Understand Electronic Filters

Owen Bishop

BUTTERWORTH
HEINEMANN

Butterworth-Heinemann
Linacre House, Jordan Hill, Oxford OX2 8DP
A division of Reed Educational and Professional Publishing

A member of the Reed Elsevier plc group

OXFORD BOSTON JOHANNESBURG
MELBOURNE NEW DELHI SINGAPORE

First published 1996

British Library Cataloguing in Publication Data
Bishop, O.N. (Owen Neville), 19–
Understand electronic filters
1 Electric filters
I Title
621.3′815′324

ISBN 0 7506 2628 3

Library of Congress Cataloguing in Publication Data
Bishop, Owen.
Understand electronic filters/Owen Bishop.
p. cm.
ISBN 0 7506 2628 3
1 Electric filters. I Title.
TK7872.F5B554 96–8190
621.3815′324–dc20 CIP

Typeset by Laser Words, Madras, India
Printed and bound in Great Britain by
Biddles Ltd, Guildford and King's Lynn

Contents

Introduction

This is a book about the way filters work. For those who want to grasp what is happening in filter circuits, this book provides relatively non-mathematical explanations backed by a wealth of graphs. It is not possible to avoid maths completely but the maths here is limited to helping the reader get the feel of filter behaviour and go on to tackle more advanced texts with assurance and understanding. A knowledge of maths up to GCSE level is assumed and any aspect slightly above that level is briefly explained.

There are dozens if not hundreds of ways of building a filter to perform a particular function but the principles on which filters operate are few. By examining a selection of filter configurations in detail and observing exactly how they operate, the reader will become sufficiently familiar with filtering principles to understand how most other types work.

It is important for the reader to master and understand each small step in what may be a long discussion. To this end, discussions are broken into short, easily managed steps with a batch of short questions at the end of each step, under the heading Keeping up? Most chapters end with a collection of slightly more demanding questions, headed Test yourself.

It is not easy for the reader to obtain hands-on experience of analogue filters without access to a wide range of precision components and test-gear. This is why the chapters on analogue filters have been so fully illustrated by printouts obtained from computer simulations. On the other hand, digital filters can readily be implemented on a computer and a number of digital filter programs in BASIC are listed in Appendix A. These provide the reader with a basis for investigating digital filters in action.

The author wishes to acknowledge the assistance so generously given by: Those Engineers Limited, Mill Hill, London, publishers of *SpiceAge*®, the electronic circuit simulation software; Wolfram Research Inc., Champaign, Illinois, publishers of *Mathematica*®, the mathematical software; Goth, Goth and Chandleri Limited, Herstmonceux, East Sussex, publishers of *Nodal*, the electronic engineering package for use with *Mathematica*®, Samuel Dick, co-author of Riddle, Alfred and Dick, Samuel, *Applied Engineering with Mathematica*®, Addison-Wesley, 1995.

Without the computing resources provided by the above, this book could not have been written.

1
Passive devices

Many of the filters described in this book are built from just three kinds of electronic device: resistors, capacitors and inductors. These are often known collectively as **passive** devices. This description is given to them because they are acted upon by voltages or currents already existing in the circuit and are not able to initiate any activity on their own account. The properties of these passive devices when acted on by voltages or currents are the subject of this chapter.

Resistors

The action of a voltage on a resistor is explored by setting up the circuit shown in Fig. 1.1. We would obviously need to connect meters to the circuit to measure what is happening but these are left out of Fig. 1.1 to keep it clear. The circular symbol on the left is a voltage generator. It produces a voltage or potential difference (pd) between its terminals. The pd is applied across the resistor. Suppose for this example that the generator produces a pd, v, that begins at 4 volts, rises steadily to 5 volts, falls steadily to 3 volts and then rises steadily back to 4 volts. This action takes 1 millisecond to happen and is repeated 1000 times per second. The upper curve in Fig. 1.2 is a graph to show how v varies with time. The graph covers a period of 5 ms, during which the pd goes through its cycle five times. A pd which varies regularly in this way is often called a **triangular** wave.

The pd makes a current flow through the resistor and this is shown by the lower curve of Fig. 1.2. Because the numerical values of the current are much

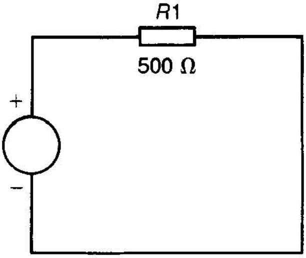


Figure 1.1 *The current through the resistor depends on its resistance and the pd across it produced by the voltage generator*

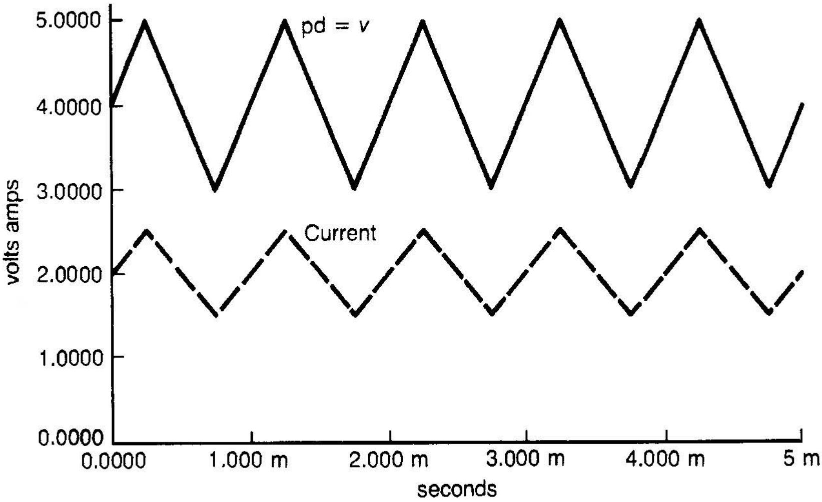


Figure 1.2 *The current flowing through the resistor in Fig. 1.1 is at all times proportional to the pd (v) across it*

smaller than those of the pd, we have had to plot the current curve on a scale that is 250 times the scale of the pd curve. The interesting result is that the current curve is very similar to the pd curve. It has the same triangular shape and it goes through 5 cycles in the same time. Also, the peaks of the current curve occur at exactly the same times as those of the pd curve.

This table lists some of the pds and currents plotted in Fig. 1.2 (remember that the current is plotted on a ×250 scale):

time (ms)	*pd* (V)	*current* (A)	*pd/current* (V/A)
0	4	0.008	500
0.25	5	0.010	500
0.5	4	0.008	500
0.75	3	0.006	500
1.00	4	0.008	500

The curves repeat after this, so there is no need for further readings. In the last column the pd at each instant has been divided by the current at the same moment. An identical value, 500, is obtained on every occasion. There is a fixed proportion between the pd across the resistor and the current flowing through it. This is known as the resistance of the resistor. Its value, in volts per amp, is more usually expressed in ohms (symbol Ω). In this example, the resistance is 500 Ω.

In general, if R is the resistance of the resistor in ohms, v is the pd across it in volts, and i is the current flowing through it in amps, then:

$$R = v/i \qquad (1)$$

We use a capital R to indicate that the resistance is a fixed quantity (for that particular resistor). Both v and i may vary from time to time, so they are printed in small (lower-case) letters. This convention is adopted throughout the book.

Example: When a pd of 48 V is applied across a resistor, the current flowing through it is 4 A. Find the resistance of the resistor.

Answer: $R = v/i = 48/4 = 12\,\Omega$.

Equation (1) can be rearranged to produce two other useful equations:

$$v = iR \tag{2}$$

$$i = v/R \tag{3}$$

Example: What is the pd across a 120 Ω resistor if the current through it is 0.25 A?

Answer: Using equation (2), $v = iR = 0.25 \times 120 = 30\,\text{V}$.

Example: What current passes through a 22 Ω resistor when a pd of 44 V is placed across it.

Answer: Using equation (3), $i = 44/22 = 2\,\text{A}$.

Equations (1), (2) and (3) allow us to calculate any one of the quantities v, i and R, given the values of the other two.

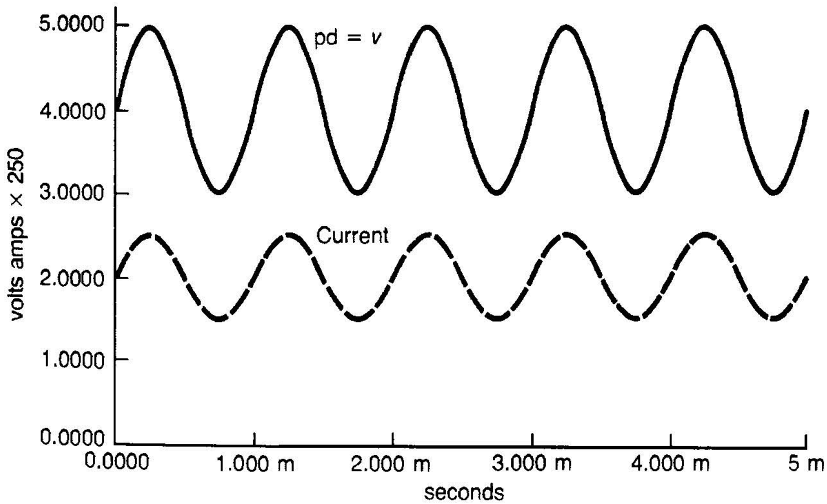


Figure 1.3 *The current flowing through the resistor in Fig. 1.1 instantly follows changes in the pd across it*

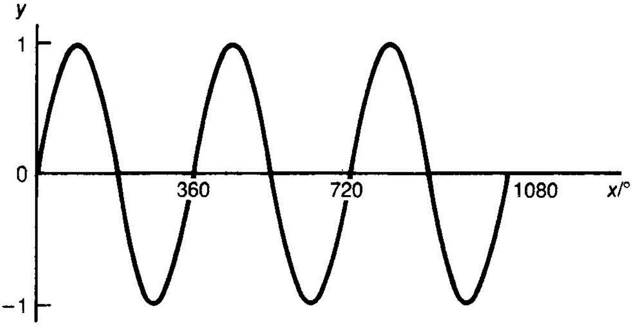


Figure 1.4 *The graph of* $y = \sin x$ *ranges between +1 and −1*

The relationship between v, i and R always holds true. In Fig. 1.3, for example, we have a sine wave instead of a triangular wave. Because $i = v/R$, the current curve also has a sine wave shape, with the same number of 'waves' in 5 ms.

Sine waves

Figure 1.4 is the graph of the sine of an angle. Here we plot it from 0° to 1080° (three complete turns). An electrical signal which has this shape when graphed is known as a sine wave, or a **sinusoid.**

Sine waves are important because it can be shown than any periodic wave of any shape can be thought of as being made up of a mixture of different sine waves (p. 22).

Keeping up?

1. In each example below, given two of the quantities, expressed in volts, amps or ohms as appropriate, calculate the third:
   (a) $R = 150$, $v = 600$, $i = ?$
   (b) $v = 4$, $i = 0.5$, $R = ?$
   (c) $R = 45$, $i = 0.1$, $v = ?$
   (d) $i = 2.5$, $v = 5$, $R = ?$
   (e) $v = 10$, $R = 80$, $i = ?$
2. When a resistor has a pd of 15 V across it, the current through it is 2 A. What pd is developed across it when the current is decreased to 0.4 A?

Capacitors

Figure 1.5 is the same circuit as Fig. 1.1 except that it has an extra component, a capacitor in series with the resistor. Current flowing through the resistor flows to

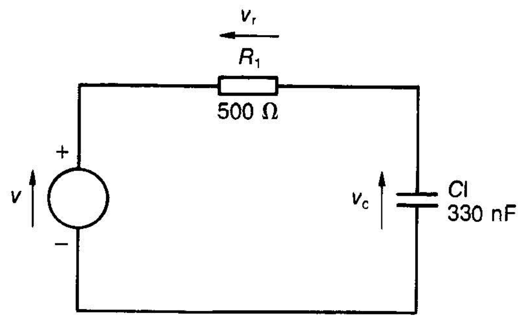


Figure 1.5 *Adding a capacitor to the circuit of Fig. 1.1 has a significant effect on the current–pd relationship. Kirchhoff's voltage law states that the sums of voltages (pds) around any loop of a circuit is zero. Here v_r and v_C are measured with polarities opposite to v, so $v - v_r - v_C = 0$, or $v = v_r + v_C$*

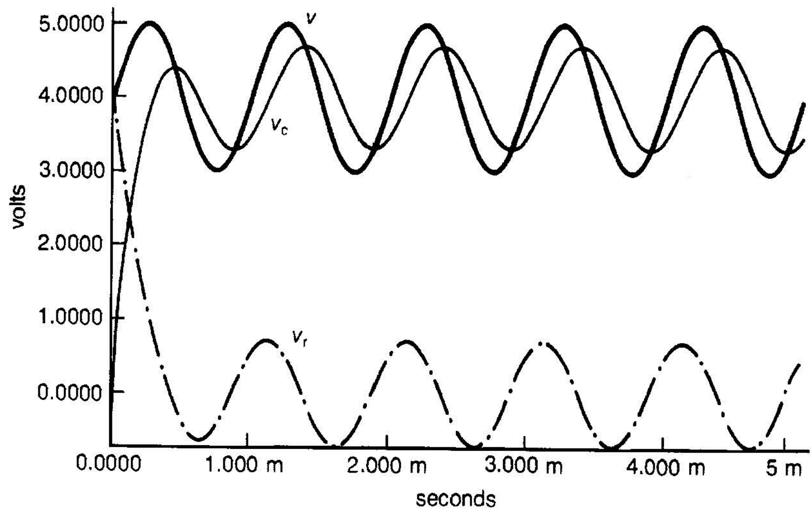


Figure 1.6 *These curves show how the pds of Fig. 1.5 vary with time*

the capacitor. Because the dielectric of the capacitor is a non-conductor, current is unable to flow through the capacitor itself, so the capacitor becomes charged. Kirchhoff's voltage law applies here. At any instant, the pd across the capacitor (v_C) plus the pd across the resistor (v_r) must be equal to the total pd (v) being produced at that instant by the generator. As an equation:

$$v = v_r + v_C \qquad (4)$$

When we measure these pds, perhaps by using an oscilloscope, we obtain the results shown in Fig. 1.6. The upper curve (v) of Fig. 1.6 is exactly the same as the pd curve in Fig. 1.3, because we are using the same sinusoidal voltage generator. At the instant the generator is switched on, it produces 4 V. There is no charge on the capacitor at this time so $v_C = 0\,V$ and, by equation (4), $v_r = v - 0 = 4\,V$. Current begins to flow through the resistor into the capacitor, charging it and causing an increasing pd to appear across it. The result is that v_C rises very rapidly. More and more of the total pd appears across the capacitor and less and less appears across the resistor. Eventually v_C rises to its maximum value after about 0.43 ms (the time when the curves for v and v_C cross). At that time $v = v_C$, which means that by that time v_r has fallen to zero, as can be seen from its curve. There is now no pd across the resistor, and no current is flowing through it to the capacitor. The capacitor has reached its peak charge. From now on, v (which is already falling) falls further and becomes less than v_C. If v is less than v_C, then v_r must be negative. As v_r becomes increasingly negative current flows through the resistor in the opposite direction, from capacitor to generator, partially discharging the capacitor. This process continues until the v and v_C curves cross again (at about 0.88 ms). Now v_r is zero again but is increasing in the positive direction. Current begins to flow through it in the original direction and the capacitor gradually regains the charge it lost during the previous 0.45 ms. It actually gains a little more charge than it previously had because this time it is already partly charged.

Current and charge

The unit of electric charge is the coulomb, which is defined as the amount of charge carried past a point in a circuit while a current of 1 A flows for a period of 1 s. If a current i flows into the capacitor, the charge accumulates at the rate of i coulombs per second. If the current flows for t seconds, the total charge q accumulating is:

$$q = it \qquad (5)$$

Accumulating charge causes an increasing pd across the capacitor proportional to the amount of charge. The number of coulombs per volt is defined as the capacitance C of the capacitor:

$$C = q/v \qquad (6)$$

In other words, a capacitor of large capacitance is able to store a large amount of charge with a relatively small pd across it. If q is measured in coulombs and v in volts, the capacitance C is in farads, symbol F.

Example: A current of 2.5 A flows for 8 s into a capacitor, producing a pd of 10 V. What is its capacitance?

Answer: The total charge stored is given by equation (5). $q = it = 2.5 \times 8 = 20$ C. The capacitance is found by equation (6). $C = q/v = 20/10 = 2$ F.

Although capacitors in the farad range do exist, the farad is too large a unit for most practical circuits. Most capacitors have capacitances rated in microfarads ($1\,\mu F = 10^{-6}$ F), nanofarads ($1\,nF = 10^{-9}$ F) or picofarads ($1\,pF = 10^{-12}$ F).

After about 2 cycles the curves settle to regular repetitions, with the capacitor alternately being charged and discharged, as described above. Charging begins when v rises above v_c, making v_r positive. Discharging begins as v falls below v_c, making v_r negative. The result is three sinusoidal waves all having the same frequency but each being in a slightly different stage of its cycle. The charging and discharging of the capacitor always lag a little way behind the increases and decreases of the generator pd. The alternating flow of current is shown in Fig. 1.7. Here we plot v, as in Fig. 1.6, together with the current (through the resistor and to or from the capacitor; it is all the same current). Since $i = v/R$ for a resistor, the current curve of Fig. 1.7 has exactly the same shape and timing as the curve for v_r in Fig. 1.6. The interesting point to notice is that, except during the first 0.75 ms when the capacitor is gaining charge, the current curve is symmetrical about the zero amp line. The mean flow of current is zero, confirming that no

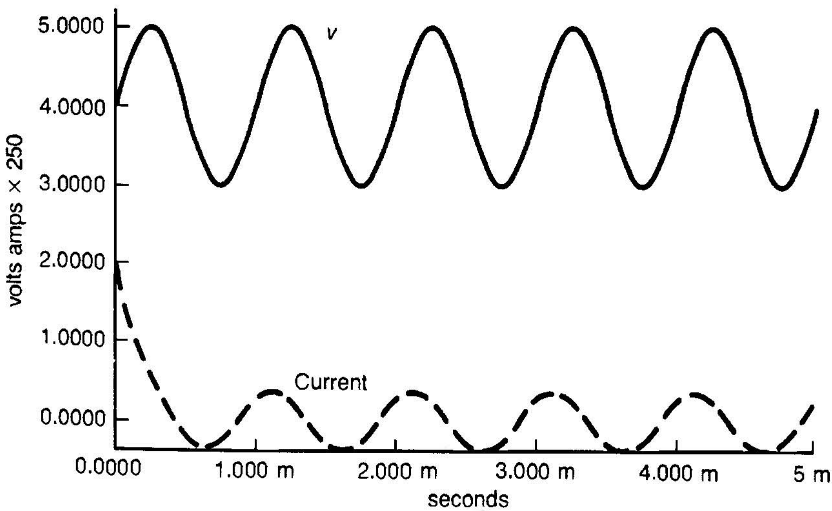


Figure 1.7 *In the circuit of Fig. 1.5, the peaks and troughs in the current lead those in the generator pd (v) by a quarter of a cycle. The mean current is zero*

current actually flows through the capacitor. On average, the capacitor behaves as an open circuit between the resistor and the 0 V line.

Measurements on the graph show that when the circuit has settled to regular repetitions, the peaks of the current are 0.25 ms in advance of the peaks of v. The current to the capacitor leads the applied pd by a quarter of a cycle.

Keeping up?

3. In each example below, given two of the quantities, expressed in amps, coulombs or seconds, as appropriate, calculate the third:

   (a) $i = 2.5$, $t = 3.5$, $q = ?$
   (b) $t = 10$, $q = 5$, $i = ?$
   (c) $i = 4$, $q = 5$, $t = ?$
   (d) $t = 0.01$, $i = 1.2$, $q = ?$
   (e) $q = 2$, $i = 5$, $t = ?$

4. In each example below, given two of the quantities, expressed in volts, coulombs or farads, as appropriate, calculate the third:

   (a) $q = 0.048$, $v = 2$, $C = ?$
   (b) $q = 2$, $C = 0.2$, $v = ?$
   (c) $C = 0.003$, $v = 1.5$, $q = ?$
   (d) $C = 0.002$, $q = 1.6$, $v = ?$
   (e) $v = 1.5$, $q = 0.03$, $C = ?$

5. A constant current of 1.5 A flows for 1 ms into an uncharged capacitor, with capacitance 0.0003 F. What pd is developed across the capacitor?

6. Describe what is happening to the capacitor in Fig. 1.6 at (a) 3.7 ms, (b) 2 ms and (c) 1.86 ms.

Cycles and angles

One cycle of a sine wave is the equivalent to a sine curve plotted from 0° to 360° (see Fig. 1.4). We can use angles to say what stage or phase a wave has got to in its cycle. We call this the **phase angle**, and it is often given the symbol ϕ. At the start of the cycle the phase angle is 0°. Half-way through a cycle, the angle is 180°. A complete cycle is 360°, which brings us back to 0°. Likewise a quarter-cycle is 90°. Instead of stating that the capacitor current is a quarter-cycle ahead of the applied pd, we can say that it leads the applied pd by 90°, or $\phi = +90°$.

Inductors

In Fig. 1.8 the capacitor of Fig. 1.5 is replaced by an inductor. Again we subject the circuit to the sinusoidal voltage from the generator and Fig. 1.9 shows the pds across the generator, the resistor and the inductor (v_l). At first glance, this graph looks exactly like Fig. 1.6. But closer inspection reveals that the curve which descends rapidly from 4 V and then oscillates about 0 V is not the pd across the resistor. It is the curve for v_l, the pd across the inductor. Conversely, the curve which looks like curve v_c of Fig. 1.6 is actually the curve for v_r.

As in the capacitor circuit, Kirchhoff's voltage law applies:

$$v = v_r + v_l$$

The reason that v_r is zero to begin with is that v starts at 4 V and this sudden rise of pd causes the inductor to generate an almost equally large back emf. Thus

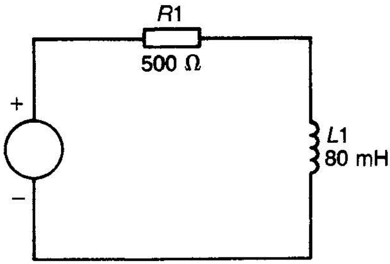


Figure 1.8 *Adding an inductor to the circuit of Fig. 1.1 causes changes in the pd–current relationship entirely different to those caused by a capacitor*

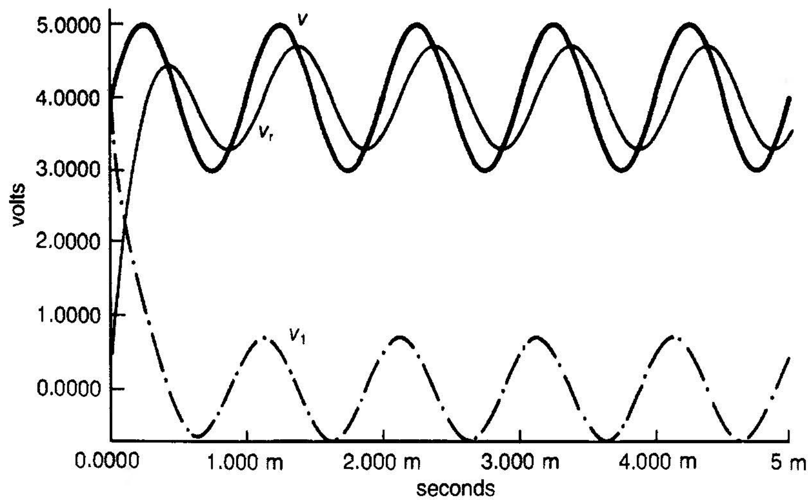


Figure 1.9 *Showing how the pds of the circuit of Fig. 1.8 vary with time*

$v_l \approx v$, and $v_r \approx 0$. As a result, the current through the resistor is practically zero. But after that sudden change of v from 0 V to 4 V, it does not rise so fast and the back emf is less than before. It falls from 4 V to zero and below. This allows some current to flow and v_r rises from 0 V to about 4 V. As in the capacitor circuit, there are two stages in the current flow. In one stage v is less than v_r, so v_l is negative. In other words, the current is decreasing and the negative back emf is acting to prevent it from decreasing. In the other stage v is greater than v_r, current is increasing and the back emf goes positive to prevent further increase. Current increases and decreases but in this circuit there is no reversal of current. Correspondingly, v_r increases and decreases but always remains positive. The figure shows that it actually oscillates around 4 V, the same as v. The average pd across the resistor equals the average pd from the generator. This is because the inductor has, on average, a negligible resistance and the pd across it is, on average, zero. Typically, the inductor behaves as a short circuit between the resistor and the 0 V line.

Current and emf

A magnetic field is created when a current passes through a coil, such as that of an inductor. When the size of the current changes, the magnetic field strength changes proportionately. According to Lenz's law, a change of field strength is opposed by a magnetic field generated within the coil. This opposing field induces in the coil an emf, known as the **back emf**. The back emf generates a current in the direction which opposes any change in the amount of current flowing through the coil. The stronger the attempted current change the more strongly it is opposed. For example, Fig. 1.9 shows the biggest effect at the beginning, where the inductor opposes the attempt to pass an initially large current through it.

The size of the back emf depends on the rate of change of current and also on L, the self-inductance of the coil. A coil is said to have a self-inductance of 1 henry (symbol H) if a back emf of 1 volt is induced in it when the current changes at the rate of 1 amp per second. The equation is:

$$v = -L \cdot \mathrm{d}i/\mathrm{d}t$$

where v is the induced emf (or back emf) in volts, L is the self-inductance in henries and $\mathrm{d}i/\mathrm{d}t$ is the rate of change of current in amps per second (see next box). The negative sign indicates that the back emf opposes the change in current. The henry is a large unit. The inductors we use in practical circuits are usually rated in millihenries ($1\,\mathrm{mH} = 10^{-3}\,\mathrm{H}$) and microhenries ($1\,\mu\mathrm{H} = 10^{-6}\,\mathrm{H}$).

Rates of change

The expression di/dt expresses the rate at which current (i) changes with respect to time (t). For example, if the current is increasing at the rate of 0.4 amps per second, $di/dt = 0.4$. Conversely, if current is decreasing at 5 amps per second, $di/dt = -5$.

In this notation, the 'd' is not an algebraic symbol (if it were, it would cancel out, leaving i/t) but is short for 'an infinitely small change'. So di/dt is the infinitely small change of current occurring during an infinitely small change of time.

In Fig. 1.10 the current increases from i_1 to i_2 while time increases from t_1 to t_2. These are not infinitely small changes. The average rate of increase between point A and B on the graph is the slope or gradient of the straight line AB:

$$\text{gradient} = \frac{\text{change in current}}{\text{change in time}} = \frac{i_2 - i_1}{t_2 - t_1}$$

It is clear from the graph that the average rate of increase between A and B is only an approximation to the actual rate at any given point on the curve between A and B. If we want to know the exact rate at a point on the curve, we must move the points closer together. At CD we have selected an infinitely short length of the curve so that C and D coincide. The gradient at CD is the infinitely small increase of i for an infinitely small increase of t, in symbolic form, di/dt.

In calculus, di/dt is called the derivative or differential of i with respect to t. Given the equation which relates i to t (for example, $i = 6 \sin 2t$), we can calculate di/dt for any given value of t and so find the rate of change of i at any specified time.

The current through the inductor and the total pd are plotted in Fig. 1.11. As is to be expected from equation 3, the shape of the current curve matches the shape of v_r in Fig. 1.9. Measurements on the graph show that when the current curve has settled to regular repetitions, the peaks of the current are 0.25 ms behind the peaks of v. In terms of angle, the current to the inductor lags behind the applied pd by a quarter of a cycle, or $\phi = -90°$.

Because of the particular values chosen for the capacitor and inductor in Fig. 1.8, the sizes of the oscillations of v_r and v_l in Fig. 1.9 are almost exactly the same as those of v_c and v_r in Fig. 1.6. With other values of capacitance or inductance, the curves still show the same overall appearance but the sizes of the oscillations are greater or smaller. For example, if in Fig. 1.8 we double the inductance to 160 mH we obtain Fig. 1.12. This has the same general pattern of curves but the oscillations of v_r are smaller and those of v_l are larger. Also the

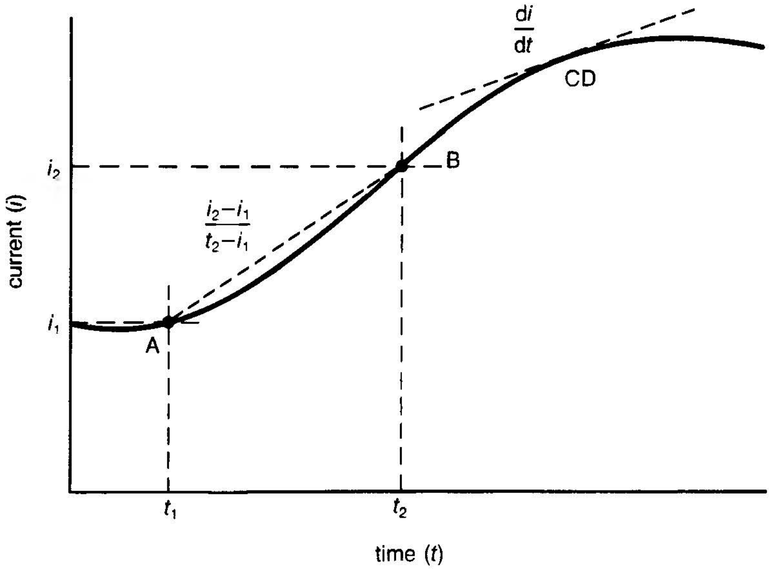


Figure 1.10 *Explaining the meaning of* di/dt

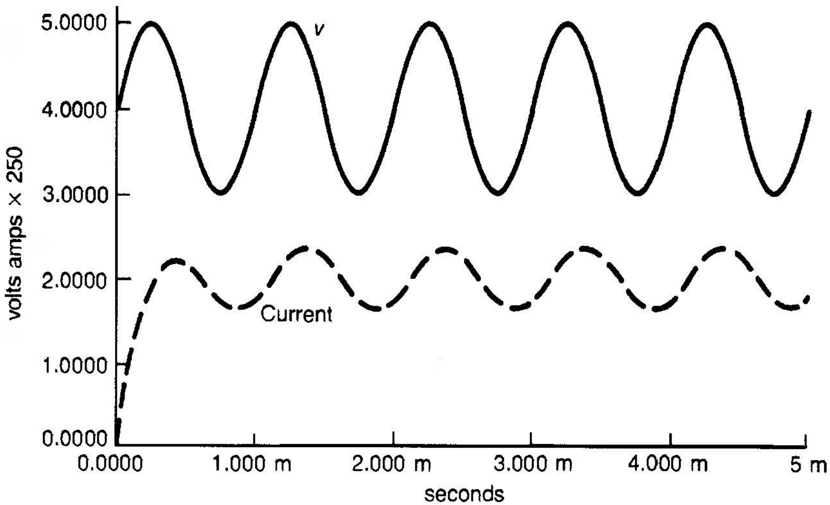


Figure 1.11 *In Fig. 1.7 the peaks and troughs in the current lag behind those in the generator pd by a quarter of a cycle. Here they lead by the same amount. The mean current is 0.008 A*

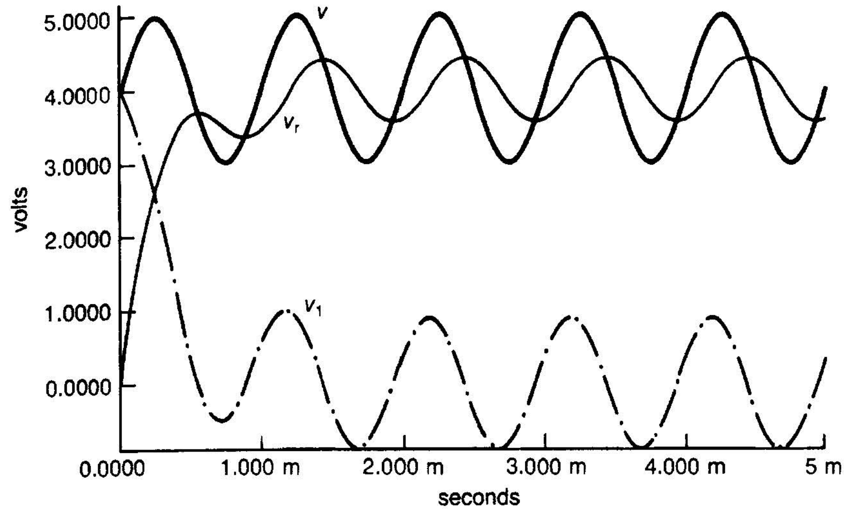


Figure 1.12 *If the self-inductance is doubled, the pds across the inductor and resistor have greater amplitude, but the overall shapes of the curve are similar to those in Fig. 1.9*

oscillations in Fig. 1.12 do not reach the steady repeating stage quite as soon as in Fig. 1.9.

Keeping up?

7. In each example below, given two of the quantities, expressed in amps per second, volts or henries, as appropriate, calculate the third:
   (a) $di/dt = 2$, $L = 0.25$, $v = ?$
   (b) $v = 5$, $L = 0.005$, $di/dt = ?$
   (c) $v = -12$, $di/dt = 500$, $L = ?$
   (d) $L = 0.6$, $di/dt = 1200$, $v = ?$
   (e) $di/dt = -25\,000$, $v = 2000$, $L = ?$
8. (a) In Fig. 1.8, the voltage generator produces a pd of 2 V when first switched on, rising at the rate of 12 000 V per second. The inductor has self-inductance of 15 mH. What is the back emf of the inductor when the generator is switched on? (b) After 0.1 ms the generator output levels off to a steady value. What is the steady value and what is the back emf?
9. Describe what is happening in Fig. 1.9 at a time 3 ms after the start.

Summary

	Resistor	*Capacitor*	*Inductor*
Opposition to current flow is due to ...	resistivity (fixed)	accumulated charge (varies with charge)	back emf (varies with change in current)
pd-current relationship	$v = iR$	$v = it/C$ (assuming constant current)	$v = -L \cdot di/dt$
Current peak occurs at ...	same time as voltage peak	leads voltage peak by 90°	lags voltage peak by 90°
Effective resistance	R	almost infinite	virtually zero

Test yourself

1. A current of 35 mA flows through a 47 Ω resistor. What is the pd across the resistor?
2. The current through a 500 μH inductor is decreasing at the rate of 35 As^{-1}. What back emf is generated?
3. A 2.2 μF capacitor has a pd of 2.5 V across it. What is the charge on the capacitor? How long would it take to charge the capacitor to this level with a constant current of 0.25 mA?
4. In Fig. 1.5, the sinusoidal generator reaches its positive peak voltage every 2 ms. What is the length of time between the voltage peak and the next positive current peak?
5. In Fig. 1.5 the generator is producing a sinusoidal voltage at 400 Hz. What is the length of one cycle? If one of the positive voltage peaks occurs at 10 ms, when is the next positive peak in the current flowing to the capacitor?
6. Explain why a circuit such as Fig. 1.8 takes longer to reach a stable oscillating state when L is 160 mH (Fig. 1.12) than when it is 80 mH (Fig. 1.9).

2
Electrical signals

Most electrical signals are **periodic**. That is to say, their variations in voltage or current are repeated regularly and exactly. A signal is defined by its:

Period, P — the length of one cycle (Fig. 2.1), in seconds.

Frequency, f — the number of cycles in 1 second ($f = 1/P$). Frequency is usually stated in hertz, symbol Hz, where 1 Hz is the same as 1 cycle per second. 1000 Hz = 1 kHz. 1000 kHz = 1 MHz.

Amplitude, A — if the signal oscillates about zero, the amplitude is the height of the peak above zero (Fig. 2.1). But some signals oscillate about a non-zero value. For example, the voltage signals in Fig. 1.3 oscillate with an amplitude of 1 V about an offset (or DC voltage) of 4 V.

Keeping up?

1. Find the frequency of these signals, given their period: (a) 2 ms, (b) 0.4 s, (c) 1 s, (d) 5 ns, (e) 2 min.
2. Find the period of these signals, given their frequency: (a) 5 Hz, (b) 2 kHz, (c) 0.1 Hz, (d) 44 MHz, (e) 56 kHz.
3. For each of the signals in Fig. 2.2, state (a) the period, (b) the frequency, (c) the amplitude, and (d) the offset.

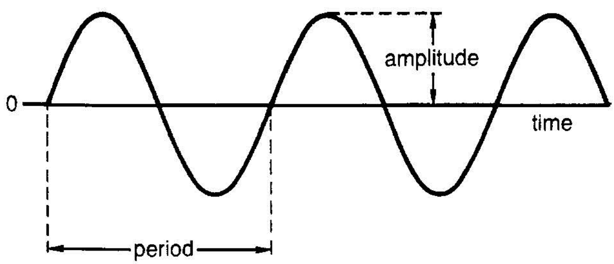


Figure 2.1 *The two most important attributes of an electrical signal are its period and its amplitude*

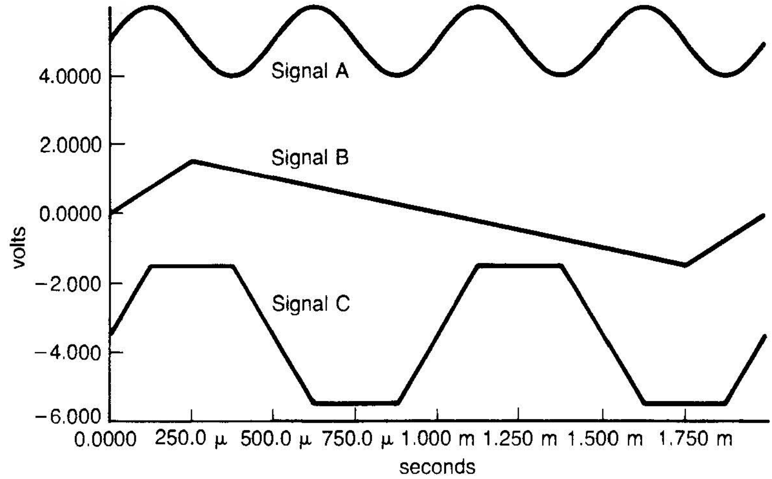


Figure 2.2 *These signals are the subject of question 3*

Degrees and radians

Phase angle and the phase differences between one signal and another are often measured in degrees (see box, p. 8). But a degree is an arbitrary unit and equations are often made simpler if we use a more fundamental unit, the radian (symbol, rad). A radian is the angle subtended in a circle radius 1 by an arc of length 1 (Fig. 2.3). It does not matter if the circle has a radius of 1 m, 1 inch or 1 km, the ratio between radius and circumference is the same and the radian is the same size. A radian equals 57.296° (to 3 decimal places). This is a little less than 60°, a fact that is worth remembering when checking calculations. A complete revolution (or cycle) is a little over 6 rad, or 6.283 rad to 3 decimal places.

To be absolutely exact, 360° equals 2π radians. Using pi is a handy way of avoiding using numbers with lots of decimal places and this is why pi appears so often in calculations connected with electrical signals. If we need to, we can substitute 3.142 (or a value with even more decimal places) for pi at the end of the calculation. Or, if the equation is a theoretical one, we can simply leave the pi as π.

Here are some useful equivalents:

Degrees	45	57.296	90	180	270	360
Radians	$\pi/4$	1	$\pi/2$	π	$3\pi/2$	2π

Angles expressed in degrees usually have the degree symbol (°) written after them. Angles expressed in radians may have the symbol 'rad' after them, but often have no symbol, especially when expressed in terms of pi.

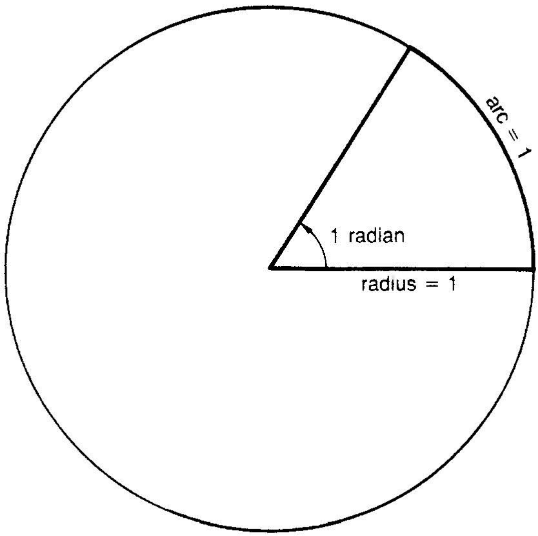


Figure 2.3 *The fundamental unit of angle is the radian, depending as it does on the geometric properties of a circle*

Phase

When the peaks of two signals of the same frequency occur at exactly the same times, we say that the signals are **in phase**. Examples are the pairs of signals in Figs 1.2 and 1.3. By contrast, the signals in Fig. 1.6, although they all have the same frequency, do not have their peaks at the same time. We say that they are **out of phase**. The peaks in v_r occur a quarter of a cycle (90°) ahead of the peaks in v. In terms of radians, the peaks of v_r lead those of v by $\pi/2$. The phase lead applies not only to the peaks but to all other stages in their cycles, such as the troughs and the points at which the signals cross their offset levels. At every stage, v_r leads v by $\pi/2$. Similarly v_c lags v by $\pi/2$, and v_r leads v_c by π.

Equations for signals

A graph of a signal helps us to visualize its exact shape but although we can take measurements of frequency and amplitude from a graph, we can not easily subject a graph to detailed analysis. At some stage we need to be able to express the signal as an equation. We need an expression that specifies the value of the voltage or current at every instant during one period.

The equation of a basic sine wave pd signal is:

$$v = \sin t$$

Compare this with the sine curve in Fig. 1.4. We have substituted pd (v) for y, and time (t) for x. We now have an equation showing how pd varies with time. The equation assumes that the amplitude is 1, but we can easily include amplitude in the equation:

$$v = V_0 \sin t$$

The amplitude V_0 multiplies every value of $\sin t$, so that the signal now swings between $+V_0$ and $-V_0$.

The next step is to bring frequency into the calculation. Introduce another constant ω, which in the basic equation had the value 1:

$$v = V_0 \sin \omega t$$

The value of ω is such that v goes through all its values (that is, a complete cycle) during one period. Let $t = 0$ at the beginning of a period:

$$v = V_0 \sin(\omega \times 0) = V_0 \sin 0 = 0$$

P is the length of a period so, at the end of that period, $t = P$:

$$v = V_0 \sin \omega P$$

But at the end of the period ωP must have reached the value 2π, so that a complete cycle has occurred:

$$\omega P = 2\pi$$

$$\omega = 2\pi/P$$

Substituting $1/P = f$:

$$\omega = 2\pi f$$

Inserting this expression into the basic equation:

$$v = V_0 \sin 2\pi f t$$

This is a more useful equation, since it takes amplitude and frequency into account. As an example, signal A of Fig. 2.4 has amplitude $V_0 = 3\,\text{V}$ and frequency $f = 2\,\text{Hz}$, so its equation is:

$$v = 3 \sin 4\pi t$$

Check this equation by calculating the value of v after 0.6 s:

$$v = 3 \sin(4\pi \times 0.6) = 3 \sin 7.540 = 2.853$$

This agrees with the value shown on the graph. For signal B, $V_0 = 1.8$, $f = 5$, so its equation is:

$$v = 1.8 \sin 10\pi t$$

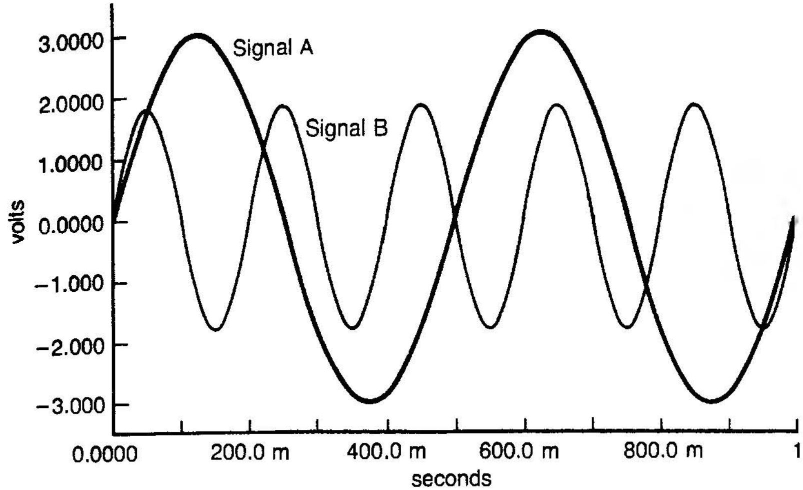


Figure 2.4 *The equations of these signals are (A)* $v = 3\sin 4\pi t$ *and (B)* $v = 1.8\sin 10\pi t$

After 0.72 s:

$$v = 1.8\sin(10\pi \times 0.72) = 1.8\sin 22.619 = -1.058$$

This too agrees with the graph.

Keeping up?

If you are using a calculator for questions 6–8, check that it is in radian mode.

4. Express these degree angles in radians: (a) 90°, (b) 45°, (c) 720°, (d) 36°.
5. Express these radian angles in degrees: (a) π, (b) 2, (c) 0.5, (d) $3\pi/2$.
6. For each of the signals in Fig. 2.5, find (a) its period, (b) its frequency, and (c) its amplitude.
7. For each of the following equations describing sinusoidal signals, find (a) its frequency, (b) its amplitude, and (c) its value after 25 ms:

   (a) $v = 6\sin 24\pi t$
   (b) $v = 0.35\sin 3000\pi t$
   (c) $v = 8.8\sin 3142t$
8. For signals with the specified characteristics, write out the equation that describes them:

   (a) $f = 7$, $A = 10$
   (b) $f = 30\,\text{kHz}$, $A = 0.25$
   (c) $f = 65\,\text{Hz}$, $A = 1$

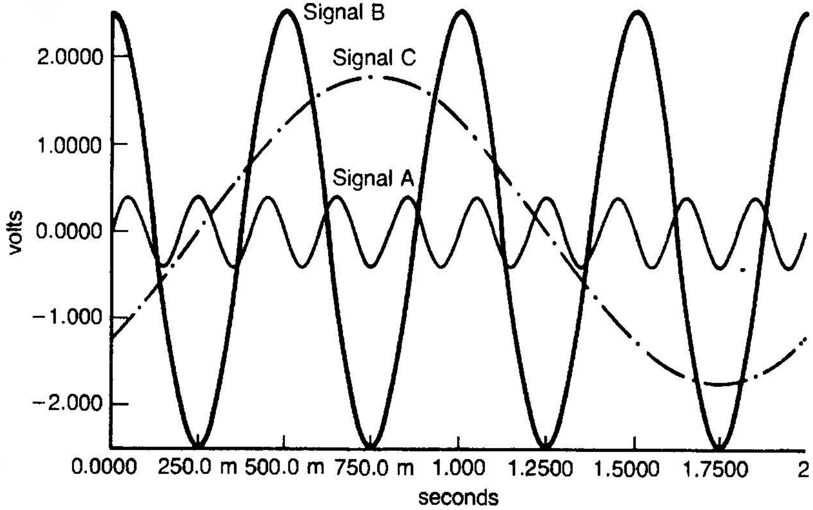


Figure 2.5 *Three signals of different amplitude and phase for question 6*

> **Pd or current?**
>
> The equations in this chapter are all pd equations, stating how pd varies with time. But we could equally well describe circuit behaviour by using current equations, such as $i = I_0 \sin 4\pi t$. It just keeps the descriptions simpler to refer only to pd equations.

