

Therefore the current I = V_{in} / R flows through the capacitor, i.e.:

$$\frac{V_{in}}{R} = -C\left(\frac{dV_{out}}{dt}\right)$$

$$V_{out} = -\left(\frac{1}{RC}\right)\int V_{in}\,(t)\;dt$$

This *avoids the restriction to small signals* that was essential to the circuit which involved only a resistor and a capacitor. In this circuit V_{out} does not need to be small, here the limit is only given by the output capability of the op-amp.

Direct RC – op-amp comparison:

The RC circuit has the restriction that: $V_{out} << V_{in}$. That limitation keeps the circuit in a regime where the *capacitor charging curve* still looks roughly straight. If this condition is violated then the exponential character of the capacitor charging will become evident. This is because the desired behaviour of the **passive integrator** circuit relies on the relationship between voltage and current for a capacitor (I = C dV/dt).

With an op-amp *the point from which the capacitor is charging stays at 0V* since it is a virtual ground. At the same time it is possible to take the signal out at the other end of the capacitor, since *the op-amp provides it*. Thus, there is no intrinsic restriction on V_{out} in this circuit.

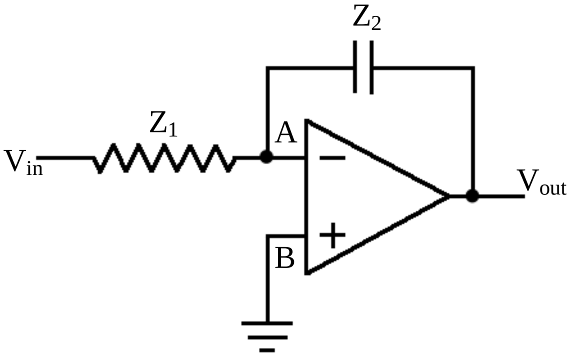
So, the **active integrator** made from an op-amp circuit *seems like a perfect integrator*. But what do you expect to happen if you put a voltage across the input for a long time? You guessed it, it will hit a limit. This limit is imposed by the power supply for the op-amp. Actually it is a voltage slightly smaller than the external supply voltage which cannot be exceeded by the output of the op-amp. It will **saturate**. Care does need to be taken to *avoid having V_{out} saturate*. The signal will distort and your measurement may be spoiled beyond recovery.

Note: if the input that needs to be integrated is already a *current* rather than a voltage then the *input resistor R is no longer required*.

Integrators / Low pass filters: (Frequency analysis)

The circuits that perform calculus can also be thought of as filters. **Filters** are circuits which *attenuate signals within a certain frequency range* while allowing signals outside this range to pass unchanged.

As with the passive circuits before we first looked at the behaviour in the time domain, and have found that *the circuit performs calculus*. We will now look at the same circuit in the frequency domain, where we will again describe its behaviour as a low pass filter.



Here we can revert to using *complex impedances*. We can then analyse the circuit using the same formalism as was used for resistors alone. This approach relies on the idea that *any time varying input voltage* can be represented as a series of *sine waves* of different *amplitudes*, *phases* and *frequencies* ω.

$$Z_1 = R$$
$$Z_2 = -\frac{i}{\omega C}$$

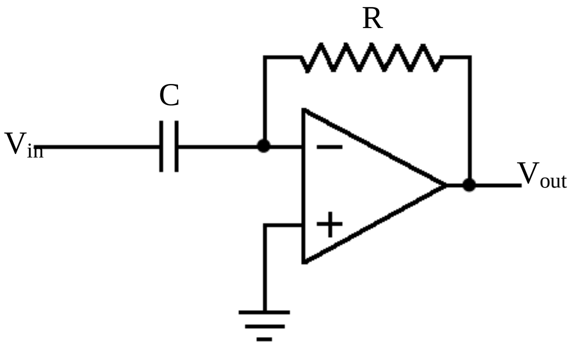
Using the golden rules we find:

- Rule 1: Since point B is at ground then point A must be too.
  Consequence: the voltage across Z_2 is V_{out} and the voltage across Z_1 is V_{in}
- Rule 2: Since no current flows into the inverting input, all current flowing through the feedback must flow towards the input.

$$\frac{V_{out}}{Z_2} = -\frac{V_{in}}{Z_1}$$
$$\frac{V_{out}}{V_{in}} = -\frac{Z_2}{Z_1} = \frac{i}{\omega RC}$$

We have a circuit which amplifies signals by varying amounts *depending on the frequency*. Low frequency signals are amplified more whereas very *high frequency signals are strongly attenuated*. The amplification factor is also *complex*, which indicates that the output of the circuit will have a *different phase to the input*. This circuit is known as a **low-pass filter**, because it allows low frequency signals through while *strongly attenuating high frequency signals*.

Differentiators: (H&H, 4.20, p. 224)
Here the positions of resistor and capacity around the op-amp circuit are swapped, the capacity in the input and feedback via the resistor. It is *only the rapidly varying part* of the input signal that will make it through to the amplifier.



From the golden rules we find:
The inverting input is a virtual ground and the op amp draws no current. Whatever current arrives at the inverting input must be cancelled by a current flowing through the feedback path. The rate of change of input voltage produces a current:

$$I = C\left(\frac{dV_{in}}{dt}\right)$$

This is balanced by the current through the feedback resistor, $I = V_{out}/R$. Thus:

$$V_{out} = -RC\left(\frac{dV_{in}}{dt}\right)$$

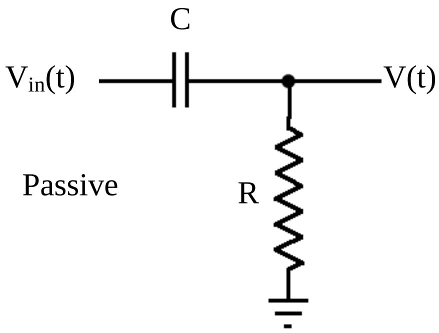
This is *perfect differentiation* – with no approximations. In contrast to the integrator there are *no problems with drift or saturation* (an integral will augment slowly over time whereas a differential does not).

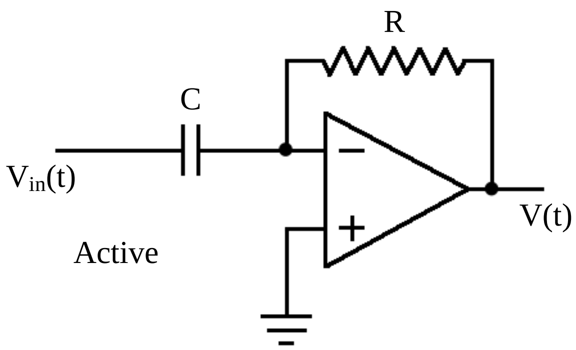
But it probably doesn't surprise you to learn that real-life op-amps *can't respond to arbitrarily fast input signals*. In the case of differentiators there will be be problems at very high frequencies due to the limitations of the op-amp. It is normal to make more additions to this circuit in order to *introduce an upper frequency limit*. We will go through that a bit later on...

Direct RC – op-amp comparison:

The RC circuit has the restriction that $\frac{dV_{out}}{dt} \ll \frac{dV_{in}}{dt}$, i.e. the components have to be chosen such that the filter circuit responds *very rapidly*. If this condition is violated then the *exponential character* of the capacitor charging and discharging will become evident.

This restriction is necessary in order to have (nearly) all of the variation in time across the capacitor. (If it was all then readout would be impossible!) With the op-amp circuit *all* the variation in time *can* be across the capacitor. The feedback loop mirrors the behaviour (because of the virtual ground). But readout remains possible as the op-amp provides the signal.