Angular velocity

In the previous section we introduced a constant ω, which is equal to $2\pi f$, but we did not say exactly what this constant represents. By definition of frequency, a signal completes f cycles in a second. A cycle is equivalent to 2π radians so, in terms of angle, f cycles per second is equivalent to $2\pi f$ radians per second. Radians per second is another way of specifying frequency. With distances, the quantity measured by metres per second is called velocity. Similarly, when frequency is measured in radians per second we call it **angular velocity**. This gives us an alternative form of the equation for a sine wave signal:

$$v = V_0 \sin \omega t$$

It has been assumed in all the equations and questions above that all sinusoidal signals have zero value when timing begins. Figure 2.5 demonstrates that this is not necessarily so. Signal A is like signals we have described previously, with zero value when $t = 0$, but signal B begins when $v = 2.5$. In other words when $t = 0$, signal B is already a quarter of the way into its cycle. It is $\pi/2$ ahead. If we use the equation $v = 2.5 \sin 4\pi t$ to describe this signal, we shall be describing a signal that lags $\pi/2$ behind signal B. To calculate the correct value of v at a given time we need to add $\pi/2$ to the angle before taking the sine. Adding $\pi/2$ to the angle gives this equation for signal B:

$$v = 2.5 \sin(4\pi t + \pi/2)$$

The phase that a signal has reached when timing begins is known as the **phase angle**. Often, when there are several signals out of phase with each other, one is taken as a reference signal and said to have zero phase angle. The phase angles of the other signals are calculated with reference to this. For example, with a set of signals such as those in Fig. 1.9, it is conventional to take v as the reference signal because it is the applied signal. It has an offset of 4 V, its frequency is 1 kHz and its amplitude is 1 V, so its equation is:

$$v = 4 + \sin 2000\pi t$$

v_l has no offset, its frequency is 1 kHz, and its amplitude is about 0.44 V. Also it leads v by $\pi/2$. Its equation is:

$$v_l = 0.44 \sin(2000\pi t + \pi/2)$$

By contrast, v_r has 4 V offset, $f = 1$ kHz, $V_0 = 0.88$ V, and it lags v by $\pi/2$. Its equation is:

$$v_r = 4 + 0.88 \sin(2000\pi t - \pi/2)$$

Sine wave equation

Given that:

v is the instantaneous pd (voltage)
V_0 is the amplitude
f is the frequency
t is the time elapsed since $t = 0$
ϕ is the phase angle (positive for phase lead, negative for phase lag)

The full equation for a sine wave signal is:

$$v = V_0 \sin(2\pi f t + \phi)$$

Or, using angular velocity ω:

$$v = V_0 \sin(\omega t + \phi)$$

Keeping up?

9. Write the equations that describe the signals of Fig. 2.5.
10. If a voltage signal has $f = 2500\,\text{kHz}$, $V_0 = 2.4$, and $\phi = 3\pi/2$, find its value when $t = 25\,\text{ms}$.
11. Calculate ω corresponding to these frequencies: (a) 1 kHz, (b) 240 Hz, (c) 5 Hz.
12. Calculate the frequency for which ω equals the value given: (a) 2, (b) 18.85, (c) 100.

Synthesizing signals

Pure sine waves seldom occur in electronic circuits. More often a signal consists of two or more sine waves added together. Combining two or more signals to make one signal is called **synthesis**. In Fig. 2.6 two signal generators are operating at the same time and their combined signal passes through R_1. We assume that the internal resistance of the generators is negligible. Because of the principle of superposition in electronic networks, the pd across the resistor is equal to the sum of the pds produced by the generators. For example, Fig. 2.7 is obtained by summing a triangular signal and a sine wave. The triangular signal has a frequency of 300 Hz and an amplitude of 5 V. Superimposed on this is the sine wave with a frequency of 4 kHz and an amplitude of 1 V.

Combining more than two signals together can produce interesting results. As an example, we begin by combining two sine waves:

$$v_1 = \sin t$$

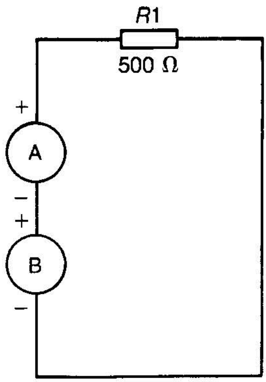


Figure 2.6 *The outputs of two or more generators are added to produce a more complicated signal*

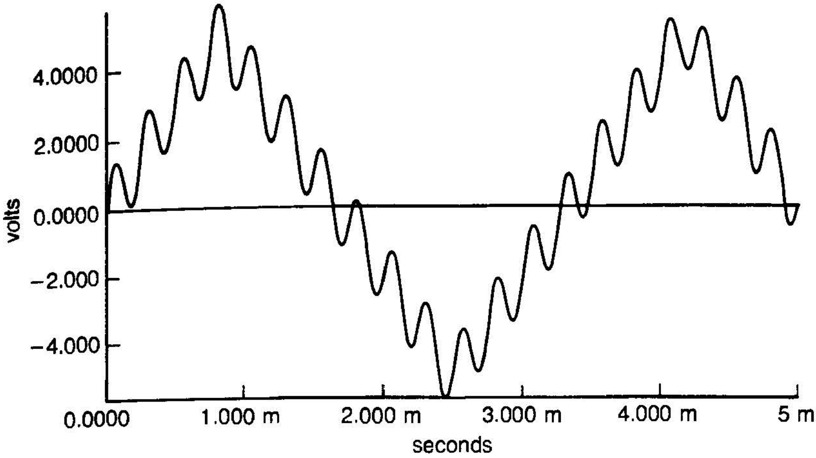


Figure 2.7 *A triangular wave and a sinusoidal wave look like this when summed*

and
$$v_2 = -\frac{\sin 2t}{2}$$

The amplitude of v_1 is 1 V and its frequency is 0.1592 Hz. The frequency of v_2 is double that of v_1, and its amplitude is half that of v_1. Note that v_2 is negative so it is subtracted from v_1. Figure 2.8(a) is the graph of the combined signal:

$$v = v_1 + v_2 = \sin t - \frac{\sin 2t}{2}$$

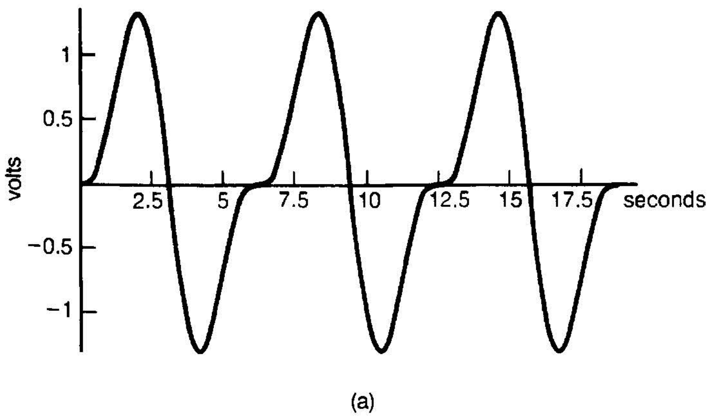


Figure 2.8 *As more and more sinusoidal signals are summed, the resulting signal approaches closer and closer to a sawtooth signal. (a) 2 signals, (b) 4 signals, (c) 12 signals, (d) 100 signals*

The signal still has a recognizable sinusoidal shape, though its peaks are sharper and it is flattened where it crosses the x-axis on the way up. The terms on the right of the equation have the general form:

$$\frac{\sin nt}{n}$$

The term is positive for odd values of n and negative for even values of n. If we continue this sequence for two more terms, we obtain the series:

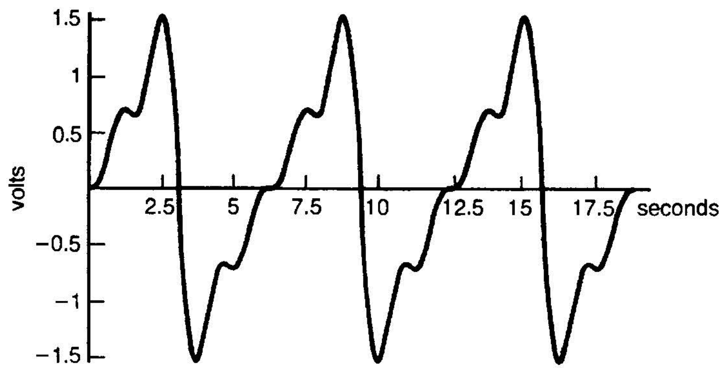
$$v = v_1 + v_2 + v_3 + v_4 = \sin t - \frac{\sin 2t}{2} + \frac{\sin 3t}{3} - \frac{\sin 4t}{4}$$

Figure 2.8(b) shows the result of combining the four signals; the signal still has the same frequency but now it shows a wavy rise followed by a very steep fall. The rising part of the signal exhibits peaks produced by the fourth signal v_4, at a frequency of four times that of v_1. These peaks are small because the amplitude of v_4 is only a quarter of that of v_1.

Many of the peaks are flattened out by taking in eight more signals. The series is now:

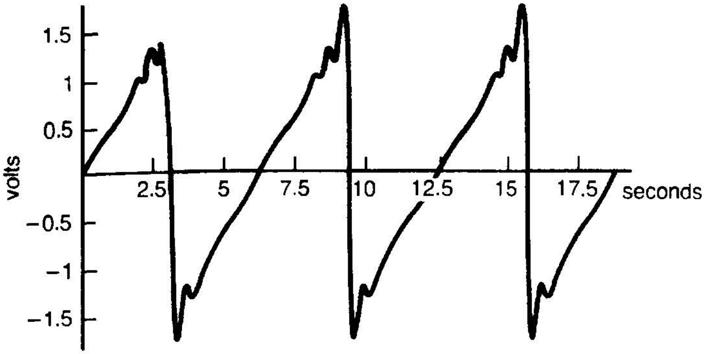
$$v = \sin t - \frac{\sin 2t}{2} + \frac{\sin 3t}{3} - \frac{\sin 4t}{4} + \frac{\sin 5t}{5} \cdots - \frac{\sin 12t}{12}$$

Except for a few slight undulations near its crests, the total signal (Fig. 2.8(c)) has the appearance of a sawtooth signal. If we continue and add signals up to v_{100}, the final term of which is $-\sin 100t/100$, the total signal is almost indistinguishable from a true sawtooth signal (Fig. 2.8(d)). There is no point in going further because the amplitude of v_{100} is only one hundredth of that of v_1 and subsequent signals make virtually no difference to the sum. The result of this operation has been to produce a signal that is far from sinusoidal in appearance, even though it is composed of nothing but sinusoids.

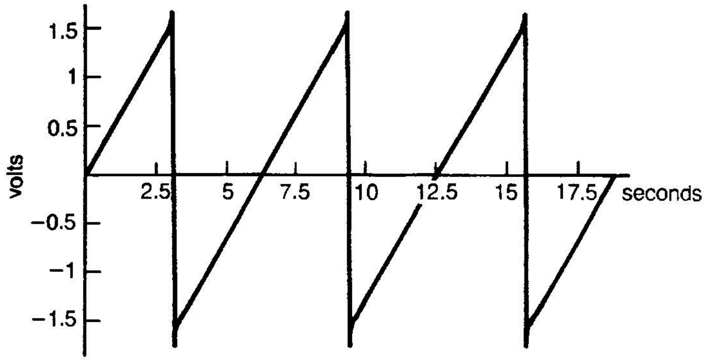


(b)

Figure 2.8 *(continued)*



(c)



(d)

Figure 2.8 *(continued)*

Why 0.1592 Hz?

The examples here and on the next few pages are all based on a signal which has $f = 0.1592$ Hz. This is the same as $1/2\pi$. The reason for choosing this frequency is that substituting $f = 1/2\pi$ into the sine wave equation and making $\phi = 0$ gives:

$$v = V_0 \sin t$$

This gives a simple equation which is made even simpler by making $V_0 = 1$. Having made $v_1 = \sin t$, we can go on to write v_2 and later terms in the series in the most easily understandable form. Using 0.1592 Hz makes things simpler, but the same principles apply and comparable results are obtained whatever basic frequency we choose.

For example, we get a sawtooth wave from:

$$v = 6\sin 50\pi t - 3\sin 100\pi t + 2\sin 200\pi t - 1.5\sin 400\pi t + \cdots$$

but now the frequency of the sawtooth is 25 Hz and its amplitude is 9.48 V. The essential point is that we have kept the same relationship between terms of the sequence, dividing the amplitude by $-n$, and multiplying the frequency by n as we proceed along the sequence.

There are many different sequences that can be summed and, as another example, take the series:

$$v = \cos t + \frac{\cos 3t}{9} + \frac{\cos 5t}{25} + \frac{\cos 7t}{49} + \cdots + \frac{\cos nt}{n^2}$$

This series is based on cosines, but a cosine curve has exactly the same shape as a sine curve. In this series all terms are positive and only odd values of n are included. The denominator of each term is n^2 instead of n.

Figure 2.9a shows the signal obtained by summing the first two terms. It has the general appearance of a cosine curve, except that its rises and falls are much less curved and its peaks are sharper. Because n skips the even values, and because the amplitude decreases as the square of n, the series converges towards virtually constant values after only a few terms. Figure 2.9b is the sum of the first 6 terms, for which n is 1, 3, 5, 7, 9 and 11. The amplitude of v_{11} is only 1/121 that of v_1, so there is no need to go further. The result is a practically perfect triangular signal.

Figures 2.8a–d and 2.9a–b demonstrate that it is possible to synthesize angular waves such as sawtooth and triangular signals from a series of rounded sinusoidal

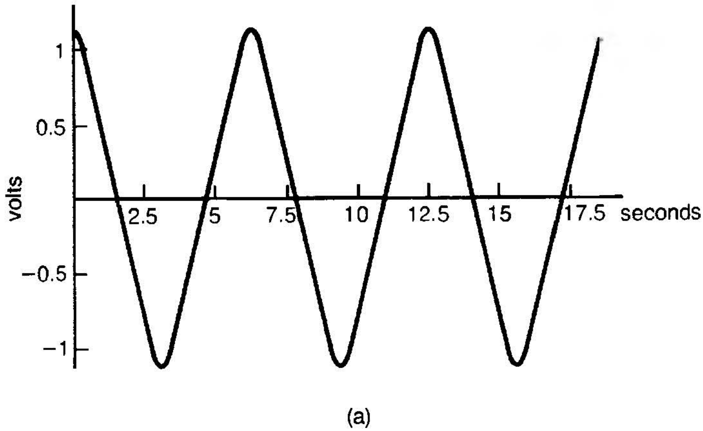
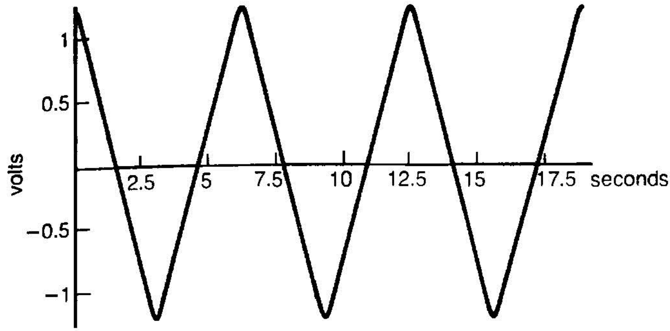


Figure 2.9 *Adding sinusoidal signals belonging to a different sequence produces a triangular signal. (a) 2 signals, (b) 6 signals*



(b)

Figure 2.9 *(continued)*

signals. It can be proved that any periodic signal, even one of very complicated shape, can be synthesized in this way.

> **Positive sines**
>
> Cosine terms and negative sine terms can all be written as positive sine terms if we change their phase. To convert cosines to sines, use the identity:
>
> $$\cos x \equiv \sin(x + 90^\circ)$$
>
> You can confirm this by checking values for a few angles using a calculator. The cosine example in the text can be rewritten:
>
> $$v = \sin(t + 90^\circ) + \frac{\sin(3t + 90^\circ)}{9} + \frac{\sin(5t + 90^\circ)}{25} + \cdots$$
>
> To convert negative sines to positive sines, use the identity:
>
> $$-\sin x \equiv \sin(x + 180^\circ)$$
>
> Check this on a calculator. The series for the sawtooth wave may be rewritten with all terms positive:
>
> $$v = \sin t + \frac{\sin(2t + 180^\circ)}{2} + \frac{\sin 3t}{3} + \frac{\sin(4t + 180^\circ)}{4} + \cdots$$

Analysing signals

The knowledge that periodic signals of all kinds can be built up from sine waves explains why there has been so much discussion of sine waves in this chapter.

Now we are going to perform the reverse process of analysing a non-sinusoidal signal to find out what sine waves are present. This is important when we are designing filters because the action of a filter is to remove or reduce the amplitude of some of the sine waves while letting others pass.

The routine for analysing a signal into its sinusoidal components is known as a **Fourier analysis**. This can be done on paper but the calculations are lengthy and tedious. Fortunately a computer can easily perform such tasks. Electronics engineers use programs specially designed for this purpose. The programs simulate the action of electronic circuits, display the signals as graphs and then perform a Fourier analysis. To illustrate this technique, we set up a circuit on the computer like that in Fig. 1.1, specifying that the voltage generator produces a sawtooth signal frequency $f = 1/2\pi = 0.1592\,\text{Hz}$, amplitude 1.58 V. Note that the signal produced by the computer is not obtained by summing sine signals of various frequencies and amplitudes. Instead, it is obtained by a routine which directly produces a sawtooth output. The Fourier analysis finds out which sinusoidal signals are present in the sawtooth signal and produces Fig. 2.10, in which a series of vertical lines are plotted to represent these sinusoidal components. Lines are plotted for each of the frequencies present and the height of each line indicates the amplitude. The tallest line plotted corresponds to v_1 which has a frequency of $1/2\pi$ Hz, equal to 159 mHz on the x-axis of Fig. 2.10. The height of the line shows that the amplitude of v_1 is 1 V. In other words the left-most line in Fig. 2.10 represents the first term of the Fourier series, $\sin t$. The next line represents v_2, the second term, $(\sin 2t)/2$. It is plotted at 318 mHz, and its height is 0.5 V. The third line represents v_3, the third term of the series, which is $(\sin 3t)/3$, and is

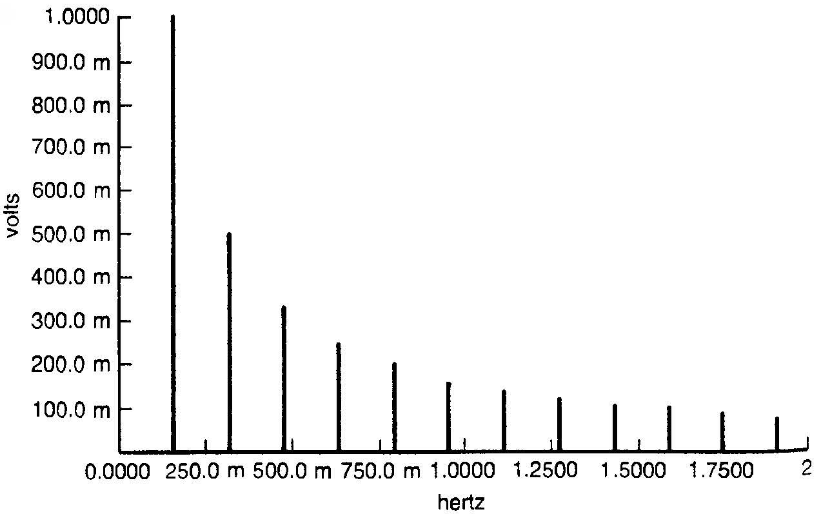


Figure 2.10 *The frequency spectrum of a sawtooth signal having the same shape as that in Fig. 2.8(d)*

plotted at 477 mHz, with height 0.333 V. The part of the analysis included in Fig. 2.10 represents all terms of the series up to the 12th term, $(\sin 12t)/12$, with frequency $159 \times 12 = 1908$ mHz, and amplitude $1/12 = 0.0833$ V or 83.3 mV.

Since Fig. 2.10 shows the frequencies found to be present in the sawtooth signal and their amplitudes, it is known as a **frequency spectrum**. A frequency spectrum is a very convenient way of representing a periodic signal and later we shall use frequency spectra to examine the ways in which filters affect signals passed through them.

Figure 2.11 shows the frequency spectrum produced by a Fourier analysis of a triangular curve similar to that shown in Fig. 2.9b. Once again, the triangular wave is generated directly by the computer, not made up by summing sine waves. The triangular wave used in the analysis has frequency $1/2\pi = 0.1592$ Hz, as before, and amplitude 1.25 V. However, the analysis shows that a regular set of sinusoids is present in the signal. The first term of the series is represented by the line on the left of the spectrum. This has the same fundamental frequency as in the previous analysis, 159 mHz and its amplitude is 1 V. The second line from the left is at 477 mHz and has amplitude 0.111 V, so this represents the term $(\cos 3t)/9$. Because the divisor in each term is the square of n, the amplitudes fall off very quickly, as can be seen in Fig. 2.11.

The Fourier analysis can also provide information about the phase of each term. In Fig. 2.12 we have the same analysis as in Fig. 2.10, but with the addition of phase information. The small squares on or above the lines of the spectrum indicate the phase, according to the scale on the right. There is a slight complication because, for reasons concerned with the analysis technique, phase is related to a

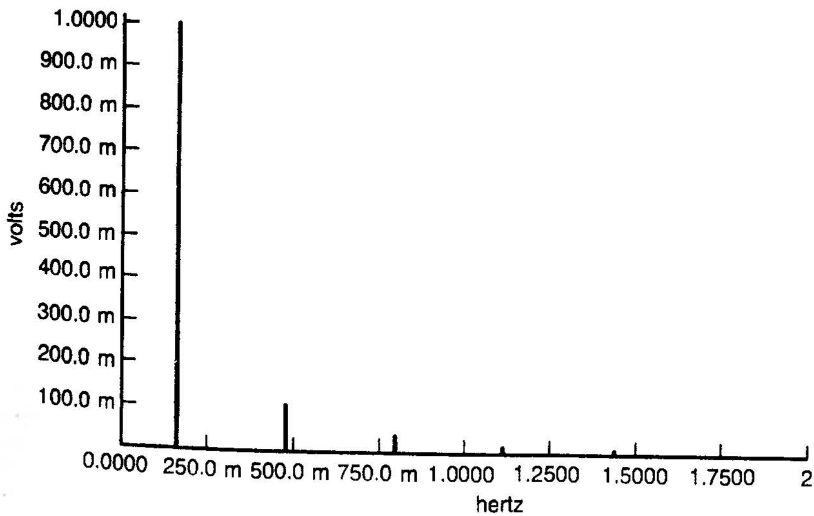


Figure 2.11 *The frequency spectrum of a triangular signal (Fig. 2.9(b)) differs in many respects from that of a sawtooth signal*

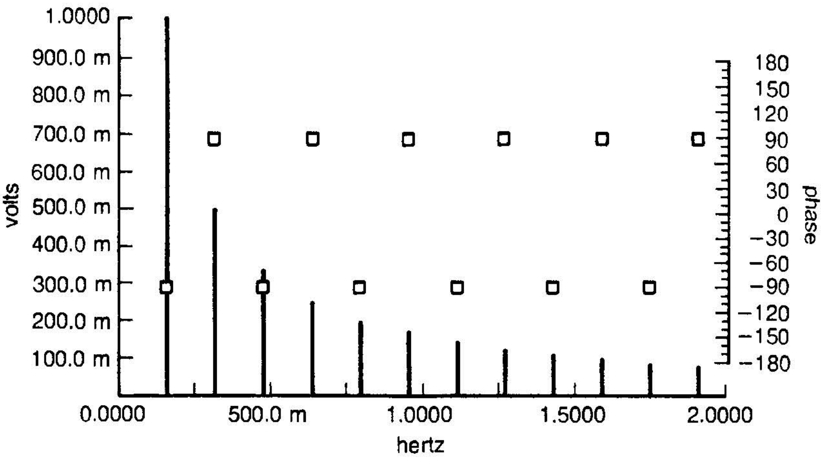


Figure 2.12 *When a phase plot is added to Fig. 2.10, it demonstrates that the terms of the equivalent Fourier series are alternately positive and negative*

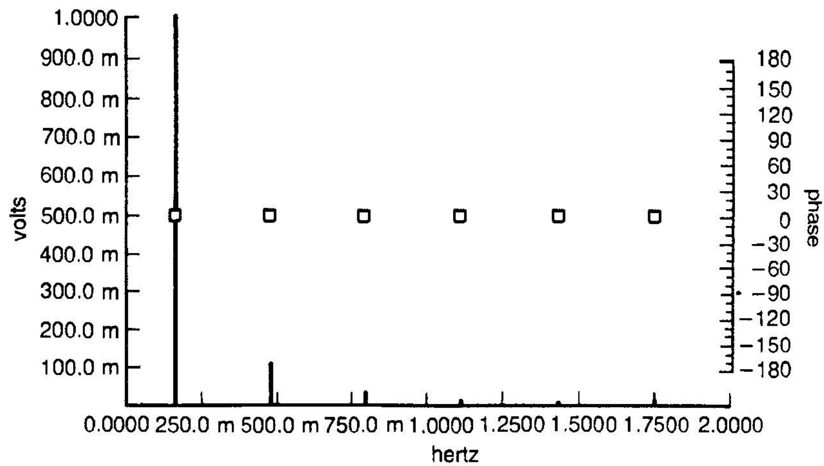


Figure 2.13 *This phase plot shows that all terms of the Fourier series of Fig. 2.11 are positive cosine terms*

sine curve that has been shifted to the left by a quarter of a cycle, so that it is symmetrical about the y-axis. Consequently, all phase values shown here and in Fig. 2.13 are 90° too small. So we must add 90° to the displayed phase angle to make them equivalent to the terms of the series. Taking this into account, it is seen that the odd terms in Fig. 2.10 have a phase angle of 0° while the even terms have a phase angle of +180°. This corresponds with the equation in the box where terms with a phase angle +180° are those that were negative sine terms in the original equation for v.

Figure 2.13 shows the analysis of the triangular wave (Fig. 2.11) with added phase information. Here all frequencies have zero phase angle, which means that the actual phase angle is 90°. A sine signal leading by 90° is the equivalent of a cosine signal (see box) so this plot corresponds to a series of cosine terms, as expected.

Fundamental and harmonics

The first term in the Fourier series (v_1) is known as the **fundamental**. It has a greater amplitude than the other signals and has the same frequency as the synthesized signal. Other components of the signal, with smaller amplitude and frequencies that are multiples of the fundamental, are known as **harmonics**. These are numbered in order. For example, v_2, represented by the term $(\sin 2t)/2$, is known as the first harmonic. v_3, represented by the term $(\sin 3t)/3$, is the second harmonic, and so on. Figure 2.12 shows the fundamental and the 1st to 11th harmonics. If harmonics are missing, for example the odd harmonics in Fig. 2.13, we still number the other harmonics in the same way. Figure 2.13 shows the fundamental and the even harmonics from the 2nd to the 10th.

Summary

A periodic signal is specified by stating these attributes:

Period — the length of a cycle, P, usually expressed in seconds.
Frequency — the number of periods in 1 second, usually expressed in hertz (Hz), where $f = 1/P$. It may also be expressed in terms of angular velocity, ω, where $\omega = 2\pi f$.
Offset — the level about which a signal fluctuates, usually expressed in volts or amps. Most of the signals we deal with have zero offset.
Amplitude — the maximum level above the offset reached by the signal, usually expressed in volts or amps.
Phase angle — The extent to which the phase of the signal leads or lags a given reference signal, sometimes expressed in degrees, more often expressed in radians. 1 rad = 57.296°.

Periodic signals can be analysed into the sum of a series of sinusoidal signals of differing amplitude and phase. Their frequencies are all multiples of a single fundamental frequency.

Test yourself

1. What are the frequencies, amplitudes, phase angles and offsets of these sinu soidal curves: (a) $v = 7.5 \sin 12.57t$, (b) $v = \sin(100\pi t + \pi)$, (c) $v = 3.6$ 0.65 $\sin(28.27t + 4.712)$?
2. List the attributes of the signals illustrated in Fig. 2.14, expressing their phase with reference to when $t = 0$.

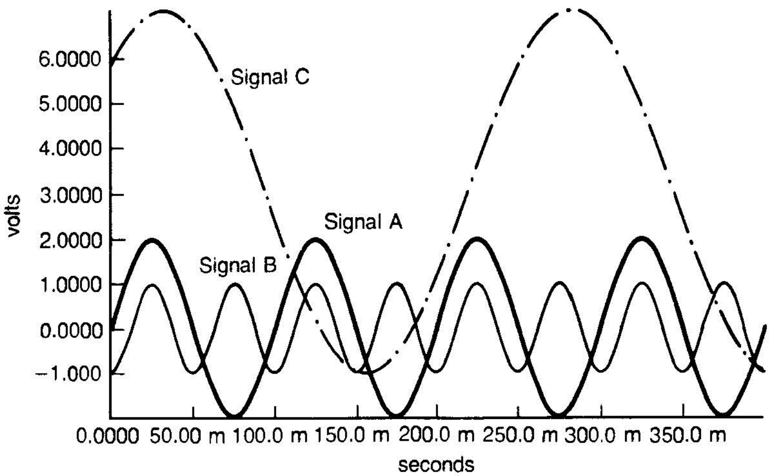


Figure 2.14 *Some signals for questions 2 and 3*

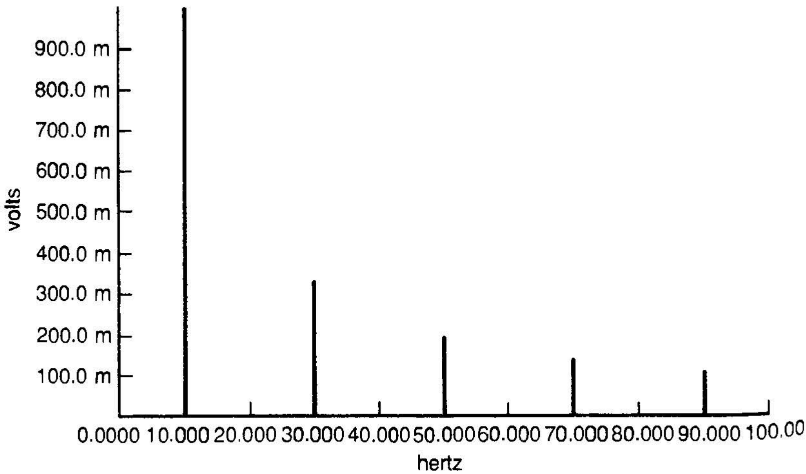


Figure 2.15 *A frequency spectrum for question 4*

3. Sketch the signal that is the sum of curves A and B in Fig. 2.14. Describe its frequency spectrum.
4. Describe the frequency spectrum shown in Fig. 2.15, write out its series and plot the curve obtained by summing the first 5 terms. What is the shape of the composite signal?

3

Reactance and impedance

This chapter begins by looking at the way capacitors behave when a sine wav pd is applied across them. Chapter 2 demonstrated that any periodic signal ca be analysed into a set of sinusoidal signals so, if we know how a capacitor reac to a given sinusoidal signal on its own, we can predict how it will behave wi a combination of sinusoids making up any periodic signal. Figure 3.1 shows capacitor being charged by a current generator. Note that, unlike the circuit Fig. 1.5, this circuit has no resistor. In Chapter 1 the capacitance C of a capacit is defined as the amount of charge (q) stored in it for a given pd (v) betwee its plates:

$$C = q/v$$

This equation may be rearranged:

$$v = q/C$$

q is the amount of charge carried by a current i flowing for time t (equation (5 p. 6) and:

$$q = it$$

assuming that i is held constant. In Fig. 3.2 a constant current i flows for a interval t. The current carried is it, the product of i and t. On the graph, th product of i and t is the shaded area beneath the curve. As a rule:

> On a current–time graph, the area below the curve represents the total charge carried by the current.

This rule can be applied to a current that is varying. In Fig. 3.3, the curve show current varying in an irregular way. Think of the area as being divided into ma narrow vertical strips. Each strip represents the current being carried during a ve short period of time, which we will call Δt. The Δ symbol is not a quantity th is multiplied by t. It is a symbol indicating that the interval is very short. Duri such a short period of time the current does not change appreciably. We c consider that the upper end of the strip is not sloping but is at right angles to t length of the strip. In other words, each strip is almost a rectangle. The area each rectangle is i multiplied by Δt, that is $i\Delta t$. The total area beneath the cur is the sum of the areas of the strips, and this represents the total charge. This

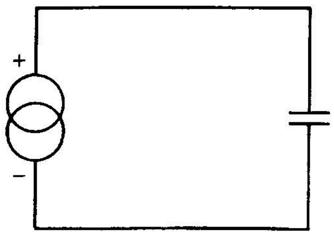


Figure 3.1 *The capacitor is being charged by a current generator*

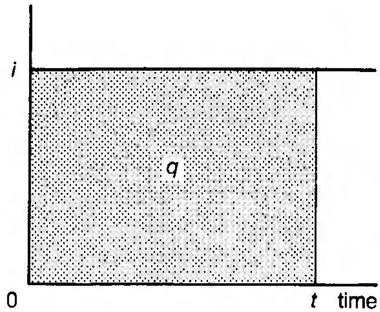


Figure 3.2 *The amount of charge is equal to the product of time and current, represented by the area below the line*

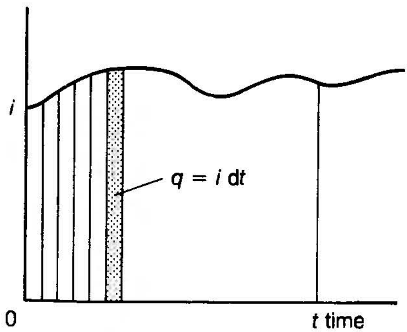


Figure 3.3 *The charge carried by an irregularly varying current may be found by summing the areas of the narrow strips*

written mathematically as:

$$q = \sum_{t=0}^{T} i\Delta t$$

In words, q equals the sum (symbol, Σ) of the areas of the strips from when $t = 0$ until $t = T$. If we make the strips narrower and narrower, so that Δt becomes infinitely small, the fact that the tops of the strips are sloping makes less and less difference to the result. In the limit, the equation becomes:

$$q = \int_0^T i \, dt$$

Now that we are dealing with infinitely small increments of t we replace (meaning 'sum') with another kind of S, the 'long S', $\int$, and we replace by dt (compare with p. 11). The discussion above follows along the same lin as the explanation of the meaning of integration. In mathematics we use the symbol to express integration, which is the operation of summing all the infinite narrow strips beneath the line of a graph to find the total area beneath the lin Given that the equation of the curve is known, there are standard maths routin for finding out the area, or integral.

Integration is a practical way of calculating the total charge carried by a varyi current. The total charge is the integral of the expression for the current over t period of time for which current flows. Charge carried to the capacitor durin that period is added to any charge already on the capacitor at the beginning. the initial charge is q_0, then:

$$q = \int_0^T i \, dt + q_0$$

As an example, consider Fig. 3.4 which shows the current produced by a gene ator according to this equation:

$$i = 2t$$

The current ramps up steadily for an interval of 6 s, increasing from zero to 12 To simplify the example we may assume that the capacitor is uncharged to beg with. To find the value of q, integrate the expression $2t$ from $t = 0$ to $t = 6$:

$$q = \int_0^6 2t \, dt = [2t^2/2]_0^6 = [t^2]_0^6 = [6^2 - 0^2] = 36$$

The charge accumulating on the capacitor in 6 seconds is 36 C. This proble could have been solved graphically, using Fig. 3.4. The current is zero to beg

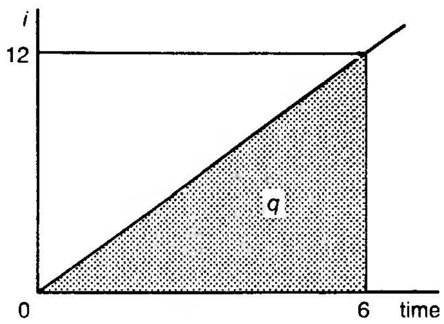


Figure 3.4 *The triangle represents the charge carried by a current $i = 2t$ during interval from 0 s to 6 s*

with and is 12 A after 6 s. The total charge is the area of the shaded triangle. Its area is $\frac{1}{2}$ × base × height = $\frac{1}{2} \times 6 \times 12 = 36$, which confirms the result obtained by integration.

Rules for integrating

If an equation contains powers of a variable such as t, the integral is found by increasing the power by 1 and then dividing by the value of the increased power. For example, if $i = 2t$, the power of t is 1. So increase this by 1 to make it 2, and divide by 2:

$$q = \int 2t \, dt = \frac{2t^2}{2} + c = t^2 + c$$

where c is an unknown amount, the constant of integration which we are able to ignore (see below). The integral above is the indefinite integral because it does not tell us the values of t to start with and finish with. In the text we have a definite integral which tells us that t starts at 0 and finishes at 6. To find the value of the definite integral, evaluate it for the finishing value of t and subtract its value for the starting value of t:

$$q = [6^2 + c] - [0^2 + c] = 36$$

The cs have disappeared.

Since most electrical signals are sinusoids, filter equations often contain the sine ratio. The integral of $\sin t$ with respect to t is $-\cos t$:

$$\int \sin t \, dt = -\cos t$$

Also, if a is a constant:

$$\int \sin at \, dt = \frac{-\cos at}{a}$$

$$\int \cos at \, dt = \frac{\sin at}{a}$$

Including phase angle, we have:

$$\int \sin(at + \phi) \, dt = \frac{-\cos(at + \phi)}{a}$$

$$\int \cos(at + \phi) \, dt = \frac{\sin(at + \phi)}{a}$$

All these integrals have the constant of integration, c, added to them, which can be ignored if we are evaluating the integral as a definite integral.

Time in terms of pi

From now on, many of the graphs have the time axis marked in pi instead of numbers of seconds. Figure 3.5, for example, shows a sine wave from time = 0 to time = 2π. This might mean that time runs from 0 s to 6.232 s, in which case $P = 0.62832$ s and $f = 1/P =$ 0.1592 Hz. But the graph covers many more cases than this. The graph shows the sine wave for 1 cycle, no matter what its frequency and how long a cycle may take. It applies equally to all sine waves, not just those for which $f = 0.1592$ Hz. This convention allows us to draw graphs and make deductions from them that apply irrespective of frequency.

A second example is illustrated in Fig. 3.5, in which:

$$i = 3 \sin t$$

The interval is from $t = 0$ to $t = \pi$, that is to say it runs for half a cycle Integrating $3 \sin t$ from $t = 0$ to $t = \pi$:

$$q = \int_0^{\pi} 3 \sin t \, dt = -3[\cos t]_0^{\pi} = -3[\cos \pi - \cos 0] = -3[-1 - 1]$$
$$= -3 \times -2 = 6$$

The charge accumulating in one half-cycle is 6 C.

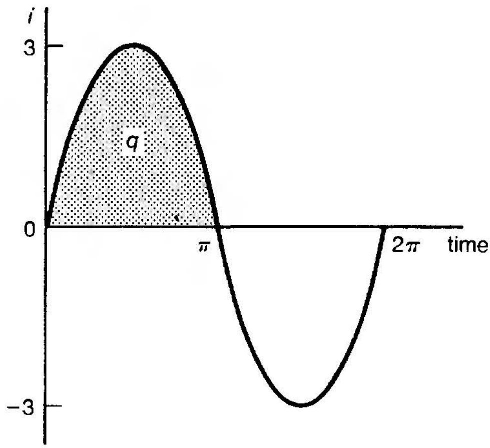


Figure 3.5 *The current equation is $i = 3 \sin t$ and the shaded area represents the charge carried during the first half-cycle*

Having calculated the charge on the capacitor, we can find the pd across it. Taking the example of Fig. 3.4 and given that the capacitance is 1.5 F, the pd across the capacitor after 6 s is:

$$v = q/C = 36/1.5 = 24\,\text{V}$$

In the example of Fig. 3.5, with the same capacitance, the pd after half a cycle is:

$$v = q/C = 6/1.5 = 4\,\text{V}$$

Keeping up?

1. In each of the examples below, which refer to the circuit of Fig. 3.1, you are given the equation for the current and the finishing time t. Assuming all timing starts from $t = 0$ and that the capacitor is uncharged at the start, sketch the curve of current against time. Find the charge accumulating during that time.
   (a) $i = 1.5$, $t = 4$
   (b) $i = 1.5t$, $t = 0.5$
   (c) $i = 2t^2$, $t = 0.1$
   (d) $i = 4$ for $0 < t < 1.2$
   $i = 10 - 5t$ for $1.2 < t < 2$
   (e) $i = 3 + 4\cos 2t$, $t = \pi/10$ (Hint: put calculator in radian mode)
2. For each of the examples in question 1, the corresponding capacitances are listed below. In each example find the pd across the capacitor at time t.
   (a) $C = 2\,\text{F}$, (b) $C = 2200\,\mu\text{F}$, (c) $470\,\mu\text{F}$, (d) $0.5\,\text{F}$, (e) $10\,000\,\mu\text{F}$.
3. In the example of Fig. 3.5, calculate the charge on the capacitor after 2πs. Explain the reason for this result.

Current and pd

This section discusses the relationship between current flowing into a capacitor and the pd appearing across the capacitor as charge accumulates there. Begin with a circuit like Fig. 3.1, except that the current generator is replaced by a sine wave pd generator. At this stage we do not know the equation for current but this does not stop us from using the equation for charge (p. 36) in which current at any (unspecified) instant is represented by i. Convert this from an equation for charge into an equation for the pd across a capacitor by dividing by the capacitance:

$$v = \frac{1}{C}\int_0^T i\,\mathrm{d}t + \frac{q_0}{C}$$

This equation contains both v and i and we will use this to find the relationship between v and i in a circuit like Fig. 3.1. The first step is to get rid of the integral by differentiating both sides of the equation:

$$\frac{dv}{dt} = \frac{1}{C}i$$

More rates of change

Differentiation and integration can be thought of as opposite operations. Therefore, derivatives of trig ratios are the inverses of the integrals of trig ratios given in the box on p. 37. The derivative of $\sin(ax + \phi)$ is $a\cos(ax + \phi)$. The derivative of $\cos(ax + \phi)$ is $a\cos(ax + \phi)$.

Differentiation has the opposite effect to integration, so differentiating the integra of i with respect to t simply produces i. Differentiating the constant value $q_0/($ produces zero, since a differential is a rate of change and, by definition, a constan has zero rate of change. This removes the initial charge q_0 from the equation demonstrating that any charge initially on the capacitor has no effect on wha happens next. Rearranging:

$$i = C(dv/dt)$$

This equation has a derivative expression in it and, if we assume the standar equation for a sine wave, we can remove the derivative by actually carrying ou the differentiation. If:

$$v = V_0 \sin \omega t$$

Differentiating with respect to t gives:

$$dv/dt = V_0\omega \cos \omega t$$

Substituting this in the equation for i:

$$i = CV_0\omega \cos \omega t$$

Comparing the form of the expression on the right-hand side of the equatio with the expression on the right-hand side of the sine wave equation (p. 21 we see that this is the equation of a sinusoidal current signal, amplitude $CV_0\omega$ frequency $\omega/2\pi$. The expression for the amplitude may be replaced by a sing symbol I_0, the maximum value (or amplitude) of the current, where $I_0 = CV_0\omega$ All three quantities C, V_0 and ω are constants, so I_0 is a constant. Substitutin I_0 for $CV_0\omega$:

$$i = I_0 \cos \omega t$$

For comparison here is the pd equation:

$$v = V_0 \sin \omega t$$

Both current and pd are sinusoidal signals with zero offset, and both have the same frequency. But the current signal is based on cosines. It has the same shape as the pd signal but it is out of phase with it. According to the box on p. 27, $\cos\theta = \sin(\theta + \pi/2)$, so we can rewrite the current equation:

$$i = I_0 \sin(\omega t + \pi/2)$$

The current into a capacitor leads the pd by $\pi/2$, a quarter of a cycle, as seen in Fig. 1.7, and as summarized at the end of Chapter 1.

Reactance

There is a relationship between the amplitudes of the current and voltage equations, that is to say between V_0 and I_0. This can be found by remembering that we substituted I_0 for $CV_0\omega$ during the discussion above. In other words:

$$I_0 = CV_0\omega$$

Rearranging:

$$\frac{V_0}{I_0} = \frac{1}{C\omega}$$

The ratio between the maximum pd and the maximum current is a constant, $1/C\omega$. There is a parallel statement to this in Chapter 1, which is expressed in equation (1). There we said that the ratio between the pd across a resistor and the current through the resistor is a constant, R, the resistance. The constant $1/C\omega$ is the equivalent of the resistance of the capacitor. Like resistance, it determines the current for a given pd, or the pd for a given current. Because its unit is volts per amp, the same as resistance, its unit is the ohm. But its action is entirely different to that of resistance, so we give it another name, **reactance**. Reactance has its own special symbol X_C, the suffix C indicating that it is capacitative reactance.

The definition of capacitative reactance is expressed in the equation:

$$X_C = 1/C\omega$$

This shows that reactance is inversely proportional to capacitance. The larger the capacitance, the smaller the reactance. It is also inversely proportional to frequency. It is in this respect that reactance differs most strongly from resistance. The resistance of a resistor is the same at all frequencies. By contrast, the reactance of a capacitor is high at low frequencies and low at high frequencies. Another important equation is:

$$\frac{V_0}{I_0} = X_C$$

We use this equation to relate the amplitudes of the pd and current signals.

A few examples will make these relationships clear. At 50 Hz, $\omega = 2\pi f = 2\pi \times 50 = 314.2$ rad/s. The reactance of a 22 μF capacitor at that frequency is:

$$X_C = 1/C\omega = 1/(22 \times 10^{-6} \times 314.2) = 145\,\Omega$$

Now increase the frequency to 1 kHz. $\omega = 2\pi \times 1000 = 6283$. The reactance of the same capacitor is:

$$X_C = 1/(22 \times 10^{-6} \times 6283) = 7.23\,\Omega$$

Having calculated reactance at a given frequency (there is no such thing as 'general' reactance; the frequency must always be stated), we can use this to calculate the amplitude of the voltage or current signal. For example, the 50 Hz signal applied to a 22 μF capacitor has an amplitude of 1.2 V. What is the amplitude of the current signal? We have already calculated that $X_C = 145\,\Omega$. Then we can say:

$$I_0 = V_0/X_C = 1.2/145 = 8.28\,\text{mA}$$

Increasing the frequency to 1 kHz reduces the reactance and the current amplitude becomes:

$$I_0 = 1.2/7.23 = 166\,\text{mA}$$

Current flow is much more at the higher frequency.

Keeping up?

4. Write out the three equations which relate pd, current and resistance.
5. Write out the three equations which relate pd amplitude, current amplitude and capacitative reactance.
6. Calculate the reactance of these capacitors at the given frequency:
   (a) $C = 0.5\,\text{F}$, $f = 500\,\text{Hz}$
   (b) $C = 2\,\text{F}$, $f = 10\,\text{kHz}$
   (c) $C = 33\,\mu\text{F}$, $f = 25\,\text{kHz}$
   (d) $C = 27\,\text{pF}$, $f = 2.2\,\text{MHz}$
   (e) $C = 18\,\text{nF}$, $f = 10\,\text{Hz}$
7. A pd signal of $v = 5.4 \sin 1000t$ is applied to a 180 nF capacitor. Calculate the amplitude of the current signal.

Inductor behaviour

The behaviour of an inductor when a varying current flows through it is described in Chapter 1. This is the result of the back emf e:

$$e = -L \cdot \text{d}i/\text{d}t$$

where L is the self-inductance of the inductor, $\text{d}i/\text{d}t$ is the rate of change of current and the negative sign indicates that the induced emf acts to oppose changes in the current. If the current is increasing, $\text{d}i/\text{d}t$ is positive, making e negative and so opposing the increase. This situation is illustrated in Fig. 3.6. If v is such that current is decreasing, the polarity of e is reversed. Taking the

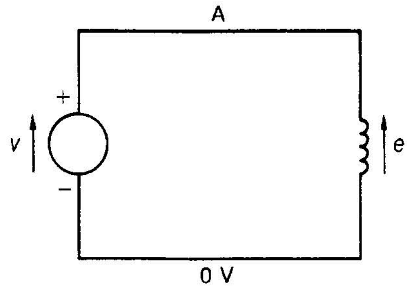


Figure 3.6 *Back emf, e, opposes the increase of current resulting from any increase in v*

potential on the bottom line of the circuit to be 0 V, then at all times e must be equal and opposite to v. If not, there would be two different potentials on the top line of the circuit (A), which would be impossible. The consequence of this is that:

$$v + e = 0$$

Assume that the generator is producing a sinusoidal signal, $v = V_0 \cos \omega t$, and is causing a varying current to flow through the inductor. This causes a back emf such that:

$$V_0 \cos \omega t - L \cdot \frac{\mathrm{d}i}{\mathrm{d}t} = 0$$

Rearranging:

$$\frac{\mathrm{d}i}{\mathrm{d}t} = \frac{V_0}{L} \cos \omega t$$

Integrating both sides of this equation gives:

$$i = \frac{V_0}{\omega L} \sin \omega t$$

Remember that integrating a derivative gives back the original expression, in this case simply i. The expression $V_0/\omega L$ is a constant which can be taken to be the amplitude I_0 of a sinusoidal current. The current through the coil is therefore:

$$i = I_0 \sin \omega t$$

Compared with the equation of the voltage generator:

$$v = V_0 \cos \omega t$$

Both current and pd have the form of a sinusoidal signal with zero offset, and both have the same frequency. But the current is based on sines so it is out of phase with the pd which is based on cosines. For comparison, we should turn the pd equation into an equation based on sines. Use the identity $\cos \theta = \sin(\theta + \pi/2)$, as explained in the box on p. 27:

$$v = V_0 \sin(\omega t + \pi/2)$$

The pd leads the current by $\pi/2$, a quarter of a cycle, as seen in Fig. 1.11. Or, taking the other viewpoint, we can say that the current lags the pd by a quarter of a cycle, as stated in the summary at the end of Chapter 1.

Inductive reactance

In the discussion of capacitance we replaced the expression $CV_0\omega$ by I_0, the amplitude of the current signal. In this discussion we replace $V_0/\omega L$ by I_0:

$$I_0 = V_0/\omega L$$

Rearranging:

$$\frac{V_0}{I_0} = \omega L$$

Once again we have a constant, this time ωL, which can be said to correspond with resistance, for it expresses the relationship between pd and current. In the case of the inductor, this constant is the inductive reactance, symbol X_L and:

$$X_L = \omega L$$

Because X_L is calculated by dividing a voltage by a current, its unit is the ohm. The equation shows that reactance is directly proportional to inductance. The larger the inductance, the larger the reactance. It is also directly proportional to frequency, the effect of frequency on the inductor being the opposite to its effect on a capacitor. The reactance of an inductor is high at high frequencies and low at low frequencies.

As with X_C, we can use X_L to relate the amplitudes of the pd and current signals. Here are some examples. At 250 Hz, $\omega = 2\pi f = 2\pi \times 250 = 1571$ rad/s. The reactance of a 100 mH inductor at that frequency is:

$$X_L = \omega L = 1571 \times 100 \times 10^{-3} = 157\,\Omega$$

Increasing the frequency to 200 kHz has the following result. At 200 kHz, $\omega = 1.257 \times 10^6$ rad/s. The reactance of the same inductor is now:

$$X_L = 1.257 \times 10^6 \times 100 \times 10^{-3} = 125.7\,\text{k}\Omega$$

The reactance has increased dramatically at the higher frequency.

As with capacitative reactance, we can use inductive reactance to calculate the amplitude of a voltage or current signal. For example given that $V_0 = 5$ V, in the first example above, the current amplitude is:

$$I_0 = V_0/X_L = 5/157 = 31.8\,\text{mA}$$

In the second example, using the same inductor but increasing the frequency to 200 kHz, the current amplitude becomes:

$$I_0 = 5/(125.7 \times 10^3) = 39.78\,\mu\text{A}$$

The current is much reduced at the higher frequency.

Keeping up?

8. Write out the three equations which relate pd amplitude, current amplitude and inductive reactance.
9. Calculate the reactance of these inductors at the given frequency:

   (a) $L = 0.5\,\text{H}$, $f = 20\,\text{Hz}$
   (b) $L = 100\,\text{mH}$, $f = 5\,\text{kHz}$
   (c) $L = 25\,\mu\text{H}$, $f = 2\,\text{MHz}$
   (d) $L = 25\,\mu\text{H}$, $f = 5\,\text{Hz}$
   (e) $L = 350\,\text{mH}$, $f = 300\,\text{kHz}$

10. A current signal of $i = 3.3 \sin 44t$ is applied to a 250 mH inductor. Calculate the amplitude of the pd signal.

Phase angle

The reactance of a capacitor or inductor affects not only the relationship between current and pd amplitude but the phase relationship between the two signals. With capacitative reactance, the current signal leads the pd signal by $\pi/2$. In question 7 above, for example, the pd signal is:

$$v = 5.4 \sin 1000t$$

We calculated that the amplitude of the corresponding current signal is $I_0 = 972\,\mu\text{A}$. Not only does the current signal have a much smaller amplitude than the pd signal but it leads the pd signal by $\pi/2$. Writing out this signal in full:

$$i = 972 \times 10^{-6} \times \sin(1000t + \pi/2)$$

In question 10 above, the current signal is

$$i = 3.3 \sin 44t$$

The amplitude of the corresponding pd signal is 5.778 V. But the pd signal lags $\pi/2$ behind the current signal, so the pd signal is:

$$v = 5.778 \sin(44t - \pi/2)$$

It is important to account for phase angles when calculating equations for pd and current. There is more about this topic in the next chapter.

Impedance

Resistance, capacitative reactance and inductive reactance all act to oppose the flow of current. Resistance opposes current flow because of the resistivity of the conductor from which it is made. Capacitance reactive opposes current flow because of charge present on the plates of the capacitor. Inductive reactance opposes current flow (when current changes) because of the electro-magnetic field within the inductor. A term that covers all these kinds of opposition to current flow is **impedance**. The impedance offered to current flow in a circuit is the sum of the resistances and reactances of its components. Impedances in parallel or in series can be summed according to the usual rules for resistances, but the result always depends on frequency.

Summary

The three kinds of impedance are summarized in this table:

	Resistance	*Capacitative reactance*	*Inductive reactance*
Symbol	R	X_C	X_L
Unit	ohm	ohm	ohm
Definition	$R = v/i$	$X_C = 1/\omega C$	$X_L = \omega L$
Effect of frequency	none	greatest at low frequency	greatest at high frequency
Phase relationship of v and i	in phase	i leads v by $\pi/2$	i lags v by $\pi/2$

Test yourself (Answers to 4 significant figures)

1. What is the reactance of a 15 µF capacitor at 2.5 kHz?
2. What is the reactance of a 25 mH inductor at 75 kHz?
3. The reactance of a capacitor is 360 Ω at 2 kHz. What is its capacitance?
4. At what frequency does a 50 µH inductor have a reactance of 35 Ω?
5. A current generator connected across a capacitor (as in Fig. 2.13) produces a signal $i = 0.25 \sin 100t$. If the pd across the capacitor varies as $v = 3.5 \sin(100t - \pi/2)$, what is the capacitance?
6. What value inductor has the same reactance as a 33 nF capacitor at 400 kHz?

4

Working with phase

Figure 4.1 is a set of curves obtained from a circuit like that of Fig. 1.5. In this example, the pd signal produced by the generator is $v = 3\sin 2000\pi t$ (that is, a 1 kHz signal with 3 V amplitude), the resistor is 680 Ω and the capacitor is 330 nF. The pds and also the current through the resistor are plotted from 2 ms onward, to give them time to settle down to steadily repeating values. The current is plotted on a ×250 scale. In accordance with the rules of circuit behaviour that we have already described:

All signals have the same frequency.
At any instant, v_c lags the current signal by 90°.
v_r is in phase with i; therefore v_c lags v_r by 90°.
At any instant, $v = v_r + v_c$.

All this can be seen in Fig. 4.1. Since v_r is out of phase with v_c, v must be out of phase with both v_r and v_c. The task is to find the phase relationships of the three signals and then calculate the amplitudes of v, v_r and v_c.

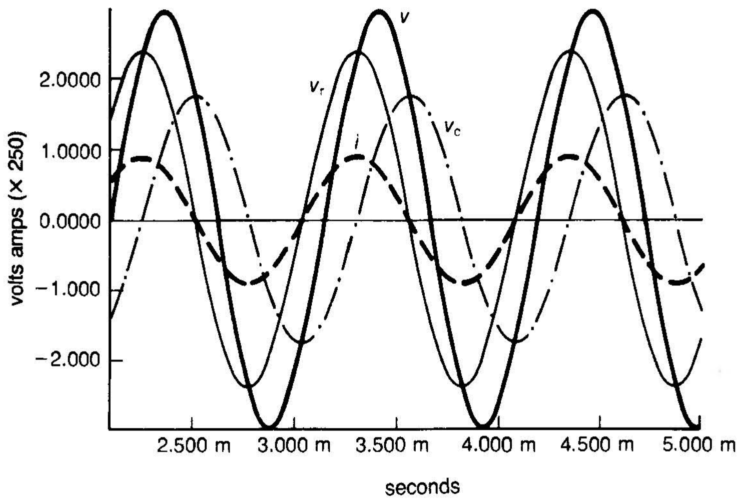


Figure 4.1 *These curves show the pds and current for a circuit like that of Fig. 1.5, as plotted by a circuit simulator. This chapter examines the relationships between these curves*

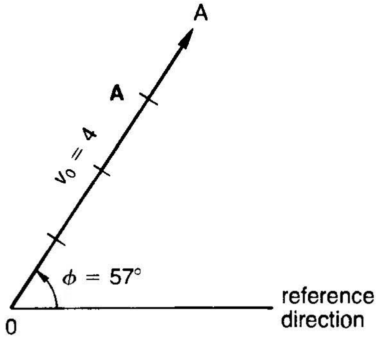


Figure 4.2 *A phasor, like a vector, has both magnitude and angle*

Phasor diagrams

One way of investigating phase is to draw a phasor diagram. A phasor is a kind of vector, which represents a sinusoidal pd or current signal in diagrammatic form. The length or magnitude of the phasor represents the amplitude of the signal. The direction of the phasor represents the phase angle of the signal. This is illustrated in Fig. 4.2, which shows a phasor **A**. The fact that the letter **A** is printed in bold type indicates that it is a phasor. The length of **A** represents the amplitude of the signal. In this example, the phasor is 4 units long, drawn to scale. The direction of the phasor is measured from a reference line running across the page to the right. In Fig. 4.2, this angle is 57°. If the phasor represents a sine wave pd signal with amplitude 4 V and phase angle 57°, then:

$$v = 4\sin(\omega t + 57^\circ)$$

One attribute that the phasor does not represent is the frequency, which is why we have written ωt in the equation instead of quoting a numerical value. The phasor in Fig. 4.2 could represent a signal of any frequency. But the phasors in any one phasor diagram all represent signals having the same frequency, which is taken into account when impedances (and hence amplitudes) are calculated.

Keeping up?

1. What signals are represented by the phasors in Fig. 4.3?
2. Sketch the phasors which represent these signals:

   (a) $v = 3.1\sin(50t + 45^\circ)$, (b) $v = 1.2\sin(100t - 75^\circ)$, (c) $v = 2.2\sin(\omega t -$ $140^\circ)$.

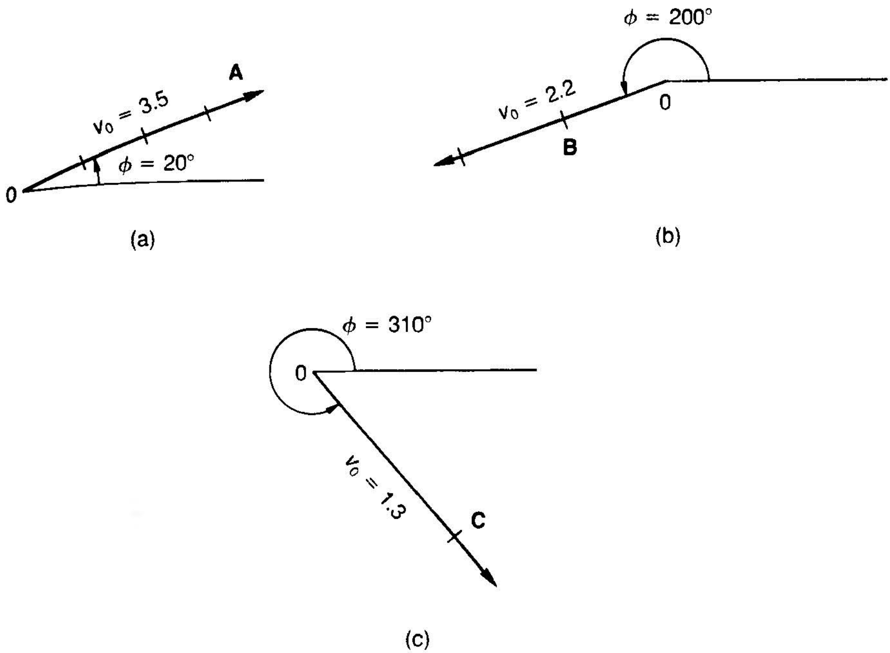


Figure 4.3 *Some phasor examples for question 1*

Adding phasors

Phasors are added in the same way as vectors. There are three cases relevant to working with filters:

1. Phasors with the same phase angle (Fig. 4.4): draw them end to end and sum their lengths. This is so simple to do that a diagram is not really necessary. Since the phase angles are equal, these phasors correspond to two signals that are in phase. Their sum is:

$$v = 3\sin(\omega t + 48^\circ) + 2\sin(\omega t + 48^\circ) = 5\sin(\omega t + 48^\circ)$$

   Sum the amplitudes and keep the same phase angle.

2. Phasors with complementary phase angles (pointing in opposite directions, Fig. 4.5):

$$\begin{aligned} v &= 5\sin(\omega t + 120^\circ) + 2\sin(\omega t + 340^\circ) \\ &= 5\sin(\omega t + 120^\circ) - 2\sin(\omega t + 120^\circ) \quad \text{(see box, p. 27)} \\ &= 3\sin(\omega t + 120^\circ) \end{aligned}$$

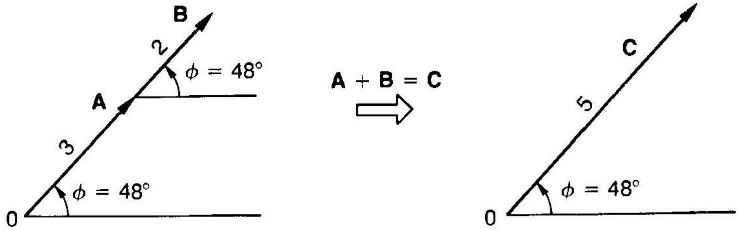


Figure 4.4 *Summing two phasors that have equal phase angles*

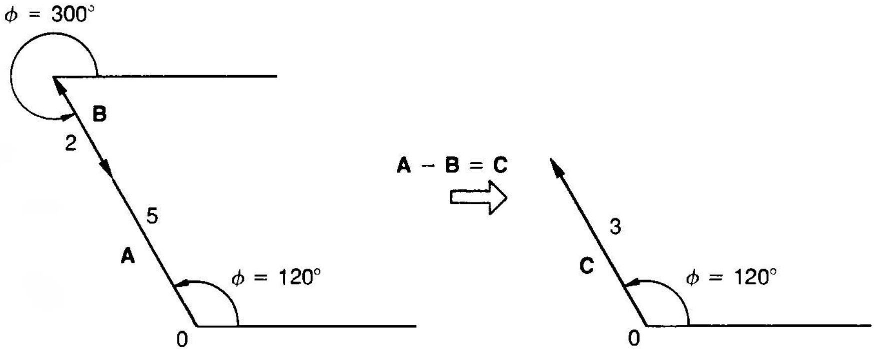


Figure 4.5 *Summing two phasors with complementary phase angles*

Find the difference between their lengths, and take the direction of the longer phasor.

3. Phasors which are 90° apart (Fig. 4.6): the two phasors are **A** and **B**; their sum is **C**, where **C** is the diagonal of the rectangle OPQR. We could use a scale diagram but it is simpler and more accurate to use geometry. Find the length of **C** by using Pythagoras' theorem, knowing that PQ = OR = 3. In the triangle OPQ:

$$OQ^2 = OP^2 + PQ^2 = 4^2 + 3^2 = 16 + 9 = 25$$
$$OQ = \sqrt{25} = 5$$

Find the direction of **C** by using tangents:

$$\tan Q\hat{O}P = QP/OP = 3/4 = 0.75$$
$$Q\hat{O}P = \tan^{-1} 0.75 = 36.9^\circ$$

The phasor sum has length 5 and phase angle 36.9°:

$$v = 5\sin(\omega t + 36.9^\circ)$$

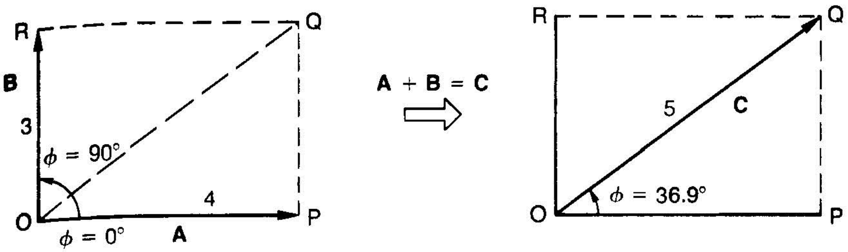


Figure 4.6 *Summing two perpendicular phasors*

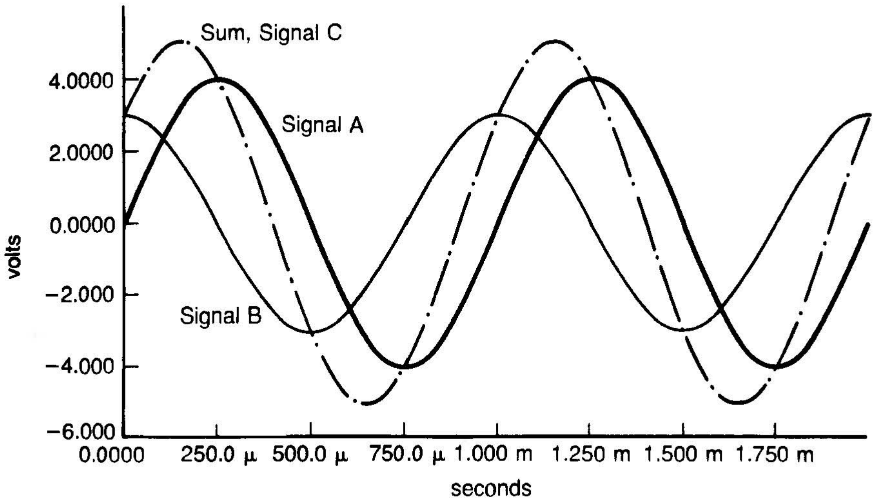


Figure 4.7 *The signals produced by rotating the phasor diagram of Fig. 4.6*

As a general rule, the amplitude is larger than that of either of the original signals and the phase angle is between their phase angles. Figure 4.7 shows the two signals of Fig. 4.6 and their sum. Inspection by eye confirms that the addition is correct.

Keeping up?

Sum these pairs of signals using phasor diagrams

3. $v_1 = 4\sin(\omega t + 20^\circ)$, $v_2 = 2.5\sin(\omega t + 20^\circ)$.
4. $v_1 = 12.6\sin(100\pi t + 15^\circ)$, $v_2 = 4.5\sin(100\pi t + 195^\circ)$.
5. $v_1 = 5\sin(25\pi t + 45^\circ)$, $v_2 = 7\sin(35\pi t + 45^\circ)$.
6. $v_1 = 3\sin\omega t$, $v_2 = 2.5\sin(\omega t + 90^\circ)$.

Hint: one of these pairs can not be summed.

Relating v_r and v_C to v

The phasor technique is used to solve the problem which began this chapter (Fig. 4.1). The problem differs from the one in the previous paragraph because we know the amplitude of v, the resultant, and want to find the magnitudes and phase angles of v_C and v_r. The stages in drawing the phasor diagram are explained in Fig. 4.8:

1. Find the impedance of C: from the definition of capacitative impedance, and given that the frequency is 1 kHz:

   $$\omega = 2\pi f = 2 \times 3.1412 \times 1000 = 6283 \text{ rad/s}$$

   The capacitance is 330 nF and:

   $$X_C = \frac{1}{\omega C} = 1/(6283 \times 330 \times 10^{-9}) = 482.3\ \Omega$$

   If I_0 is the amplitude of the current, the amplitude of the pd across the capacitor is $482.3I_0$, and the pd across the resistor is $680I_0$ (compare with equation (2), $V = IR$).

2. Take the phasor v_r to be the reference phasor, with a phase angle of 0°, draw OP to represent this, pointing across the page to the right.

3. We know that v_C lags v_r by 90°. Sketch in OR to represent v_C at right angles to and clockwise of v_r. This is not necessarily a scale diagram, but at least draw v_C so that the lengths of v_r and v_C are roughly in the ratio 680 to 480. Complete the rectangle OPQR. This establishes the shape or proportions of the rectangle (Fig. 4.8).

4. The pd phasor v is represented by the diagonal OQ. Its length is found by Pythagoras' theorem:

   $$OQ^2 = OP^2 + OR^2 = (680I_0)^2 + (482.3I_0)^2 = 695\,013.29I_0^2$$

   $$OQ = \sqrt{695\,013.29I_0^2} = 833.6I_0$$

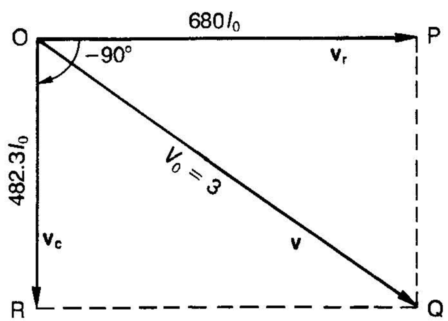


Figure 4.8 *This phasor diagram establishes the proportions of the rectangle OPQR*

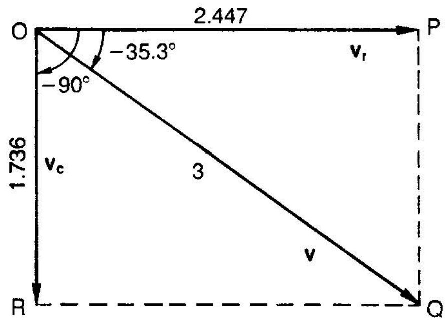


Figure 4.9 *The amplitudes and angles calculated after the scale of Fig. 4.8 has been established*

5. The length OQ represents an amplitude of $V_0 = 3$ volts. This fact makes it possible to determine the scale of the diagram (Fig. 4.9). In proportion, OP represents:

$$\frac{3 \times 680}{833.6} = 2.447\,\text{V}$$

This is the amplitude of v_r. Similarly, the OR represents the amplitude of v_c:

$$\frac{3 \times 482.3}{833.6} = 1.736\,\text{V}$$

6. The angle ϕ between v and v_r is found by using tangents:

$$\tan\phi = -1.736/2.447 = -0.7090$$

$$\phi = \tan^{-1} -0.7093 = -35.3^\circ$$

This is the angle by which v lags v_r. Figure 4.9 shows that v_c lags v by $-90 + 35.3 = -54.7^\circ$. We now have all the information we need to write out the equations of the signals, the frequency of all three signals being 1 kHz:

Given that $\quad v_r = 2.447 \sin 2000\pi t$

then $\quad v_c = 1.736 \sin(2000\pi t - 90^\circ)$

and $\quad v = 3 \sin(2000\pi t - 35.3^\circ)$

To simplify the diagram we have taken v_r to be the reference phasor, with zero phase angle. Often we prefer to make the generator pd the reference phasor, as in Fig. 4.1. We change the reference simply by rotating the phasor diagram, as in Fig. 4.10. This makes no difference to the relative lengths of the phasors or to the sizes of the angles between them. We have simply added 35.3° to the phase of each phasor and now the equations are:

$$v_r = 2.447 \sin(2000\pi t + 35.3^\circ)$$

$$v_c = 1.736 \sin(2000\pi t - 54.7^\circ)$$

$$v = 3 \sin 2000\pi t$$

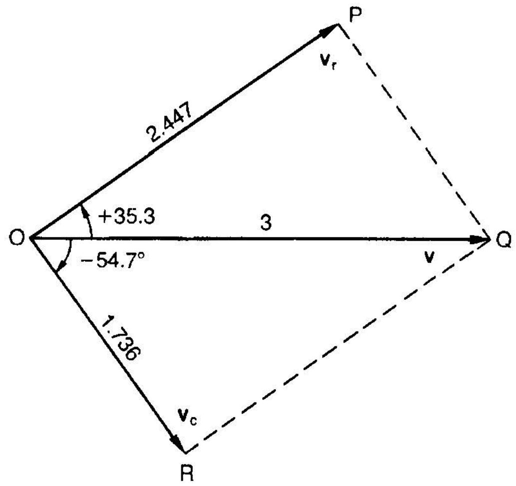


Figure 4.10 *Rotating Fig. 4.9 makes* **v** *the reference signal. The dimensions of the rec angle are unaltered but the phase angles are all increased by 35.3°*

Current equation

Applying equation (3) and knowing that the current signal is necessarily in phase with the pd across the resistor, calculate:

$$I_0 = V_0/R = 2.447/680 = 3.599\,\text{mA}$$

Write the current equation (in milliamps):

$$i = 3.599\sin(2000\pi t + 35.3^\circ)$$

This gives the phase of the current signal with reference to v.

Phasor solutions

A summary of the stages in finding the relationship between v, v_r and v_c.

References: Figures 1.5 and 4.8–4.10

Given: The sine wave equation for the voltage generator, and the values of R and C.

Calculate: Impedance X_c of capacitor.
Pds across R and C, in terms of I_0; their ratio establishes the shape of the rectangle.

Sketch:	Phasors v_r and v_c at right angles to each other. Complete the rectangle. Draw the diagonal to represent v.
Calculate:	Pd across generator, in terms of I_0 (Pythagoras); its ratio to V_0 establishes the scale of the rectangle. Actual amplitudes of signals across R and C. Angle between v and v_r (using tangents). Angle between v and v_c (by subtraction).
Write:	Equations for v, v_r, and v_c, with v_r as the reference.
If preferred:	'Rotate' the diagram to make v or v_c the reference.

Keeping up?

In the circuit of Fig. 1.5, given the equation for v and the values of R and C, calculate the equations for v_r and v_c with v as the reference.

7. $v = 4\sin 1000\pi t$, $R = 820\,\Omega$, $C = 470\,\text{nF}$.
8. $v = 2.5\sin 10\pi t$, $R = 12\,\text{k}\Omega$, $C = 1.5\,\mu\text{F}$.
9. $v = 15\sin 400\pi t$, $R = 1\,\text{M}\Omega$, $C = 22\,\text{nF}$.

Root mean square values

The equations of sinusoidal pds and currents specify their magnitudes in two different ways. As an example, take this standard equation for a pd signal:

$$v = V_0 \sin \omega t$$

v tells us the size of the pd at a given instant in time; one of this infinite number values of v is picked out in Fig. 4.11. V_0, also pictured in Fig. 4.8, tells us the largest and smallest values the pd attains during a cycle. Neither of these values are really typical of the signal. We need a more representative value for the pd, a sort of average value. Unfortunately the average value of any sinusoid is either zero or the value of its offset, if any. This is because v oscillates equally above and below zero (or the offset) and all the positive values are exactly cancelled out by all the negative values.

One way around the problem is to calculate the **root mean square** pd, v_{rms}. An approximate method of calculating v_{rms} is to take several values of v (say, 20 values) evenly spaced in time during one cycle. Square them (so making them all positive) and sum them, then divide by 20 to obtain the mean square. Finally take the square root of the mean square. In reality a sine curve is continuous and v takes infinitely many values during a cycle, so the method is not really

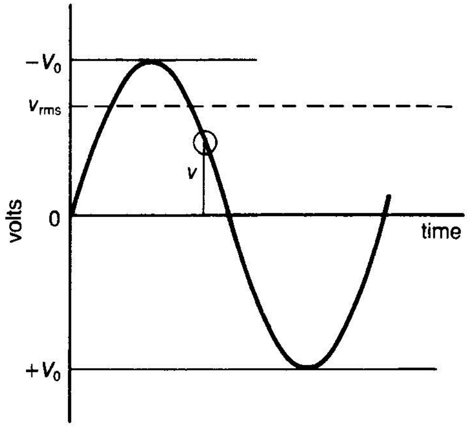


Figure 4.11 *Various ways of specifying the magnitude of a sinusoidal signal*

practicable. Fortunately it can be shown that:

$$v_{rms} = \frac{V_0}{\sqrt{2}}$$

v_{rms} is shown as the dashed line in Fig. 4.8. To 4 significant figures, $v_{rms} = V_0/1.414 = 0.7071V_0$. For example, given a signal $v = 4.5 \sin 200t$, the value of V_0 is 4.5, so $v_{rms} = 0.7071 \times 4.5 = 3.182\,\text{V}$.

Because there is a fixed ratio between amplitude and rms, we can use either of these in phasor diagrams. Whether we use amplitude or rms affects the scale of the diagram, but not its shape. The ratios between the lengths of phasors, or the angles between phasors, are not affected by whether we use amplitude or rms. The essential point is to keep to amplitude or rms for all stages in the calculations.

Pd and current values

Given a pd signal, $v = V_0 \sin \omega t$, or a current signal, $i = I_0 \sin \omega t$, we declare the size of the signal by using:

the instantaneous value	v	i
the amplitude	V_0	I_0
the root mean square	$v_{rms} = 0.7071\text{V}_0$	$i_{rms} = 0.7071I_0$

Keeping up?

10. Convert from amplitude to rms: (a) 25 V, (b) 17 A, (c) 7.5 V.
11. Convert from rms to amplitude: (a) 7 V, (b) 230 V, (c) 13 A.

Phasor shorthand

It is helpful to draw sketches of phasor diagrams, but the essential calculations can be done without sketches. Instead of drawing the phasor as a line of specified length pointing in a certain direction, we can simply write its length and direction in a standard way, which is usually known as polar form. For example, the phasor in Fig. 4.2 is written:

$$4\underline{/57^\circ}$$

The angle may also be expressed in radians. The phasor in Fig. 4.3(c) is written either as $1.3\underline{/310^\circ}$ or as $1.3\underline{/-50^\circ}$. Polar form is a convenient way of writing phasors and it is easy to multiply phasors written in this way.

Complex numbers

The reason for using phasor diagrams is that a phasor is a way of representing both amplitude and phase at the same time. Another way of representing two quantities at the same time is to use complex numbers. The two techniques have much in common. A complex number consists of two parts, a **real** part and an **imaginary** part. As an example, take the complex number $3 + j2$. The real part is the 3, which simply is an ordinary number. The imaginary part is the j2. The imaginary part always begins with a j, so here is a way of recognizing a complex number.

There are two ways of thinking about j. One way is to consider it as the square root of negative 1. Another way of thinking about j is to say that it is an operator. An operator is a symbol for an action, such as $\int$, the symbol for integration. j is the symbol for a quarter-turn anticlockwise. A complex number can be taken as a set of directions for moving from one place to another in a plane. Imagine yourself in such a plane. Begin at the origin, 0 (Fig. 4.12), facing right. The complex number $5 + j2$ can be interpreted as '5 steps forward, quarter-turn anticlockwise, two steps forward'. This brings you along the path of the dashed line to point A. Similarly, the number $3 + j4$ takes you along the path of the dotted line to point B. Taken in this way, a complex number is nothing more puzzling than a set of directions for getting from the origin to any point on the plane.

Complex numbers can have negatives in them, the equivalent of backward steps. In Fig. 4.12, the number taking you to point C is $-2 + j5$, and to get to D the number is $4 - j5$. In the latter example the negative sign before the imaginary part can be taken to mean a quarter-turn clockwise, followed by 5 forward steps, or the sign may be linked with the number 5, so that the imaginary part is a quarter-turn anticlockwise followed by 5 backward steps. In either case you get to point D.

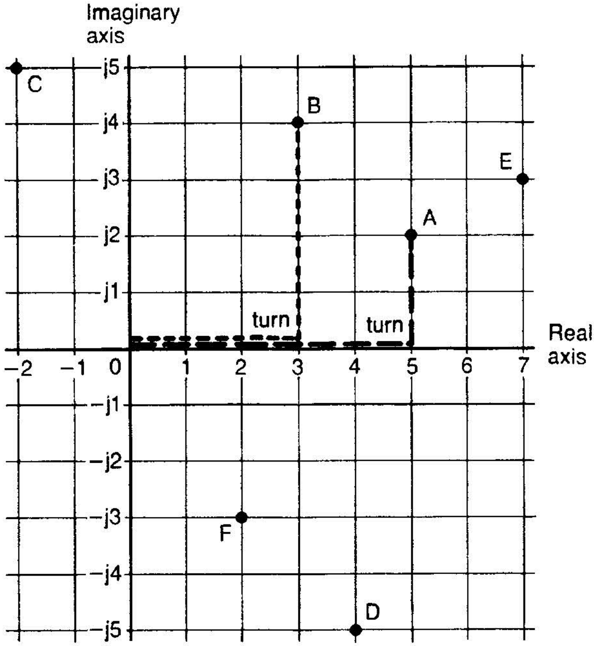


Figure 4.12 *Complex numbers give instructions for reaching any point on the compl plane*

Keeping up?

12. Express the phasors of Fig. 4.3(a) and (b) in polar form.
13. Write the complex numbers that direct you to points E and F in Fig. 4.12.

Complex numbers and phasors

Complex numbers are used as another way of representing phasors. The phasor in Fig. 4.13 is described by the complex number 4 + j7. This number defines th point at which the tip of **A** is located. Similarly, phasor **B** is described by −3 + j and phasor **C** by 4 − j2. These two-part numbers simply give the co-ordinate of the tips of the phasors. If you find j difficult to understand or impossible t believe in, just think of the real part of the complex number to mean steps i an east–west direction and the imaginary part to mean steps in a north–sout direction.

Phasors are easy to add when expressed as complex numbers. Simply follo the general rule for adding complex numbers, which is to add real parts to rea parts and imaginary parts to imaginary parts. Since real and imaginary part correspond to perpendicular directions, we can never add them to each other

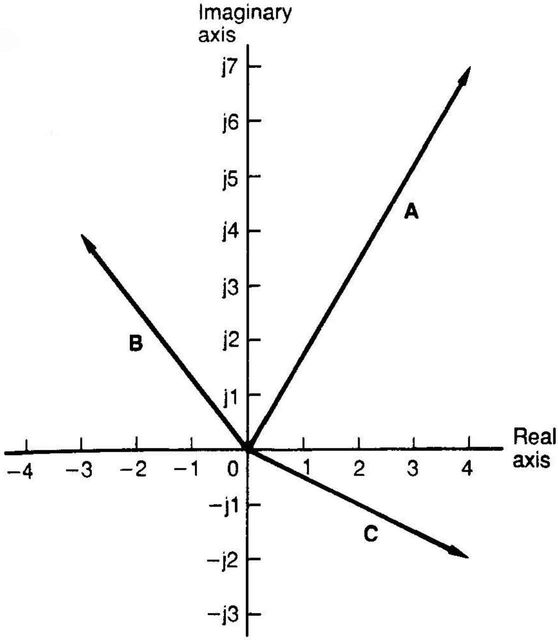


Figure 4.13 *Complex numbers represent phasors in rectangular form*

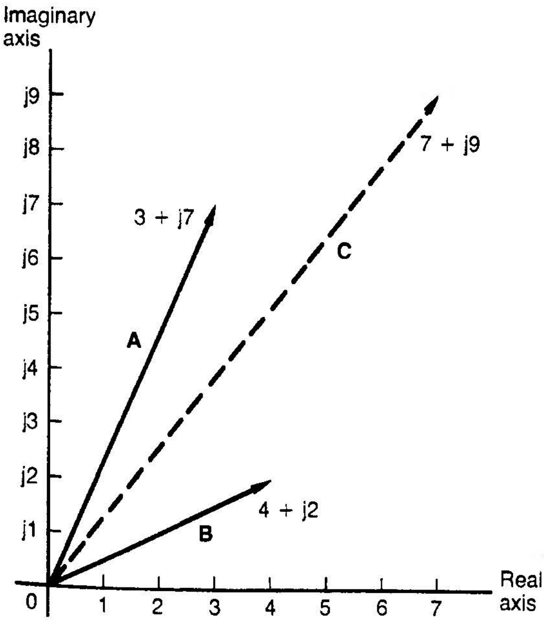


Figure 4.14 *Phasors can be summed by summing the corresponding complex numbers*

Take the example in Fig. 4.14. The phasor **A** is $3+j7$, and phasor **B** is $4+j2$. Their sum is:

$$\begin{array}{rr} & 3+j7 \\ + & 4+j2 \\ \hline & 7+j9 \end{array}$$

The result is the phasor **C**. We can never add the 7 to the j9 because one is real (east–west) and the other is imaginary (north–south).

When a phasor is represented by a complex number, we have the two co-ordinates needed to plot it on a grid. This way of representing a phasor is known as rectangular form. In Fig. 4.15, the phasor **A** has magnitude r and phase angle ϕ. In polar form it is $r\angle\phi$. Use these equations to find the co-ordinates of P:

$$a = r\cos\phi$$

$$b = r\sin\phi$$

The rectangular form is:

$$a + jb = r\cos\phi + jr\sin\phi$$

For example, a phasor has polar form $3\angle 50^\circ$, in which $r = 3$ and $\phi = 50^\circ$. From these two values we calculate:

$$a = r\cos\phi = 3\cos 50^\circ = 1.928$$

$$b = jr\sin\phi = j3\sin 50^\circ = j2.298$$

The rectangular form is:

$$1.928 + j2.298$$

When converting from rectangular form to polar form, we are given a and b and have to find r and ϕ. The calculation is the same as on pp. 52–3.

$$r = \sqrt{(a^2 + b^2)}$$

$$\phi = \tan^{-1}(b/a)$$

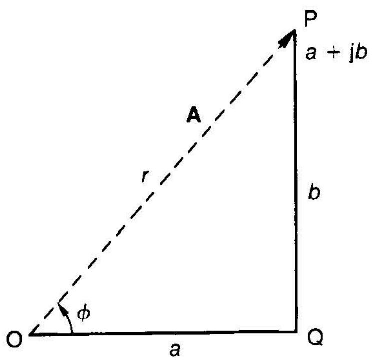


Figure 4.15 *This triangle establishes the relationships between the polar form and the rectangular form of a complex number*

If, for example, $a = 5$, $b = 4$:

$$r = \sqrt{(5^2 + 4^2)} = \sqrt{(25 + 16)} = \sqrt{41} = 6.403$$

$$\phi = \tan^{-1}(4/5) = 38.66^\circ$$

The result of converting $5 + j4$ to polar form is $6.403\underline{/38.66^\circ}$. It is useful to draw a sketch when converting from rectangular to polar form using a calculator. This is because a calculator does not necessarily refer a result to the correct quadrant when evaluating arctan ($\tan^{-1}$).

Keeping up?

14. Convert from rectangular form to polar form: (a) $5 + j2$, (b) $3 - j7$, (c) $-6 - j$.
15. Convert from polar form to rectangular form: (a) $5\underline{/36.87^\circ}$, (b) $7.07\underline{/225^\circ}$, (c) $2.24\underline{/206.57^\circ}$.

Impedances as complex numbers

In a phasor diagram it is convenient to assume that a current phasor has zero phase angle. Since the pd across a resistor is in phase with the current, the phasor of resistor pd ($\mathbf{v_r}$) also has zero phase angle. We also know that the pd across a capacitor lags the current by 90° and that the pd across an inductor leads the current by 90°. The typical phasor diagram comprises the phasors shown in Fig. 4.16. There will be at least one other phasor to represent the pd applied from the voltage generator. This equals the sum of the other phasors, and its magnitude and phase angle depend on the relationships between the other phasors.

The diagram shows the pd phasors across three different kinds of impedance:

Impedance type	*Signal*	*Phase*
R	v_r	0°
X_C	v_c	−90°
X_L	v_l	+90°

The information about the phase of pds across reactances may be incorporated into the definitions of impedances by using j, as in Fig. 4.16.

Resistance, R. The equation $v_r = i/R$ expresses the effect that resistance has in producing a pd signal when a current signal passes through it. The phasor diagram shows that current phasor i lies on the real axis, as is usually taken to be the case. The pd phasor $\mathbf{v_r}$ is necessarily in phase with i so neither of these phasors has an imaginary part.

Capacitative reactance, X_c. The equation $v_c = iX_c$ expresses the effect that reactance has in producing a pd signal when a current signal passes through it. The phasor diagram shows that $\mathbf{v_c}$ lags by 90°. To incorporate the 90° clockwise

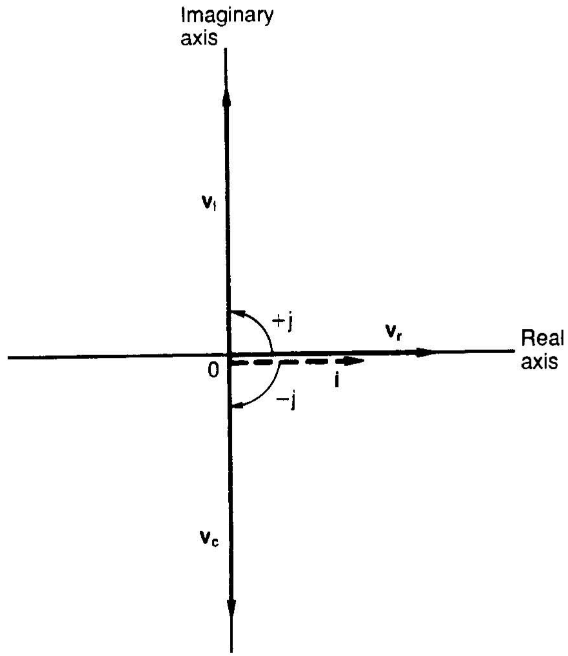


Figure 4.16 *The phasors which represent different kinds of impedance*

turn, we include −j in the definition of X_C:

$$X_C = -j/\omega C$$

It is allowable, as we have done before, to use the equation $X_C = 1/\omega C$ to calculate numerical values of X_C, if we are not concerned about phase. But the definition which includes j takes the phase effect of the capacitor into account.

−j and j

Given that $X_C = -j/\omega C$, multiply the numerator and denominator by j:

$$X_C = \frac{j \times -j}{j \times \omega C}$$

This does not change the value of the quotient. Multiplying out:

$$j \times -j = -(-1) = 1$$

$$\Rightarrow \qquad X_C = \frac{1}{j\omega C}$$

This is often a more convenient definition of capacitative impedance. In general, j and −j are reciprocals of each other.

Inductive reactance, X_L. In a corresponding way, the equation for X_L incorporates the anticlockwise turn by including +j:

$$X_L = j\omega L$$

Writing impedances as complex numbers takes phase into account in circuit calculations. One example will be given here, leaving the main applications of this technique until the next chapter, which describes some simple filters.

In the circuit with which we began this example, we have a 500 Ω resistor in series with a 330 nF capacitor. Given that the frequency is 1 kHz (it is always essential to state the frequency if there is a reactance present), $\omega = 6283$. As a complex number the capacitance is:

$$-j/\omega C = -j/(6283 \times 330 \times 10^{-9}) = -j/(2.073 \times 10^{-3}) = -j482.3$$

When we have two resistors in series their combined resistance is their sum. Similarly, for two impedances in series. The total impedance Z of the circuit is their sum:

$$Z = 680 - j482.3$$

The signal from the generator is $v = 3 \sin 2000\pi t$, which has no phase shift, so its phasor is 3 + j0. Ignore the term j0 and calculate the current phasor:

$$\mathbf{i} = \mathbf{v}/Z = \frac{3}{680 - j482.3}$$

The task is to find the size and direction of the current phasor. Dividing by a complex number is difficult, but there is a trick way to do it. Multiply both top and bottom of the quotient by the expression obtained when the sign of j is reversed:

$$\mathbf{i} = \frac{3(680 + j482.3)}{(680 - j482.3)(680 + j482.3)}$$

This has no effect on the value of the expression but multiplying −j by +j gives +1, so the j disappears from the denominator. Multiplying out, top and bottom:

$$\mathbf{i} = \frac{2040 + j1447}{462\,400 + 232\,613} = \frac{2040 + j1447}{695\,013}$$

Dividing the numerator by 695 013 gives:

$$\mathbf{i} = 0.00294 + j0.00208$$

This is the current phasor in rectangular form. Converting it to polar form, we obtain:

$$\mathbf{i} = 0.0036\underline{/35.3^\circ}$$

In milliamps, this becomes $\mathbf{i} = 3.6\angle 35.3^\circ$, which is the result obtained earlier, in the box on p. 54.

This example illustrates that when we write impedances as complex numbers, we handle them by applying the usual rules for resistances in series or in parallel. There is no need to draw a phasor diagram because phase is taken into account by the complex representation of reactances.

Test yourself

1. A circuit has a voltage generator, a $47\,\text{k}\Omega$ resistor and a $220\,\text{nF}$ capacitor in series. The generator produces a pd signal, $v = 8\sin 100\pi t$. Use the phasor technique to find the pd signals across the resistor and the capacitor, with phase relative to the generator signal.
2. A circuit has a voltage generator, a $120\,\Omega$ resistor and a $10\,\text{mH}$ inductor in series. The generator produces a signal, $v = 4.5\sin 6000\pi t$. Use the phasor technique to find the pd signals across the resistor and the inductor, with phase relative to the generator signal.
3. In the circuit of question 1 the pd signal is changed to $v = 6\sin 200\pi t$. Express the impedances as complex values and use these to find the total circuit impedance, the current signal and the pd signal across the capacitor. Express the results as phasors in rectangular form, phasors in polar form, and as equations for current or pd.
4. In the circuit of question 2 the pd signal is changed to $v = 5\sin 4000\pi t$. Express the impedances as complex values and use these to find the total circuit impedance, the current signal and the pd signal across the inductor. Express the results as phasors in rectangular form, phasors in polar form, and as equations for current or pd.

5

Passive filters

The function of a filter is to allow signals of a given band of frequencies to pass, while obstructing or reducing the amplitude of others. In this context we are referring to sinusoidal signals, since signals of other forms are considered to be a mixture of sinusoids of different frequencies and different amplitudes. Figures 2.12 and 2.13 show the composition of typical sawtooth and triangular waves. If these signals are passed through a filter, some of the frequencies may be stopped altogether from passing while others may be attenuated (reduced in amplitude). The frequency spectrum has a markedly different appearance after the signals have passed through a filter.

Before considering the effects of filters on triangular and other composite waveforms, we will look at what happens to a pure sine wave of a single frequency. The essentials of this have already been described in earlier chapters but now we represent the facts in the context of a filter circuit. This chapter deals with passive filters, and the simplest type of filter is built from two passive components, a resistor and a capacitor. The circuit of the filter illustrated in Fig. 5.1 is exactly the same as that of Fig. 1.5. The pd source is drawn in dashed lines because it is not part of the filter. It could be any device or circuit that produces a varying pd, $\mathbf{v}_{in}$. It might be a microphone, a photo-cell, a circuit based on a thermistor, a radio antenna, a pair of electrodes in an electrocardiograph, or one of thousands of other sources of varying pd. For the purpose of this discussion, it is the source of a sine wave signal of a single frequency.

The other special feature of this circuit is that there are connections on either side of the capacitor to convey a signal $\mathbf{v}_{out}$ to an external circuit. It is essential that the following circuit draws as little current as possible from the filter. Otherwise, the action of the filter is degraded. The following circuit must have a high input impedance and we assume in the discussions below that its input impedance is so high as to draw no appreciable current from the filter. It is also assumed that the pd source has a suitably low output impedance so that it is able to supply as much current as the filter can accept at any instant.

The filter of Fig. 5.1 is labelled with complex impedances, as discussed at the end of the last chapter. Only the capacitor has a complex impedance in this circuit. The total impedance of the filter is the sum of the impedances of the resistor and capacitor in series:

$$Z = R - \mathrm{j}/\omega C$$

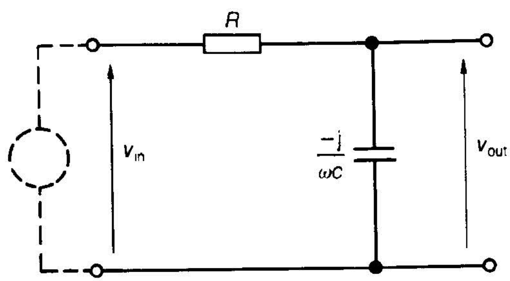


Figure 5.1 *The circuit of Fig. 1.5 redrawn as a low-pass passive filter*

For example, if the resistance is 220 Ω, the capacitance is 1 μF and the frequency is 1 kHz, then $\omega = 2\pi f = 6283$ rad/s and:

$$Z = 220 - j159$$

The impedance is in ohms, as usual. Take $\mathbf{v_{in}}$ and $\mathbf{v_{out}}$ as symbols for the input and output phasors. When $\mathbf{v_{in}}$ is applied to the filter, the resistor and capacitor act as a potential divider. Therefore $\mathbf{v_{out}}$, the pd signal across the capacitor, bears the same proportion to $\mathbf{v_{in}}$ as the impedance of the capacitor bears to the total impedance.

$$\frac{\mathbf{v_{out}}}{\mathbf{v_{in}}} = \frac{-j159}{220 - j159}$$

or

$$\mathbf{v_{out}} = \frac{\mathbf{v_{in}} \times -j159}{220 - j159}$$

If the source signal is $\mathbf{v_{in}} = 2 \sin 2000\pi t$, and is taken to be the reference signal with $\phi = 0°$, $\mathbf{v_{in}} = 2 + j0 = 2$, and:

$$\mathbf{v_{out}} = \frac{-j318}{220 - j159} = 0.686 - j0.949$$

The technique for dealing with a complex divisor is explained at the end of Chapter 4. Converting $\mathbf{v_{out}}$ into polar form:

$$\mathbf{v_{out}} = 1.17\angle{-54.1°}$$

Writing this as an equation for the output signal:

$$\mathbf{v_{out}} = 1.17 \sin(2000\pi t - 54.1°)$$

Figure 5.2 shows the curves for $\mathbf{v_{in}}$ and $\mathbf{v_{out}}$. After the initial stage while the capacitor gains charge, it is clear that $\mathbf{v_{out}}$ has the same frequency as $\mathbf{v_{in}}$. Measurements on the graph show that the amplitude of $\mathbf{v_{out}}$ is 1.17 V, and that its phase lag is 54°. The gain of the filter is $1.17/2 = 0.585$. It is a characteristic of all passive filters that the output signal has a smaller amplitude than the input signal; in other words, that it has a gain less than unity. This can be seen on the phasor diagram of the filter (Fig. 5.3).