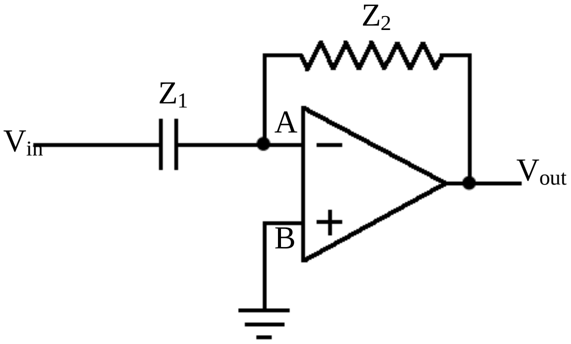




Reminder: the op-amp circuit is and active circuit, it draws power from a power supply to generate its output signal. By contrast circuits with resistors and capacitors alone are passive.

Differentiators / High pass filters: (Frequency analysis)

We can repeat our analysis in the frequency domain again by making use of *complex impedances* for the feedback networks. You may be beginning to see how this form of analysis



could be extended to more complicated input and feedback networks: all you need to know is the impedance of each. Here the complex impedances are:

$$Z_1 = -\frac{i}{\omega C}$$
$$Z_2 = R$$

Using the golden rules:

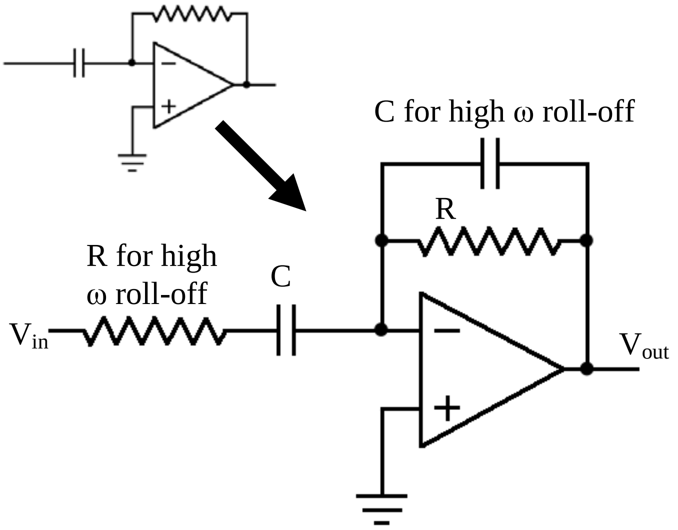
- Rule 1: Since point B is at ground then point A must be too.
  Consequence: the voltage across Z_2 is V_{out} and the voltage across Z_1 is V_{in}
- Rule 2: Since no current flows into the inverting input, all current flowing through the feedback must flow towards the input.

$$\frac{V_{out}}{Z_2} = -\frac{V_{in}}{Z_1}$$
$$\frac{V_{out}}{V_{in}} = -\frac{Z_2}{Z_1} = -i\omega RC \text{ (ideally!)}$$

This time we have a circuit which amplifies the input signal *in proportion to its frequency*. Hence *low frequency signals are relatively attenuated* while high frequency signals are magnified. Again there is a *change in phase* indicated by the fact that the amplification is *complex*. This type of circuit is called a **high-pass filter** because it lets high frequency input signals go past.

Coping with op-amp limitations:

As mentioned above op-amps have their *own performance limitations* especially at *very high frequencies*. These can be dealt with via the addition of one or two more components.



The performance of the op-amp differentiator begins to *deteriorate* at high frequencies. This is due to the sub-optimal performance of the *op-amp itself*. The best way forward is to take control of the frequency at which the circuit ceases to operate successfully. An additional R and C are added to create a *well controlled maximum frequency* at which the differentiator will cease to operate.

The RC product for the new components will be *smaller* than for the original differentiator components, to set the frequency limit at a higher value. The new circuit actually will become a **band-pass filter** with controlled low and high roll-off.

Active Filters:

In analysing circuits in the frequency domain we again have in mind the idea that *any input signal can be represented as a superposition of sine waves* of varying amplitudes, phases and frequencies, ω. For circuits involving more than just resistors the amplitude and phase typically are *frequency dependent*. The characteristics of a circuit can be plotted on graphs showing how the *amplitude and phase* of the output signal varies with the *frequency* of the input signal.