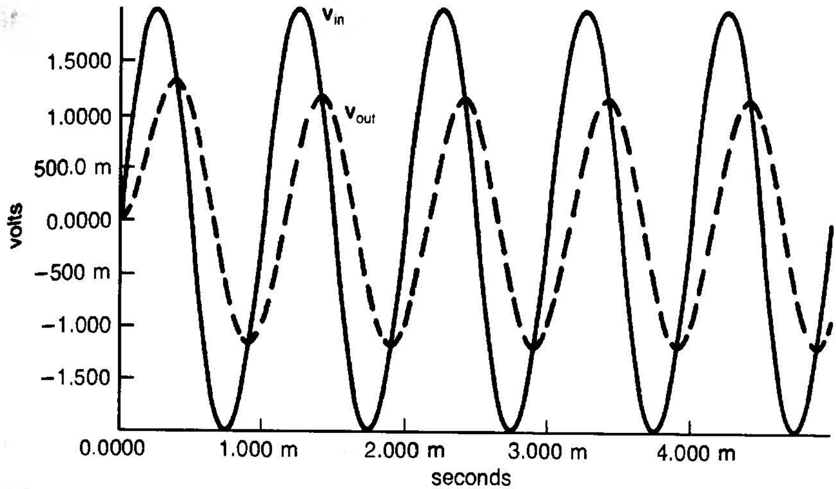


Figure 5.2 *When the frequency is 1 kHz, the output of the filter of Fig. 5.1 has reduced amplitude and appreciable phase lag*

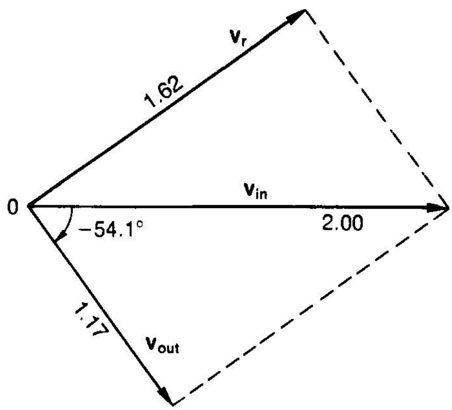


Figure 5.3 *A phasor diagram shows the relationship between the input and output of the filter of Fig. 5.1 when the frequency is 1 kHz*

Keeping up?

1. Given a low-pass filter built from a 470 Ω resistor and a 2.2 μF capacitor, with an input signal $\mathbf{v}_{\mathbf{in}} = 4 \sin 300\pi t$, calculate the output signal $\mathbf{v}_{\mathbf{out}}$ and express it as an equation, in complex (rectangular) form and in polar form.
2. If the capacitor and resistor in Fig. 5.1 are exchanged, so that $\mathbf{v}_{\mathbf{out}}$ becomes the pd across the resistor, what kind of filter does this produce? Given the

values listed in question 1, calculate the output signal $\mathbf{v}_{out}$ and express it as an equation, in complex (rectangular) form and in polar form.

Variable impedance

The filter circuit of Fig. 5.1 can be thought of as a potential divider. The capacitor is equivalent to a variable resistor controlled by frequency (Fig. 5.4). The higher the frequency, the lower the impedance of the capacitor and the smaller the amplitude of $\mathbf{v}_{out}$. The action is that of a low-pass filter.

General equation

Having looked at an actual example, we will work through the same problem but without inserting actual component values. For the input signal, let:

$$\mathbf{v}_{in} = V_0 \sin \omega t$$

Then the total circuit impedance is:

$$R - j/\omega C$$

The corresponding output signal is:

$$\mathbf{v}_{out} = \frac{\mathbf{v}_{in} \times -j/\omega C}{R - j/\omega C}$$

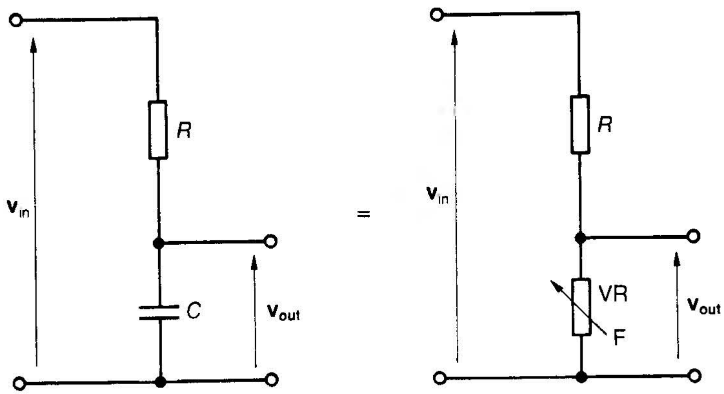


Figure 5.4 *Another way of looking at a low-pass resistor/capacitor filter*

Multiply top and bottom by $j\omega C$:

$$\mathbf{v}_{\mathbf{out}} = \frac{\mathbf{v}_{\mathbf{in}}}{j\omega RC + 1}$$

We leave the equation in this form. If you substitute values of $\mathbf{v}_{\mathbf{in}}$, ω, R and C, you can confirm the equation by obtaining the same results as above. In its general form, the equation shows that the output is larger if:

- input is larger
- frequency is smaller
- R is smaller
- C is smaller

Obviously we obtain a larger output with a larger input, otherwise the outgoing signal would not be a replica of the ingoing signal. Making ω smaller increases the impedance of the capacitor and hence the pd across it. Making R smaller increases the current flowing in the circuit and therefore increases the pd across the capacitor. Making C smaller causes larger changes of pd for a given change of charge ($v = q/C$, see equation (6) in Chapter 1).

If input, R and C are held constant, the amplitude of the output depends on ω. The equation shows that the output signal amplitude increases as ω decreases. This is seen in Fig. 5.5 where the results of reworking the numerical example above for four different frequencies are shown side by side. The amplitude of the input signal and the values of R and C are left unchanged. Looking at the diagrams in order of frequency we see that, as frequency decreases, the proportions of the rectangle OPQR change, while the length of the diagonal (source phasor) remains constant. With decreasing frequency the signal across the resistor becomes smaller while that across the capacitor becomes larger. In other words $\mathbf{v}_{\mathbf{out}}$ increases. At the same time the output phasor swings round nearer and nearer to the source phasor. It lags less and less far behind the source phasor. Continuing this trend it is possible to imagine the output phasor swinging round to coincide with the source phasor when frequency is very low. At the limit, the amplitude of the output signal becomes equal to that of the input signal and is in phase with it. The signal passes unaffected through the filter. This effect can be recognized in the equation above as ω approaches a limit of 0:

$$\mathbf{v}_{\mathbf{out}} = \frac{\mathbf{v}_{\mathbf{in}}}{j0RC + 1} = \frac{\mathbf{v}_{\mathbf{in}}}{1} = \mathbf{v}_{\mathbf{in}}$$

Since $\mathbf{v}_{\mathbf{in}}$ has no imaginary part, neither does $\mathbf{v}_{\mathbf{out}}$ and the phase angle is zero. A zero-frequency signal is a constant DC level. If a DC pd is applied to the filter, the capacitor quickly charges to that level. There is no fall in pd across the resistor (we are assuming that the following circuit is drawing no current) so $\mathbf{v}_{\mathbf{out}} = \mathbf{v}_{\mathbf{in}}$.

At the other extreme, at very high frequency, $\mathbf{v}_{\mathbf{out}}$ approaches zero, with $-90°$ phase angle. A diagram to summarize these changes can be plotted by using a circuit simulator.

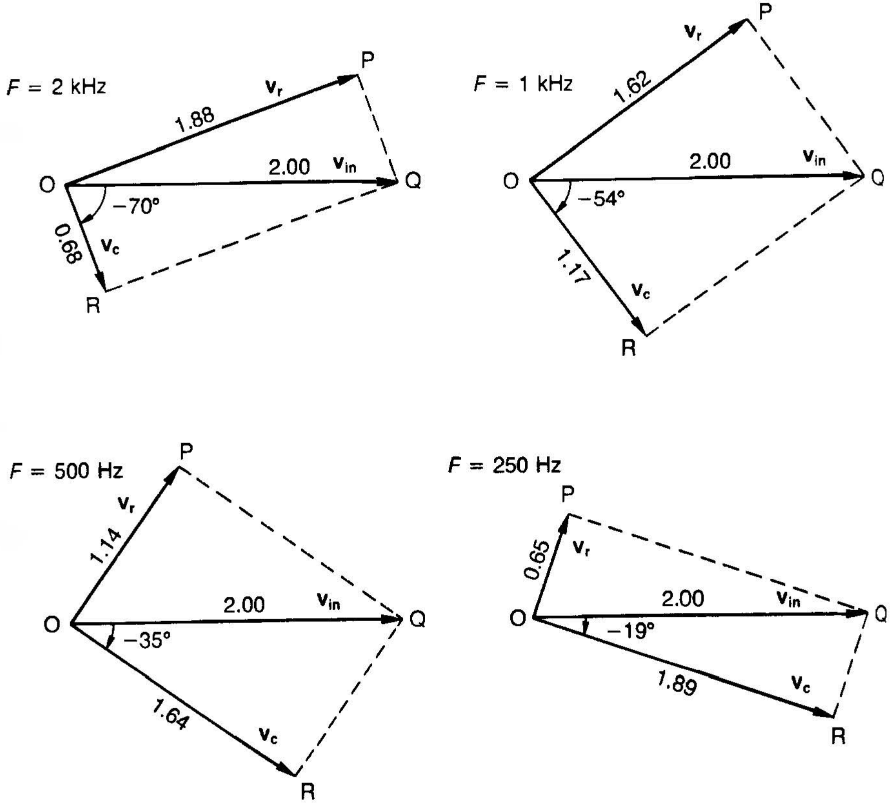


Figure 5.5 *Phasor diagrams for different frequencies show that the* $\mathbf{v_c}$ *phasor lags less and becomes longer as frequency decreases. This is the characteristic of a low-pass filter*

Keeping up?

3. A high-pass filter can be made by interchanging the resistor and capacitor of Fig. 5.1. Derive a general equation for the output $\mathbf{v_{out}}$ of such a high-pass filter when the input signal is $\mathbf{v_{in}} = V_0 \sin \omega t$.
4. In question 3, to what limits does $\mathbf{v_{out}}$ tend as ω approaches (a) zero (DC) (b) infinity?

Output signals

Figure 5.6 graphs the output signals illustrated by the phasor diagrams in Fig. 5.5. The graph is plotted from 4 ms onward at which time the input signal is at the beginning of a cycle. The output signals all lag behind the input and do not begin their cycles until a fraction of a millisecond later. The graph clearly shows that the lower the frequency, the greater the amplitude of the output signal.

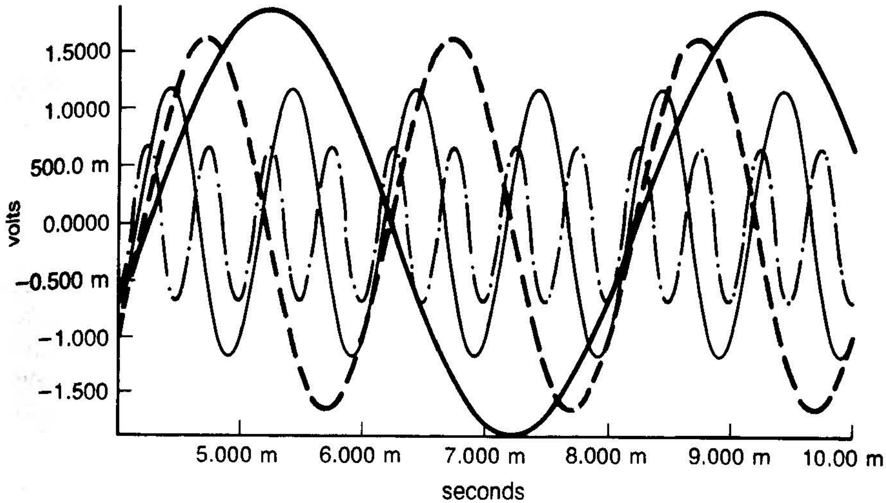


Figure 5.6 *These are the output signals corresponding to the phasor diagrams in Fig. 5.5. The higher the frequency the lower the amplitude*

Transfer function

For a given circuit, its transfer function expresses the relationship between the input and output signals. For a filter circuit of the type shown in Fig. 5.1, the transfer function is:

$$\frac{\mathbf{v}_{\mathbf{out}}}{\mathbf{v}_{\mathbf{in}}} = \frac{1}{\mathrm{j}\omega RC + 1}$$

Remember that the terms $\mathbf{v}_{\mathbf{out}}$ and $\mathbf{v}_{\mathbf{in}}$ are not just simple voltages. Each represents a sinusoidal voltage with a given frequency, amplitude and phase angle. They have the same frequency ($f = \omega/2\pi$) and the transfer function expresses the relationship between their amplitudes and phase angles. The reason that a single function is able to express the relationship between two different quantities (amplitude and phase) is that the transfer function contains a complex term (in this example, $\mathrm{j}\omega RC + 1$), which represents a phasor, and which in its turn represents both amplitude and phase.

Bode plot

The relationship between frequency and output amplitude can be plotted as a graph (Fig. 5.7). The range of Fig. 5.7 is from 0 Hz (DC) to 3 kHz. At 0 Hz, the output amplitude is 2 V, equal to the input amplitude. As frequency increases, amplitude falls steadily, until it reaches about 280 mV at 3 kHz. Amplitudes at 250 Hz, 500 Hz, 1 kHz and 2000 Hz are the same as are drawn in Fig. 5.5. The graph also shows phase lag falling from 0° at 0 Hz to about −78° at 3 kHz. In Fig. 5.7, amplitude is plotted against linear voltage and frequency scales. More often we use logarithmic scales, as in Fig. 5.8, such a graph being known as a **Bode plot**. The values plotted are as before but the shape of the curve is changed by plotting it logarithmically. A logarithmic frequency scale is often used because it is good for displaying the effects of relative frequency changes (for example, doubling or halving frequency) over a wide frequency range.

The pd scaling is not only logarithmic but is expressed in a different unit, the **decibel**. The scale runs from 0 dB at the top where 0 dB corresponds to 2 V, taken as the reference pd for this plot. The amplitude at other frequencies is measured relative to the reference level, in decibels. For example, at 1 kHz, the graph shows that $\mathbf{v}_{out}$ is approximately −4.7 on the decibel scale. From the values quoted in Fig. 5.7, where the amplitude of $\mathbf{v}_{out} = 1.17\,\text{V}$:

$$20\log(2/1.17) = -4.65\,\text{dB}$$

As with the frequency scale, the logarithmic plot of pd in decibels emphasizes relative values, which are usually more important than absolute ones.

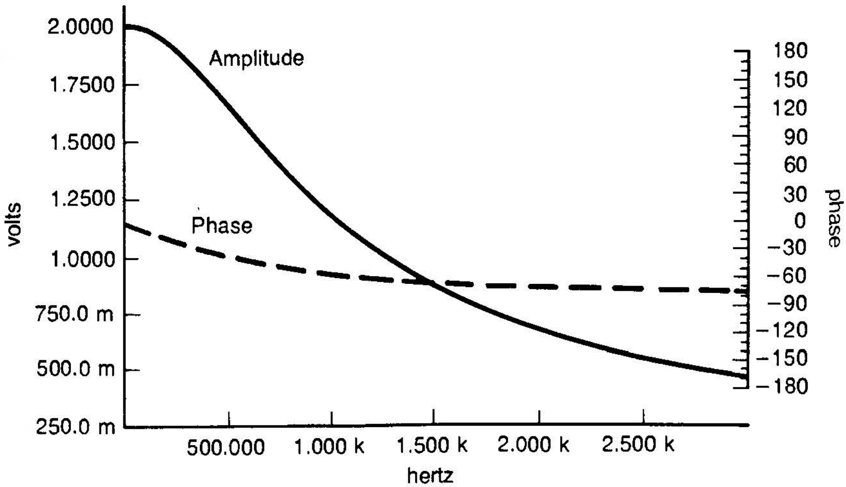


Figure 5.7 *These curves show how the low-pass filter responds to a range of frequencies from DC (0 Hz) to 3 kHz*

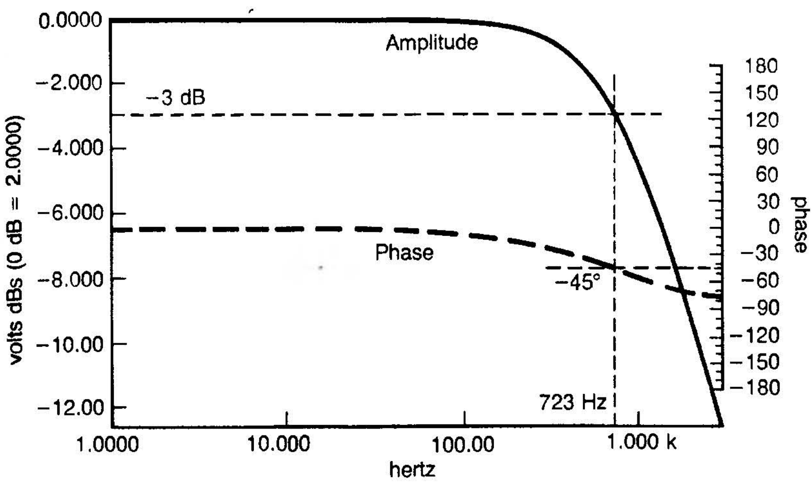


Figure 5.8 *This Bode plot has a logarithmic frequency scale and a decibel output pd scale. At the –3 dB point the frequency is 723 Hz and the phase angle is –45°*

Decibels

Decibels are a way of expressing the ratio between two quantities, x_1 and x_2. If n is the ratio in decibels:

$$n = 10 \times \log_{10}(x_2/x_1)$$

However, in the case of filters and other electronic circuits, as well as in some other applications, the most important consideration is the ratio of powers. Power is proportional to the square of pds or currents. Given pds or currents, the formula for the ratio of powers is:

$$n = 10 \times \log_{10}({v_2}^2/{v_1}^2)$$

The logarithm of a squared number is obtained by doubling the logarithm of the unsquared number, so the easiest formula to use for power ratios is:

$$n = 20 \times \log_{10}(v_2/v_1)$$

In filters, v_1 is the amplitude of the input signal and v_2 is the amplitude of the output signal. There are 3 cases:

Power gain	$v_2 > v_1$	n is positive
Power equality	$v_2 = v_1$	$n = 0$
Power loss	$v_2 < v_1$	n is negative

One way of producing a Bode plot is to take measurements on an actual circuit. A sinusoidal signal of given amplitude, frequency and phase is applied to the input and the amplitude and phase of the sinusoidal output signal are measured. The frequency remains unchanged. Measurements are made at various frequencies, from which the Bode plot is drawn. As we shall see later (Chapter 7) it is also possible to construct a Bode plot from calculated values, given the transfer function of the circuit.

Keeping up?

5. Find the power loss in decibels when (a) $v_1 = 3.5$ and $v_2 = 2.5$, (b) $v_1 = 5$ and $v_2 = 0.5$, (c) $v_1 = 1.2$ and $v_2 = 1.1$.
6. The input to an amplifier has amplitude 0.2 V. The output has amplitude 5.6 V. What is the power gain, in decibels?
7. The open-loop gain of an operational amplifier is said to be 106 dB. If the amplitude of the input signal is 2.5 μV, what is the amplitude of the output signal?

Filter characteristics

The Bode plot of Fig. 5.8 illustrates the main features of a low-pass filter. Towards the left side of the diagram $\mathbf{v}_{out}$ is almost equal to $\mathbf{v}_{in}$. This is the pass band. Toward the right, and on toward even higher frequencies, $\mathbf{v}_{out}$ is very much smaller than $\mathbf{v}_{in}$. This is the stop band. Between the pass band and the stop band is the transition region, in which $\mathbf{v}_{out}$ falls steeply with increasing frequency.

By definition, the pass band of a low-pass filter extends from 0 Hz up to a frequency at which the power of the signal is half that of v_{in}. If the power is half, then $v_{out}^2/v_{in}^2 = 0.5$ and:

$$n = 10\log_{10} 0.5 = 10 \times -0.3010 = -3.010\,\text{dB}$$

This level is marked in Fig. 5.8 as the '−3 dB' line. It is usually referred to as the −3 dB point, the cut-off point, the half-power point, or sometimes as f_c. The graph shows that for this filter, the −3 dB point is at 723 Hz. Phase angle is 0° at 1 Hz, falling through −45° at the −3 dB point and eventually reaching −90° at high frequencies, beyond the right-hand edge of the graph. Figure 5.9 shows the corresponding phasor diagram at the −3 dB frequency, for comparison with those in Fig. 5.5. By definition of the −3 dB point, $\mathbf{v}_{out}^2/\mathbf{v}_{in}^2 = 0.5$. In the phasor diagram, Pythagoras theorem tells us that:

$$\mathbf{v}_{in}^2 = \mathbf{v}_{out}^2 + \mathbf{v}_r^2$$

Dividing by $\mathbf{v}_{in}^2$:

$$1 = \mathbf{v}_{out}^2/\mathbf{v}_{in}^2 + \mathbf{v}_r^2/\mathbf{v}_{in}^2$$

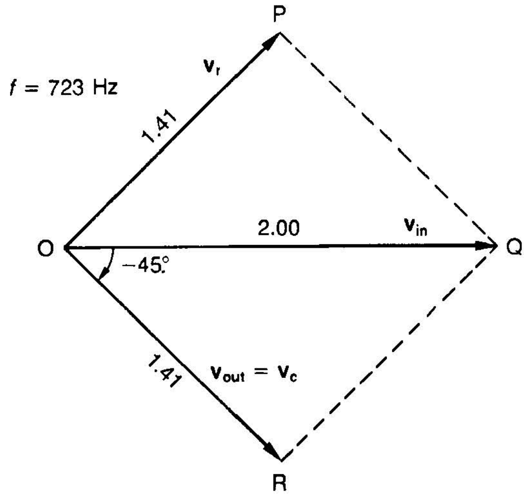


Figure 5.9 *Another phasor diagram to include with those in Fig. 5.5. At 723 Hz,* $\mathbf{v_r} = \mathbf{v_c}$, *the rectangle becomes a square, the phase angle is −45°. This is when the output power is 3 dB below input level*

Substituting the value for $\mathbf{v_{out}}^2/\mathbf{v_{in}}^2$ at the −3 dB point:

$$1 = 0.5 + \mathbf{v_r}^2/\mathbf{v_{in}}^2$$

$$\mathbf{v_r}^2/\mathbf{v_{in}}^2 = 0.5$$

But it has already been stated that:

$$\mathbf{v_{out}}^2/\mathbf{v_{in}}^2 = 0.5$$

Therefore

$$\mathbf{v_r}^2/\mathbf{v_{in}}^2 = \mathbf{v_{out}}^2/\mathbf{v_{in}}^2$$

$$\mathbf{v_r}^2 = \mathbf{v_{out}}^2$$

$$\mathbf{v_r} = \mathbf{v_{out}}$$

At the −3 dB frequency, the pds across the resistor and capacitor have equal amplitude. This means that their impedances are equal. So the −3 dB is not just a conveniently selected point. It is the frequency at which the resistor and capacitor have equal impedance. In Fig. 5.9, the lengths of the two phasors are equal, and $\mathbf{v_{out}}$ lags 45° behind $\mathbf{v_{in}}$. Ignoring their phase angles, the equality of the sizes of the impedances means that:

$$R = X_c = 1/\omega C$$

Rearranging:

$$\omega = 1/RC$$

But $\omega = 2\pi f$

Rearranging $f = \omega/2\pi$

Substituting $f = 1/2\pi RC$

This equation is used for calculating the -3 dB frequency, given the values of R and C. For example, in Fig. 5.8, $R = 220\,\Omega$, $C = 1\,\mu F$ and:

$$f = \frac{1}{2\pi \times 220 \times 10^{-6}} = 723\,\text{Hz}$$

This agrees with the value of the -3 dB frequency obtained from the graph.

Cut-off frequency

At the cut-off frequency, or -3 dB point, of a resistor/capacitor low-pass filter:

Impedances of resistor and capacitor are equal in size.
Phase angle is $-45°$
$f = 1/2\pi RC$

Roll-off

The rate at which amplitude falls off in the transition region is an important characteristic of a filter. The steeper the slope of the curve, the more sharply does the filter distinguish between those frequencies it passes at almost full amplitude and those that it strongly blocks. The equation on p. 71 shows that:

$$\mathbf{v}_{\mathbf{out}} = \frac{\mathbf{v}_{\mathbf{in}}}{\mathrm{j}\omega RC + 1}$$

If frequency is reasonably high, as it is above the cut-off point, we can ignore the 1 in the denominator and, since we are not concerned with phase, we can ignore j too:

$$\mathbf{v}_{\mathbf{out}} = \frac{\mathbf{v}_{\mathbf{in}}}{\omega RC}$$

$\mathbf{v}_{\mathbf{out}}$ is inversely proportional to ω. If frequency is doubled, the amplitude of $\mathbf{v}_{\mathbf{out}}$ is halved. A doubling of frequency is often referred to as an octave, a word borrowed from musical terminology. Hence, the ratio of two outputs for two signals an octave apart is 0.5. In decibels, this is $20\log 0.5 = -6$ dB. For a doubling of frequency, the output falls by 6 dB. This rate of roll-off is the general rule in most simple filters.

Filter action

The action of a filter on a single sine wave is to reduce its amplitude and cause a phase lag. If we filter a more complicated signal, such as a sawtooth wave (p. 28), each sinusoidal component of the signal is affected in amplitude and phase depending on its frequency. Figure 5.10 shows a sawtooth wave before and after passing through a low-pass filter. Like the signal of Fig. 2.8(d), the wave has a frequency of 0.159 Hz and an amplitude of 1.58 V. Because this is a low-frequency signal, we are using a filter with a low −3 dB point. The capacitor is 1 μF as before but the resistor is increased to 390 kΩ. The −3 dB point is:

$$f_c = 1/2\pi RC = 1/(2\pi \times 390\,000 \times 10^{-6}) \approx 0.4\,\text{Hz}$$

The effect of the filter on the triangular wave is visually obvious. The sharp downward-pointing corners of the signal become rounded off. Some of its high-frequency components are being lost or, at least, reduced in amplitude.

The filter action is investigated further by a Fourier analysis (Fig. 5.11). This has a similar appearance to the analysis of the original unfiltered signal (Fig. 2.12). But both figures have been plotted with the same vertical scale to make it clear that the amplitudes of the fundamental and harmonics are reduced

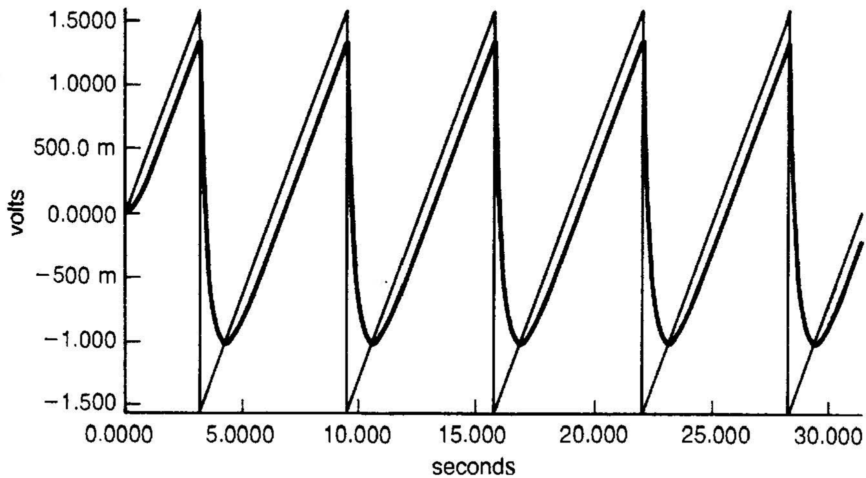


Figure 5.10 *The effect of a low-pass filter on a sawtooth signal is to reduce its amplitude, round off some of the corners and introduce a phase delay*

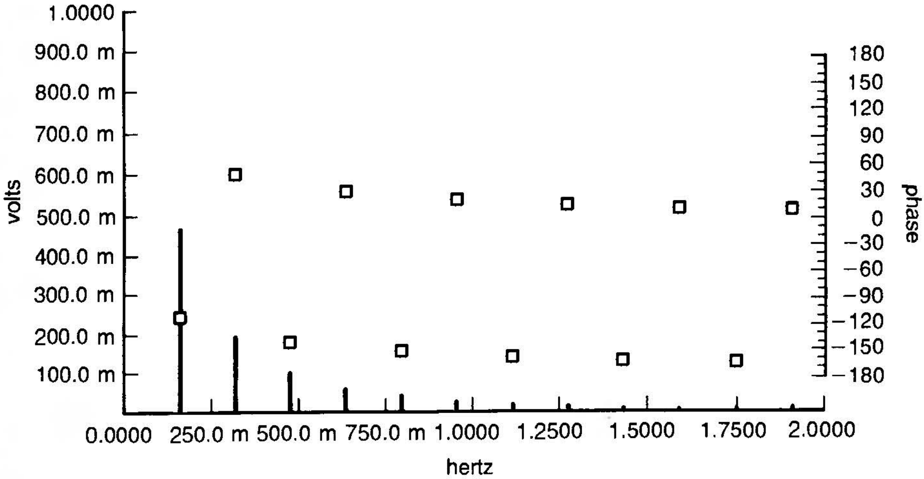


Figure 5.11 *A Fourier analysis reveals the relatively greater attenuation of high-frequency components after filtering a sawtooth signal through a low-pass filter (compare with Fig. 2.12). Phase lag increases with increasing frequency*

by filtering. The effect is least for the fundamental and the first few harmonics. Measurements on the graphs confirm this effect:

Frequency	*Amplitudes* (mV) v_{in} *(unfiltered)*	v_{out} *(filtered)*	v_{out}/v_{in}
Fundamental	1000	465	0.465
1st harmonic	500	196	0.392
- - - - - -	- - - -	- - - -	- - - -
2nd harmonic	333	108	0.324
3rd harmonic	250	69	0.276
4th harmonic	200	45	0.225
5th harmonic	167	30	0.180
6th harmonic	143	24	0.168

The dashed line in the table shows where the cut-off frequency is located. The last column of the table shows that attenuation of amplitude increases progressively with increase in frequency. But, as illustrated in Fig. 5.7, this is a gradual effect. There is no sharp edge to the pass band. Ideally the pass band is flat and a sudden and steep roll-off begins at the −3 dB frequency. To obtain this we need a filter of more elaborate design.

Figure 5.11 shows a further effect of filtering. In the original signal (Fig. 2.12) the phase of each component is either +90° or −90°. In the filtered signal there is a gradual increase in phase lag as frequency increases. This is in accordance with the effect of frequency on filter characteristics. The result is that the components of a signal are each delayed by a different amount. This is an additional source of distortion in the filtered signal for, not only are the harmonics attenuated by different amounts, but they arrive at the output of the filter at slightly different stages in their cycles. This effect is known as group delay. It is particularly important with high-frequency pulsed signals, such as are often found in digital circuits.

High-pass filters

The action of high-pass filters has already been the subject of some of the questions in Keeping up? We found that the transfer function of a resistor/capacitor high-pass passive filter is:

$$\frac{\mathbf{v}_{\mathbf{out}}}{\mathbf{v}_{\mathbf{in}}} = \frac{j\omega RC}{j\omega RC + 1}$$

Once the action of a low-pass filter has been studied and understood, there is little new to be learned about high-pass filters. It is easy to convert one sort into the other. In the case of a passive low-pass resistor/capacitor filter, simply exchanging the resistor for the capacitor turns it into a high-pass filter. The Bode plot for a high-pass filter is similar to that for a low-pass filter but reversed from left to right. Compare Fig. 5.12 with Fig. 5.8, both of which are based on the same values of R and C. In Fig. 5.12 the frequency range has been extended up to 100 kHz to allow more of the pass band to be plotted. Another difference in the plots is that the decibel scale extends down to −56.58 dB to cover the curve down to 1 Hz. This represents an attenuation of ×0.06.

The transition region has a slope of +6 dB per octave and reaches the −3 dB point at 723 Hz, the same frequency as in a low-pass filter. In the pass band, above

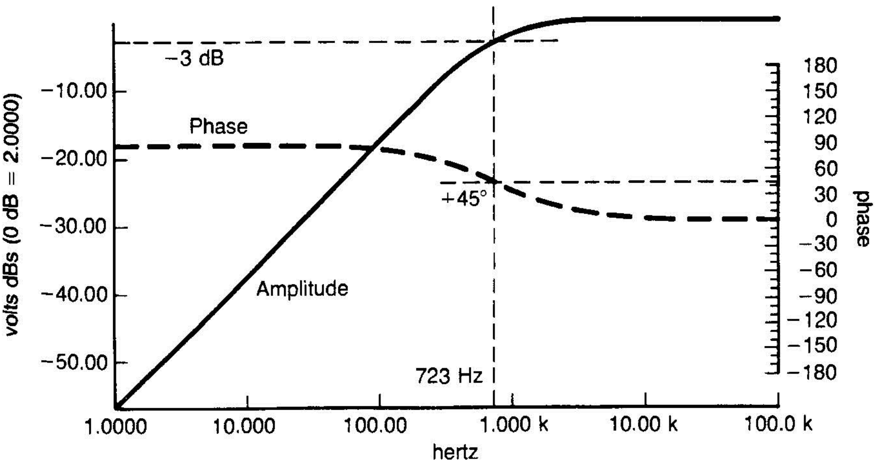


Figure 5.12 *Compare these frequency response curves of a high-pass filter with those in Fig. 5.8*

about 10 kHz, the filter transmits the signal with virtually no loss. Changes of phase are the opposite to those found in a low-pass filter because phase angles are positive instead of negative. Phase angle is +90° at 1 Hz, falling to exactly +45° at the −3 dB frequency. As frequency increases beyond that, the phase angle gradually levels out to 0°.

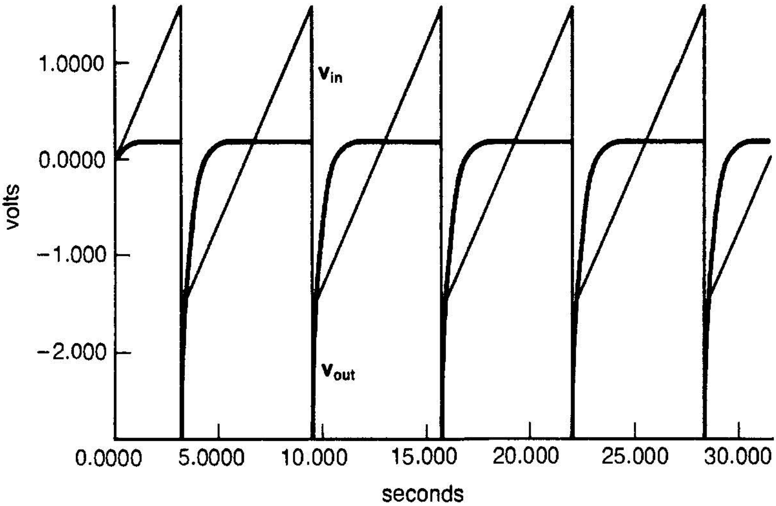


Figure 5.13 *The effect of filtering a sawtooth signal through a high-pass filter*

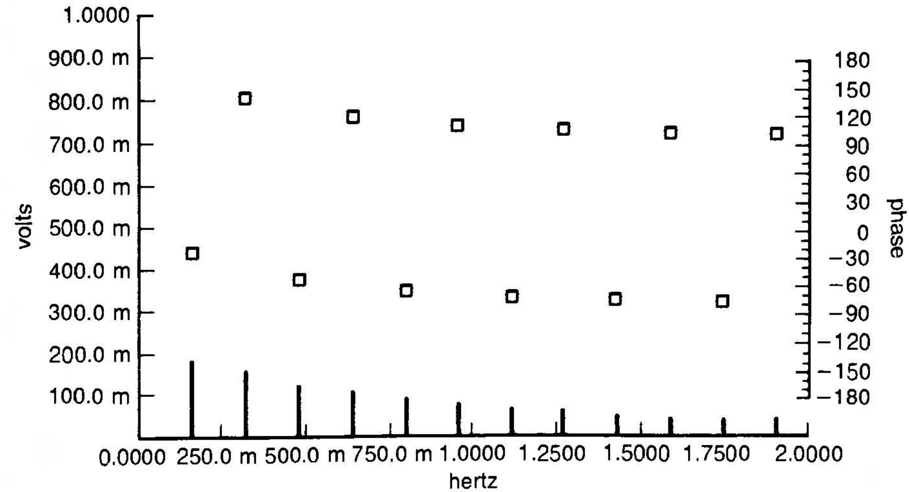


Figure 5.14 *The frequency spectrum loses much of its lower-frequency components after high-pass filtering. Compare this spectrum with Figs 2.12 and 5.11*

The effect of the high-pass filter on a sawtooth wave is a complete distortion of its shape (Fig. 5.13) due to the reduction in amplitude of the fundamental and lower harmonics. In the frequency spectrum (Fig. 5.14), the fundamental is reduced to about 0.18 of its value in an unfiltered signal (Fig. 2.12), and there is a general 'flattening' of the spectrum, so that the higher harmonics become relatively more significant and the waveform exhibits more pronounced 'spikes'.

Inductive filters

The filtering action of resistor/capacitor filters depends upon the way the impedance of the capacitor varies with frequency. The impedance of a capacitor is inversely proportional to frequency, in other words, increasing frequency leads to reducing impedance. Depending on the arrangement of the resistor and capacitor, we obtain a low-pass or high-pass filter (Fig. 5.15(a) and (b)).

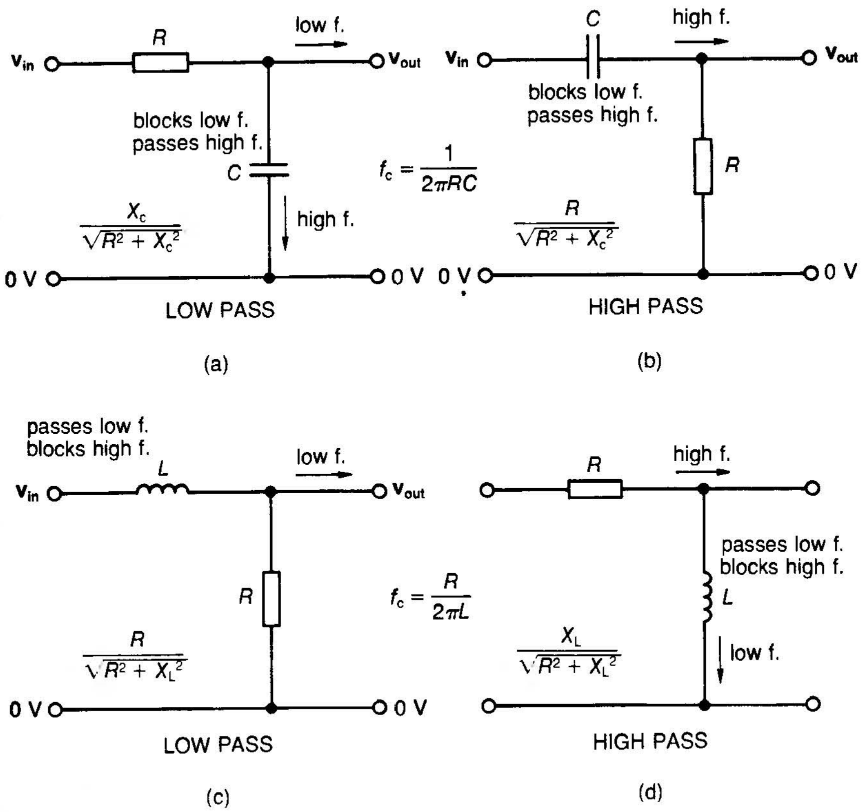


Figure 5.15 *A summary of the four types of first-order passive filters. The formula at the bottom left of each diagram gives the transfer ratio* $\mathbf{v}_{out}/\mathbf{v}_{in}$ *at any given frequency*

An inductor is also frequency dependent, but its impedance is directly proportional to frequency. Increasing frequency leads to increasing impedance. This property can be used in building a low-pass or high-pass filter from a resistor and an inductor (Fig. 5.15(c) and (d)). The main reason why inductors are seldom used in practical filters is that filtering frequencies in the audio range and below requires the inductor to be unduly large and heavy. This goes against the present-day trend toward light, small and portable equipment. Another reason for the unpopularity of inductors is that they generate magnetic fields which may interfere with nearby circuits. Conversely, they may also pick up magnetic interference unless they are very thoroughly shielded. It is only in high-frequency circuits, such as high-frequency radio or microwave circuits, that inductors can be small enough to be practicable.

Figure 5.15 summarizes the structure and properties of filters in which there is just one reactance, either a capacitor or an inductor. These are known as first-order filters. The figure includes the formulae for f_c and for the ratio $\mathbf{v}_{out}/\mathbf{v}_{in}$ for each type of filter. These ratios may be deduced from the geometry of phase diagrams like those of Figs 5.5 and 5.9.

Test yourself

1. Design a passive low-pass resistor/capacitor filter with a −3 dB point of 500 Hz, using a 10 kΩ resistor. Calculate the attenuation of a 600 Hz sine wave signal when passed through this filter.
2. Design a passive high-pass resistor/capacitor filter with a −3 dB point of 2 kHz, using a 2.2 nF capacitor. Calculate the attenuation of a 1 kHz sine wave signal when passed through this filter.
3. Design a passive high-pass resistor/inductor filter with a −3 dB point of 10 MHz, using a 10 kΩ resistor. Calculate the attenuation of a 9 MHz sine wave signal when passed through this filter.

6

Second-order passive filters

The filters described in Chapter 5 each have a single reactive component, either a capacitor or an inductor. Because of this, they are known as **first-order filters.** Although a first-order filter is adequate for many purposes, it has two important drawbacks:

1. The pass-band merges with the transition region, so that there is no sharp differentiation between frequencies that are to be passed and those that are to be attenuated.
2. The slope of the response curve in the transition region is only −6 dB per octave, with the result that frequencies several octaves away from the nominal cut-off point are present in the output from the filter.

Improved filtering is obtained by using a **second-order filter**, which contains two reactive components. The 'knee' of the response curve between the pass-band and the transition region can be made sharper in such a filter. A second-order filter usually has a steeper response curve in the transition region. A further possibility in a second-order filter is to combine the low-pass function with the high-pass function to produce band-pass and band-stop filters, as described later in this chapter.

Two-capacitor filters

Two low-pass resistor/capacitor filters may be connected **in cascade** (Fig. 6.1). The output from the first filter is fed to the second filter. Suppose that $R1 = R2 = 220\,\Omega$ and $C1 = C2 = 1\,\mu$, making both stages the same as the circuit analysed in Fig. 5.2. We also use the same input signal, $v = 2\sin 2000\pi t$. Figure 6.2 plots the result of filtering. In order of decreasing amplitude, the curves are the original signal (v_{in}), the output of the first stage (v_1), and the output of the second stage (v_{out}). Although v_{in} has an amplitude of 2 V, the same as that in Fig. 5.2, the output of the first stage v_1 of Fig. 6.2 is only 0.81 V contrasted with 1.17 V for v_{out} of Fig. 5.2. This is because the first stage of the two-stage filter has had to supply current to the second stage, causing a drop in pd across $C1$. There

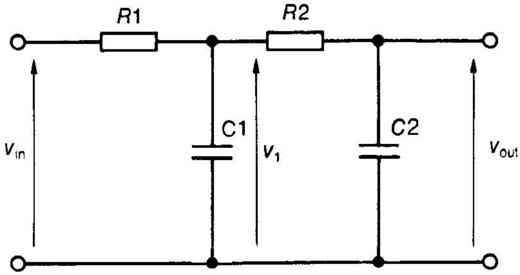


Figure 6.1 *Cascading two low-pass filters increases roll-off but reduces output amplitude*

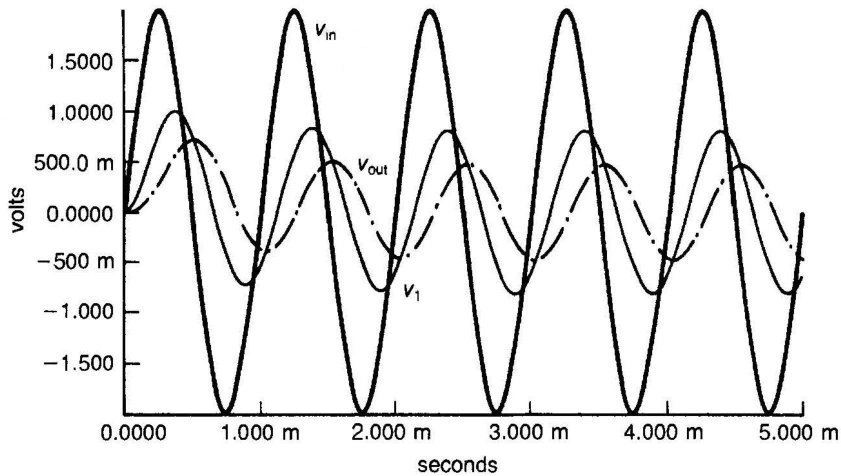


Figure 6.2 *In a cascaded low-pass filter (Fig. 6.1) the amplitude is reduced at each stage and the phase angle is increased*

is a further drop in amplitude in the second stage so that v_{out} for Fig. 6.2 is only 0.47 V. Overall, the second-order filter causes amplitude to fall from 2 V to 0.47 V. This cascading effect makes the behaviour of the filter depart from the theoretical predictions, which assume that the output of a filter stage is fed to a high impedance input at the next stage. This is something that cannot be done with passive components.

Looking at phase changes, the first stage of the second-order filter produces a phase lag of 51°. This is slightly less than that produced by the first-order filter. But the overall effect of the second-order filter is a lag of 106°. Summing up, the second-order filter has greater attenuation and bigger phase lag at 1 kHz than the first-order filter. But the action of the filter at any particular frequency is not of primary interest. The essential point about a filter is that it operates over a wide range of frequencies, and we should study its action over an appropriate range.

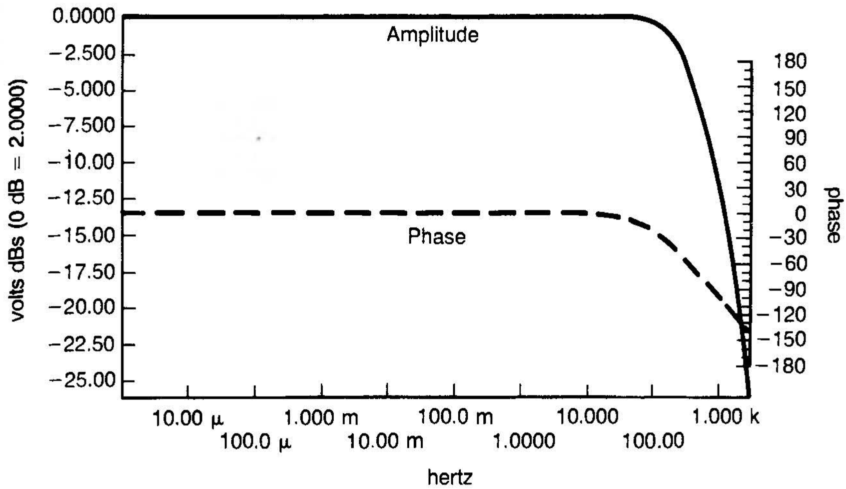


Figure 6.3 *A cascaded low-pass filter has a roll-off of −12 dB per octave*

To do this, we use a Bode plot (Fig. 6.3). This demonstrates that the second-order filter is an improvement on the first-order filter in that the rate of fall in the transition region is doubled to −12 dB per octave or slightly more, due to the cascading effect. This effect also causes the −3 dB point to occur at a lower frequency.

When the resistors and capacitors of Fig. 6.1 are exchanged, the filter becomes a second-order high-pass filter. Again there is greater attenuation than in a first-order filter but the cut-off is sharper. It is also possible to cascade resistor/inductor filters with similar results.

Capacitor/inductor filters

When a capacitor and an inductor are present in the same filter, as in Fig. 6.4, there are two reactive devices responding in opposite ways to changes of

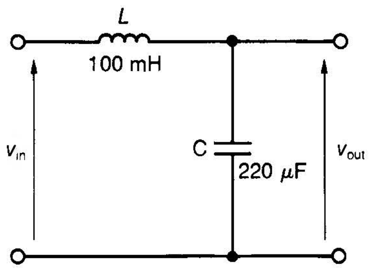


Figure 6.4 *This low-pass filter comprises two devices which have opposite reactions to frequency*

frequency. The inductor blocks high frequencies and passes low frequencies. The capacitor passes high frequencies to the 0 V line but blocks low frequencies. Their combined effect in this circuit is to produce a low-pass filter with enhanced action. The Bode plot produces some surprises (Fig. 6.5). Far from there being a gentle 'knee' on the frequency response curve, there is a very sharp spike at 33.9 Hz. To gain a full appreciation of this, we replot the response with actual voltages on the y-axis (Fig. 6.6). Output amplitude remains close to input amplitude (2 V) at low frequencies, but rapidly rises to 181 V at 33.9 Hz. Then it falls very steeply with increasing frequency. Phase shows similarly striking behaviour. At low frequencies, the phase angle is 0°, but at 33.9 Hz, it swings very rapidly to −180°.

A clue to this behaviour may be understood by calculating the reactances at 33.9 Hz. First of all, $\omega = 2\pi f = 213$ rad/s. For the capacitor:

$$X_C = -j/(213 \times 220 \times 10^{-6}) = -j21.3$$

For the inductor:

$$X_L = j \times 213 \times 100 \times 10^{-3} = j21.3$$

When the frequency is 33.9 Hz, the inductances are equal in magnitude, but opposite in direction. Because of the sign of j, pds across them are exactly 180° out of phase. At any instant the pds are of opposite signs and the total pd across the capacitor and inductor is zero. In this state, the circuit is **resonant**. The frequency at which this occurs is known as the resonant frequency.

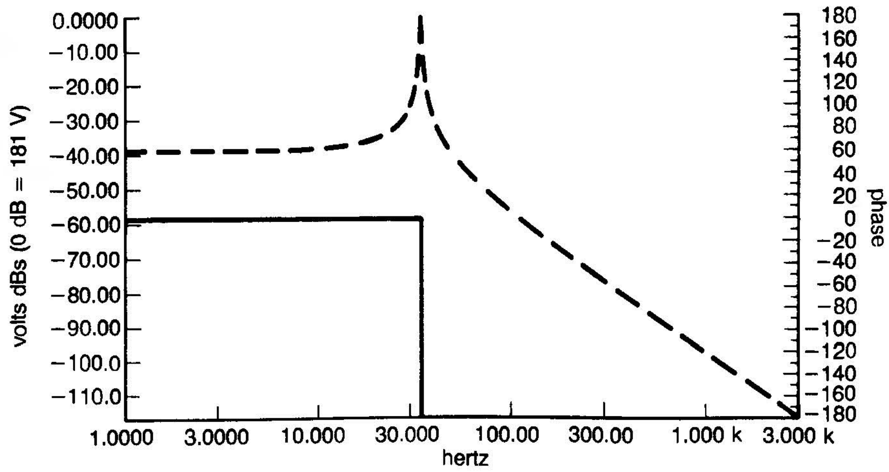


Figure 6.5 *The filter of Fig. 6.4 has a dramatic response to frequency, both in the amplitude and the phase of the output signal*

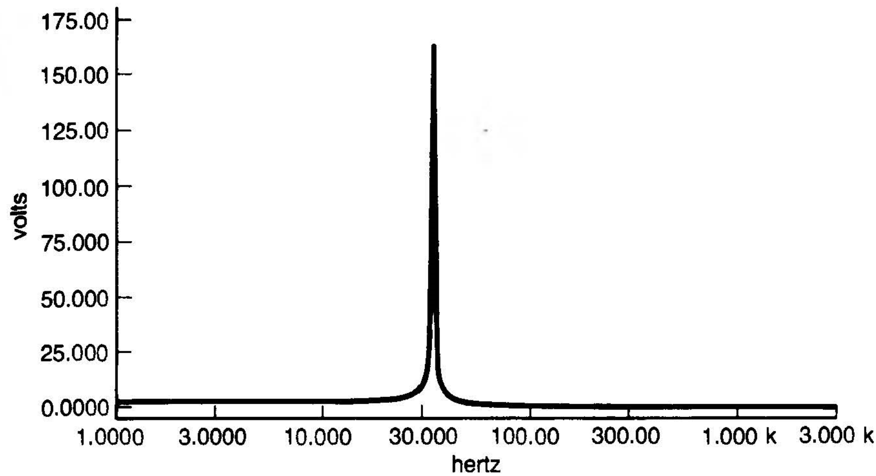


Figure 6.6 *Plotting Fig. 6.5 on linear scales makes the filter's amplitude response seem even more impressive*

Keeping up?

1. In a capacitor/inductor filter there is a peak in the response at 500 Hz. What can we say about the capacitor and inductor?
2. In what ways does the frequency response of two cascaded first-order low-pass filters differ from that of a first-order filter?

Resonance

The effect is like that of pushing a child on a swing. A swing, being a pendulum, has a natural frequency of swinging. Even if we push the swing gently every time the child is swinging away from us, we supply a little energy to the swing at each push. This energy, though small, is greater than that lost from the swing by air resistance and friction. Gradually the amount of energy in the system increases and the child swings higher and higher. The swing resonates at its natural frequency. If we push at a different rate, there are times when we are adding energy to the swing, but there are also times when we push the swing as it is coming towards us. Then we are reducing the energy of the swing and it swings less high; there is no resonance.

A similar effect is noticed in a room, such as a bathroom, when the walls, floor, ceilings and furnishings are reflective of sound. There are several frequencies at which the air of the room vibrates. If we sing a note at one of these frequencies, we supply energy to the air in phase with its natural vibrations. The amplitude of the vibrations builds up and the air 'booms' loudly at that frequency.

In the filter, a small amount of energy supplied from the input source at each cycle soon builds up to a large oscillating pd.

At resonance we have:

$$X_C = X_L$$

or

$$\frac{1}{\omega C} = \omega L$$

Here we are concerned only with the magnitudes of the signals, so we can omit j. Rearranging this equation:

$$\omega^2 = \frac{1}{LC}$$

$$\Rightarrow \qquad \omega = \frac{1}{\sqrt{(LC)}}$$

$$\Rightarrow \qquad f = \frac{1}{2\pi\sqrt{(LC)}}$$

Substituting $C = 220\,\mu\text{F}$ and $L = 100\,\text{mH}$ into this equation yields the result $f = 33.9\,\text{Hz}$, which confirms the value in Fig. 6.5. At this frequency the current drawn from the source is limited only by the output impedance of the source, the resistances of the inductor coil and the leads of the capacitor. Such impedance is likely to be low; call it R. The current to the capacitor is $i = v_{in}/R$. The current i is the same through both capacitor and inductor and we are interested in the pd across the capacitor, for that is where we are acquiring the output from the filter. If the current through the capacitor is i and its impedance at resonance is X_{Cr}, the pd across it is:

$$\mathbf{v}_{out} = iX_{Cr} = \mathbf{v}_{in}/R \cdot X_{Cr} = \mathbf{v}_{in}/(\omega_r CR)$$

where ω_r is the value of ω at the resonant frequency. The gain of the filter is:

$$\frac{\mathbf{v}_{out}}{\mathbf{v}_{in}} = \frac{1}{\omega_r CR}$$

At the resonant frequency and for a given value of C, the gain of the filter is inversely proportional to R. If R is small enough, the gain is more than 1. In other words, $\mathbf{v}_{out}$ is greater than $\mathbf{v}_{in}$. This is why it is possible to obtain an output amplitude of 181 V when the input amplitude is only 2 V, as in Fig. 6.6. Remember that Fig. 6.6 shows only the amplitude of the pd signal across the capacitor. At any instant the pd across the inductor is equal and opposite to the pd across the capacitor, so this large pd never appears across the pd generator.

The combination of capacitor and inductor has introduced a new feature into the filter. This is the possibility of producing resonance at a required frequency and so sharpening the knee of the frequency response curve. In Fig. 6.6 the knee is far too sharp to make a satisfactory filter but, by increasing the value of R, we can fashion the knee to the shape we require. Since ω_r is fixed by the value of L and C and since C itself is fixed, the gain can be controlled by selecting a suitable value for a resistor placed in series with the inductor (Fig. 6.7). In Fig. 6.8 the curves show the frequency response with resistors of different values. From top to bottom, the values of R are 0, 20, 40, 60, 80 and 100 ohms. The resistor introduces damping into the system. There is virtually no damping when $R = 0\,\Omega$, and we say that the filter is **underdamped**. When

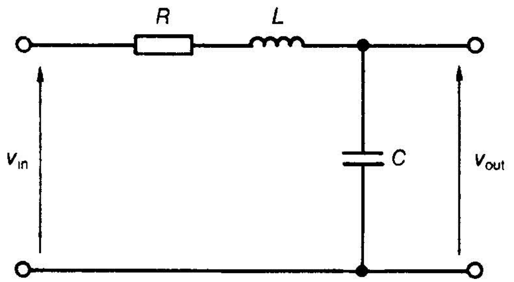


Figure 6.7 *Adding a resistor to the inductor/capacitor filter of Fig. 6.4 allows critical damping to be achieved*

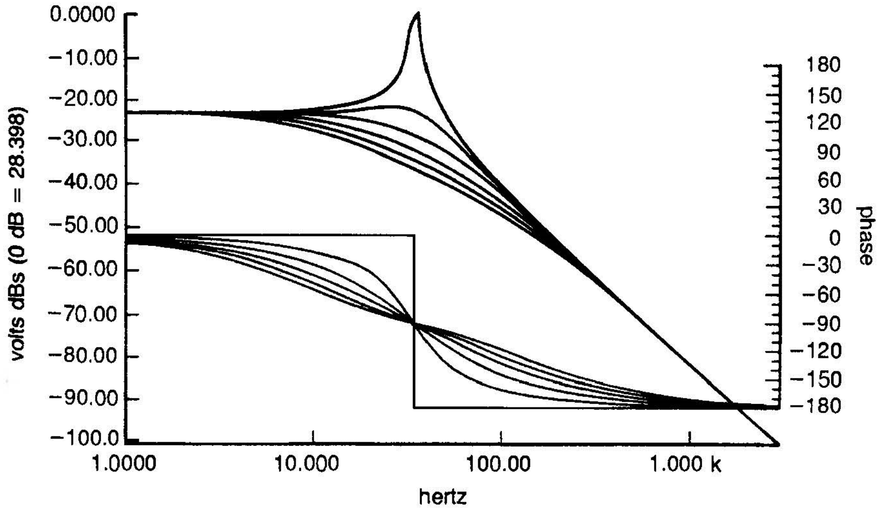


Figure 6.8 *These curves illustrate the effects of sweeping the damping resistor of Fig. 6.7 through a range of values from 0 Ω (top curves) to 100 Ω in steps of 20 Ω*

$R = 100\,\Omega$ the knee of the curve is no more prominent than in an ordinary resistor/capacitor filter (Fig. 5.15(a)) and the beneficial effect of the inductor is lost. The filter is **overdamped**. Somewhere between these extremes there is a value for R which produces **critical damping**. Judging from the figure, it looks as if the filter is slightly underdamped when $R = 20\,\Omega$, and is overdamped when $R = 40\,\Omega$. Further experimenting within this range shows that $R = 30\,\Omega$ results in a curve that remains level in the lower frequencies, then drops sharply, but without a hump. The filter is **critically damped**. Measurement on the transition region of the curve reveals that the −3 dB point is at 34 Hz and that amplitude falls by 12 dB per octave, which is typical of second-order filters.

The combination of capacitor and inductor, together with a resistor for the control of damping, produces a filter with a sharper knee and a more rapid roll-off in the transition region. The same principles can be applied to filters of other designs. We can build high-pass filters, and can cascade several low-pass filters or several high-pass filters for even sharper cut-off and steeper roll-off (though with considerable attenuation). In the case of cascaded filters containing inductors, the chief limitation is the physical size (and often weight) of the inductors required. It is only at high frequencies that capacitor/inductor filters are really practicable (see p. 82).

Keeping up?

3. What is the resonant frequency of a capacitor/inductor low-pass filter in which $C = 22\,\mu F$ and $L = 15\,\mu H$?
4. What is the reason for including a series resistor in a capacitor/inductor filter?

Band-pass filters

The idea of cascading two filters makes it possible to build a band-pass filter. Basically, all we need to do is to cascade a low-pass filter (for example, Fig. 5.15(a)) with a high-pass filter (Fig. 5.15(b)). If the cut-off points are chosen correctly, the combination cuts out low frequencies and high frequencies, leaving a pass-band of intermediate frequencies. There are several ways in which this can be done, with combinations of resistors, capacitors and inductors in various configurations. The same principles apply to all, so we shall study only one, which is similar to the low-pass filter we have just studied. It has an inductor, capacitor and resistor in series as in Fig. 6.7, but now the output is the pd across the resistor (Fig. 6.9). Giving the inductor and capacitor the same values as before, and making $R = 30\,\Omega$ to produce critical damping, the frequency response shows a peak at 33.9 Hz (Fig. 6.10). In a band-pass filter, this is known as the centre frequency. The roll-off is +6 dB per decade on the low-frequency side and −6 dB per decade on the high-frequency side. Second-order low-pass and high-pass filters normally have a roll-off of −12 dB per octave but in a band-pass filter the

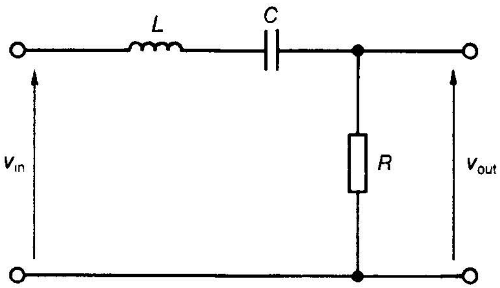


Figure 6.9 *An inductor and capacitor in series produce a band-pass filter. The resistor acts to dampen the response at the resonant frequency*

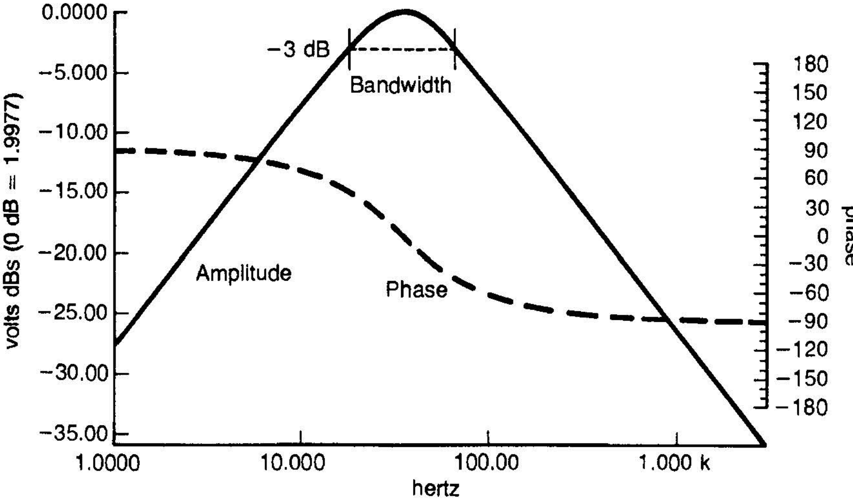


Figure 6.10 *A Bode plot of the action of the band-pass filter of Fig. 6.9. The −3 dB line cuts the amplitude curve at 19.4 Hz and 59.1 Hz, giving a bandwidth of 39.7 Hz*

roll-off is distributed equally on both sides of the centre frequency, −6 dB on each side.

The two important characteristics of a band-pass filter are **bandwidth** and the **selectivity**. By definition, the bandwidth is the difference between the upper and lower frequencies at which the output is 3 dB below the input. The bandwidth gauges the extent of the pass-band. In Fig. 6.10 the −3 dB points are at 59.1 Hz and 19.4 Hz, so the bandwidth is given by:

$$BW = 59.1 - 19.4 = 39.7\,\text{Hz}$$

The size of the bandwidth does not indicate how effective the filter is in any given application. A bandwidth of 39.7 Hz is very narrow if the frequencies we are dealing with are of the order of hundreds of kilohertz. We would rank such performance as highly selective. But with frequencies of only a hundred or so hertz, a pass-band of 39.7 Hz is relatively wide and the filter is considered to be unselective.

The selectivity of the filter, Q, relates bandwidth to the resonant frequency f_r of the filter:

$$Q = \frac{f_r}{\text{BW}}$$

For example, in the filter of Fig. 6.9, BW = 39.7 and $f_r = 33.9$, and:

$$Q = \frac{33.9}{39.7} = 0.85$$

Since Q is a ratio it has no units. Q appears in various guises in electronic circuits, often being referred to as the **quality factor**.

Although the response curve of Fig. 6.10 looks symmetrical, this is because it is plotted on a logarithmic frequency scale. With actual values, the resonant frequency does not lie half-way between the lower and upper −3 dB points. In other words f_r is not the arithmetic mean of the lower and upper points. Instead, it is the geometric mean:

$$f_r = \sqrt{(19.4 \times 59.1)} = 33.9\,\text{Hz}$$

Keeping up?

5. The upper and lower −3 dB points of a band-pass filter are 2420 kHz and 2300 kHz. What are its bandwidth and its selectivity or quality factor?
6. A band-pass filter has a bandwidth of 250 Hz and a Q of 15. What are its resonant frequency and its upper and lower −3 dB points?

Band-stop filter

When a capacitor and inductor are wired in series, they act as a band-pass filter. When wired in parallel (Fig. 6.11) the capacitor passes high frequencies and the inductor passes low frequencies. All frequencies are passed except those which are too low to pass through the capacitor and too high to pass through the inductor. In this way we obtain a filter which passes all frequencies except those within an intermediate band. This is a **band-stop filter**, sometimes known as a **notch filter** if the band is narrow. A typical response is illustrated in Fig. 6.12, using components of the same values as used in the previous band-pass filter. The output amplitude is constant at 2 V over most of the frequency range but plunges sharply to reach a minimum value of −25 dB at the resonant frequency 33.9 Hz. Calculations of bandwidth and Q are the same for this filter as for a band-pass filter.

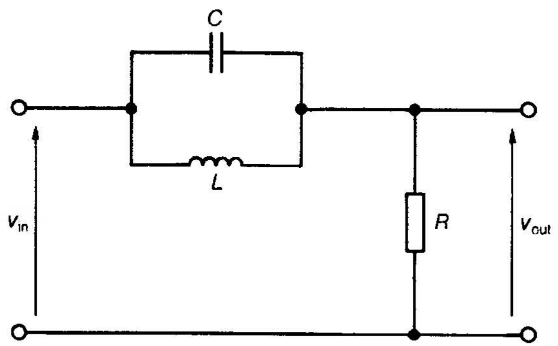


Figure 6.11 *This band-stop filter is based on a capacitor and inductor connected in parallel*

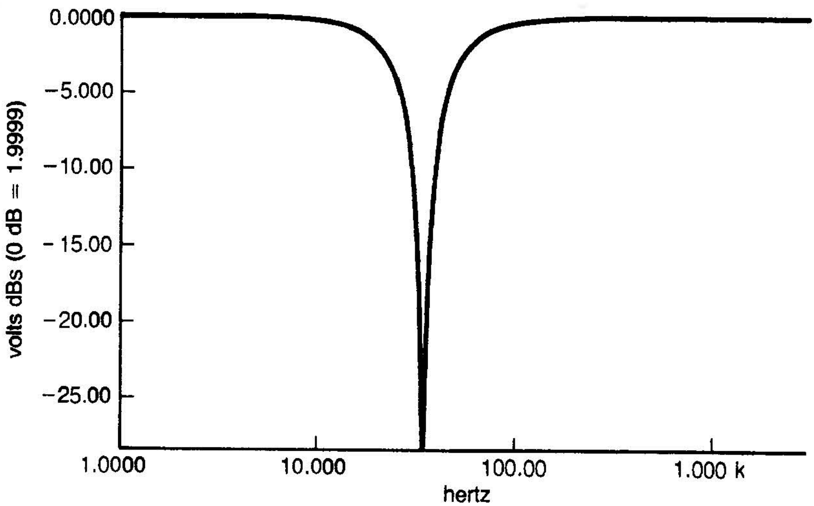


Figure 6.12 *The Bode plot of the filter in Fig. 6.11. If* $C = 220\,\mu F$, $L = 100\,\mu H$ *and* $R = 30\,\Omega$ *the curve dips sharply to −25 dB at 33.9 Hz*

The band-stop combination of a capacitor and inductor in series may also be used for band-pass filtering. As an example, take the filter of Fig. 6.13. This has two band-pass sections in cascade. Between these two sections there is a band-stop section. Very high and very low frequencies are passed through this section to the 0 V line, thus enhancing the action of the two band-pass filters. There are many other filter designs based on this principle.

Filters operating at frequencies of a few hundred megahertz may be built from capacitors of a few picofarads and inductors of a few tens of nanohenries. Although their capacitances and inductances are so small, the high frequency

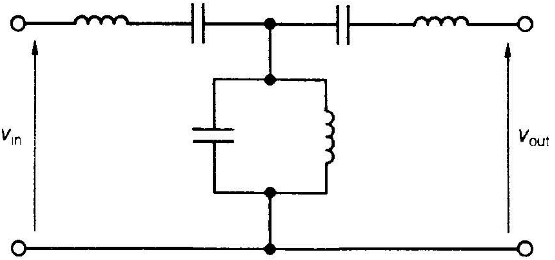


Figure 6.13 *The series-connected capacitors and inductors act as band-pass filters. Their action is enhanced by the parallel-connected capacitor and inductor which constitute a band-stop filter, which prevents intermediate frequencies from passing through to the 0 V line*

results in reasonably high reactances. For example, at 200 MHz, an inductor of 50 nH has a reactance of 63 Ω. In the gigahertz range it is not possible to manufacture capacitors and inductors of suitably low value. Even if it were, the inductance of component leads and stray capacitances in inductors would play an unduly large part in affecting circuit behaviour. Instead of inductors and capacitors, we use a shaped **microstrip**. A microstrip is a strip of conductor coated on a board of insulating material, with a continuous coating of conductor (the ground plane) on the reverse side. It is similar in appearance to one of the tracks on a printed circuit board. Microstrips are used as transmission lines for conveying microwave signals from one part of a circuit to another. If the width of the microstrip varies abruptly, as in Fig. 6.14, the variations in width cause distortions in the electrical field between the microstrip and the ground plane. The interaction between this distorted field and the currents within the microstrip produce effects that are analogous to capacitance and inductance of

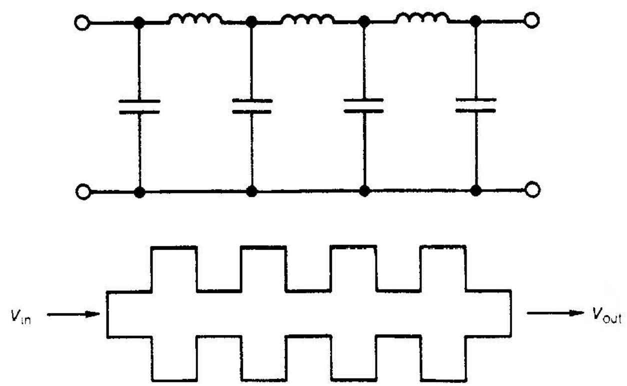


Figure 6.14 *The action of this seventh-order low-pass capacitor/inductor filter can be duplicated at ultra-high frequency by a length of microstrip shaped as shown*

very low magnitude. In Fig. 6.14, the shape of the microstrip produces an effect equivalent to that of a low-pass inductor/capacitor filter. The inductances and capacitances are determined by the dimensions of the microstrip. In this way we can produce filters with the very small inductances and capacitances required for filtering at ultra-high frequencies.

Summary

When a passive filter includes two or more reactive devices, it makes it possible to design band-pass and band-stop filters in addition to low-pass and high-pass filters. A filter that comprises one or more capacitors and one or more inductors is capable of resonance. This effect can be used to sharpen the response of the filter, and a resistor is generally used to dampen the response to its critical level. The performance of a band-pass or band-stop filter is characterized by its bandwidth and its selectivity, or quality factor. The main disadvantage of multi-stage passive filters is that the signal is severely attenuated.

Test yourself

1. Identify the filter types in Fig. 6.15(a) and (b).
2. Given that the capacitors in Fig. 6.15(a) are 47 nF, what inductances are needed to produce resonance at 25 kHz?
3. What roll-off would you expect in the filter of Fig. 6.15(b)?
4. Describe the behaviour of (a) capacitors and (b) inductors when signals of low and high frequency are passed through them.

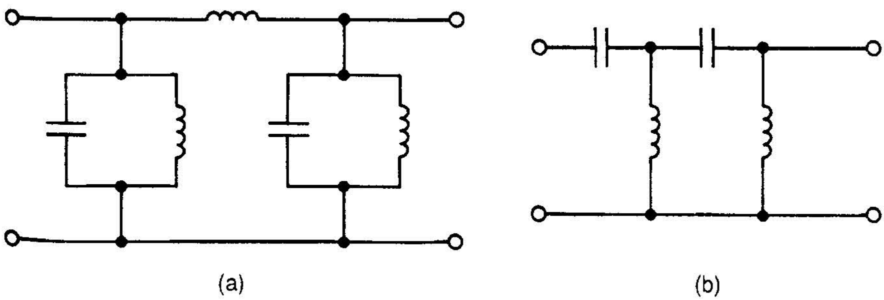


Figure 6.15 *These filter designs are the subjects of question 1*

7

Transfer functions and *s*

This chapter goes a little further into the maths behind transfer functions. In particular, it introduces a variable known as the **complex frequency variable** or, more briefly, *s*. You will often come across *s* in later chapters in this book and in the corresponding chapters of other books. In this chapter an attempt is made to help you understand it, without getting too involved in the details. In many instances *s* may be replaced in equations and transfer functions by $j\omega$, which is something that we have already come across. For example, when you find the reactance of a capacitor expressed as $1/sC$, you may consider this as replacing the more familiar expression $1/j\omega C$. If you are content with this explanation, you may skip this chapter and go straight on to Chapter 8, where we look at the more practical aspects of active filters. You may come across *s* in Chapter 8 but, when you do, mentally replace it with $j\omega$. In most cases this will not affect the relevance of the equations. On the other hand, if you would like to know more about *s*, continue reading.

The Laplace transformation

If we have a dynamic system, whether it be an electrical circuit (see the example below), the braking system of a vehicle, the flow of water in a catchment area, the financial economy of a country, or a population of elephants in a game park, it can usually be modelled mathematically by one or more differential equations. These are equations which contain terms that are derivatives or integrals, usually with respect to time. Although differential equations are relatively easy to write and understand, solving them can be very difficult. Yet solve them we must if we are ever to use the equations to explain past and present activities of a system and to predict future ones. Several techniques have been evolved for solving differential equations, one of the more recent ones being the Laplace transformation.

The aim of a transformation is to make calculations simpler. Before pocket calculators became available, we would often use the logarithmic transformation as a computational aid. Instead of multiplying two or more numbers together, we would look up their logarithms in a table, add the logarithms, then use another table to turn the sum of the logarithms back into ordinary numbers. This would give us the product of the numbers. Or we could square numbers by simply doubling their logarithms. We could even find, for example, the 5.47th power of a number very easily by multiplying its logarithm by 5.47. Similarly, the Laplace transformation makes solving differential equations easier because, instead of differentiating an expression, we multiply it by s. Instead of integrating an expression, we divide it by s. It is very useful to be able to do this because some expressions are extremely difficult to differentiate or integrate. This indicates the benefit of the transformation in broad terms and there are several details that we shall not go into here.

In practice, the transformation technique has four stages:

1. Write a differential equation to model the system. In electronics, this is an equation in the time domain (see example below).
2. Transform every term of the equation into its Laplace equivalent. Tables of transform pairs are available which list expressions and their transforms. The equation has now been transformed to the s domain.
3. Simplify and rearrange the transformed equation to produce an equation for the function we are interested in, for example the function which describes how current varies with s. The rules of ordinary algebra can be used with s, just as if it were an ordinary variable.
4. Reverse-transform the equation back into the time domain to obtain the solution of the original differential equation. The same transform-pair table is used.

As an example of using the Laplace transform, consider the low-pass filter of Fig. 5.1, in which $R = 220\,\Omega$ and $C = 1\,\mu\text{F}$. Instead of prescribing a sinusoidal input we will simplify the example by applying a DC voltage, say 4 V. The DC voltage is switched on when time is zero ($t = 0$) and we assume that there was no charge on the capacitor at that time. The current through the filter at any particular instant is $i(t)$, in other words, current $i(t)$ is a function of time t. We can use $i(t)$ to express the pds across the components:

pd across the resistor = resistance × current = $220i(t)$

pd across the capacitor = (charge accumulated since $t = 0$) divided by capacitance = $\frac{1}{10^{-6}} \int_0^t i(t)\,\mathrm{d}t$

pd across the voltage source = 4

These pds can be added together using Kirchhoff's voltage law, as in Fig. 1.5, to form a differential equation which models the action of the filter:

$$220i(t) + \frac{1}{10^{-6}} \int_0^t i(t)\,\mathrm{d}t = 4$$

The next step is to solve this equation in order to obtain an expression for $i(t)$. The difficulty is that the equation contains an integral, which makes it difficult to solve as it stands. Moreover, we want to find an equation for $i(t)$ yet here we are having to integrate it before we can find out what it is. This is a seemingly impossible task so we call on the Laplace transformation to lessen the complications. Transform each term using a table of transform pairs:

$$220I(s) + \frac{1}{10^{-6}s}I(s) = \frac{4}{s}$$

There is no space here to explain how the transforms are obtained, but it can be seen that the transformed equation is not vastly different from the original. The most notable difference is that, instead of an integral, we now have a term divided by s. The reader who prefers not to take this step in the description on trust should refer to a table of transform pairs in an electrical engineering textbook. Instead of i as a function of t, we now have the transformed equivalent I as a function of s. Solving this equation for $I(s)$, we treat s as an ordinary variable:

$$I(s)\left[220 + \frac{1}{10^{-6}s}\right] = \frac{4}{s}$$

$$\Rightarrow \qquad I(s)\frac{[220 \times 10^{-6}s + 1]}{10^{-6}s} = \frac{4}{s}$$

$$\Rightarrow \qquad I(s) = \frac{4}{s} \times \frac{10^{-6}s}{220 \times 10^{-6}s + 1} = \frac{4 \times 10^{-6}}{220 \times 10^{-6}s + 1} = \frac{0.01818}{s + 4545}$$

In the last stage above we have divided throughout by 220×10^{-6} to make the coefficient of s equal to 1. This puts the expression in a form that occurs in the table of transform pairs. In general, this stage of the calculation is one in which we try to cast the expression for $I(s)$ into a form that can be found in the pairs table, so that the next stage of transforming back to the time domain is just a matter of referring to the table. The reverse transform of $1/(s + a)$ is e^{-at}, so:

$$i(t) = 0.01818e^{-4545t}$$

This is the solution of the differential equation. It tells us that the current is 0.01818 A to start with and then decreases exponentially as charge accumulates and the pd across the capacitor gradually increases to 4 V. We can use this result to calculate other quantities. For example, the pd across the resistor is:

$$v_r = 220i(t) = 220 \times 0.01818e^{-4545t} = 4e^{-4545t}$$

From these equations we can plot graphs showing how current and the pd across the resistor vary with time.

Keeping up?

1. Given that current is determined by the function $i(t)$, what terms are used in a differential equation to model the pd across (a) a 470 Ω resistor, (b) a 2.2 μF capacitor, (c) a voltage generator delivering 5 V DC?

2. What are the Laplace transforms of the three quantities in question 1?
3. If the components mentioned in question 1 are joined together so as to make a low-pass filter like that in Fig. 5.1, find the equation for the current $i(t)$.

Another transform example

Instead of making the source pd a constant 4 V, we can make it sinusoidal. To do this, change the right-hand side of the differential equation to a sine function such as $2\sin(6283t)$, for example. Now the equation models a filter receiving the 1 kHz signal we have used in several previous discussions. From the table we find that the transform of $\sin\omega t$ is $\omega/(s^2+\omega^2)$, so the right-hand side of the transformed equation becomes $6283/(s^2+6283^2)$. We then proceed to solve the equation for $I(s)$. The reverse transformation yields an expression which includes an exponential term which attains a negligible value in less than 1 second. This corresponds to the initial transient state of the filter before the capacitor becomes fully charged, and may be ignored. The remainder of the equation is:

$$i(t) = 0.004317\cos(6238t) + 0.005968\sin(6283t)$$

Thus the current is the sum of two 1 kHz sinusoids, so it is itself a 1 kHz sinusoid. Calculations show that its amplitude is 0.006 A and it has a phase lead of about 6°. The results of using the Laplace transform are of the same order as those obtained using a circuit simulator or by observing the sinusoid with an oscilloscope in a bench test.

Interpreting *s*

In electrical and electronic calculations, s is a complex number with frequency as its dimension. As a complex number, it has two parts, and therefore can represent two attributes of a signal. This is a concept we have met before (p. 58) when we used complex numbers to represent the amplitude and the phase of a sinusoidal signal. Likewise, we use s to represent the following two aspects of a signal:

σ, the rate at which the amplitude of the signal is changing
ω, the frequency of the signal

They are combined in s for, by definition:

$$s = \sigma + j\omega$$

In a way similar to plotting the amplitude and phase of a signal on a complex plane to obtain phasor diagrams (p. 58), we plot the rate of change of amplitude and the frequency of a signal on a different complex plane, the s-plane. Plotting on this plane can be helpful in visualizing signals of certain types, so helping us to understand them. This technique for representing a waveform allows for a wide range of signal types to be investigated. We are no longer limited to pure sinusoids and mixtures of sinusoids.

We shall not follow this aspect of *s* any further except to give a few examples of what *s* can do. Figures 7.1–7.3 represent just three of the wide variety of signal types that may be analysed using *s*. The signal illustrated in Fig. 7.1 is a cosine wave of increasing frequency and exponentially decreasing amplitude. A signal like this in the audio range is a short 'chirp' (increasing frequency) with decreasing loudness. In this particular example, the rate of decrease is e^{-t}, which means that $\sigma = -1$. Its frequency is proportional to time, and in this example

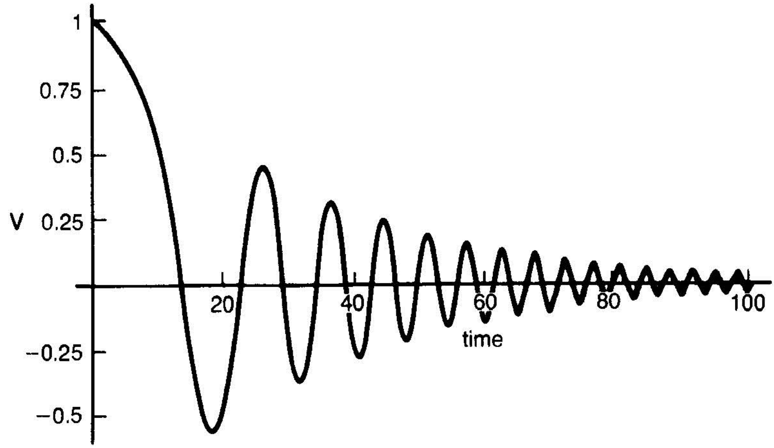


Figure 7.1 *A waveform such as this with increasing frequency and decreasing amplitude can be described by using the complex variable, s*

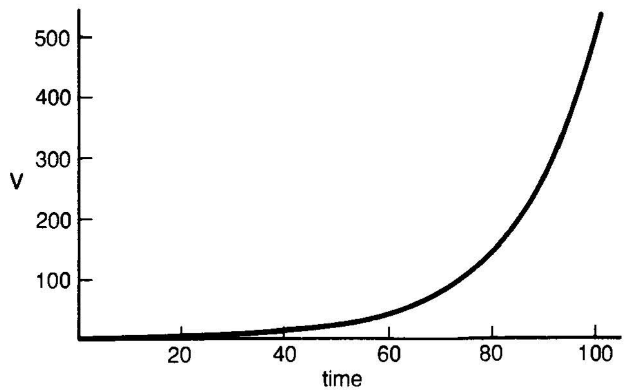


Figure 7.2 *We can also use s to describe a non-periodic signal, such as this exponentially increasing potential*

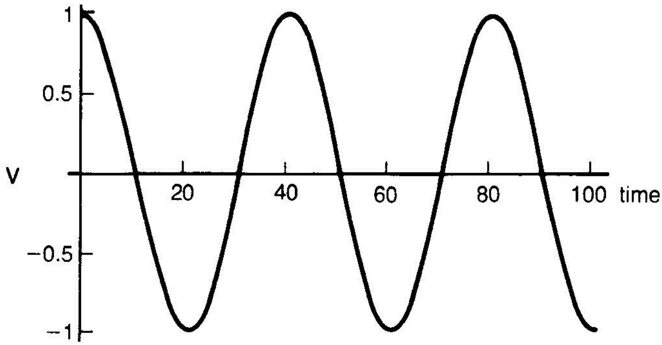


Figure 7.3 *Ordinary sinusoids can also be described using s*

$\omega = 10t$. Combining these values, $s = -1 + j10t$. This contains both amplitude and frequency information.

To visualize the signal we need to plot it in the time domain, which can be done by plotting e^{st} against time. The expression e^{st} is the inverse of the factor e^{-st} which is used to multiply a time-domain expression when transforming it to the *s* domain, so calculating values of e^{st} and plotting their real parts gives us a plot of the signal in the time domain. What happens is made clearer if we separate e^{st} into its two components:

$$e^{st} = e^{(-1+j10)t} = e^{-t} \times e^{j10t}$$

e^{-t} acts as a coefficient to determine the amplitude of the waveform. e^{j10t} generates the waveform itself, since it contains ω, the angular frequency. In fact $e^{j\omega t}$ is a third way (the exponential form) of representing a complex number. It can be proved that:

$$e^{j\omega t} = \cos \omega t + j \sin \omega t$$

The real part of this ($\cos \omega t$) is plotted in Fig. 7.1. In this example, $\omega = 10t$, so we are plotting $\cos 10t^2$, which gives a cosine curve but with steadily increasing frequency. Simultaneously, the amplitude decreases exponentially. The horizontal scale is graduated in units of $\pi/100$ seconds. Thus the full equation for the curve plotted in Fig. 7.1 is:

In the *s* domain	$v = e^{st} = e^{-1+j\omega 10t}$	(only the real part plotted)
In the time domain	$v = e^{-t} \cos 10t^2$	

It is allowable for either of the components of *s* to have zero value. If $\omega = 0$, then $e^{j\omega t} = e^0 = 1$. The signal is non-periodic and its magnitude is determined by σ. In Fig. 7.2 $\omega = 0$ and $\sigma = 2$, making the amplitude coefficient e^{2t}, and leading to an exponentially increasing signal. There is no imaginary part to this signal.

The converse case occurs when $\sigma = 0$. This makes the amplitude coefficient $e^{0t} = e^0 = 1$ and amplitude remains constant. But the signal is periodic with a frequency depending on the value of ω. This produces waveforms like the one in Fig. 7.3, for which $\omega = 5$, giving $2\frac{1}{2}$ cycles in π seconds. This is just a sinusoid signal of the type we have met so often. In most applications we shall be dealing only with filtering sinusoidal signals of constant amplitude, for which $\sigma = 0$. As a result, s and $j\omega$ may be substituted for each other in the equations and transfer functions. Often, using s rather than $j\omega$ makes the expressions a little simpler to look at and to understand.

Keeping up?

4. For what type of signal is s equal to (a) j5, (b) −3, (c) 2 + j6, (d) 0?
5. If a 1 kHz signal has constant amplitude, what is the value of s?

Transfer function of a low-pass filter

The transfer function of a first-order low-pass resistor/capacitor filter (Fig. 5.15(a)) was stated as:

$$\frac{\mathbf{v}_{\mathbf{out}}}{\mathbf{v}_{\mathbf{in}}} = \frac{1}{j\omega RC + 1}$$

Now that we have looked at s we should add that this function applies only to signals of constant amplitude and frequency. To simplify this function and at the same time make it applicable to a wider range of signal types, substitute s for $j\omega$:

$$\frac{\mathbf{v}_{\mathbf{out}}}{\mathbf{v}_{\mathbf{in}}} = \frac{1}{sRC + 1}$$

The function includes R and C as separate variables. For a filter, we can combine these into a single variable because:

$$RC = 1/\omega_0$$

where ω_0 is the −3 dB frequency (see p. 76). The transfer function becomes:

$$\frac{\mathbf{v}_{\mathbf{out}}}{\mathbf{v}_{\mathbf{in}}} = \frac{1}{s/\omega_0 + 1}$$

This form of the transfer function applies to any combination of R and C that produces a given value of ω_0. Recalling from p. 99 that $s = \sigma + j\omega$, it is interesting to plot the value of the transfer function for different values of σ and ω. Note the distinction between ω (the signal frequency) which may vary over a wide range, and ω_0 (the −3 dB frequency) which is fixed by the values selected for R and C. Figure 7.4 is a three-dimensional plot of the value of the transfer function (tf) against the components of s. This is for a low-pass filter for which $\omega_0 = 10$. The horizontal plane of the plot is the s-plane. The highest plotted

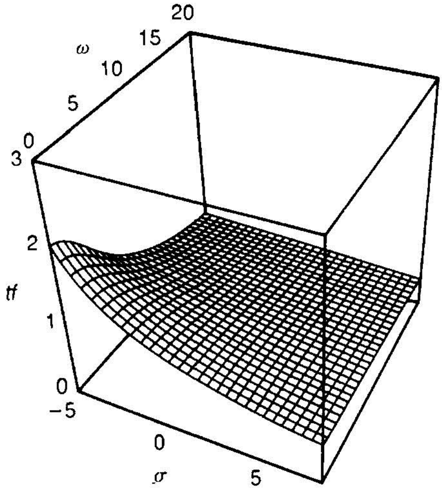


Figure 7.4 *A 3-dimensional plot of a low-pass transfer function against σ and ω shows how the filter responds to signals of the types illustrated in Figs 7.1–7.3*

value for the transfer function is 2, when $\sigma = -5$ and $\omega = 0$. This is for a non-periodic signal decreasing exponentially at the rate e^{-5t}. The shape of the curve on the rear left surface (for $\sigma = -5$, $\omega = 0$ to 30) shows the response of the filter to increasing frequency, with an exponentially decreasing amplitude. The shape is very similar to that of Fig. 5.7, and is characteristic of a low-pass filter. In Fig. 7.5 we have sliced vertically through the plot of Fig. 7.4 along the vertical plane for which $\sigma = 0$. We have also turned the plot around to take a better look

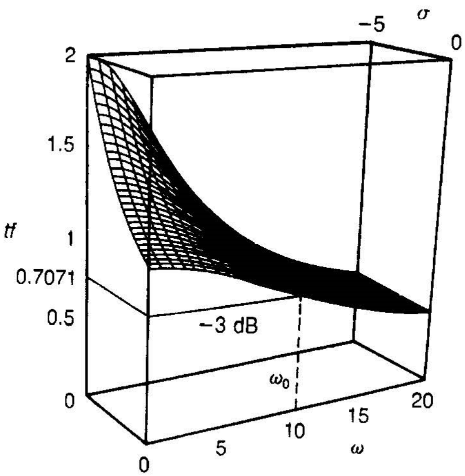


Figure 7.5 *This view of a portion of the plot of Fig. 7.4 depicts the low-pass action on a sinusoid of constant amplitude*

at this plane. The intersection between the plane for $\sigma = 0$ and the plotted surface is the graph for the transfer function for sinusoidal signals of fixed amplitude. It has the same form and the same interpretation as the response curve plotted in Fig. 5.7. At the left-hand end of this curve, $\sigma = 0$ and $\omega = 0$, which makes $s = 0$ and we have a constant DC level. At this point the term s/ω_0 disappears from the transfer function, the value of which becomes 1, as can be seen on the graph. Increasing ω, and hence increasing s, makes the term s/ω_0 increase, causing the value of the transfer function to decrease. This too can be seen in the graph, in which tf falls as ω rises, just as expected in a low-pass filter. Further, when $\omega = 10$, the value of the -3 dB point, the transfer function takes the value 0.7071 (or $1/\sqrt{2}$), the half-power value (see box).

Returning to Fig. 7.4, we see that a similar low-pass action occurs when σ is negative, for signals which are decreasing in amplitude. On the other hand if σ is positive, meaning that amplitude is increasing, the filter response becomes much flatter and has much reduced overall gain.

Evaluating the transfer function

As an exercise in evaluating the transfer function, take Fig. 7.4, in which $\omega_0 = 10$. There are three cases to consider:

1. If s is entirely real (that is, $\omega = 0$), the transfer function is real. For example, if $\sigma = -5$ and $\omega = 0$, then:

$$tf = \frac{1}{-5/10+1} = \frac{1}{-1/2+1} = \frac{1}{1/2} = 2$$

2. If s is entirely imaginary (that is, $\sigma = 0$), the transfer function is complex. Taking as the example the -3 dB point where $\sigma = 0$ and $\omega = \omega_0 = 10$, then:

$$tf = \frac{1}{\mathrm{j}10/10+1} = \frac{1}{\mathrm{j}+1} = \frac{1-\mathrm{j}}{1+1} = \frac{1-\mathrm{j}}{2} = 0.5 - \mathrm{j}0.5$$

   Division by expressions such as $\mathrm{j}-1$ is explained in Chapter 4. The transfer function is complex with both real and imaginary components. This is the rectangular form; converting into polar form (Fig. 4.15), the magnitude of their sum is $tf = \sqrt{(0.5^2 + (-\mathrm{j}0.5)^2)} = 0.7071 = 0.707$. The phase angle is $\tan^{-1}(-0.5/0.5) = \tan^{-1}(-1) = -45°$.

3. If s is complex, the transfer function is complex. For example, if $s = 8$ and $\omega = 20$, then:

$$tf = \frac{1}{(8+\mathrm{j}20)/10+1} = \frac{1}{1.8+\mathrm{j}2} = \frac{1.8-\mathrm{j}2}{19.24}$$
$$= 0.09356 - \mathrm{j}0.1040$$

Converting into polar form, magnitude is 0.14 and phase angle −48°.

All of these results for magnitude are displayed in Fig. 7.4.

Transfer function of a high-pass filter

By contrast, Fig. 7.6 shows the s-plane plot of the transfer function of a high-pass filter of the type illustrated in Fig. 5.15(b). The transfer function of this is:

$$tf = \frac{v_{out}}{v_{in}} = \frac{j\omega RC}{1 + j\omega RC} = \frac{sRC}{1 + sRC} = \frac{s}{s + \omega_0}$$

Substitutions have been made as in the case of the low-pass filter. The intersection of the surface with the near right plane (for $\sigma = 8$, $\omega = 0$ to 30) shows a high-pass response. This response is stronger when $\sigma = 0$, for then the response drops all the way to zero when $\omega = 0$, that is for a DC signal with constant amplitude. This is not surprising because a DC signal does not pass through a capacitor. The transfer function predicts this result because substituting $s = 0$ in the function makes it take zero value. Substituting zero for s in the numerator of a quotient such as the transfer function makes it equal to zero. This condition is said to be a **zero** of the transfer function. We can also infer the response to high frequencies. As ω increases, so does s, but ω_0 remains constant. Consequently, large values of s mean that ω_0 plays a relatively smaller part in determining the value of tf.