First we will look at just the amplitude of the output for both high-pass and low-pass filters. This is the key characteristic of these circuits. To completely characterise a filter both the amplitude and the phase information is required. The standard way to present this is via a **Bode plot**, a specific choice of display using logarithmic quantities detailed further below.

Active high-pass filters:

For the sake of variety let's consider a circuit with a resistor R_1 and a capacitor C_1 arranged in series in the input circuit and a resistor R_2 in the feedback circuit. By determining the impedance of the input circuit and of the feedback circuit it is straight forward to calculate the amplitude and phase response (we just have to copy what has been done above).
So, what is the amplitude of V(t) as a function of frequency, ω? Remember, for complex numbers we need to calculate:

$$|V| = \sqrt{V\,V^*}$$

With the impedances:

$$Z_1 = -\frac{i}{\omega C_1} - R_1$$
$$Z_2 = R_2$$

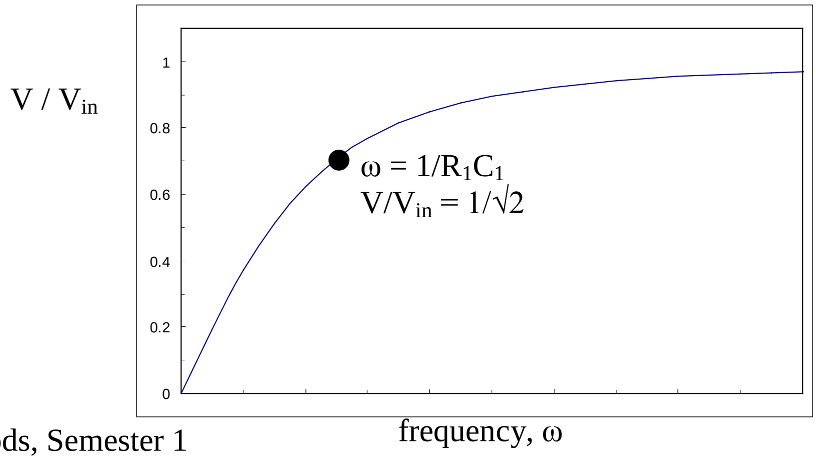
we get:

$$\frac{V_{out}}{V_{in}} = -\frac{Z_2}{Z_1} = \frac{R_2}{\frac{i}{\omega C_1} + R_1} = \frac{R_2}{R_1}\frac{\omega C_1 R_1}{i + \omega C_1 R_1}$$

and yield:

$$|V| = \sqrt{V\,V^*} = |V_{in}|\left(\frac{R_2}{R_1}\right)\sqrt{\left(\frac{\omega C_1 R_1}{i+\omega C_1 R_1}\right)\left(\frac{\omega C_1 R_1}{-i+\omega C_1 R_1}\right)} = |V_{in}|\left(\frac{R_2}{R_1}\right)\frac{1}{\sqrt{\left(1 + \frac{1}{\omega^2 {R_1}^2 {C_1}^2}\right)}}$$

From this expression it can easily be seen that when the frequency of the input signal ω becomes very small (or zero) the gain of the circuit also becomes very small (or zero). For large ω the gain approximates the ratio R_2/R_1, i.e. *the gain of the op-amp*. Plotted quantitatively for the case ($R_1 = R_2$) this looks like:



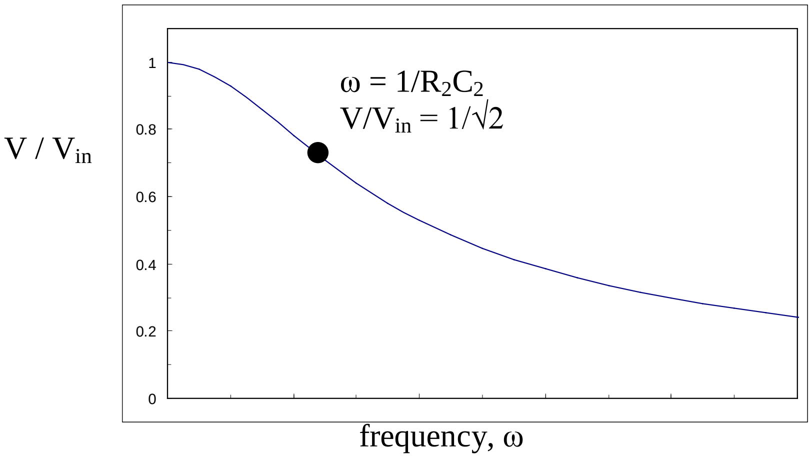
At the frequency $\omega=1/R_1C_1$ the output is $1/\sqrt{2}$ of the maximum output signal. This is taken as the **characteristic frequency** of a filter circuit. For reasons which we will encounter below when looking at the *Bode plot* this is also called the *-3dB point* for the filter.

Active low-pass filters:

Here we will consider a similar circuit where only the capacitor has been moved. In this case the input circuit consists of a resistor R_1 while the feedback circuit consists of a resistor R_2 and a capacitor C_2 in parallel. By calculating the impedances of the input and feedback circuit the gain of the circuit can be determined in the same way as above. The amplitude of V(t) as a function of frequency, ω, here is:

$$|V| = \sqrt{V\,V^*} = |V_{in}|\left(\frac{R_2}{R_1}\right)\frac{1}{\sqrt{(1+\ \omega^2 {R_2}^2 {C_2}^2)}}$$

Again the behaviour for low and high ω can be considered. Small ω leads to relatively unchanged amplitudes at gain R_2/R_1. High ω gives significant attenuation. Below is the graph plotting the gain (again for the case $R_1 = R_2$). Again the *-3dB point* for the filter is pointed out where the gain is $1/\sqrt{2}$ of the maximum:



Bode plots:

You may have noticed the *exponential behaviour* of the last two gain curves. In **Bode plots** this is linearised by plotting the amplitude and the phase of the output as a function the *logarithm of the frequency* of the input signal. On the y-axis the phase is plotted linearly, but the *gain also plotted in logarithmic form*.

The *logarithmic scale* used for the gain is called **Decibels** (dB) and is defined as:

$$\text{Gain } dB = 20\log\frac{V_{out}}{V_{in}}$$

That may look odd to you to begin with. Let's look at some values to get a feeling for it:

linear gain: V_{out}/V_{in}	log gain (dB)		linear gain: V_{out}/V_{in}	log gain (dB)
10	20 dB		10000	80 dB
2	6 dB		1000	60 dB
√2	3 dB		100	40 dB
1	0 dB		10	20 dB
1/√2	-3 dB		1	0 dB
1/2	-6 dB		0.1	-20 dB
1/10	-20 dB		0.01	-40 dB

So, the Decibel scale is designed to display *huge gain ratios* but also to make small changes from unity visible, using the very nature of logarithms. Why then the odd factor 20? Remember, the power in a system is *the square* of the amplitude. Taking the logarithm turns that into a *factor 2*. The decibel scale therefore is *linear in the exponent of the power* handled in a system: a step of 10 dB corresponds to a change of a *factor of 10* in the power of a signal.

Now think about your ear, which dynamic range in power it has to work with. Here are some examples related to the power (in Watts) of noise levels and to the Decibel scale:

noise power: (Watt)	noise level (dB)	examples (at 1m distance)
1000000 W	180 dB	rocket engine
10000 W	160 dB	jet engine
1000 W	150 dB	klaxton (warning hooter)
100 W	140 dB	truck diesel engine
10 W	130 dB	machine gun
1 W	120 dB	pneumatic hammer
0.3 W	115 dB	trumpet
10^{-1} W	110 dB	power saw
10^{-3} W	90 dB	shouting voice
10^{-5} W	70 dB	human conversation, typewriter
10^{-7} W	50 dB	fridge
10^{-8} W	40 dB	human whisper
10^{-12} W	0 dB	nominal hearing threshold

Human hearing clearly works on a logarithmic scale in power. But it is also clear that high powered noise will physically damage the hearing apparatus, especially when extended over time. Since the advent of mobile sound equipment in the 80's the rate of hearing loss has – especially in young people – increased dramatically. Why? People consume the music dominantly via ear phones and tend to crank up the volume to block out environmental sounds, like that of traffic. This way they end up with sound levels of up to 110 dB directly delivered to their ears. Also the sound level in discos has gradually been increased over the last decades, in average by 10 dB. Have sound levels of 110-115 dB been the norm in the 80's recent surveys generally found levels of above 120 dB, with a peak value of 128 dB. The same tendency has been found for live music performances. The way to enjoy these and still to be friendly to the own ears is to use some kind or ear protection, which – sadly – only few people do. So the need for hearing aids will further rise.

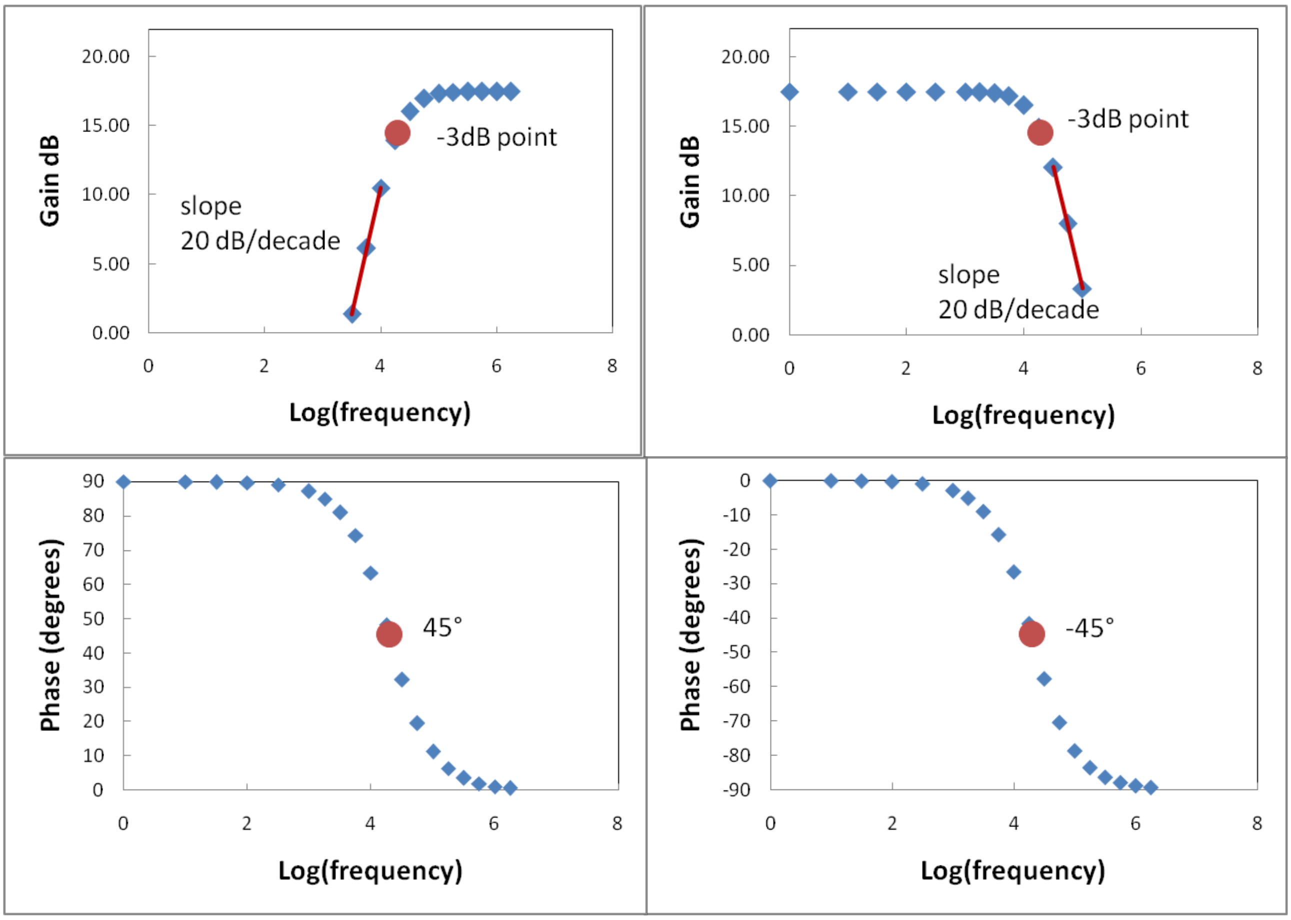
Let's return to the characteristics of the Bode plot of our electrical signals. Above the *-3 dB point* was mentioned as important. It is the point where the output voltage of the filter circuits is attenuated to the factor $1/\sqrt{2}$ of the input voltage:

$$-3dB = 20\log\frac{1}{\sqrt{2}}$$

The output power at this point is attenuated to the factor 1/2. The -3 dB point is also called **roll-off point**. Beyond this point the attenuation of the gain of the filter generally will be *exponential*, i.e. *linear in the Decibel scale*. This is one of the main characteristics of a filter circuit: the ability to *remove unwanted frequency components of a signal* beyond a given point. To describe this ability the *slope* of the attenuation beyond the -3 dB point is recorded in dB/decade: the sharper the slope, the better the filter.

In the example Bode plots below for a RC high-pass and a RC low-pass filter the -3dB point is highlighted at 10 kHz. The gradient of the attenuation of the voltage gain beyond the roll-off point is -20dB per decade in frequency.

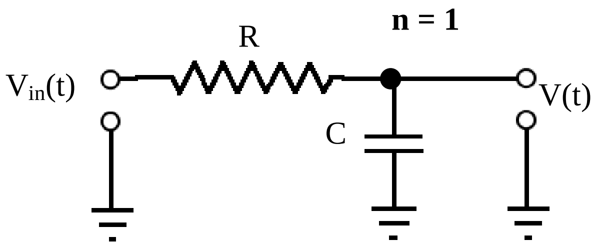
High-pass filter Low-pass filter



Below the gain curves for the two filters the corresponding phase curves are given on the same frequency scale. This allows studying the phase behaviour of the circuits in correlation with the gain behaviour, i.e. to fully understand the behaviour of the RC circuits.

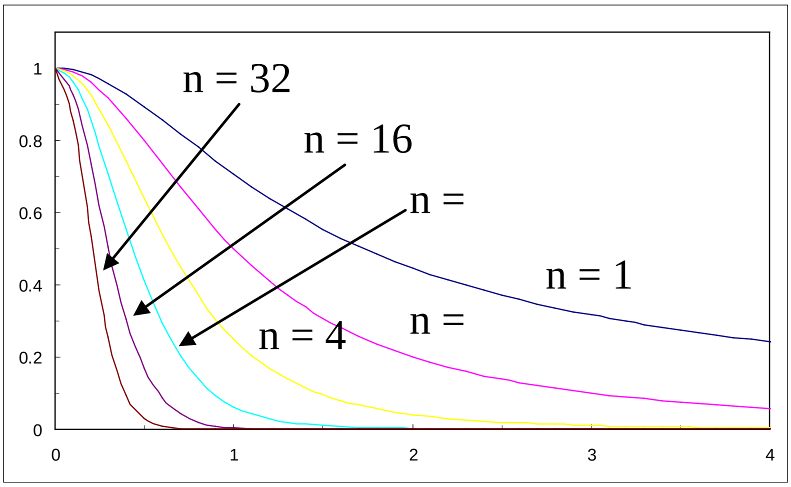
It turns out that the phase shift at the -3 dB point is ±45° with respect to the low and high frequency limits. The phase turns symmetrically around this point on the logarithmic frequency scale. This is another reason for locking at the behaviour on the logarithmic scale. In the region where the circuits let the signal pass the phase shift is small. But already one decade in frequency off the -3 dB phase shifts become visible, coinciding with attenuation effects starting to rise.