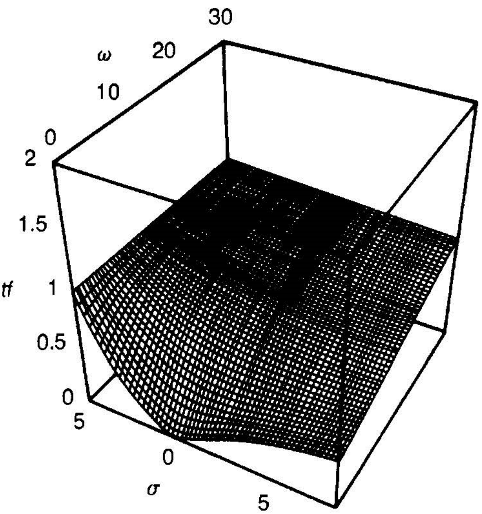


Figure 7.6 *The plot of the response of a high-pass filter has a zero when* $\sigma = 0$ *and* $\omega = 0$

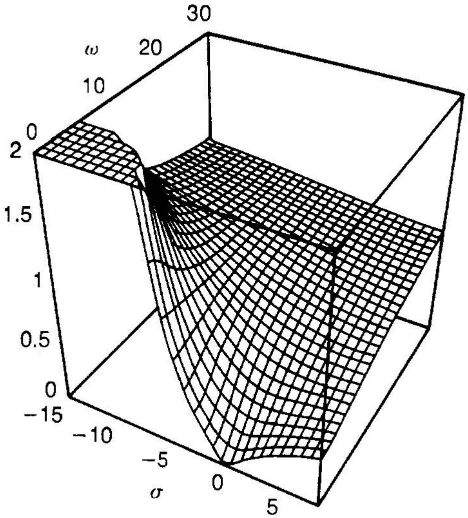


Figure 7.7 *Extending the plot of the high-pass response reveals a pole at* $\sigma = -10$ *and* $\omega = 0$

As s tends to large values, tf tends toward s/s, which equals 1. This can be seen in Fig. 7.6.

If we expand the range of the plot of Fig. 7.6 to cover more negative values of σ, another feature of the transfer function 'landscape' appears (Fig. 7.7). This was hinted at on the left of Fig. 7.6 but now it is seen as a massive upward sweep, known as a **pole**. Here we see only the base of the pole for it has been truncated by plotting only as far up as $tf = 2$. The axis of this pole is at the point where $\sigma = -10$ and $\omega = 0$. At these values, $s = -10$, and equals $-\omega_0$. Substituting $s = -10$ and $\omega_0 = 10$ in the transfer function makes the expression $(s + \omega_0)$ equal to zero, so the denominator as a whole is zero and the transfer function becomes infinitely great. A pole in this position means that when the input is a non-periodic signal decreasing exponentially at the rate e^{-10t} the output at any instant is infinite. In practice there is always some damping due to the resistance of connections and terminal wires and also to power losses in dielectric of the capacitor, but at least the output of the filter is many times the input. This high gain is explainable. The input is falling so rapidly that the capacitor is unable to discharge through the resistor fast enough to keep up with it. The same applies to periodic signals decreasing at similar rates and oscillating at low frequencies, such that the rate of exponential decrease swamps signal changes caused by its periodicity.

Locating poles and zeros is important in filter design as we are able to predict regions of instability and regions of high attenuation. The pole in this filter is not on the $s = 0$ axis, so will not interfere with the action of the filter for constant-amplitude signals.

Transfer function of a band-pass filter

The filter chosen for this example is the resistor/capacitor/inductor filter illustrated in Fig. 6.9. Its transfer function is:

$$\frac{v_{out}}{v_{in}} = \frac{sCR}{1 + sCR + CLs^2}$$

This is a second-order filter and its transfer function contains both s and s^2. By substituting $\omega_0 L/Q$ for R, and $1/\omega_0^2$ for LC, the transfer function can be re-expressed in terms of Q, the quality factor, ω_0 the centre frequency, and s:

$$\frac{v_{out}}{v_{in}} = \frac{s}{Q\omega_0 + s + s^2/\omega_0}$$

Examination of this function shows that it has a zero when $s = 0$, that is when $\sigma = 0$ and $\omega = 0$. This can be seen in the plot in Fig. 7.8, where the plane dips down to zero. It can also be shown that the expression tends toward zero as s approaches infinity. A filter for which gain is low around zero frequency and also as frequency approaches infinity (perhaps well before it approaches infinity) is obviously a band-pass filter. Finding the poles involves finding the roots of the quadratic expression in the denominator when it is put equal to zero. The expression has two roots, so we would expect there to be two poles. The calculations are rather complicated but Fig. 7.8 shows one of the poles. This is its location when $\omega_0 = 10$ and $Q = 20$. The pole is situated where $\sigma = 0$ and $\omega = 10$. It is centred on the $\sigma = 0$ axis, so is most effective for signals of constant amplitude. This corresponds to the pass-band of the filter. As in Fig. 7.7, the pole is truncated because the plot limits tf to 0.1, so as to make the form of the lower levels clearer. In fact, it peaks sharply at the centre frequency. In practice, a resistor would be used to dampen the response. The other pole is situated symmetrically at $s = 0$ and $\omega_0 = -10$, which is a region not plotted in Fig. 7.8.

Knowing the locations of poles and zeros is important in predicting circuit behaviour. If a pole is situated to the right of the $s = 0$ axis, so that s is positive, it is in the region of exponentially increasing output signals, which is an inherently unstable situation. In practice, dissipation of energy in resistors will soon dampen the response of a passive filter but with the active filters described in the next chapter the output is likely to go out of control, no longer being dependent on the input signal. In contrast, if all the poles of a filter are to the left, so that s is negative and the output falls exponentially after input has ceased, the filter is stable. The low-pass and high-pass filters demonstrated in Figs 7.6 and 7.8 are therefore stable filters. If the filter has a pair of poles symmetrically placed on the $s = 0$ axis, the circuit is marginally stable. In practice its action is damped in a passive filter because of loss of energy in the resistors and other components. But in an active filter (see next chapter), damping may be countered by the amplification. Such a filter has the behaviour of an oscillator. In fact, the circuits of band-pass filters and oscillators have many features in common.

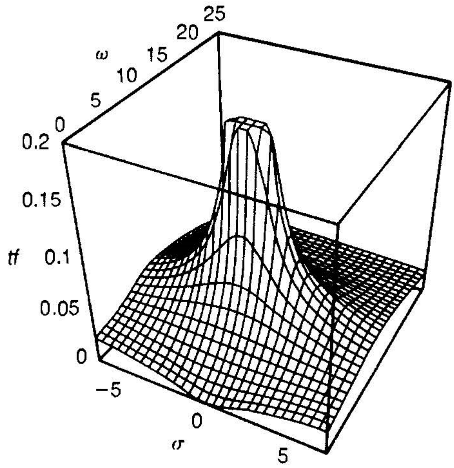


Figure 7.8 *A second-order band-pass filter has a pronounced pole in the centre of its pass-band*

Summary

The Laplace transformation is used in electronics to convert functions in the time domain to functions in the complex frequency (s) domain. The purpose of this is to make calculations easier.

s consists of a real part σ, the exponential rate of change of amplitude, and an imaginary part jω, where ω is the frequency, which may or may not be constant.

If amplitude and frequency are constant, s may be replaced in transfer functions by jω.

By finding values of s which make the transfer function equal to zero or have an infinite value, we locate the positions of poles and zeros in the s-plane.

Information about poles and zeros is used to locate regions of strong response and to determine if a circuit is stable or not.

Test yourself

1. Rewrite these transfer functions in terms of jω: (a) $4s/(4s+1)$. (b) $0.005s/(1+0.005s+0.01s^2)$, (c) $0.1/(s^2-s-2)$.
2. Find the poles and zeros for the transfer functions of question 1. Which of these, if any, represents an unstable circuit?
3. A band-pass filter (Fig. 8.2) has the transfer function $s/(Q/\omega_0+s+s^2)$. Given that $Q=2$ and $\omega_0=10$, find the location of its poles and zeros. Is this a stable circuit?

8
Active filters

The passive filters we have described in Chapters 5 to 7 all have one characteristic in common. The amplitude of the output signal is less than that of the input signal. With a filter of several stages the loss in amplitude becomes very serious. An active filter incorporates one or more stages that restore or even enhance the amplitude. The active stage may be a transistor but most often it is an **operational amplifier**. The amplifier draws power from an external source and uses this to increase the power of the output signal. An equally important outcome of introducing operational amplifiers is that they can be used to shape the response of the filter, particularly to sharpen the knee of the response curve, so making it possible to dispense with weighty and bulky inductors.

Operational amplifiers

Without going into their internal details, we must outline certain features of operational amplifiers (from now on referred to as op amps) which are relevant to their use in filters. An op amp has five, possibly more, terminals (Fig. 8.1). Two of these, +V and −V, are the power supply terminals which draw power from a pair of supply rails held at voltages equally above and below the 0 V rail. Occasionally the −V terminal may be connected to the 0 V rail, but this is not

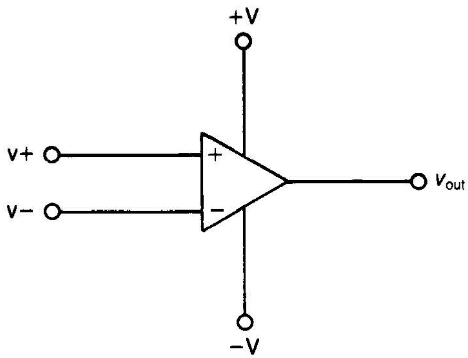


Figure 8.1 *An operational amplifier accepts two input signals and produces one output signal*

of concern here. There are two input terminals, v+ and v−, the function and properties of which will be described shortly. There is a single output terminal, v_{out}. There may be one or two other terminals, to which external components are connected for biasing or balancing the amplifier but these are used only in circuits of the highest precision.

There are several ways in which an op amp may be connected into a circuit. Figure 8.2(a) and (b) show the two most popular ways of using op-amps as amplifiers. In Fig. 8.2(a) the op amp is wired as a **non-inverting amplifier**. The input signal is applied to the v+ terminal. One important feature of op amps is that v+ and v− have a very high input impedance. In op amps with BJT inputs, the input impedance is often about 2 MΩ. In op amps with JFET or MOSFET inputs, the input impedance is of the order of 1 teraohm, that is 10^{12} ohms. In either case the input impedance is so high that it draws virtually no current, a property of great importance in filters, because we often rely on one stage of a filter supplying no current to the next stage.

When a positive pd (measured with reference to the 0 V line) is applied to the v+ terminal, the output potential rises above zero. R_F and R_A act as a potential divider and a proportion of the output pd is fed back to the v− terminal. The action of the op amp is to raise v_{out} until v− and v+ are at the same potential, which is v_{in}. As an equation:

$$\text{Potential at } v- = v_{out} \times \frac{R_A}{R_F + R_A} = v_{in}$$

The gain of the amplifier is:

$$K = \frac{v_{out}}{v_{in}} = \frac{R_F + R_A}{R_A}$$

For example, if $R_F = 470\,\text{k}\Omega$ and $R_A = 10\,\text{k}\Omega$, then $K = 480\,\text{k}/10\,\text{k} = 48$. Note that the gain is determined only by the values of the resistors, so it is possible to set the gain very exactly by using high precision resistors. There are two limitations on gain. There is an upper limit determined by the characteristics of the op amp circuit. This is usually about 200 000, and a gain of this size is well above the requirements of a filter. More significant is the fact that the gain falls off with increasing frequency. Typically, gain remains constant up to 10 kHz, then falls steadily to reach 1 at 1 MHz. In this way the op amp itself acts as a low-pass filter. For many applications, including audio-frequency circuits, this makes no practical difference.

The output impedance is low, often in the region of 75 Ω, so it is able to supply ample current for driving the next stage of a filter circuit.

The non-inverting amplifier circuit of Fig. 8.2(a) simply amplifies the input signal. The **inverting amplifier** of Fig. 8.2(b) both amplifies and inverts it. In Fig. 8.2(b) the input is fed to the v− terminal by way of R_A. Since the v+ terminal is connected to the 0 V rail, its potential is very close to zero. As explained earlier, the op amp settles when its v+ and v− inputs are at the same potential. In the inverting amplifier it settles with both inputs at or exceedingly close to zero. For

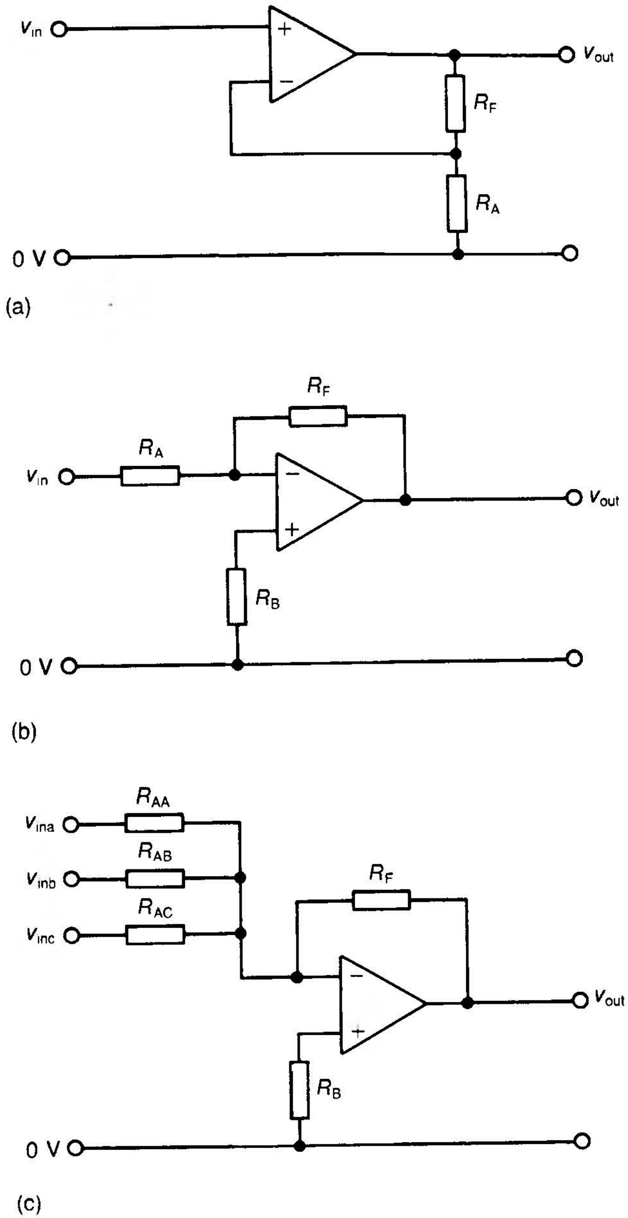


Figure 8.2 *Operational amplifiers can be connected to make (a) a non-inverting amplifier, (b) an inverting amplifier and (c) a summing amplifier. In these and subsequent figures, the connections to the power lines, +V and −V are omitted for clarity*

this reason, the v− terminal in an inverting amplifier is referred to as a **virtual ground**, or **virtual earth**. As a result of this, the input impedance of the amplifier equals the value of R_A, which may be only 10 kΩ. This is fairly high, but not as high as that of the v+ terminal in the non-inverting amplifier. When a positive pd is applied to the amplifier in Fig. 8.2(b), current flows through R_A to the virtual ground at the v− input. It apparently flows into that input but actually it flows on through R_F to the output terminal. This now has a negative potential so that the current flows into the output terminal. As v_{in} increases positively, v_{out} falls by exactly the amount needed to keep the v− terminal at 0 V. R_F and R_A act as a potential divider in which the potential at the junction of the resistors is always zero. Apart from the minute current entering v−, the current i flowing through R_A equals the current flowing through R_F:

$$i = v_{in}/R_A = -v_{out}/R_F$$

The negative sign is needed because v_{out} and v_{in} have opposite signs. Calculating the gain:

$$K = \frac{-v_{out}}{v_{in}} = \frac{-R_F}{R_A}$$

As with the non-inverting amplifier, the gain is dependent only on the values of R_F and R_A. A significant effect of the inverting action of the op amp is that the output and input signals are 180° out of phase. The limitations regarding maximum op amp gain and the effects of frequency apply also to this amplifier. For the highest accuracy, R_B should be equal to R_A and R_F in parallel, but for low-precision circuits, R_B may be omitted and the v+ terminal connected directly to the 0 V line.

Op amps are also used as **summers**, for adding two or more values together. Figure 8.2(c) explains how three values are added. There are three inputs at which the three values to be added are represented by three voltages V_{AA}, V_{AB} and V_{AC}. The key to the operation of this circuit is the virtual ground at the inverting input of the op amp. No matter what currents are flowing through the resistors, their terminals at the op amp end are all at 0 V. The amount of current flowing in through one resistor has no effect on the currents flowing through the other resistors. If all three input resistors are equal, the currents flowing through them are proportional to the values to be added. The currents flowing on through R_F is the total (or sum) of the currents flowing through the input resistors. The output potential falls to draw the total current in through the output terminal of the op amp:

$$v_{out} = -(V_{AA} + V_{AB} + V_{AC}) \times \frac{R_F}{R_A}$$

where R_A is the resistance of each input resistor, on the assumption that they are equal. In this way we take the output pd as the inverse of the sum of the inputs. Usually we make R_F equal to R_A so there is a gain of 1 and the output is the numerical equivalent of the sum of the inputs. Alternatively R_F can be more or less than R_A so as to multiply or divide the sum by a required amount. There are occasions when we wish to multiply one or more of the inputs before adding them. If so, we reduce the value of its input resistor proportionately.

Keeping up?

1. A non-inverting amplifier is connected as in Fig. 8.2(a), using an op amp with BJT inputs, with $R_A = 22\,k\Omega$ and $R_F = 560\,k\Omega$. What are the input impedance and the gain of the circuit?
2. An inverting amplifier is connected as in Fig. 8.2(b), using an op amp with MOSFET inputs, with $R_A = 12\,k\Omega$ and $R_F = 1\,M\Omega$. What are the input impedance and gain of the circuit?
3. In an op amp summer connected as in Fig. 8.2(c), all input resistors and R_F are $10\,k\Omega$. The input voltages are 1.3 V, 0.9 V and 2.24 V. What is the output voltage?

First-order active filter

This consists of a passive filter followed by an amplifier, usually an op amp (Fig. 8.3). In this example the passive filter is a resistor/capacitor low-pass filter as in Fig. 5.15(a), connected to a non-inverting amplifier as in Fig. 8.2(a). The output of the passive section is connected to the v+ input of the op amp which has a high input impedance and so exerts very little influence on the action of the filter. The frequency response of the circuit is the same as that of the passive filter, except that signal amplitude is increased in the ratio $(R_F + R_A)/R_A$. As in the passive filter, the $-3\,dB$ point is at $f = 1/2\pi RC$, and the roll-off is $-6\,dB$ per octave.

If the gain of the op amp is K and we adapt the general equation derived in Chapter 5 (p. 69), the transfer function of the filter is:

$$\frac{\mathbf{v_{out}}}{\mathbf{v_{in}}} = \frac{K}{j\omega RC + 1}$$

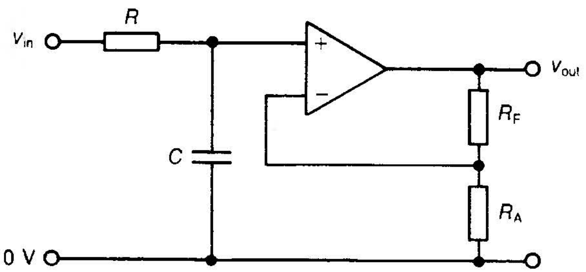


Figure 8.3 *A first order low-pass active filter, consists simply of a passive filter followed by a non-inverting amplifier*

In this equation $K = (R_F + R_A)/R_A$. The equation describes the action of a first-order low-pass active filter. Given the values of K, R and C, we can calculate the amplitude and phase of the output signal at any frequency.

Second-order active filter

The second-order low-pass filter in Fig. 8.4 shows how the performance of a filter may be modified by **feedback**. A filter of this type is also known as a **VCVS filter** because the op amp is used as a voltage-controlled voltage source. At first glance, the circuit looks like two cascaded low-pass filters followed by an op amp, but closer inspection reveals that $C1$ is not connected to the 0 V line but to the amplifier output. In this way the output signal is fed back to an early stage in the input chain. More significant is the fact that this feedback goes to the v+ input. It is positive feedback.

There are two types of feedback in this circuit. There is feedback from the junction of $R3$ and $R4$ to the v− terminal. This is negative feedback and, like all negative feedback, its effect is to limit the voltage swings at the output. In op amp circuits it is used to reduce gain to a precisely limited amount. In other words, it gives the amplifier stability. In contrast the feedback through $C1$ tends to make the output unstable. As output rises, part of the signal is fed back into the resistor chain through $C1$. This passes through $R2$ to the v+ input of the op amp, causing the output to rise further. Conversely, as the output falls, positive feedback acts to make it fall still further. With too much feedback the amplifier goes into continuous oscillation at the resonant frequency of the circuit. We met resonance before, in connection with capacitor/inductor filters (p. 87) and this suggests that a capacitor/op amp filter may be expected to have some of the features of capacitor/inductor circuits, perhaps the sharpened knee of the frequency response curve.

Like resonance, the positive feedback effect only occurs at a given range of frequencies. At low frequencies, signals are not passed through $C1$ but are

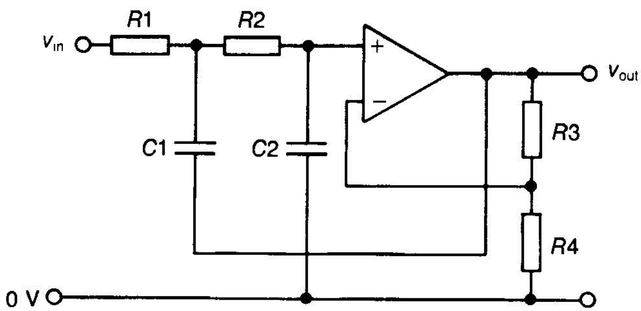


Figure 8.4 *A second-order low-pass active filter has positive feedback through $C1$ and negative feedback to the $-V$ terminal*

more easily passed on to the output, so there is minimal feedback. At high frequencies the signal is passed strongly through $C1$, but the stabilizing action of the negative feedback comes into play. The result is a tendency for the circuit to resonate at some intermediate frequency between low and high. Figures 8.5 and 8.6 illustrate this effect. These are the frequency response curves, or Bode plots, when the component values are $R1 = R2 = 10\,\mathrm{k}\Omega$, $R3 = 22.85\,\mathrm{k}\Omega$, $R4 = 39\,\mathrm{k}\Omega$, $C1 = C2 = 0.016\,\mu\mathrm{F}$. These produce a filter with 1 kHz as the −3 dB point. In Fig. 8.5 the 'lower' terminal of $C1$ is connected to the 0 V line instead of to the op amp output terminal as in Fig. 8.4. This converts the filter into two cascaded low-pass passive filters with amplification. The curve for v_{out} shows the typically gradual fall of amplitude as frequency increases. This is shown in the other curve, which is the pd across $C1$ and simply represents the output of the first stage of the cascaded filters. Figure 8.5 demonstrates what happens when the 'lower' terminal of $C1$ is connected to the op amp output, as in Fig. 8.4. The amplitude of the pd across $C1$ peaks around the −3 dB point. The effect of this signal fed back positively into the system is to sharpen the knee of the curve for v_{out}. This demonstrates that an op amp circuit with positive feedback can have a similar frequency response curve to a capacitor/inductor second-order passive filter. But, whereas an inductor is often bulky and massive, is expensive, is subject to electromagnetic interference, and cannot easily be manufactured to a high degree of precision, an op amp takes up only a fraction of a square millimetre on a silicon chip, is very cheap, is not subject to interference and can give high-precision performance at the relatively low cost of a pair of precision resistors.

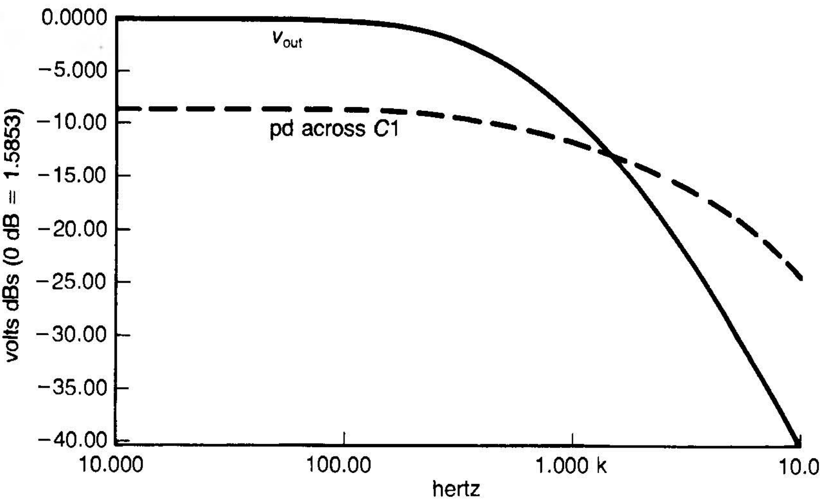


Figure 8.5 *Pds in the second-order filter (Fig. 8.4) when C1 is connected to the 0 V line, so that there is no feedback through it*

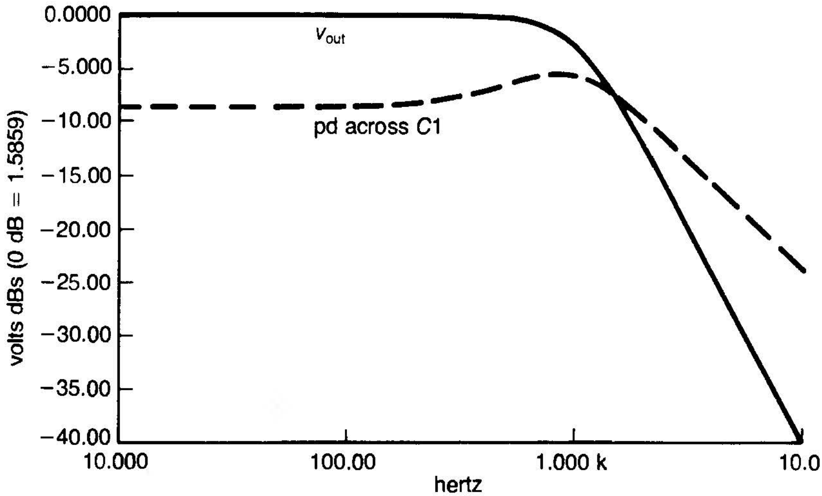


Figure 8.6 *Pds in the second-order filter when there is feedback through C1. The peak in the pd across C1 makes the knee in the output curve more prominent*

The low-pass filter described above is readily convertible to a high-pass filter by exchanging the resistors with the capacitors.

Shaping the response curve

By using op amps we can do more than mimic the behaviour of an inductor. We can exercise much more precise control over the shape of the response curve. The filter of Fig. 8.4 is a version of the well-known Sallen-Key filter. In the true Sallen-Key filter, *R*4 is omitted so that the feedback to the v− input is through *R*3. As a result the filter has a gain of 1. With *R*4 present, it is possible to adjust the gain. If we are using the op amp simply to amplify the output from cascaded passive filters, it is feasible to set the gain to any value that we require. But in the version of the Sallen-Key filter depicted in Fig. 8.5, the gain determines the effect of feedback and therefore has a significant effect on the frequency response of the filter. For stability, the gain must not be more than 3 and is best expressed as $K = 3 - d$, where d is the damping factor. d can take any value between zero and 2, and is the reciprocal of q, the quality factor (p. 92). Thus gain lies between 1 and 3. The equation for the gain of a non-inverting amplifier is:

$$K = \frac{v_{out}}{v_{in}} = \frac{R_A + R_F}{R_A}$$

$$\Rightarrow \qquad R_F = R_A(K - 1) = R_A(2 - d)$$

R_A can be given any reasonable value, such as 39 kΩ and R_F is then calculated given the required value of d. The effect of d is demonstrated in Fig. 8.7, where

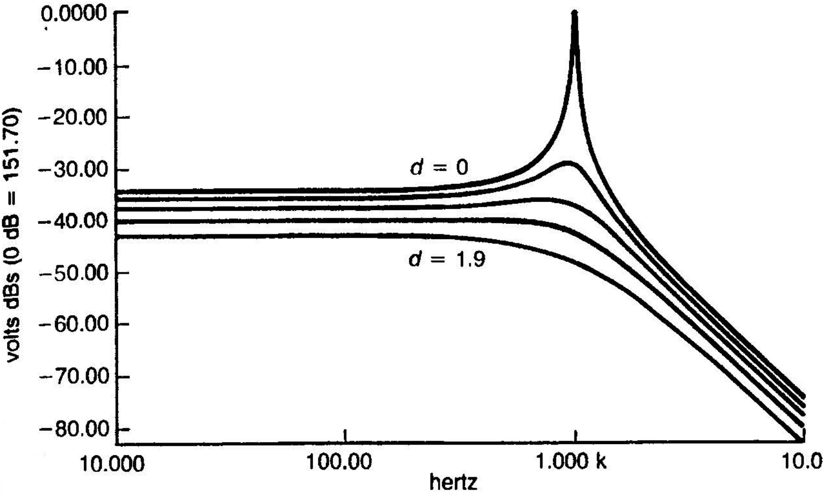


Figure 8.7 *Sweeping the damping factor from zero to 1.9 has a marked effect on the response of the second-order filter*

d is varied from 0 to 1.9. When $d = 0$, the output peaks very sharply at 1 kHz and the filter is underdamped. This is a pole of the filter. As damping is increased the gain of the filter is decreased. This is not a problem as, if the filter needs to be so damped that its gain is too small, its gain is made good by wiring another amplifier to the output. The more important effect is that damping reduces the peaking in a controlled way. For $d = 1.9$ the filter is overdamped but, somewhere between $d = 0$ and $d = 1.9$, there is a point where the response is critically damped. Visually, this appears to be shown by the fourth curve down, the one for which $d = 1.425$. The value of $R3$ which produced Fig. 8.7 was such as to make $d = 1.414 = \sqrt{2}$. This value produces the flattest response in the pass-band coupled with a sharp but not exaggerated knee. As we shall see later, this is the characteristic of the Butterworth response.

Transfer function

The transfer function of the filter of Fig. 8.4 is

$$\frac{K\omega_0{}^2}{s^2 + d\omega_0 s + \omega_0}$$

where ω_0 is the -3 dB frequency (p. 74). Remember that we can replace s by $j\omega$ for a sinusoidal signal of constant amplitude. As frequency increases, the denominator increases but the numerator remains the same; this shows that it is a low-pass filter. The poles of the filter are found by finding the values of s for which the denominator equals zero. These are calculated by using the quadratic equation formula to find the roots of the expression in the denominator:

$$s = \frac{-d\omega_0 \pm \sqrt{(d^2\omega_0{}^2 - 4\omega_0{}^2)}}{2}$$

If $d = 0$, for the undamped circuit (Fig. 8.7, top curve):

$$s = \frac{0 \pm \sqrt{(0 - 4\omega_0{}^2)}}{2}$$

If $\omega_0 = 6283$ rad/s, as in the example (equivalent to 1 kHz):

$$s = \frac{\pm\sqrt{(-4 \times 6283^2)}}{2} = \frac{\pm j12566}{2} = \pm j6283$$

The poles are on the ω-axis at $\omega = +j6283$ and $\omega = -j6283$. The $\omega = +j6283$ pole corresponds with the peak at 1 kHz.

Classifying shapes

There are an infinite number of shapes possible for a low-pass response, and likewise for high-pass and other categories of filter but, to help avoid confusion, responses are classified under a few main types. The classification is based on various characteristics:

Pass-band flatness. Ideally all frequencies in the pass-band should be passed with the same amplitude, that is, the pass-band should be flat. In certain types of filter the pass-band has ripples, that is, has regions of slightly increased or decreased response. Overall the ripples are usually less than 3 dB deep.

Phase response. Filters differ in the degree to which signals of different frequencies are delayed as they pass through the filter. Phase differences result in distortion of the signal.

Roll-off. This is the steepness of the transition region. The steeper the curve in this region, the more effectively the filter separates out the frequencies that are to be passed and those that are to be stopped.

Overshoot. When the input level changes abruptly (especially with squarewave or pulsed signals) the output may swing too far, oscillating for a short while before settling to the correct new level (Fig. 8.8). This type of distortion occurs also with a smoothly changing sinusoid, though its effects are less obvious.

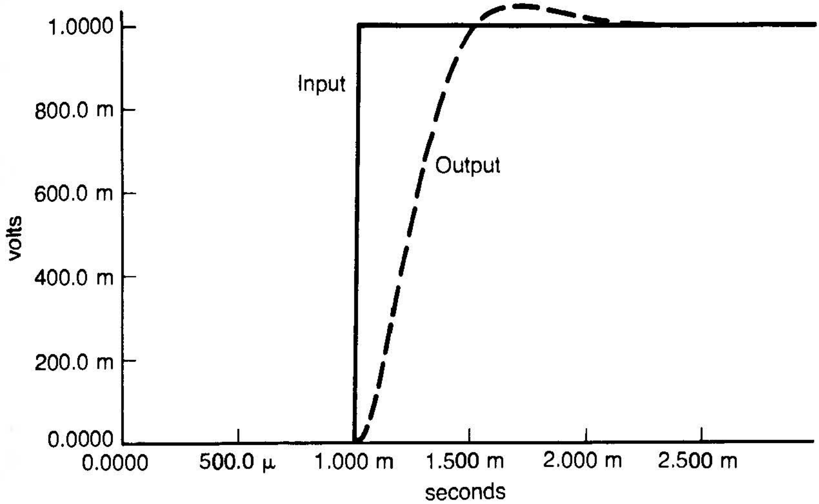


Figure 8.8 *The effect of a step input on a second-order Butterworth filter. The output rises relatively slowly, overshoots, and takes about 1.5 ms to settle*

The features of the most commonly used filter types are summarized in this table:

Type	*Pass-band flatness*	*Phase response*	*Roll-off*	*Overshoot*
Butterworth (flattest amplitude filter)	The flattest of all types, provided that high-precision components are used	Poor	Moderate	Reasonable
Chebyshev	Has ripples (Figs 8.10–8.11) and sharp knee	Poor	The steepest of all	Reasonable
Bessel	Sloping, with rounded knee	Least delay	Flattest of all	Smallest

As can be seen in the table, selecting a filter for a particular application is usually a matter of compromise. No type of filter is ideal in all respects. Within the Chebyshev group there is further compromise between ripple in the pass-band and steepness of roll-off. Chebyshev filters with small pass-band ripple have a relatively slow roll-off. To obtain steep roll-off you have to accept deep ripples in the pass-band.

The responses of different filter types are obtained by selecting appropriate values for the resistors and capacitors. The calculations required are complicated but there are published tables to make the task easier. Alternatively a circuit simulator or filter design package may be used with a computer.

The difference between the types of filters is the result of differing component values, not the way they are connected. As an example take the second-order filter of Fig. 8.4. This can be configured in many different ways by selecting values for the resistors or capacitors. Some tables assume that the resistors all have the same values and provide the user with suitable values for capacitors. But in practical filter building it is generally more convenient to have capacitors of equal value and select differently valued resistors. Given that the capacitors are all to be 100 nF, we obtain different responses by selecting the resistors. In Fig. 8.4 values of $R1$ and $R2$ are equal and are selected to obtain the required cut-off frequency. Values of $R3$ and $R4$ are selected to obtain the feedback necessary to obtain the required amount of damping. For a Butterworth filter, the tables give $R1 = R2 = 1.6\,\text{k}\Omega$. The damping factor is 1.414 which means that if $R4 = 47\,\text{k}\Omega$, $R3$ must be $27.5\,\text{k}\Omega$. The response is plotted in Fig. 8.9. It has the typical flat pass-band and rounded (but not over-rounded) knee. The −3 dB point is at 1 kHz as expected, and output falls to −40.2 dB at 10 kHz. Phase lag increases from 0° at low frequencies to −175° at 10 kHz. The filter has a gain of 1.59.

Next we look at the Chebyshev response. Figure 8.10 demonstrates the result of changing the value of $R1$ and $R2$ to $1.758\,\text{k}\Omega$ and at the same time reducing the damping factor to 0.89. Damping is reduced by changing $R3$ to $51.7\,\text{k}\Omega$. As with the Butterworth filter, the −3 dB point is close to 1 kHz, but the pass-band is not at full amplitude. In fact most of it is at −1.8 dB, for this is a Chebyshev filter with a 2 dB ripple. The ripple curves upward as frequency approaches 1 kHz. After this there is a fall, steeper than with the Butterworth filter, so that response reaches −43.7 dB at 10 kHz. The phase curve shows the lag increasing from 0° to

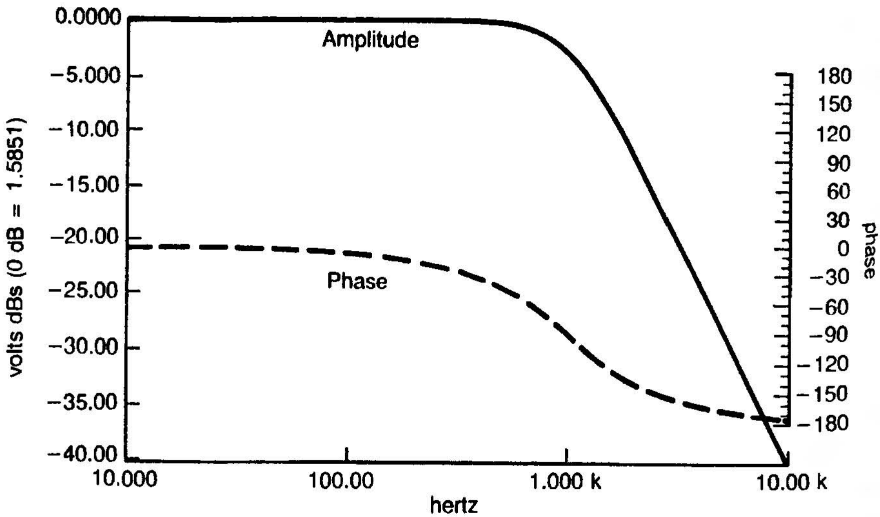


Figure 8.9 *A Butterworth filter has a very flat pass-band and a reasonably conspicuous knee*

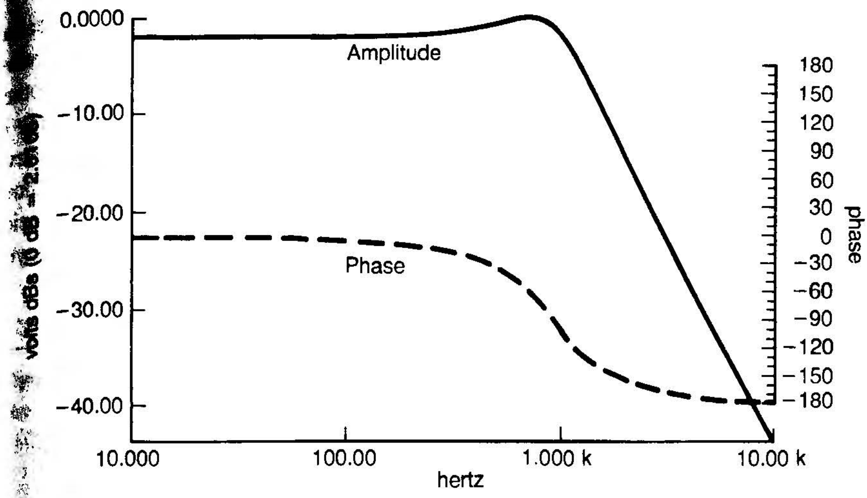


Figure 8.10 *A second-order Chebyshev filter has a ripple in the pass-band which serves to sharpen its knee. Phase response slopes steeply down in the higher part of the pass-band*

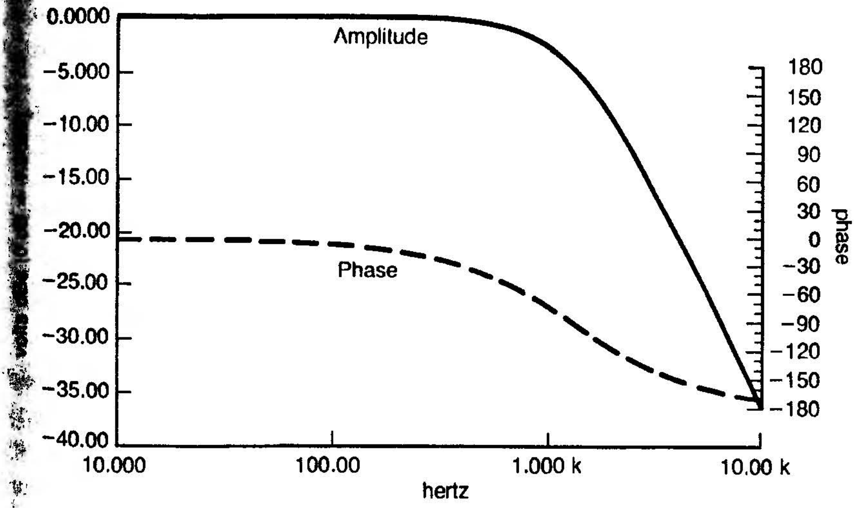


Figure 8.11 *A Bessel filter has a rounded knee, but its phase response is more evenly spread over the frequency range*

180°, which is not much more than in the Butterworth filter, but the fall is less smooth, there being a very sharp drop around the cut-off point which produces less distortion of signals in the 800 Hz to 2 kHz region.

Finally, the Bessel filter (Fig. 8.11) displays the rounded knee of the Butterworth filter, but its roll-off is appreciably less, amplitude falling only to −36 dB at

10 kHz. In this version of the filter, $R1 = R2 = 1.26\,\mathrm{k\Omega}$ and damping is increased to 1.73 by making $R3$ and $R4$ $27\,\mathrm{k\Omega}$ and $100\,\mathrm{k\Omega}$ respectively. The phase curve is the smoothest of the three. The overall increase in lag is less than in the other two curves, though not much less. The important feature is that the phase lag is spread as evenly as possible over the frequency range, in other words the curve should be an evenly sloping line, as straight as possible. The graph covers a thousandfold frequency range but, in any given application, the range of frequencies is usually more restricted. Signals of reasonably close frequencies are delayed by approximately the same amount so distortion will be minimal.

Other filter types

The Butterworth, Chebyshev and Bessel filters are the most commonly used but other types are sometimes needed. One of these the **elliptical** or **Cauer filter**. This is a relative of the Chebyshev filter and has a very steep roll-off. Like the Chebyshev filter, it has ripples in the pass-band. Also, because its transfer function has zeros, which the other types do not, it has ripples in the stop-band. This means that the low-pass elliptic filter may pass certain higher frequencies but these frequencies are too high to be significant in the applications for which the elliptic filter is used. Although the Cauer filter has the benefit of steep roll-off, its phase response is very irregular and it cannot be used in situations in which phase is important.

The name **all-pass** filter sounds like a contradiction of what filtering is about. All-pass filters have a relatively flat pass-band over a wide range of frequencies and their function is to selectively change the phase lag at different frequencies. Usually this is to correct for phase delays introduced by other filters or amplifiers, for which reason all-pass filters are also known as equalizers.

For steeper roll-off we must use filters of a higher order. A third-order filter is built by cascading a first-order filter (Fig. 8.3) and a second-order filter (Fig. 8.4). This has a roll-off of -18 dB per octave. The next filter in the series is a fourth-order filter built by cascading two second-order filters (Fig. 8.12). The output of this shows a steep roll-off (Fig. 8.13) of -24 dB per octave. The pass-band begins with an amplitude of -2 dB below the peak (this is a 2 dB ripple Chebyshev filter), and the filter has two ripples in the pass-band. In general, Chebyshev filters of order n have $n/2$ ripples. If the order is even, the amplitude at the low-frequency end of the pass-band is attenuated, but the response rises to a peak just before the cut-off frequency. If the order is odd, the low end is not attenuated and the odd half-ripple brings the response to a peak, again just before the cut-off frequency. The phase response of the Chebyshev fourth-order filter is very poor.

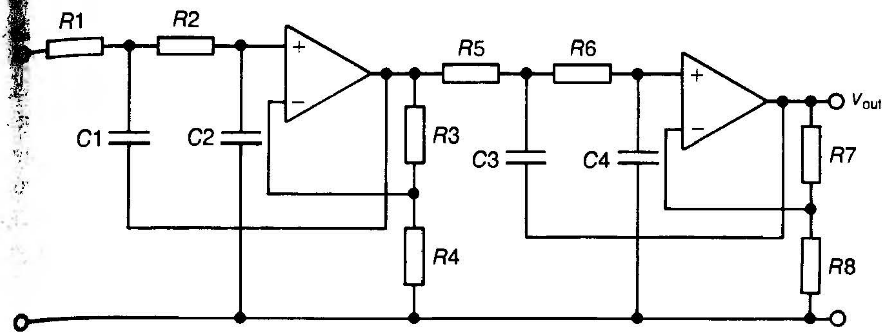


Figure 8.12 *A fourth-order active filter is built by cascading two second-order sections. Here both are low-pass sections to achieve steeper roll-off of −12 dB per octave*

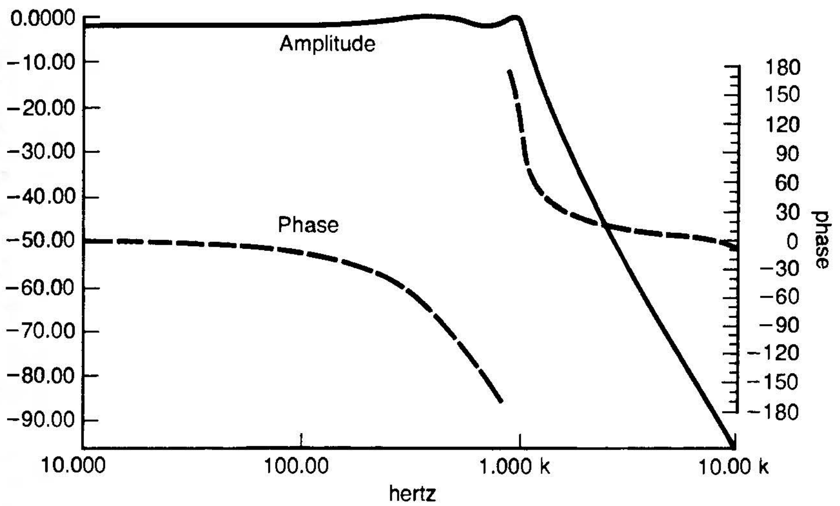


Figure 8.13 *The resistor and capacitor values in Fig. 8.12 have been selected to give this fourth-order Chebyshev response. Note the two ripples in the pass-band, the sharp knee and the steep roll-off*

Figure 8.13 shows the phase lag reaching −180° at the cut-off frequency. Then the curve jumps suddenly from −180° to +180°, but this is to be interpreted as its continuing to decrease until it reaches about −360° at 10 kHz.

nd-pass filters

way to produce an active band-pass filter is to cascade a low-pass active with a high-pass active filter. The cut-off frequency of the low-pass filter

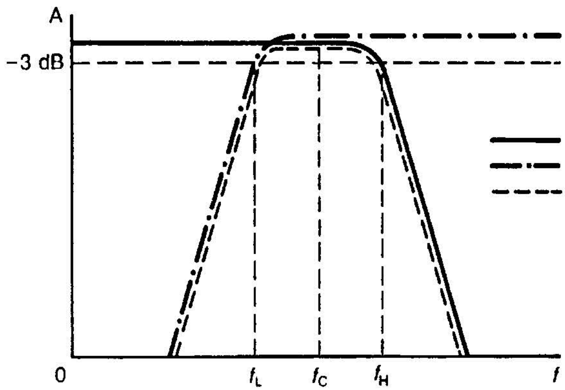


Figure 8.14 *Overlapping responses of a low-pass filter cascaded with a high-pass filter produce a band-pass response*

f_H is set to be higher than that of the high-pass filter f_L (Fig. 8.14). The width of the pass-band is the difference between the two cut-off frequencies, $f_H - f_L$. Note that we are making use of overlapping response curves so that the low-pass filter determines the high cut-off point f_H and the high-pass filter determines the low cut-off point f_L. Such filters are suitable if the pass-band is reasonably wide. As we move f_H and f_L closer together to obtain a narrow pass-band the region of overlap includes more and more of the knees, so that the amplitude at the centre frequency is attenuated. For a narrow pass-band we make use of the feedback feature of op amps.