Improvements of the frequency response of RC filters: (H&H, 5.01, p. 263)



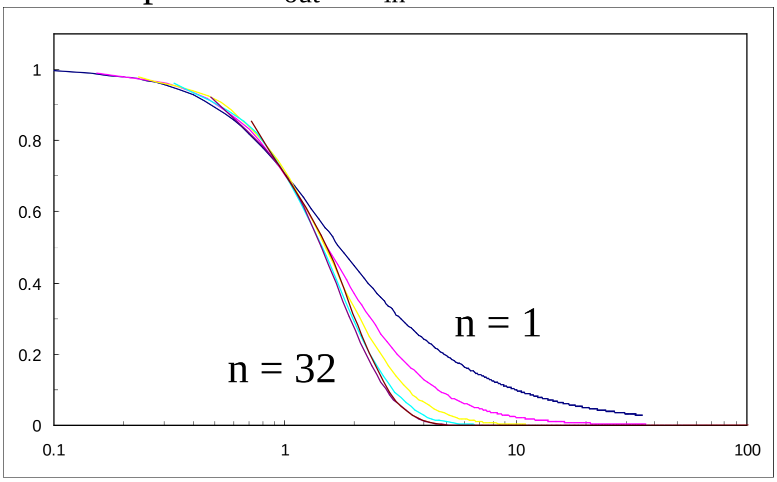
The RC circuits we have dealt with so far are the simplest filters available. And you have seen that low-passes as the one above also act as integrators (and high-passes as differentiators).

The performance in these functions can be improved, even using passive components only. The simplest means is to cascade several circuits to a sequence of filters (although they will have to be interspersed with *follower* circuits to avoid loading problems due to input impedance). *Cascading* several circuits as the above to a sequence causes the “knee” in the amplitude response of the output to *become sharper*. Below this is shown on the linear scale as well as in the Bode plot for a selected number (n) of chain elements.



Frequency

Amplitude response V_{out} / V_{in}



Log(normalized frequency)

However, as n increases the improvement in the filter properties slows down. To achieve an even sharper fall-off *different filters* must be used. There are numerous different sorts of filters available, which will not be described here. Each type represents a different sort of *compromise*:

- some have very sharp cut-offs in frequency
- some offer a very flat gain as a function of frequency for un-attenuated signals
- still others are very well behaved in terms of phase and so don’t distort the unfiltered signal.

Many of these ideas have their foundation in all-passive circuits. The use active components generally improves further the quality and range of the scope of the application.

Filter performance criteria – ω: (H&H, 5.04, p. 267)
The following criteria are used to qualify the performance of a filter with respect to the frequency ω.

Reminder: In comparing the amplitude of two signals the *Decibel scale* is often used:

$$dB = 20\log_{10}\frac{A_2}{A_1}$$

hence 3dB is a factor of √2, 20 dB is a factor of 10 and 40dB is a factor of 100.

Cut-off frequency: this is another name of the *roll-off point* at -3dB.

Passband: the region of frequencies where the *amplitude and phase* of the signal are *relatively un-attenuated* by the filter. The passband is limited by the point (or points) where the gain falls below the -3dB limit. The *phase shifts* which do occur within the passband are important as they will *distort* un-attenuated signals.

Stopband: the region of frequencies where the *attenuation exceeds some minimum* amount. This minimum could be 40dB of attenuation, i.e. a suppression factor of 1/100 in amplitude and 1/10000 in power. Signal components with frequencies that fall in the stopband are so strongly suppressed that they don't play any significant role in the output signal.

Attenuation gradient: the slope of the attenuation well beyond the cut-off frequency giving the rate of the increase of the attenuation (in dB) with every decade in frequency.