A tuned active filter circuit (Fig. 8.15) has a frequency rejection circuit in the negative feedback loop. There are several networks that may be used but we have chosen a twin-T network, named after the configuration of the resistors and capacitors. One of the Ts consists of two resistors, value R and a capacitor value $2C$. This acts as a low-pass filter. The other T, consisting of two capacitors value C and a resistor value $R/2$, is a high-pass filter. These are connected in

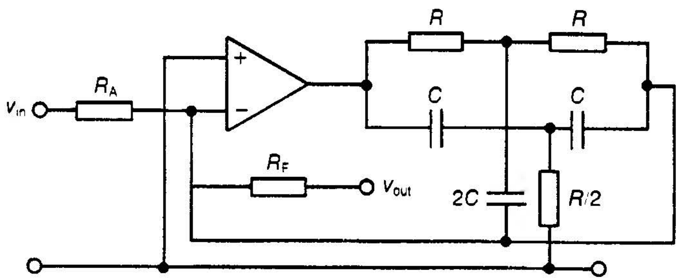


Figure 8.15 *Another way of constructing a band-pass filter is to use a twin-T frequency rejection network in the feedback loop of an operational amplifier*

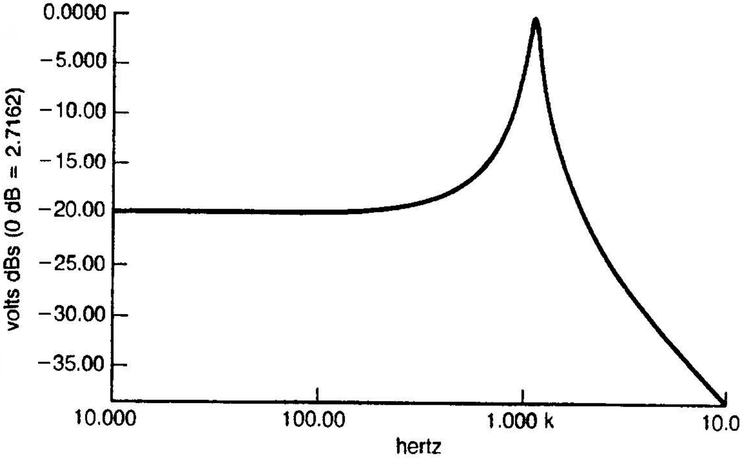


Figure 8.16 *The circuit of Fig. 8.15 is good for producing a very narrow pass-band as this Bode plot shows*

parallel so that one passes low frequencies, the other passes high frequencies and only intermediate frequencies are attenuated by the network. This is why it is called a frequency rejection network. The signal from the network is fed back to the negative input of the op amp. At low frequencies the signal is passed back inverted to the negative input and thus partly or wholly cancels out the signal. Similarly, at high frequencies, negative feedback cancels out the signal wholly or partly. But with intermediate frequencies the feedback is small and the signal coming from the op amp output is not attenuated. Thus we have a band-pass filter. The frequency response (Fig. 8.16) has a sharp peak at 1 kHz ($f = 1/(2\pi RC)$, where $R = 1.592\,\mathrm{k\Omega}$ and $C = 100\,\mathrm{nF}$). The stop-band below the peak is at −20 dB. Above the peak the response drops at −6 dB per octave. In this filter the signal goes to the inverting input, so the filtered signal is the inverse of the input signal. In other words there is a phase difference of 180° in addition to the phase changes associated with the capacitors.

The response shown in Fig. 8.16 is that obtained when $R_F = 20\,\mathrm{k\Omega}$. This feedback resistor is in parallel with the rejection network and acts to dampen the response. If R_F is omitted, the output peaks very sharply to a high level, as there is a pole when $f = 1$ kHz. The smaller the value of R_F the greater the damping and the broader the pass-band.

Bandwidth and another definition of *Q*

There are two −3 dB points in a band-pass filter (Fig. 8.14). The difference between these is the bandwidth:

$$\text{Bandwidth} = f_H - f_L$$

The central frequency of the pass-band is the geometric mean of the two −3 dB frequencies:

$$f_C = \sqrt{(f_H \times f_L)}$$

If $f_H = 1050$ Hz and $f_L = 950$ Hz, the bandwidth is 100 Hz and the central frequency is 998.7 Hz. An identical bandwidth when the central frequency is, say, 100 kHz represents a very precisely tuned filter with a narrow pass-band. By contrast, a bandwidth of 100 Hz represents a very broad pass-band if the central frequency is only 300 Hz. We need to quote the bandwidth relative to the central frequency:

$$\frac{\text{Bandwidth}}{f_C} = \text{fractional bandwidth}$$

The fractional bandwidths of the three filters described above are 0.1, 0.001 and 0.333 respectively.

The quality factor *Q* (p. 92) is a measure of the sharpness of response and, in the case of band-pass filters, of the width of the pass-band (bandwidth) relative to the central frequency. The greater the bandwidth, the lower the quality factor. This gives another definition for *Q*:

$$Q = \frac{1}{\text{fractional bandwidth}}$$

The values of *Q* for the three filters above are 10, 1000 and 3 respectively.

Another approach to band-pass filtering is shown in Fig. 8.17. This is known as a **multiple feedback** band-pass filter because there are two feedback loops. The loop through $C1$ is through a high-pass filter consisting of $C1$ and $R1$. High-frequency signals enter the circuit through $R1$, then pass easily through $C2$ to the op amp, they are inverted and then fed back through the $C1/R1$ high-pass filter, so that they cancel out. The loop through $R2$ is a low-pass filter consisting of $R2$ and $C2$. Any low frequency signals that manage to pass through $C2$ on their way to the op amp are inverted and fed back through the $R2/C2$ low-pass filter, to cancel out low-frequency signals. Only signals of intermediate frequency can pass through the circuit without attenuation. The advantage of this filter is that it is tuneable. The centre frequency can be adjusted by varying $R3$. In Fig. 8.18 the response curves are shown as $R3$ is swept from 100 Ω to 1 kΩ. Note that the centre frequency is changed but the amplitude of the output and the sharpness of the pass-band (*Q*) are unaltered. Actually the bandwidth appears to decrease as frequency increases, but this is the effect of plotting the response on a logarithmic scale.

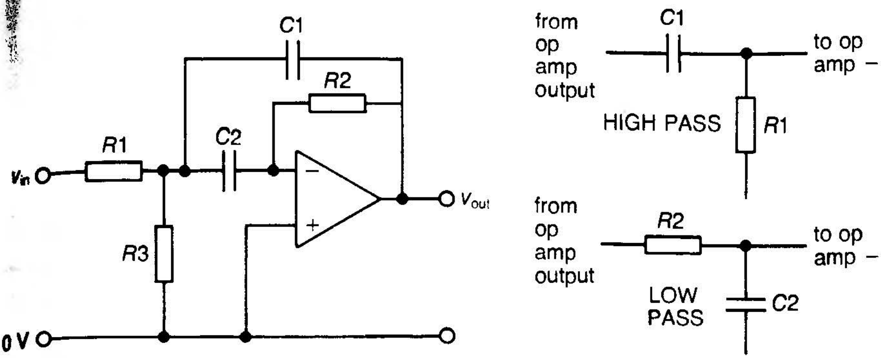


Figure 8.17 *A third way of obtaining a band-pass response is a multiple feedback circuit. The circuit contains both a high-pass and a low-pass filter, which provide negative feedback of high and low frequencies, so allowing intermediate frequencies to pass*

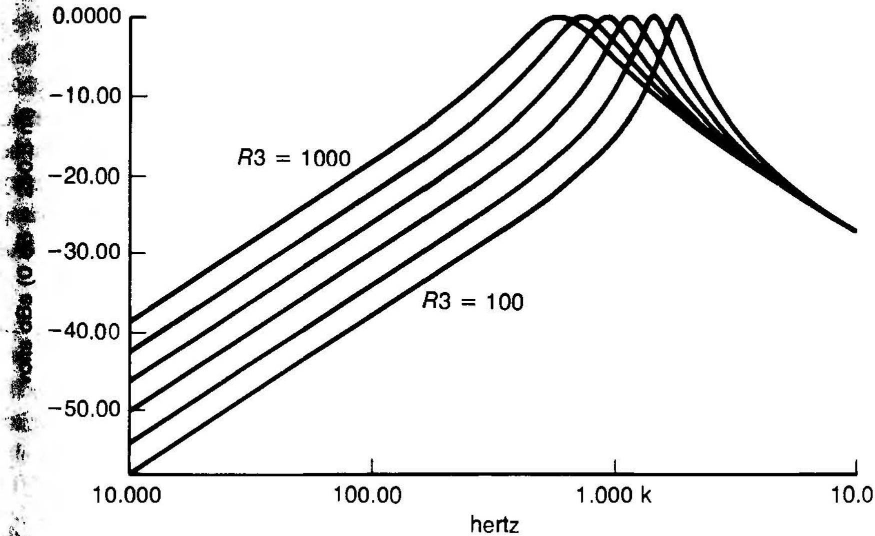


Figure 8.18 *The centre frequency of a multiple feedback band-pass filter (Fig. 8.17) may be adjusted by setting the value of R3*

Using the principles of frequency rejection or of multiple feedback, we can design band-pass filters of higher order to produce steeper roll-off. By choosing suitable resistor or capacitor values they may be designed to have any one of the conventional response curves such as Butterworth or Chebyshev. However, these introduce no new aspects of filter operation so we do not pursue this subject further.

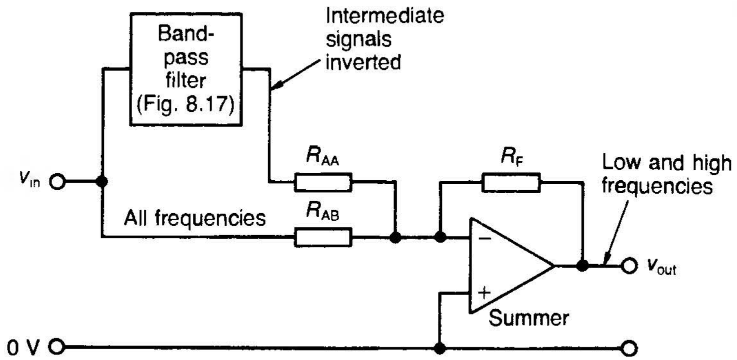


Figure 8.19 *One way of building a notch filter is to sum the signal with the inverted output from a band-pass filter*

Band-stop filters

Another and more descriptive name for these is **notch filters**. The frequency response curve of these filters is an inverted version of the band-pass curve. One way of obtaining the notch function is to take a band-pass filter (for example a multiple feedback filter such as that illustrated in Fig. 8.17) and mix its inverted output with the original signal. The result of this is to subtract the output (the passed band of frequencies) from the original, creating a notch in the frequency response curve. In Fig. 8.19 the subtraction is done by using a second op amp as a summer. The depth of the notch is determined by the relative values of R_{AA} and R_{AB}.

Test yourself

1. A first-order high-pass filter is made by exchanging the resistor and capacitor in Fig. 8.3. What is its transfer function?
2. What is meant by a VCVS filter?
3. For a given order of filter, which response type produces the steepest roll-off?
4. Which response type produces the flattest pass-band?

9 State variable filters and others

State variable filters require either three or four op amps and produce low-pass, high-pass and band-pass outputs. It is not so often that we require all these different responses simultaneously but, by combining them in various ways, we are able to obtain frequency responses that are not so readily obtainable with other circuits. The term 'state variable' refers to capacitor pds and (though they are not applicable to these filters) to inductor currents which satisfy the state equations derived from Kirchhoff's voltage and current laws. The filter action depends upon the accumulation of charge on capacitors and the pds which result from this.

A state variable filter with four op amps is pictured in Fig. 9.1. It consists of a summer, two integrators and an inverting amplifier. The latter acts as a damping subcircuit. The integrators produce an output which is:

$$\mathbf{v}_{\mathbf{out}} = \frac{-1}{RC}\int_0^t \mathbf{v}_{\mathbf{in}}\,\mathrm{d}t + k$$

where k is the initial pd across the capacitor. To see how the integrators function, refer to integrator 1 in Fig. 9.1. As current flows to the integrator, charge accumulates on the capacitor and the pd on plate X would rise. But, because the inverting input is a virtual ground (p. 112), this is not allowed to happen. The output falls, taking the potential at plate Y down, so keeping the potential of plate X at zero. Thus, given a constant input voltage and therefore a constant flow of current, the output from the op amp falls steadily. This action is modified if the input is varying but at any stage, v_{out} is falling according to the value of v_{in}. Acting in this way as an integrator, the op amp circuit has the same effect as a low-pass filter. This is because high-frequency sinusoids have little effect on output. The highs and lows of v_{in} alternate many times, perhaps many thousands of times, per second. Individual rises and falls of v_{in} do not have time to take effect. By contrast, at low frequencies there is more time for charge to accumulate as input pd rises, and more time for charge to diminish as input pd falls. So the integrators act as cascaded low-pass filters, producing a steep roll-off and giving a $-3\,\mathrm{dB}$ point at $f = 1/(2\pi RC)$. Since there are two filters, the inverting

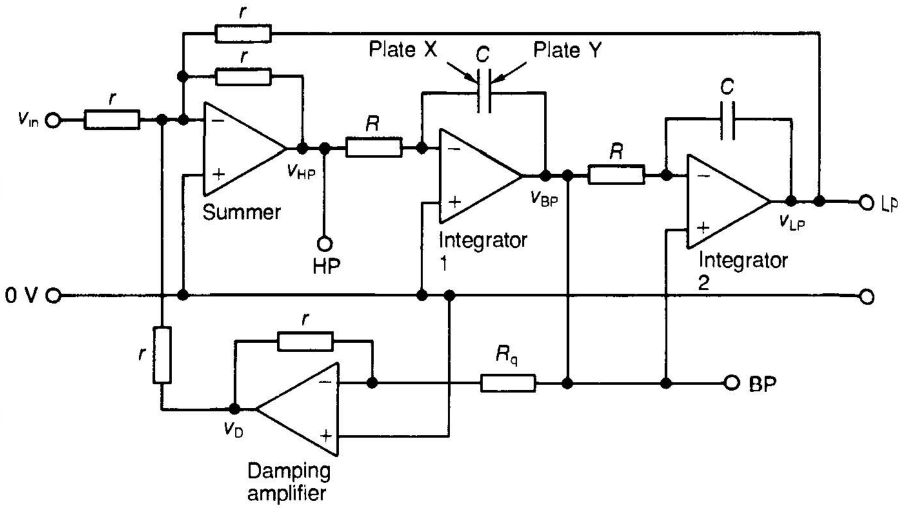


Figure 9.1 *A state variable filter is built from 3 or 4 op amps. The 4-amplifier version is illustrated above. The 3-amplifier version has a resistor network instead of the damping amplifier*

actions of each cancel out and the output is in phase with input, apart from phase differences due to capacitance.

Considering the integrator as an inverting amplifier, but with a capacitor in the feedback loop instead of a resistor, we may insert the reactance of the capacitor $1/j\omega C$ in place of R_F in the gain equation (or transfer function) of p. 112:

$$\frac{\mathbf{v_{out}}}{\mathbf{v_{in}}} = \frac{-R_F}{R_A} = \frac{-1}{j\omega R_A C}$$

Applying this function to integrator 2, R replaces R_A, $\mathbf{v_{BP}}$ replaces $\mathbf{v_{in}}$, and $\mathbf{v_{LP}}$ replaces $\mathbf{v_{out}}$, to conform with the symbols used in Fig. 9.1. The transfer function of the second integrator is:

$$\mathbf{v_{LP}} = \frac{-\mathbf{v_{BP}}}{j\omega RC}$$

This is the transfer function of a low-pass filter (compare with p. 71). Replacing $j\omega$ by s, and making $R = 1$ and $C = 1$ to simplify the expression:

$$\mathbf{v_{LP}} = \frac{-\mathbf{v_{BP}}}{s}$$

Similarly, the transfer function of the first integrator is:

$$\mathbf{v_{BP}} = \frac{-\mathbf{v_{HP}}}{s}$$

r the inverting amplifier:

$$v_D = -Dv_{BP}$$

where D is the damping factor which equals r/R_q. At the summer:

$$\begin{aligned} v_{HP} &= -(v_{LP} + v_D + v_{in}) \\ &= -\left(\frac{-v_{BP}}{s} - Dv_{BP} + v_{in}\right) \\ &= -\left(\frac{v_{HP}}{s^2} + \frac{Dv_{HP}}{s} + v_{in}\right) \end{aligned}$$

At each stage we have substituted from one of the functions listed above.

$$\Rightarrow \qquad v_{HP}\left(1 + \frac{1}{s^2} + \frac{D}{s}\right) = -v_{in}$$

$$v_{HP}\left(\frac{s^2 + 1 + Ds}{s^2}\right) = -v_{in}$$

$$\Rightarrow \qquad \frac{v_{HP}}{v_{in}} = \frac{-s^2}{s^2 + Ds + 1}$$

This is the transfer function of a second-order high-pass inverting filter. At low frequencies the gain is low and falls to zero at $s = 0$, because there is a zero (p. 105) at that point. At high frequencies the function gradually approaches the value -1. By similar reasoning, it can also be shown that:

$$\frac{v_{BP}}{v_{in}} = \frac{s}{s^2 + Ds + 1}$$

The value falls to zero when $s = 0$ and also falls to zero as s approaches infinity. The function is positive because the inverted signal produced by the summer is reinverted by the first integrator.

It can also be shown that:

$$\frac{v_{LP}}{v_{in}} = \frac{-1}{s^2 + Ds + 1}$$

The value falls as s increases, but is 1 when s is zero. The function is negative because there are three inversion stages between v_{in} and v_{LP}.

Figure 9.2 shows the frequency responses at the three outputs of the filter when R and C are chosen to give 1 kHz as the -3 dB point. The low-pass, band-pass and high-pass responses are clearly seen and symmetrical. In this display, $r = 10\,\text{k}\Omega$ and $R_q = 8\,\text{k}\Omega$, so $D = 0.8$ or $q = 1/0.8 = 1.25$. This is quite a low value for q so the band-pass response is relatively broad. In Fig. 9.3 we have added the phase plots to the graph. It can be seen that although phase varies with frequency as is normal in a filter, the three outputs remain 90° apart from each other over the whole frequency range.

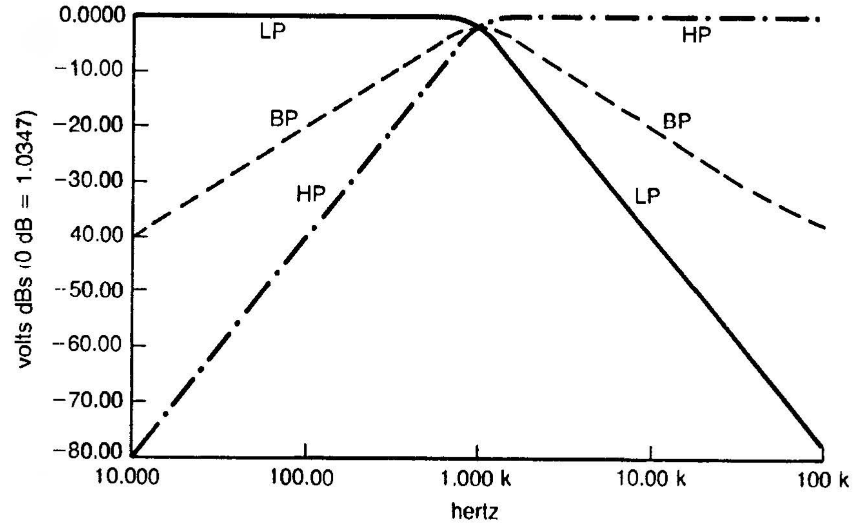


Figure 9.2 *The three outputs of a state variable filter provide low-pass, band-pass and high-pass signals, in this case with a centre or cut-off frequency of 1 kHz*

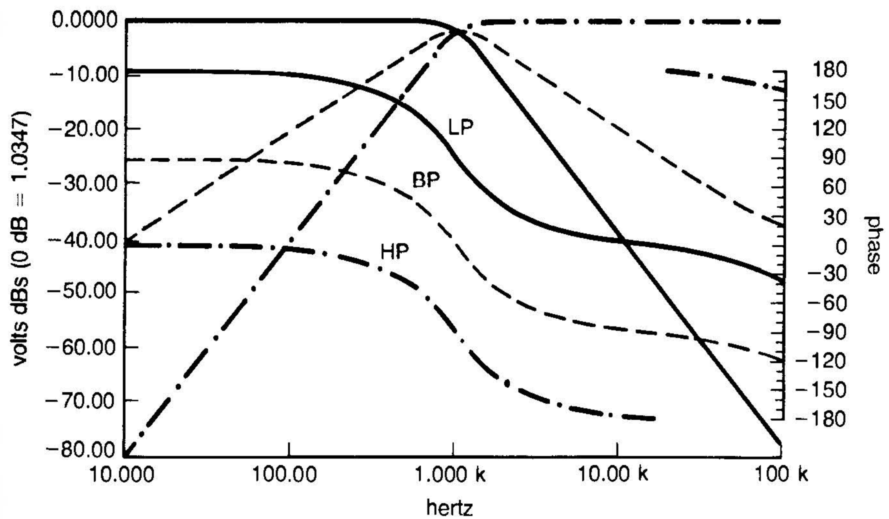


Figure 9.3 *This shows the phase response of the three outputs of a state variable filter, each 90° apart at all frequencies*

Sinusoids and the integrators

The standard tables of integrals tell us that $\int \sin t\,dt = -\cos t$, and that $\int \cos t\,dt = \sin t$. In this discussion we make $R = 1$ and $C = 1$ so that the gain at each stage is -1, and we can follow phase changes more clearly. If the input to integrator 1 of a state variable filter is a sinusoid:

$$\mathbf{v}_{\mathbf{in}} = \sin t$$

$$\mathbf{v}_{\mathbf{out}} = -\int \sin t\,dt = \cos t$$

$$= \sin(t + 90^\circ)$$

The output signal ($\mathbf{v}_{\mathbf{BP}}$) leads the input ($\mathbf{v}_{\mathbf{HP}}$) by 90°. At integrator 2:

$$\mathbf{v}_{\mathbf{in}} = \cos t$$

$$\mathbf{v}_{\mathbf{out}} = -\int \cos t\,dt = -\sin t$$

$$= \cos(t + 90^\circ)$$

The output of integrator 2 ($\mathbf{v}_{\mathbf{LP}}$) leads the input ($\mathbf{v}_{\mathbf{BP}}$) by 90°. As a result of this $\mathbf{v}_{\mathbf{HP}}$ and $\mathbf{v}_{\mathbf{LP}}$ are 180° out of phase. These results can be seen in Fig. 9.3.

One of the features of a state variable filter is that very high values of q are attainable. In effect the circuit acts as an oscillator unless damping is applied. What happens is that the two integrators together invert the signal coming from the summer (see box). This signal is then fed back to the summer, which inverts it again. The signal is now in phase with itself and reinforces itself. Signal amplitude increases steadily until the circuit goes into oscillation at a frequency determined by R and C.

When being used as a filter, the band-pass signal from the first integrator is 90° out of phase with that from the other two outputs. This is fed back through the inverting amplifier which produces a total phase shift of 270°. This signal is mixed in at the summer and cancels out a proportion of the signal being fed back from the second integrator. It acts as a damping signal. The amount of damping, and hence the value of Q, depends on the ratio r/R_Q.

The state variable filter is the basis of a high-Q band-reject filter. This is better described as a notch filter because the bandwidth is usually very narrow. As Fig. 9.3 demonstrates, the signals of the low-pass and high-pass outputs are of equal magnitude but exactly 180° out of phase at the -3 dB point. If these two signals are fed to an op amp summer (a fifth op amp) through resistors of equal value, they cancel each other out and there is zero output at this frequency. As shown in Fig. 9.4, this produces the response of a notch filter. If the circuit is adjusted to give high q, this can be a response with a very narrow notch.

The outputs of a state variable filter can be summed in various combinations with their relative contributions depending upon the weighting given by the resistors of the op amp summer. A good example is depicted in Fig. 9.5 in which the

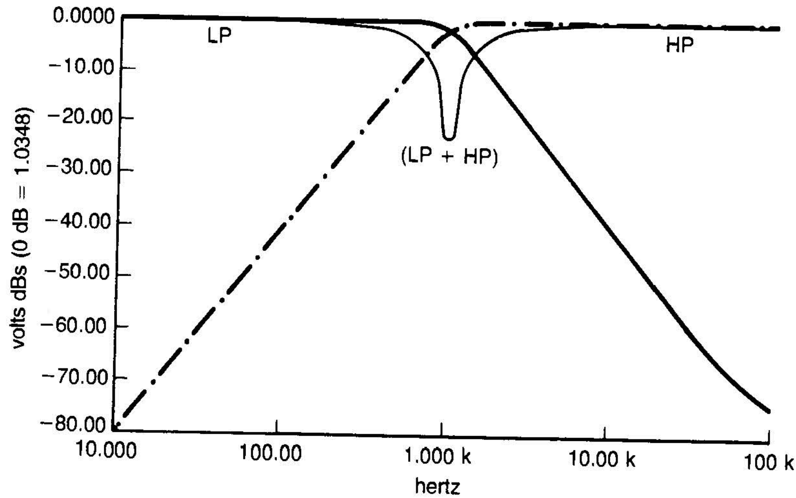


Figure 9.4 *Summing the high-pass and low-pass outputs of a state variable filter in equal amounts yields a notch response*

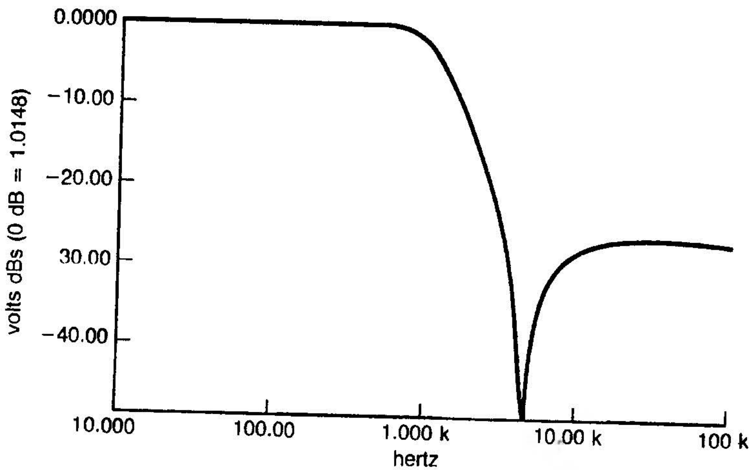


Figure 9.5 *By summing the high-pass and low-pass outputs of a state variable filter we produce a low-pass response with very steep roll-off*

low-pass and high-pass outputs are summed in the ratio 20LP to 1HP. The result is a 'lop-sided' notch filter with the notch reaching down almost to −60 dB. The effect is also that of a low-pass filter with extremely rapid roll-off, about −30 dB per octave. The fact that response increases slightly above the notch may or may not be significant, depending on the application. Another course is to sum all three outputs in equal amounts, to produce an all-pass filter (p. 122).

Frequency-dependent negative resistor

This is not a single resistor but a subcircuit which behaves as if it has negative resistance. The negative resistance depends on signal frequency. It is one of a category of circuits known as **gyrators**. In Fig. 9.6(a) the op amps are being used as subtractors and there is both feedback and feed-forward in the circuit. As a result the action of the circuit is extremely difficult to describe. The overall effect is that when an alternating pd is applied to the terminal v, the current flowing into the terminal is 180° out of phase with the pd. When the pd is increasing the current is decreasing and the other way about. Because of the capacitors, the effect is dependent on frequency. This means that the FDNR has applications in filter

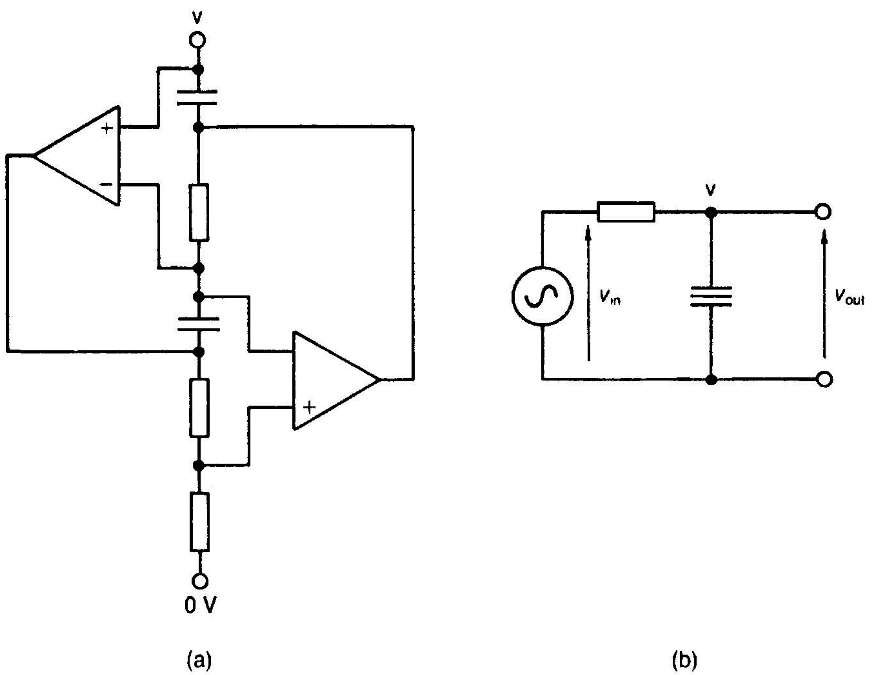


Figure 9.6 *This op amp gyrator circuit is known alternatively as a frequency-dependent negative resistor, or as a super-capacitor*

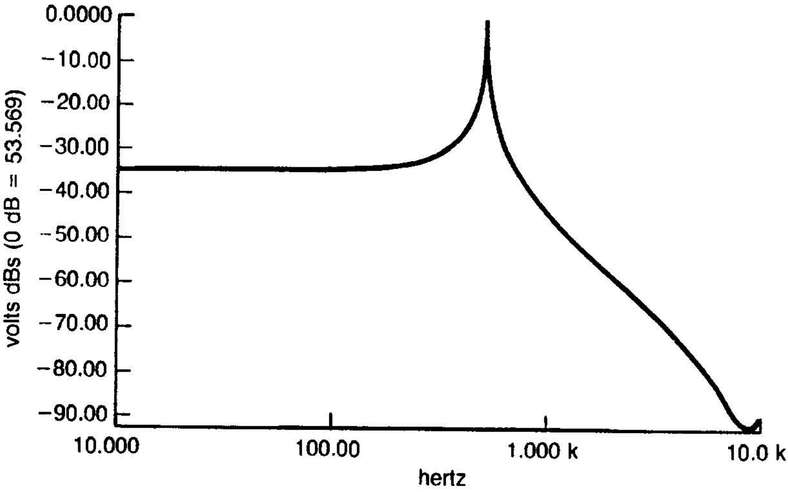


Figure 9.7 *The response of a simple low-pass filter based on a super-capacitor (Fig. 9.6) shows a sharp peak at the cut-off frequency and an initially steep roll-off*

circuits. In Fig. 9.6(b), simply to demonstrate its action, we have connected the FDNR into a low-pass filter circuit in place of the capacitor, and it is represented by the capacitor-like symbol. This has the appearance of two capacitors close together and is symbolic of the fact that the FDNR behaves also as a **super-capacitor.** Its frequency-dependent behaviour gives it capacitor-like properties but they exceed those of capacitors in two ways. The reactance of a capacitor is inversely proportional to frequency ($X_C = 1/\omega C$) but the reactance of the FDNR is inversely proportional to the square of the frequency. Consequently, halving the frequency makes the reactance four times as great. Also the phase lead of an FDNR is 180° instead of only 90° as with a normal capacitor.

A simulation run on the low-pass filter of Fig. 9.6(b) yields Fig. 9.7. $\mathbf{v}_{\mathbf{out}}$ peaks very sharply at 500 Hz, indicating that the circuit has very high q. This is a property that can be made good use of in certain filtering applications. There is no phase change below the peak point and a 180° phase lead at higher frequencies. This response is obtained with a certain combination of resistors and capacitors (resistors all 10 kΩ, capacitors all 100 nF) but is different with other combinations. FDNRs have many and varied applications as units for filter building.

Voltage-controlled filters

The most usual way of controlling the cut-off frequency of a filter is to use a variable resistor. It is also feasible to have a number of capacitors which may be switched into the circuit to provide for different frequency ranges. It may also be useful to be able to control a filter directly by electronic means. A device which makes this possible is the **transconductance amplifier**, which is available in integrated circuit form. The output of such an amplifier is a current which is proportional in size to the pd between its two inputs. It has the same characteristic as a resistor, for the current through a resistor is likewise proportional to the pd between its two ends. The transconductance of the amplifier varies according to the potential at the control input. We may think of the transconductance amplifier as a voltage-controlled resistor.

Transconductance amplifiers may be used in most types of filter, replacing the resistor upon which the cut-off frequency depends. Figure 9.8 illustrates a voltage-controlled state variable filter. There are two transconductance amplifiers replacing the two resistors labelled *R* in Fig. 9.1. Taking the link between the summer and the first integrator as an example, the output from the summer is fed to a voltage divider RA/RB. The negative input of the transconductance amplifier is held at 0 V but the positive terminal is connected to the junction between RA and RB. The alternating signal across the terminals of the amplifier results in an alternating current from its output. The amplitude of this signal depends on the voltage being fed to its control input. The alternating current flows on toward the capacitor of the integrator and the integrator behaves just as if it is receiving current directly from the summer as it is in Fig. 9.1. The difference is that the amount of current is influenced by the control voltage. In this circuit both transconductance amplifiers are controlled from the same voltage so both integrators have the same frequency settings.

Another type of voltage-controlled filter is described in the next section.

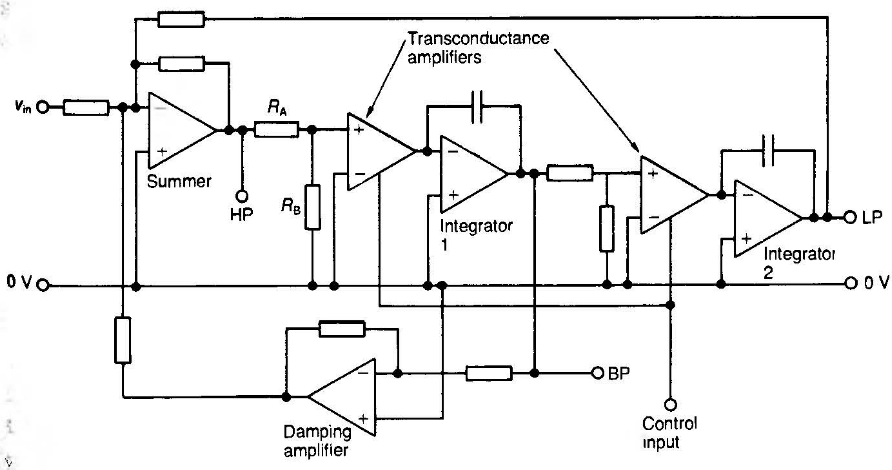


Figure 9.8 *Including a pair of transconductance amplifiers in a state variable filter makes it possible to control the cut-off point by applying a suitable voltage*

Switched capacitor filter

This uses an entirely different technique for filtering. The filter consists of two capacitors, two switches and an operational amplifier connected as a unity gain buffer (Fig. 9.9). Ordinary mechanical switches could not operate rapidly or reliably enough for service in this filter. Instead, we use analogue switches or transmission gates, which are built from MOSFET transistors (Fig. 9.10) and are under the control of a clock circuit. The clock produces a squarewave output at high frequency, usually several hundred kilohertz or even a few megahertz. When the clock output is at logical high, the switch is 'on' and acts as a resistor of about 100 ohms. When the clock output is at logical low the resistance between input and output is very high and the switch is off. The circuit in Fig. 9.9 is connected so that one switch receives the clock output direct and the other switch receives its inverse. When one switch is on, the other is off. It is obvious from the foregoing description that the switched capacitor circuit is implemented on silicon as an integrated circuit, including the capacitors. This is no problem as the capacitance required is extremely small, usually less than 1 pF.

To outline how the filter operates, we begin at the instant when $S1$ turns on and $S2$ turns off. Current flows from the input, through $S1$, and charges $C1$. Because $C1$ has a very small capacitance there is ample time while $S1$ is on for $C1$ to charge to the full value of the input voltage, v_{in}. Also, since the clock frequency is about 100 times greater than the maximum frequency that the filter is required to work with, there is insufficient time for v_{in} to change significantly while charging takes place. Next $S1$ is turned off and $S2$ is turned on. The charge on $C1$ is now shared with that already on $C2$ from the previous cycle. Since the capacitors are fabricated with high precision and are close together on the same chip, they are practically equal in capacitance and share the charge equally. A current flows either way through $S2$ until the potential across them is the mean of the charges prior to $S2$ being turned on. Then $S2$ is turned off and $C2$ is

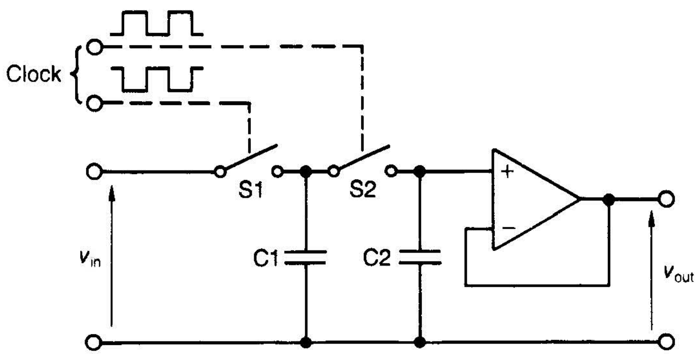


Figure 9.9 *The cut-off frequency of a switched-capacitor filter is determined by the clock rate*

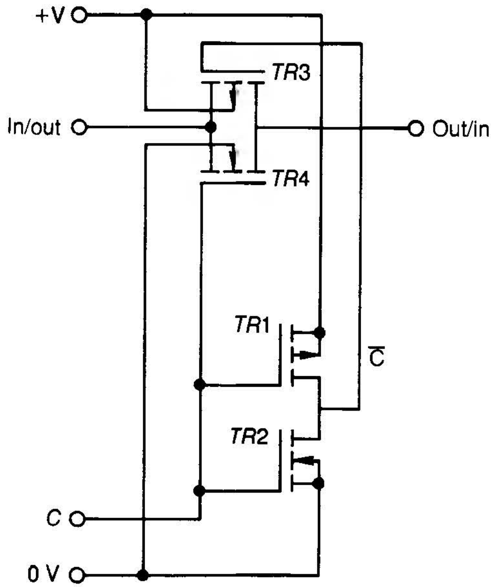


Figure 9.10 *An analogue switch, as used in a switched-capacitor filter (Fig. 9.9). TR_1 and TR_2 are a complementary pair of MOSFETs used to invert the control voltage C to produce $\overline{C}$. Together, C and $\overline{C}$ control another complementary pair of MOSFETs, TR_3 and TR_4. When C is high (and $\overline{C}$ low), TR_3 and TR_4 conduct signals in either direction. When C is low (and $\overline{C}$ high) the transistors are not conductive*

isolated while $C1$ is being charged to the new value of v_{in}. The pd across $C2$ is detected by the non-inverting input of the op amp and its output changes to the same level so that the potential at both inputs is the same. The op amp is acting as a unity gain voltage follower. The purpose of using an op amp at this stage is that it has very high input impedance so it removes virtually no charge from $C2$ yet provides sufficient current to drive any circuit connected to its output.

The overall effect of one switching cycle is to average the present value of v_{in} with its previous value. For a constant (DC) input, the output eventually reaches the input value. With repeated samplings of a waveform, the output varies with the same frequency as the waveform but with lesser amplitude. Both of these responses correspond to the action of a low-pass filter. The action is most easily simulated by the BASIC program listed in Appendix A (p. 155). Given $D = 20$, so that the input alternates between +10 and −10 at each sampling, the output alternates between +3.333 and −3.333. In other words the gain of the 'circuit' is 0.3333. This is for a high-frequency signal. For a lower frequency, with $D = 5$ so that the input signal is 10, 5, 0, −5, −10, −5, 0, 5, and back to 10, the output is a sequence of 8 values ranging from −5.59 to +5.59. The output signal has the same frequency as the input but the gain is now 0.559. As D is made smaller again to simulate reducing frequency, the output amplitude increases, as might be expected with a low-pass filter.

The action of this filter is affected by the clock rate. If the clock rate is made faster there is not sufficient time for $C1$ to charge fully, or for the charge to equilibrate between $C1$ and $C2$. This has the effect of increasing the resistance of the filter and affects its response. We have the equivalent of a voltage-controlled filter.

Another form of switched capacitance filter is a switched capacitor integrator. A single chip may have on it all that is needed to build a filter of the state variable type, including the clock circuit. By connecting a few external capacitors for timing and by joining the switched capacitors in various ways by external connections, it is possible to build second-order filters with any required output (high-pass, low-pass, band-pass, band-reject, and all-pass) with any required response (Butterworth, Chebyshev, Bessel and others). The cut-off frequency is determined by the clock rate which may be voltage controlled. Typically the clock rate is 100 times the cut-off frequency so that, although the switching action introduces a signal at clock rate, this does not corrupt the output.

Summary

The active filter circuits described in this chapter are:

State variable filters, which have high-pass, band-pass and low-pass outputs, and can be summed in various ways using an additional op amp to produce a wide range of responses. Filters of high Q are realizable.

Frequency-controlled negative resistors, also known as super-capacitors, in which reactance varies as the square of the frequency. They provide a building block for high-q filters.

Voltage-controlled filters based on transconductance amplifiers can be tuned electronically.

Switched capacitor filters can be configured in many ways. Having on-chip capacitors of very low value and small size, they are less bulky than conventional active filters. Their cut-off or centre frequency is externally controlled by setting the clock rate.

10
Digital filters

Digital signalling has become dominant in more and more fields of electronics including audio, radio and TV. It offers the advantages of reproducibility, robust circuitry, relative freedom from interference and noise, dependable reconstitution of degraded signals and the high-speed transmission of data of all kinds (including audio and video data) over long distances. Without digital signal transmission, the Internet would be impracticable. There is also the advantage that digital signal processing makes it possible to modify, compress and, in general, process signals in ways that are not possible with analogue circuits.

Figure 10.1 illustrates the difference between analogue and digital signals. One is a smoothly varying voltage or current, the other is a sequence of binary values consisting of a given number of bits. In Fig. 10.1 there are only 3 bits, to keep the diagram simple, and they represent the 8 values from 0 to 7, but digital signals are commonly expressed in 8, 12, 16 or even more bits. Before a signal can be processed digitally it must be in digital form. The analogue signal is sampled at regular intervals and each sample in converted into its digital equivalent by an analogue to digital converter, or ADC.

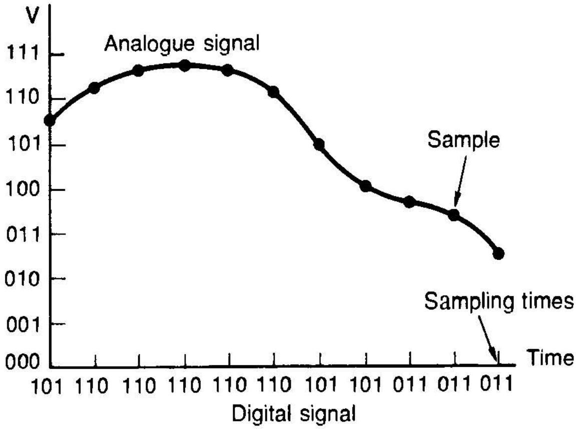


Figure 10.1 *An analogue signal varies smoothly with time. In analogue to digital conversion, samples are taken at regular intervals and converted into binary numbers*

Analogue to digital converters

It is clear from Fig. 10.1 that sampling of the analogue signal must be frequent, in order to follow the changes in the signal level, and that samples must contain sufficient bits to accurately represent the signal level at any instant. There are several different types of ADC of which the **flash converter** is one of the fastest. It works on the principle shown in Fig. 10.2. A standard reference potential is applied to one end of a chain of resistors. Tapping the chain between adjacent resistors yields a sequence of potentials against which the input potential is compared by a set of comparators. The outputs of the comparators are decoded

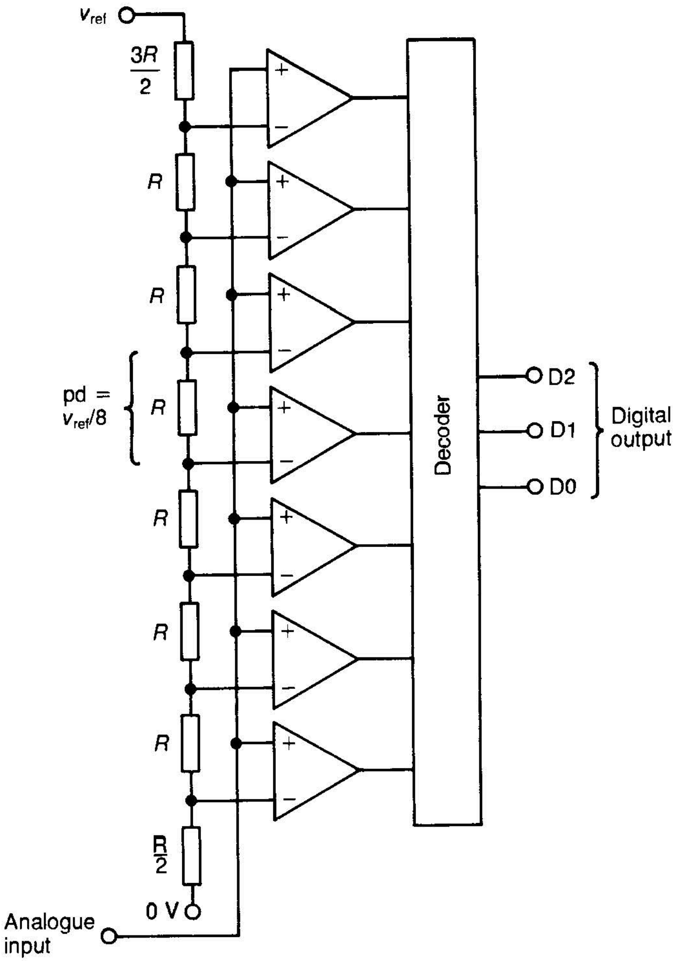


Figure 10.2 *A flash ADC operates by comparing the input voltage against a sequence of reference voltages produced by a resistor chain*

by logic circuits to provide a range of digital values from 000 (for a 3-bit ADC as illustrated) to 111. The exact value of R does not matter but the ratios between resistor values must be exact. Since the resistors are all produced on the same chip, this is relatively easy to achieve. Conversion time is very short, being limited to the time required for the comparators to settle and the propagation delay of the decoder logic. But high conversion speed is bought at the price of circuit complexity. For n-bit conversion the flash ADC requires 2^n resistors and $2^n - 1$ comparators. Thus an 8-bit ADC requires 256 resistors and 255 comparators. A 12-bit converter would need 4096 resistors and 4095 comparators, so 8 bits are the limit for a practicable flash converter of reasonable price. Other types of ADC include the **successive approximation** converter, which is reasonably fast and can provide a 16-bit output, and the **dual slope integration** type which is slower but has high precision.

Whatever type of ADC is used, the output potential at any given instant is necessarily restricted to one of 2^n values. For example, in a 8-bit converter which is set to cover a range of 0 V to 2.56 V, adjacent values are 2.56/256 = 0.01 V apart. As the input ramps up smoothly, the output increases by steps of 0.01 V. The step-like nature of the output introduces a high-frequency component into the signal, known as quantization noise. The higher the rate of sampling, the higher the frequency of the quantization noise and the easier it is to filter this out at a later stage.

Figure 10.3 demonstrates another aspect of conversion which must always be considered. Here is shown a sequence of 10 samples taken from an analogue signal. It appears to be a sinusoid with a period of approximately nine sampling periods. Figure 10.4 reveals that these samples do not come from such a sinusoid but from one which has a much higher frequency. The sampling period is very slightly longer than the period of the signal so the signal is sampled at a slightly later instant in each successive cycle. The effect is similar to that sometimes seen

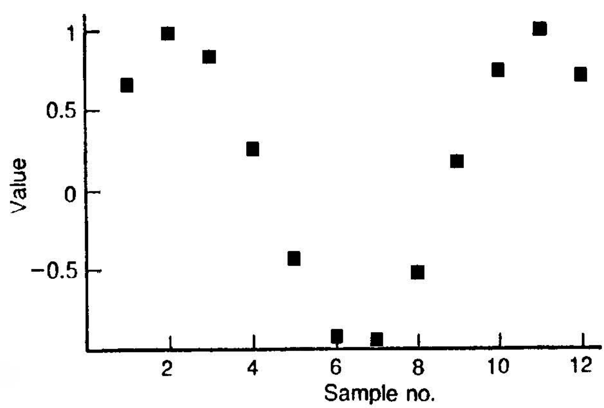


Figure 10.3 *These samples appear to be taken from a sinusoid, with about 9 samplings per cycle, but see Fig. 10.4*

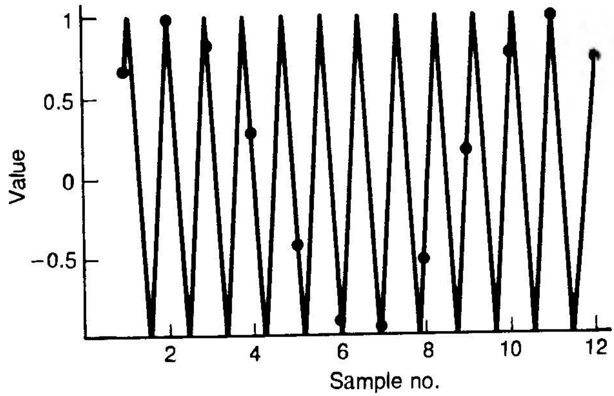


Figure 10.4 *The original signal is actually sampled slightly less than once during each cycle. Its frequency is higher than suggested by Fig. 10.3*

in the cinema or on TV when a spoked carriage wheel or the rotor of a helicopter turns at such a rate that successive samples (frames of the film) catch the wheel at slightly earlier or later positions. Between one frame and the next a spoke turns to almost the same position formerly occupied by an adjacent spoke. The result is to make the wheel apparently turn much more slowly or even to be turning backward. The effect is known as **aliasing** and gives a completely false signal. The only way to avoid this is to increase the sampling rate. It has been shown that aliasing does not occur if the sampling rate is double the highest frequency that is of interest. This is known as the **Nyquist frequency**. In audio circuits it is assumed that the highest frequency that the human ear can detect is 20 kHz. The corresponding Nyquist frequency is 40 kHz. This is the minimum sampling frequency adopted for high-quality audio circuits. In practice, the basic sampling frequency for compact discs is standardized at 44.1 kHz, with 16-bit samples. Higher rates are often used and this is known as **oversampling**. Sampling at double or four times the Nyquist frequency means that quantization noise is of such a high frequency that it is readily filtered out of the processed signal. In addition, the noise power is spread over double or four times the frequency range so that a higher proportion of it is removed with filtering. For video signals the standard sampling rate is 13.5 MHz, with 8-bit samples.

Sigma delta ADC

This is a more recent development in ADCs which takes an approach contrary to that of the conventional one. Instead of trying to improve conversion by increasing the number of bits to 16 and above, we increase the sampling rate. At the same time we reduce the sample size to one bit. This means that the conversion circuitry is much simpler, for only one comparator is required. The function of this is to compare the present sample with the previous sample, to detect if the signal level has increased or decreased between samples. This is the reason for the term 'delta', where delta is the Greek letter 'δ', standing for the difference between one sample and the next. With a high sampling rate the difference between one sample and the next is very small and may be represented by a single bit, either 0 for 'less than' or 1 for 'more than'. The principle of this is illustrated in Fig. 10.5. As the analogue input rises and falls, the digital output alternates between high (1) and low (0). The proportion of 0s to 1s in each segment of the digital signal corresponds with the value of the analogue signal. If this method is to give an accurate conversion, the sampling rate needs to be much greater with respect to the frequency of the analogue signal than is shown in Fig. 10.5.

A 1-bit signal can easily be converted back to an analogue signal. All that is needed is an integrator (see Fig. 9.1) followed by a low-pass filter. For as long as the digital signal is high, charge accumulates on the capacitor and the output level falls. For as long as the digital signal is low, charge is removed from the capacitor

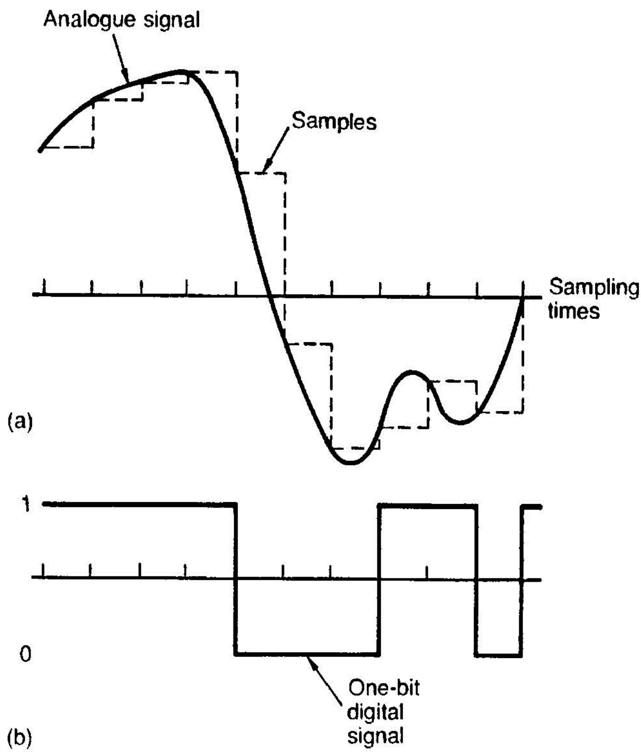


Figure 10.5 *Sampling at very high speed using a 1-bit quantizer produces a bit stream signal*

and the output rises. In this way we obtain an inverted replica of the original analogue signal. The low-pass filter is needed to smooth off quantization noise.

Figure 10.6 is a block diagram of a sigma delta converter. The analogue input is summed (this is the 'sigma' part of the process) with the feedback of the output of the circuit. As the + and − signs show, the operation is one of subtraction to detect changes in input level. This is passed to an integrator which is actually the integrator referred to in the previous paragraph. It is more convenient to place it here, and makes no difference to the operation of the system as a whole. The 1-bit quantizer produces a series of 1s and 0s in rapid succession. This is referred to as a **bit stream.**

The bit stream from the sigma delta ADC is fed to a digital **decimation filter**. The function of this is to convert the bit stream into a sequence of n-bit digital values at lower frequency. For example, suppose that a segment of the bit stream is:

1 0 1 1 0 0 0 0 1 1 1 0 1 0 1 1 0 1 1 1 0 0 0 1 1 0 0 0 1 0 1 0 1 1 1

This stream of 35 bits might be broken by the decimator into five groups of 7 bits each:

1 0 1 1 0 0 0   0 1 1 1 0 1 0   1 1 0 1 1 1 0   0 0 1 1 0 0 0   1 0 1 0 1 1 1

Each group is analysed to find which level, 1 or 0, is in the majority and is replaced by the majority bit:

3 ones, 4 zeros $\Rightarrow 0$ 4 ones, 3 zeros $\Rightarrow 1$ 5 ones, 2 zeros $\Rightarrow 1$ 2 ones, 5 zeros $\Rightarrow 0$

5 ones, 2 zeros $\Rightarrow 1$

This yields a 5-bit digital value 01101, which may then be passed to a digital filter.

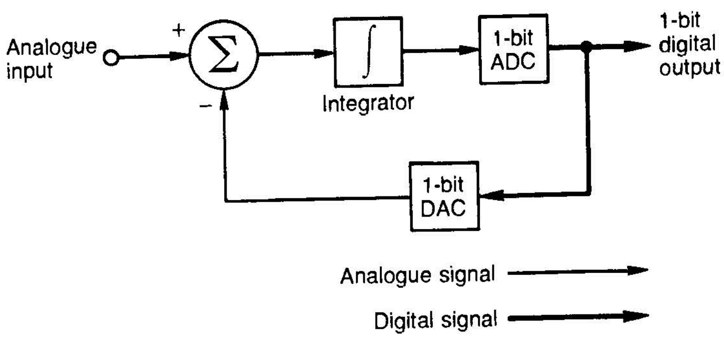


Figure 10.6 *A sigma delta ADC is used to turn an analogue signal into a bit stream*

Low-pass filter

To appreciate the essential differences between analogue filters and digital filters we examine the operation of a typical low-pass digital filter (Fig. 10.7). One notable difference is the fabrication of the filter. There are no resistors, capacitors, inductors or op amps. Instead, there are registers, multipliers and summing blocks. The arrowed lines in the diagram do not represent single electrical connections. They represent data buses, consisting of 8 or more parallel connections for transferring digital data from one part of the filter to another. In another sense, Fig. 10.7 represents not an electronic circuit but a flow-chart representing a mathematical algorithm. It could be, and often is, realized completely by software similar to that used to program a computer.

For the moment, think of the filter as a chain of data registers (8 bits or more) linked to form a multi-bit shift register (Fig. 10.7). The first register periodically receives a multi-bit digital value x_n. Think of the *n* as standing for 'now', when that particular sample is received. The registers hold a sequence of samples from x_{n-1}, the previous sample or the 'one before now' to x_{n-5}, the 'fifth before now'. As each sample is received, the sample values are shifted along the row of registers. Each time a new sample enters the register, sample x_{n-5} is lost and is replaced by x_{n-4}. The present input value and the values held in each register are passed to a row of multipliers. These multiply each value by a fixed coefficient, a_0 to a_5. The multiplied values are added by a row of summers to produce the output y_n.

In the simplest case, the value of *a* is 1/6 for each multiplier. Then y_n is the mean value of x_0 to x_5. In other words, the output of the filter is the moving average of the present and most recent five values of the input. This has a filtering effect, for a moving average reduces the impact of short-term variations in the value of *x*. It reduces the high-frequency (short-term) variations while transmitting long-term trends. It acts as a low-pass filter.

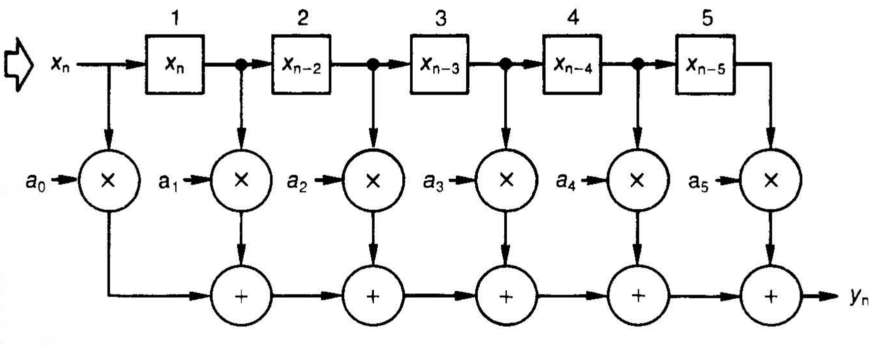


Figure 10.7 *An FIR digital filter consists of a chain of registers, tapped by multipliers, with summers to generate the output signal*

The multipliers in Fig. 10.7 are often referred to as **taps**, since they tap the values of previous inputs as they are shifted along the chain of registers. Figure 10.7 is a 6-tap filter. Quite often a filter may consist of a hundred or more registers with taps. It is interesting to note that the output of the filter is obtained by summing the outputs from the taps. This corresponds to the action of an analogue filter in which the present charge on a capacitor is the sum of charges recently arriving at the filter.

It is possible to use pencil, paper and calculator to investigate the action of this filter with other values of a and with differing types of input signal. But the calculations are quicker with a BASIC program as listed in Appendix A. The 'registers' hold values from 0 to 4095 so the program simulates a 12-bit filter. Using this program, we set up a 3-tap filter with $a_0 = 0.25$, $a_1 = 0.5$ and $a_2 = 0.25$. Figure 10.8 is the graph drawn for a run of 100 samples. The sampling frequency is such that each cycle of the sine wave is sampled 8 times. We say the period is 8. The graph shows the last 8 values calculated. It is necessary not to take values from the first one or two cycles because, just as the capacitors in an analogue begin uncharged and take a few cycles to reach a steady state, so the registers in the digital filter start containing zeros and take a few cycles to reach steady values. The output is a sine wave of the same period as the input, but lagging one sampling period behind. The amplitude is 1478 compared with 2048 for the input. The gain is 0.72. The interesting thing to do is to vary the frequency of the input signal and note the amplitude of the output (the difference between the largest and smallest output value). A second BASIC program does

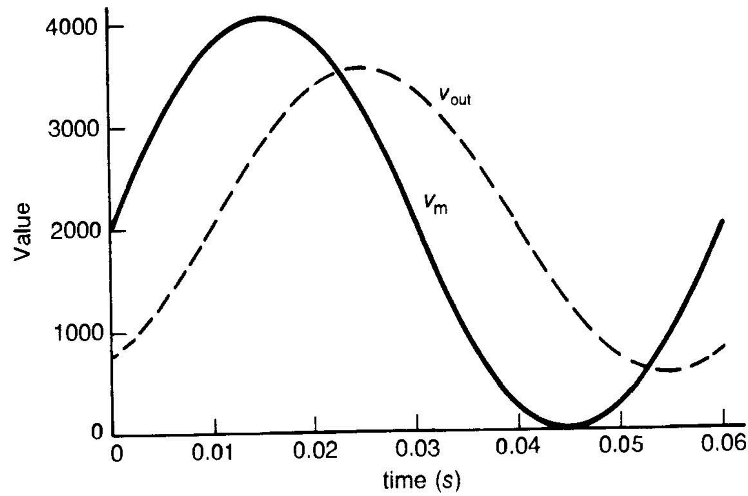


Figure 10.8 *The result of a typical run of the BASIC FIR program (p. 155) shows a sinusoidal input producing a sinusoidal output at the same frequency but with reduced amplitude*

this. It runs the filter program repeatedly for a range of periods and picks out the maximum and minimum outputs obtained. From this it calculates the output amplitude.

A graph of the output amplitude against frequency (1/period) is shown in Fig. 10.9, plotted with a logarithmic scale for frequency and a decibel scale for output. This is in fact a Bode plot of the filter's transfer function. The graph shows that gain falls with increasing frequency. This is the response typical of a low-pass filter. Roll-off is gradual (about −4 dB per octave) and, by running at other frequencies, it is also found that there is no change of phase lag with change of frequency. This filter has the characteristics of a Bessel filter (p. 121). This characteristic is the result of the symmetrical values given to the coefficients, with the largest value for the middle tap and reducing values on either side of centre. Note that the output curve in Fig. 10.8 is close to the −3 dB value as shown in Fig. 10.9, which is as anticipated, since it has a gain of 0.72 compared with the 0.71 expected at the −3 dB point.

Quite often the coefficients are symmetrical, as in the examples above. In that case processing time may be shortened by adding pairs of register contents before multiplying them. In Fig. 10.7, with symmetrical components, $a_0 = a_5$, $a_1 = a_4$, and $a_2 = a_3$. Rather than form the products a_0x_n and a_5x_{n-5} and then add them, we sum x_n and x_{n-5} and then multiply the sum by a_0. But coefficients do not have to be symmetrical and negative values are allowed, with possibly some zero values too. By a suitable choice of the number of taps and the coefficients

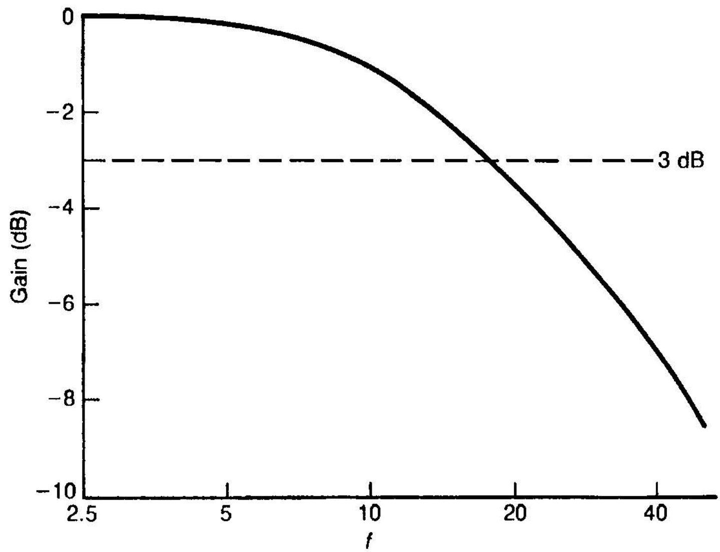


Figure 10.9 *Repeating the analysis of Fig. 10.8 over a range of frequencies produces this Bode plot of the frequency response of an FIR low-pass filter*

associated with them it is possible to produce all the popular filter types such as Butterworth and Chebyshev as well as many types that cannot be produced in analogue form.

One of the advantages of digital filters is that they do not rely on the effects of reactance of capacitors or inductors, which inevitably produce phase delays that often vary markedly with frequency (for example, see Fig. 8.13). Such phase distortion is very perceptible in audio signals. Digital filters can be designed so that delays are unaffected by frequency and phase distortion is absent.

By using suitable values for *a* it is also possible to turn Fig. 10.7 into a high-pass filter. In one trial run we set up a 5-tap filter with the coefficients 0.5, −0.5, 1, −0.5, 0.5. The result was a filter in which increasing frequency led to higher output amplitude. Another set of coefficients in a 9-tap filter produced a band-pass action. This demonstrates the fact that the behaviour of a digital filter can be radically altered simply by configuring it differently. A general-purpose filter ic can be programmed to produce a wide range of transfer functions. Contrast this with an analogue filter where producing a different transfer function usually means substituting components of different values and sometimes completely rewiring the filter.

The BASIC programs in Appendix A are a good way of getting to understand digital filtering and can readily be adapted to filters of other designs. It is possible to add routines to produce other waveforms such as squarewaves, sawtooth waves and pulses and to study the effects of digital filtering on these. The mathematics of converting a given transfer function into a set of coefficients is beyond the scope of this book, but a lot can be discovered by trial and error.

Integrated filters

Digital filters are necessarily fabricated as integrated circuits for they are too complex to make it feasible to build them from discrete components. There are ics which have dedicated filter circuits that can be programmed to perform a range of functions. Also there are general-purpose digital signal processor (DSP) ics which perform an even wider range of functions. These have the complexity of micro-processors but are superior to microprocessors because they have specially built multipliers. These multiply in one clock cycle whereas microprocessors need up to 10 clock cycles for the same operation.

IIR filters

The filter described in the previous section is known as a **finite impulse response** (FIR) filter. The reason for the word 'impulse' in this name is that the samples

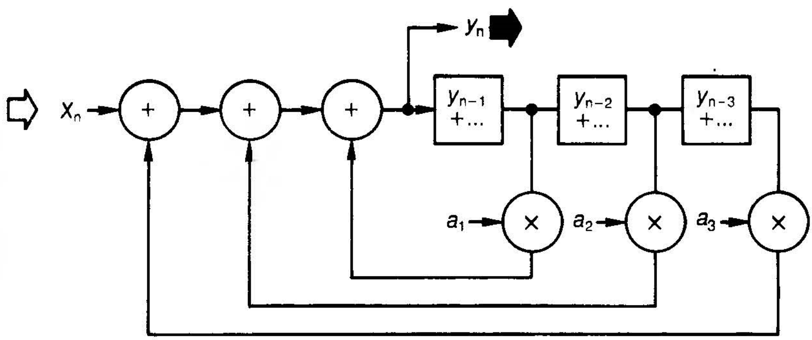


Figure 10.10 *In an IIR filter the values in each register, after multiplying by their coefficients, are fed back to be summed with the input signal*

are regarded as impulses. They are instantaneous values of the changing analogue signal, like the dots in Fig. 10.3. The samples pass along the chain of registers, appearing at each tap in succession and eventually are lost. Their life is finite. Figure 10.10 shows a different type of digital filter, which has feedback. As the samples pass along the chain of registers they are multiplied by the various coefficients and the products are fed back to be added to subsequent samples. In this way the value of a sample, repeatedly modified by coefficients circulates forever or at least until the filter is switched off. This type of filter is known as an **infinite impulse response** (IIR) filter.

As might be expected, different architecture gives IIR filters different properties. Since the action of the filter is fed back to the beginning of the chain the filtering effect is accentuated. Steeper roll-off is obtainable with fewer taps. Unfortunately feedback may lead to instability so greater care has to be taken when designing the filter, that is, deciding the number of taps and the values of the coefficients. IIR filters are also subject to phase distortion. A BASIC program for simulating an IIR is given in Appendix A.

Flexible filtering

A digital signal processor contains all the subcircuits needed to build up digital filters of many types. The basic types are the FIR and IIR filters already described. These two types are often combined into a single filter as in Fig. 10.11. This is a simple second-order filter unit that may be cascaded to produce filters of fourth and higher orders. Digital filtering allows a tremendous diversity of filtering regimes.

It has already been pointed out that digital filters depend as much, if not more, on their coefficients than on the detailed structure of the integrated circuit with which they are implemented. This makes it possible to produce a new

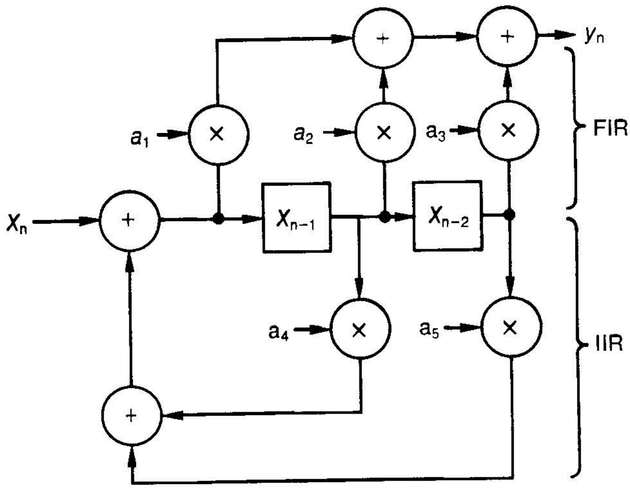


Figure 10.11 *FIR and IIR filters may be combined in many different configurations to produce a wide variety of responses*

kind of filter, the **adaptive filter**. You might imagine a hi-fi enthusiast sitting at the controls of the CD player altering the bass and treble responses to suit the passages of the music currently being played. Similarly the output of a filter may be monitored automatically and the values of the coefficients changed accordingly while the filter is still running.

Errors in digital filters

The main practical advantages of digital filters from the point of view of the manufacturer are that their circuits are entirely integrated which means that they have small size and require little power. Being digital, there is little chance of a '0' turning into a '1' or a '1' turning into a '0'. This gives greater reliability and reproducibility, for the filter will always behave in the way that is intended. There are none of the inaccuracies associated with component tolerances or ageing.

With regard to noise, the digital filter is immune to the kinds of noise that plague analogue systems, but has noise problems of its own. The chief source of this is quantization noise, which is due to values being expressed to the nearest single bit. This is introduced at the stage of analogue to digital conversion, and can be minimized by having either a large number of bits (say 12-bit resolution), or fewer bits combined with high-frequency sampling. Another source of noise, this time originating in the filter, is due to the quantization of the coefficients. Inescapably, the coefficients are represented as binary numbers. A binary number of a given number of bits is not necessarily exact since it represents the required value of the coefficient only to the nearest bit. Suppose, for example, that the intended value of the coefficient is 0.23. This is represented as an 8-bit binary fraction, 00111010. The actual value of this is 0.2266, so an error of 1.49% is introduced. Multiplying by this binary value increases the error, particularly in IIR filters in which the values circulate round and round reintroducing the error at each cycle. One way to reduce errors of this kind is to increase the number of bits to, say, 16. Another is to use coefficients that require only a few digits, for example 0.5 (1), 0.25 (01) and 0.75 (11). If such coefficients are used it is easier to perform the multiplications by shifting the bits. For example, if the coefficient is 0.625, shift the bits one place to the right to multiply by 0.5. Then shift two more places to the right to multiply by 0.125. Finally, sum the results of multiplying by 0.5 and 0.125. Filters based on 'easy' coefficients may not give exactly the response required but the advantage gained from simpler circuitry and faster processing may outweigh a slight lack of precision. It may also be possible to design two or more of these simpler filters which when cascaded produce a response very near to that required.

Another difficulty with digital filters is that the registers may overflow (exceed their maximum value of all 1s) or underflow (fall below their minimum value of all 0s). Like the odometer on a car, the overflow produces an abrupt change from a very high value to a very low one. When this occurs the output of the filter is seriously in error. The situation is worse in an IIR because the erroneous value circulates indefinitely. This is something that has no counterpart in analogue filters. A capacitor is usually rated to withstand a voltage many times, perhaps many hundreds of times, the maximum voltage likely to occur in the system. Similarly, underflow in a capacitor merely charges the capacitor in the opposite direction. Some types of digital filter have subcircuits to check the contents of the registers and, once a register has reached its maximum of all 1s, it is not allowed to increase further. Conversely, it is not allowed to fall below its minimum value of all 0s. This reduces the extent of overflow and underflow errors but may still allow appreciable errors to occur. This may be avoided by making sure that input levels are held within reasonable bounds.

Summary

Digital filtering requires analogue data to be converted into digital form for filtering and usually to be reconverted to analogue form afterward.

Digital filtering is a mathematical process that can be performed on a microprocessor, a special-purpose programmable digital filter ic or a general-purpose digital signal processor. Essentially it involves passing samples along a chain of registers, and multiplying and summing their contents in various ways to produce the output.

The two main types of digital filter are finite impulse response (FIR) filters in which samples are eventually lost from the filter, and infinite impulse response (IIR) filters.

Test yourself

1. What are the main advantages of digital filters when compared with analogue filters?
2. What sources of error are there in digital filters?
3. What is the difference between FIR and IIR filters?

Appendix A BASIC programs

These programs are written in GWBASIC but are readily adaptable to other dialects of BASIC. They are minimal programs with no input checking or other refinements, but it is easy to adapt or expand them.

```
10 REM *** SWITCHED CAPACITOR FILTER**
20 CLS
30 INPUT" D=";D
40 A = 10
50 FOR J = 1 TO 300
60 B = A
70 C = (C+B)/2
80 PRINT C;"    ";
90 IF A = 10 OR A = -10 THEN D=-D
100 A = A + D
110 NEXT
```

A is the input voltage, *B* is the voltage on *C*1 (see Fig. 9.9), *C* is the voltage on *C*2, and also the output voltage, *D* is the amount by which the input voltage changes at each sampling. When requested enter a value for *D*, which should be a factor of 20, or a decimal fraction such as 0.25 that is non-recurring. When the result is displayed (successive values of *C*) look for the maximum and minimum values to find the amplitude.

```
10 REM  *** FIR DIGITAL FILTER ***
20 DIM X(100), A(100)
30 CLS: INPUT"Number of taps";T
40 FOR J = 0 TO T-1
50 PRINT "Coefficient, a";J;:INPUT"=  ";A(J)
60 NEXT
70 INPUT"Number of samples";N
80 INPUT"Period";P
90 INPUT"Printout required? (y/n)";K$
100 FOR SAMPLE= 0 TO N
110 XS = SAMPLE MOD P
120 X = INT((SIN(XS*6.283/P))*2047+2047)
```

```
130 FOR J= T-1 TO 1 STEP -1
140 X(J) = X(J-1)
150 NEXT
160 X(0) = INT(X)
170 Y=0
180 FOR J=0 TO T-1
190 Y = INT(Y+X(J)*A(J))
200 NEXT
210 PRINT X;Y;"      ";
220 IF K$="y" THEN LPRINT X;Y;"      ";
230 NEXT
240 PRINT:INPUT"Repeat with same filter? (y/n)";R$
250 IF R$="y" THEN 70
260 END
```

The period is the number of samples in each cycle of the sine wave. Enter the number of samples to cover several periods so that values have time to settle. A suitable value is 200, though a larger number is required for periods over 80. The display consists of pairs of values (running on across the lines to display more values before scrolling occurs). Usually only the final screenful is needed, which will show the results of the final few periods.

```
10 REM  *** FIR DIGITAL FILTER - frequency response ***
20 DIM X(100), A(100)
30 CLS: INPUT"Number of taps";T
40 FOR J = 0 TO T-1
50 PRINT "Coefficient, a";J;:INPUT"=  ";A(J)
60 NEXT
70 INPUT"Period from";PF
80 INPUT"Period to";PT
90 INPUT"Step";ST
100 INPUT"Printout required? (y/n)";K$
110 FOR P = PF TO PT STEP ST
120 FOR J= 0 TO 100:X(J) = 0:NEXT
130 YMA=0;YMI=5000
140 FOR SAMPLE= 0 TO 500
150 XS = SAMPLE MOD P
160 X = INT((SIN(XS*6.283/P))*2047+2047)
170 FOR J= T-1 TO 1 STEP -1
180 X(J) = X(J-1)
190 NEXT
200 X(0) = INT(X)
210 Y=0
220 FOR J=0 TO T-1
230 Y = INT(Y+X(J)*A(J))
240 NEXT
250 IF Y>YMA THEN YMA=Y
260 IF Y<YMI THEN YMI=Y
```

```
270 NEXT
280 GAIN = (YMA-YMI)/4096
290 PRINT 100/P;"      ";GAIN;"      ";8.6858*LOG(GAIN)
300 IF K$="y" THEN LPRINT 100/P;"    ";GAIN;"    ";8.6858*LOG(GAIN)
310 NEXT
320 PRINT:INPUT"Repeat with same filter? (y/n)";R$
330 IF R$="y" THEN 70
340 END
```

This program repeats the routines of the first FIR program above for a range of periods. It does not print out the input and output values but looks for the maximum and minimum output values at each period. From these it calculates gain, both absolute and on a decibel scale. The formula used for decibels takes into account that GWBASIC works with natural logarithms, not logarithms to base 10, as usually required for decibels. Instead of displaying the period, it displays frequency, assuming that the data is sampled 100 times per second. Values of the maximum and minimum for a given period length are less extreme when the expression on line 60 does not evaluate $\sin X = 1$ and $\sin X = -1$, which may happen when the period is short, as the STEP does not take the routine to these points. The error is negligible except when P is 1-3 or 5-7.

```
10 REM  *** IIR DIGITAL FILTER ***
20 DIM X(100), A(100)
30 CLS: INPUT"Number of taps";T
40 FOR J = 0 TO T-1
50 PRINT "Coefficient, a";J;:INPUT"=   ";A(J)
60 NEXT
70 INPUT"Number of samples";N
80 INPUT"Period";P
90 INPUT"Printout required? (y/n)";K$
100 FOR SAMPLE= 0 TO N
110 XS = SAMPLE MOD P: S=0
120 X = INT((SIN(XS*6.283/P))*2047+2047)
130 FOR J= T-1 TO 1 STEP -1
140 X(J) = X(J-1):S = S+X(J)*A(J)
150 NEXT
160 X(0) = INT(X+S)
170 Y=X(0)
180 PRINT X;Y;"      ";
190 IF K$="y" THEN LPRINT X;Y;"      ";
200 NEXT
210 PRINT:INPUT"Repeat with same filter? (y/n)";R$
220 IF R$="y" THEN 70
230 END
```

This is similar to the first program listed above but simulates the feedback architecture of Fig. 10.10. Output values can become very large (unstable filter?) unless some of the coefficients are negative.

Appendix B Answers to Keeping up?

Chapter 1

1. (a) 4 A, (b) 8 Ω, (c) 4.5 V, (d) 2 Ω, (e) 0.125 A or 125 mA.
2. $R = 15/2 = 7.5\,\Omega$, so $v = 0.4 \times 7.5 = 3$ V.
3. (a) 8.75 C, (b) 0.5 A, (c) 1.25 s, (d) 0.012 C, (e) 0.4 s.
4. (a) 0.024 F, (b) 10 V, (c) 0.0045 C, (d) 800 V, (e) 0.02 F
5. $q = 1.5 \times 0.001 = 0.0015$ C. $v = 0.0015/0.0003 = 5$ V.
6. (a) Discharging, (b) charging, (c) just finished discharging and about to begin charging.
7. (a) −0.5 V, (b) $-1000\,\text{As}^{-1}$, (c) 0.024 H, (d) −720 V, (e) 0.08 H.
8. (a) Initial voltage is not relevant; rise of $12\,000\,\text{Vs}^{-1}$, means current would be rising at $24\,\text{As}^{-1}$. Back emf is $0.015 \times 24 = -0.36$ V. (b) Steady value is $2 + 0.0001 \times 12\,000 = 3.2$ V. Current is constant so there is no back emf.
9. v_r has just been overtaken by v, so v_I is positive and increasing. The magnetic field in the inductor is increasing and so is the back emf.

Chapter 2

1. (a) 500 Hz, (b) 2.5 Hz, (c) 1 Hz, (d) 200 MHz, (e) 0.00833 Hz.
2. (a) 0.2 s, (b) 0.5 ms, (c) 10 s, (d) 22.73 ns, (e) 17.86 μs.
3. Curve A: (a) 500 μs, (b) 2 kHz, (c) 1 V, (d) +5 V.
   Curve B: (a) 2 ms, (b) 500 Hz, (c) 1.5 V, (d) 0 V.
   Curve C: (a) 1 ms, (b) 1 kHz, (c) 2 V, (d) −3.5 V.
4. (a) $\pi/2$, (b) $\pi/4$, (c) 4π, (d) $\pi/10$.
5. (a) 180°, (b) 114.6°, (c) 28.65°, (d) 270°.
6. A (a) 0.2 s, (b) 5 Hz, (c) 0.4 V.
   B (a) 500 ms, (b) 2 Hz, (c) 2.5 V.
   C (a) 2 s, (b) 0.5 Hz, (c) 1.75 V.

7. A (a) 12 Hz, (b) 6 V, (c) 5.706 V.
   B (a) 1.5 kHz, (b) 0.35 V, (c) 0 V.
   C (a) 500 Hz, (b) 8.8 V, (c) −0.0896 V.
8. A $v = 10 \sin 14\pi t$, or $v = 10 \sin 44t$.
   B $v = 0.25 \sin 60000\pi t$, or $v = 0.25 \sin 188496t$.
   C $v = \sin 130\pi t$, or $v = \sin 408.4t$.
9. A $v = 0.4 \sin 10\pi t$.
   B $v = 2.5 \sin(4\pi t + \pi/2)$.
   C $v = 1.75 \sin(\pi t - \pi/4)$.
10. $v = 2.4 \sin(5000\pi t + 3\pi/2) = 2.4\,\text{V}$
11. (a) 2000π, or 6283, (b) $480\pi = 1508$, (c) $10\pi = 31.42$.
12. (a) 0.3183, (b) 3.000, (c) 15.92.

Chapter 3

1. (a) 6 C, (b) 187.5 mC, (c) 667 μC, (d) 6.4 C, (e) 2.118 C.
2. (a) 3 V, (b) 85.2 V, (c) 1.418 V, (d) 12.8 V, (e) 211.8 V.
3. 0 V; during the second half of the cycle the area between the curve and the axis is below the axis, so it is negative and exactly cancels out the charge accumulated during the first half-cycle.
4. $R = v/i$, $i = v/R$, $v = iR$.
5. $X_C = V_0/I_0$, $I_0 = V_0/X_C$, $V_0 = I_0 X_C$.
6. (a) 636.6 μΩ, (b) 7.958 μΩ, (c) 0.1929 Ω, (d) 2679 Ω, (e) 884.2 kΩ.
7. $\omega = 1000$ rad, $X_C = 5556\,\Omega$, $I_0 = 5.4/5556 = 972\,\mu\text{A}$.
8. $X_L = V_0/I_0$, $I_0 = V_0/X_L$, $V_0 = I_0 X_L$.
9. (a) 62.83 Ω, (b) 3142 Ω, (c) 314.2 Ω, (d) 785.4 μΩ, (e) 659.7 kΩ.
10. $\omega = 44$ rad, $X_L = 11\,\Omega$, $V_0 = 3.3 \times 11 = 36.3\,\text{V}$.

Chapter 4

1. (a) $v = 3.5 \sin(\omega t + 20°)$, (b) $2.2 \sin(\omega t + 200°)$, (c) $v = 1.3 \sin(\omega t + 310°)$.
2. See Fig. B1.
3. $v = 6.5 \sin(\omega t + 20°)$.
4. $v = 8.1 \sin(\omega t + 15°)$.
5. They have different frequencies so they cannot be summed.
6. $v = 3.905 \sin(\omega t + 39.81°)$.
7. $v_r = 3.084 \sin(1000\pi t + 39.55°)$, $v_c = 2.547 \sin(1000\pi t - 50.45°)$.
8. $v_r = 1.231 \sin(10\pi t + 60.50°)$, $v_c = 2.176 \sin(10\pi t - 29.50°)$.

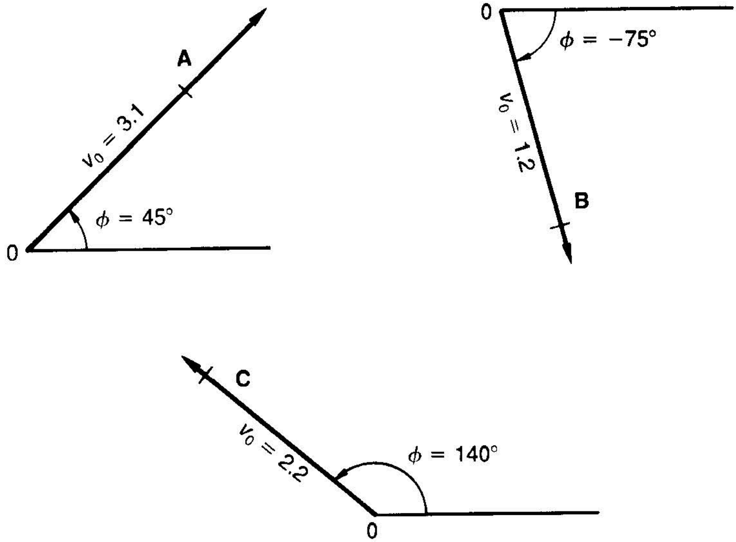


Figure B1 *Answers to keeping up?, Chapter 4, question 2*

9. $v_r = 14.99\sin(400\pi t + 2.07°)$, $v_c = 0.5422\sin(400\pi t - 87.93°)$.
10. (a) 17.68 V, (b) 12.02 V, (c) 5.30 V.
11. (a) 9.899 V, (b) 325.3 V, (c) 18.38 V.
12. (a) $3.5\angle 20°$, (b) $2.2\angle 200°$.
13. (a) $E = 7 + j3$, (b) $F = 2 - j3$.
14. (a) $5.39\angle 21.80°$, (b) $7.62\angle 293.20°$, (c) $6.08\angle 189.46°$.
15. (a) $4 + j3$, (b) $-5 - j5$, (c) $-2 - j$.

Chapter 5

1. $\mathbf{v}_{\text{out}} = 2.86\sin(300t - 44.25°)$, $\mathbf{v}_{\text{out}} = 2.0517 - j1.999$, $\mathbf{v}_{\text{out}} = 2.86\angle -44.25°$.
2. It forms a high-pass filter. $\mathbf{v}_{\text{out}} = 2.79\sin(300t + 45.74°)$, $\mathbf{v}_{\text{out}} = 1.948 + j1.999$, $\mathbf{v}_{\text{out}} = 2.79\angle 45.74°$.
3. $\mathbf{v}_{\text{out}} = (\mathbf{v}_{\text{in}} \times j\omega RC)/(j\omega RC + 1)$.
4. (a) $\mathbf{v}_{\text{out}} = (\mathbf{v}_{\text{in}} \times 0)/(0 + 1) = 0/1 = 0$, (b) dividing top and bottom by ω, $\mathbf{v}_{\text{out}} = (\mathbf{v}_{\text{in}} \times jRC)/(jRC + 1/\omega)$. But $1/\omega$ approaches 0, so $\mathbf{v}_{\text{out}} = (\mathbf{v}_{\text{in}} \times jRC)/jRC = \mathbf{v}_{\text{in}}$.
5. (a) −2.92 dB, (b) −20 dB, (c) −0.76 dB.
6. 28.9 dB.
7. 0.5 V.

Chapter 6

1. Their impedances have equal magnitudes at 500 Hz.
2. Steeper roll-off in the transition region, greater phase lag, greater attenuation.
3. 8761 Hz.
4. To reduce the peak at the resonant frequency, and so critically dampen the frequency response.
5. BW = 120 kHz, $f_r = 2359$ Hz, $Q = 19.7$.
6. $f_r = 3750$, upper −3 db point = 3877, lower −3 dB point = 3627.

Chapter 7

1. (a) $470i(t)$, (b) $\frac{1}{2.2 \times 10^{-6}} \int_0^t i(t)\,dt$, (c) 5.
2. (a) $470I(s)$, (b) $\frac{I(s)}{2.2S \times 10^{-6}}$, (c) $5/s$.
3. $i(t) = 0.01064e^{-967.1i}$
4. (a) Sinusoid of constant amplitude, $\omega = 5$; (b) non-periodic signal amplitude decreasing, e^{-3t}; (c) sinusoid at 6ω, amplitude increasing, e^{2t}; (d) constant DC level, its value undefined.
5. $s = j6283$.

Chapter 8

1. 2 MΩ, 26.45.
2. 12 kΩ, −83.33.
3. −4.44 V.

Appendix C Answers to Test yourself

Chapter 1

1. 1.645 V.
2. 0.0175 V.
3. $q = 5.5\,\mu C$, $t = 22$ ms.
4. 1.5 ms.
5. Period is 2.5 ms. $90° \equiv 2.5/4 = 0.625$ ms. Current peak occurs 0.625 ms before next peak. Next peak is at 12.5 ms, next current peak is at 11.875 ms.
6. If L is larger, the back emf is greater, more strongly opposing the flow of current, and so increasing the time to reach steady oscillations.

Chapter 2

1. (a) $f = 2$ Hz, $A = 7.5$ V, $\phi = 0$, offset = 0.
   (b) $f = 50$ Hz, $A = 1$ V, $\phi = \pi$, offset = 0.
   (c) $f = 4.5$ Hz, $A = 0.65$ V, $\phi = 4.712$ (or $3\pi/2$), offset = 3.6.
2. A $f = 10$ Hz, $A = 2$ V, offset = 0, $\phi = 0°$.
   B $f = 20$ Hz, $A = 1$ V, offset = 0, $\phi = -90°$.
   C $f = 4$ Hz, $A = 4$ V, offset = 3 V, $\phi = 45°$.
3. For curve of $A + B$, see Fig. C1. The frequency spectrum has a vertical line at 10 Hz, height 2 and a line at 20 Hz height 1.
4. The fundamental frequency is 10 Hz and there are the first 4 even harmonics at 3f, 5f, 7f and 9f, with amplitudes 1/3, 1/7, 1/5 and 1/9. $v = \sin t + (\sin 3t)/3 + (\sin 5t)/5 + (\sin 7t)/7 + (\sin 9t)/9$. The graph is shown in Fig. C2, and appears to be a squarewave signal, $f = 10$ Hz, $A = 0.75$ V, offset = 0.

Chapter 3

1. 4.244 Ω.
2. 11.78 kΩ.
3. 221.0 nF.

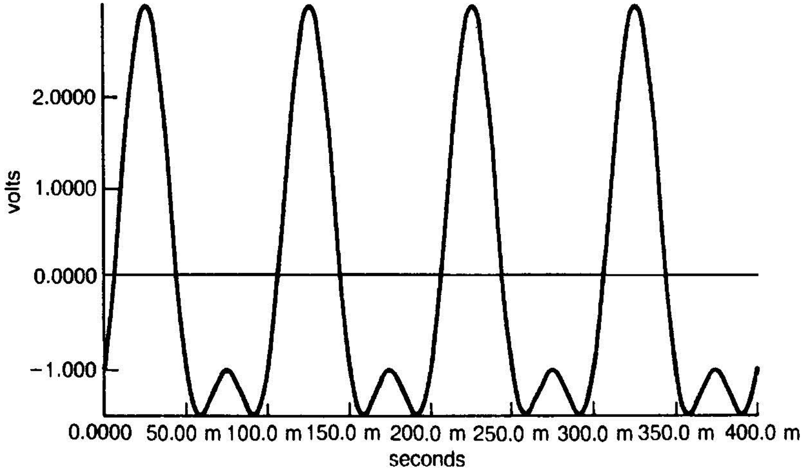


Figure C1 *Answer to Test yourself, Chapter 2, question 3*

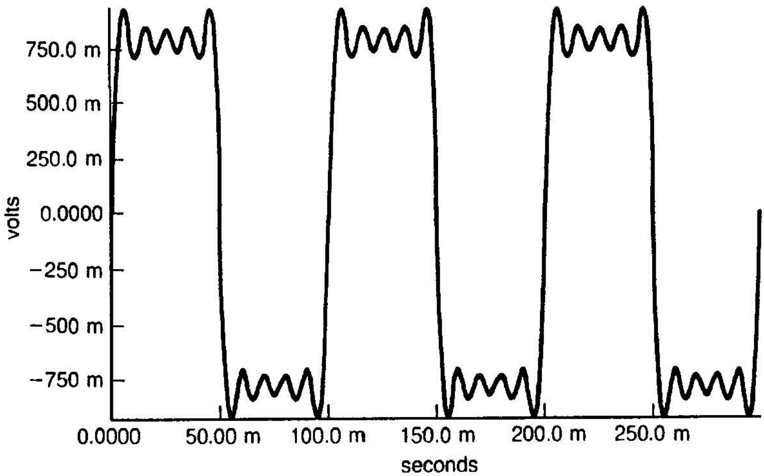


Figure C2 *Answer to Test yourself, Chapter 2, question 4*

4. 111.4 kHz.
5. 71.43 μF.
6. 4.797 μH.

Chapter 4

1. $v_r = 7.646\sin(100\pi t + 17.11^\circ)$, $v_c = 2.354\sin(100\pi t - 72.89^\circ)$.
2. $v_r = 2.416\sin(6000\pi t - 57.52^\circ)$, $v_l = 3.795\sin(6000\pi t + 32.48^\circ)$.

3. $Z = 47\,000 - j7234$, currents in microamps are $i = 124.7 + j19.19$ (rectangular form), $i = 126.2\angle 8.751^\circ$ (polar form) or $i = 1.262\sin(200\pi t + 8.751^\circ)$. $v_c = 0.1389 - j0.9022$, $v_c = 0.9128\angle -81.25^\circ$, or $v_c = 0.9128\sin(200t - 81.25^\circ)$
4. $Z = 120 + j125.66$, currents are $0.01987 - j0.02081$, $0.02877\angle -46.32^\circ$, or $i = 0.02877\sin(4000\pi t - 46.32^\circ)$, $v_l = 2.615 + j2.497$, $v_l = 3.616\angle 43.68^\circ$, or $v_l = 36.16\sin(4000\pi t + 43.86^\circ)$.

Chapter 5

1. Circuit as in Fig. 5.15(a), $C = 31.83\,\text{nF}$, v_{out}/v_{in} at 600 Hz is 0.64.
2. Circuit as in Fig. 5.15(b), $R = 36.17\,\text{k}\Omega$, v_{out}/v_{in} at 2 kHz is 0.45.
3. Circuit as in Fig. 5.15(d), $L = 159\,\mu\text{H}$, v_{out}/v_{in} at 9 MHz is 0.67.

Chapter 6

1. (a) 5th order band-pass, (b) 4th order high-pass.
2. 862 μH.
3. −12 dB per octave.
4. (a) Block low frequencies, pass high frequencies, (b) pass low frequencies, block high frequencies.

Chapter 7

1. (a) $\dfrac{j\omega 4}{j\omega 4 + 1}$, (b) $\dfrac{j\omega 0.005}{1 + j\omega 0.005 - 0.01\omega^2}$, (c) $\dfrac{0.1}{-\omega^2 - j\omega - 2}$.
2. (a) Zero at $s = 0$, pole at $s = -0.25$; (b) zero at $s = 0$, poles at $s = -0.25 \pm j0.1999$; (c) zero at $s = \pm\infty$, poles at $s = -1$ and $s = 2$; (c) is unstable.
3. Zero at $s = 0$. Poles at $s = -0.2764$ and -0.7236. This is a stable circuit.

Chapter 8

1. $\dfrac{j\omega RC}{1 + j\omega RC} \times \dfrac{R_F + R_A}{R_A}$.
2. A filter that is based on a voltage-controlled voltage source, such as an op amp.
3. Chebyshev.
4. Butterworth.

Chapter 10

1. Reliability, reproducibility, independent of component tolerances or ageing, take up small space on circuit board, low power requirements, programmable to produce a wide range of conventional transfer functions plus many not possible with analogue filters, reprogrammable by changing coefficients, even while running.

2. Quantization errors in A to D conversion, and in coefficient values, overflow, underflow.
3. In FIR samples pass along chain of registers and then lost; acts as a weighted moving average calculator, FIR filters are always stable. In IIR filters there is feedback of values, giving steeper roll-off for a given number of stages but with the risk of instability.

Index