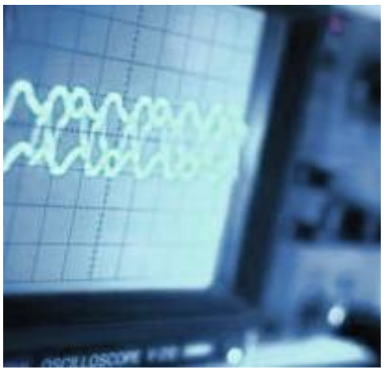
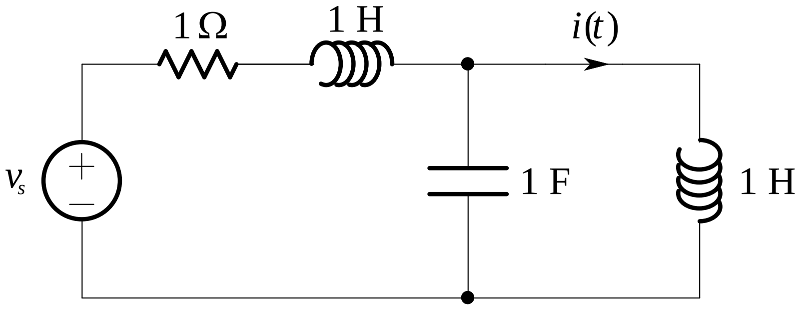
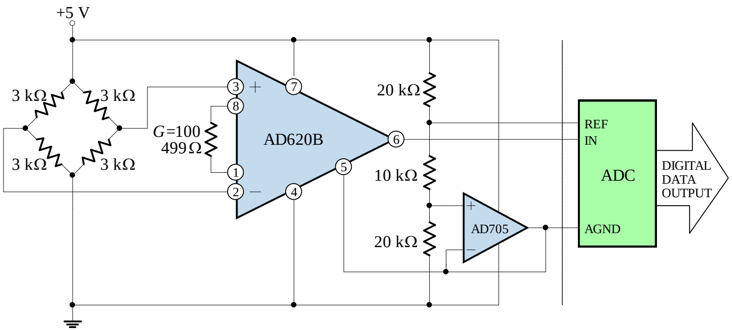
48520
Electronics and Circuits

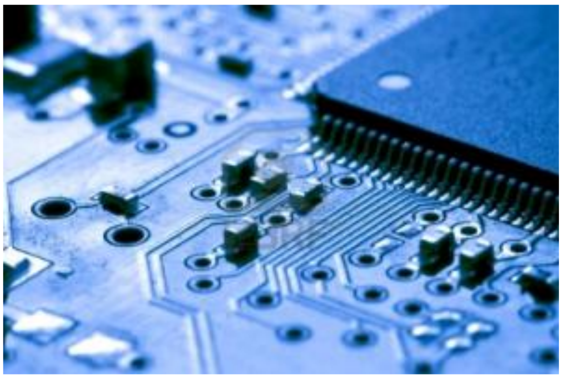
Topic Notes

2018









PMcL

Preface

These topic notes comprise part of the learning material for *48520 Electronics and Circuits*. They are not a complete set of notes. Extra material and examples may also be presented in the face-to-face activities.

Using the electronic version of these notes

These notes are hyperlinked. All green text is a link to somewhere else within this document. For example, the contents page links to the appropriate page in the text, and the page numbers in the header on each page link back to the contents page. There are also some internal linked words that take you to the relevant text.

Links to external web pages are red in colour. Provided your PDF reader (e.g. Adobe Acrobat Reader) is set up correctly these links should open the appropriate page in your web browser.

Contact

If you discover any errors or feel that some sections need clarifying, please do not hesitate in contacting me:

Peter McLean
School of Electrical and Data Engineering
Faculty of Engineering and Information Technology
University of Technology Sydney

Office: CB11.11.405 - Building 11 (Broadway), Level 11, Room 11.405
Voice : +61-2-9514-2339
Email : peter.mclean@uts.edu.au
Web : http://www.uts.edu.au/staff/peter.mclean

Contents

1 Basic Laws

2 Amplifiers

3 Nodal and Mesh Analysis

4 Circuit Analysis Techniques

5 Linear Op-Amp Applications

6 Reactive Components

7 Diodes and Basic Diode Circuits

8 Source-Free RC and RL Circuits

9 Nonlinear Op-Amp Applications

10 The First-Order Step Response

11 Op-Amp Imperfections

12 The Phasor Concept

13 Circuit Simulation

14 The Sinusoidal Steady-State Response

15 Amplifier Characteristics

16 Frequency Response

17 First-Order Op-Amp Filters

18 The Second-Order Step Response

19 Waveform Generation

20 Second-Order Frequency Response

21 Second-Order Op-Amp Filters

22 Complex Frequency

23 Specialty Amplifiers

24 Transfer Functions

25 Sensor Signal Conditioning

26 System Modelling

27 Revision

Matrices - Quick Reference Guide

Answers

1 Basic Laws

Contents

Introduction

Electric circuit theory and electromagnetic theory are the two fundamental theories upon which all branches of electrical engineering are built. Many branches of electrical engineering, such as power, electric machines, control, electronics, communications, and instrumentation, are based on electric circuit theory. Circuit theory is also valuable to students specializing in other branches of the physical sciences because circuits are a good model for the study of energy systems in general, and because of the applied mathematics, physics, and topology involved.

Electronic circuits are used extensively in the modern world – society in its present form could not exist without them! They are used in *communication systems* (such as televisions, telephones, and the Internet), *digital systems* (such as personal computers, embedded microcontrollers, smart phones), and *industrial systems* (such as robotic and process control systems). The study of electronics is therefore critical to electrical engineering and related professions.

One goal in this subject is to learn various analytical techniques and computer software applications for describing the behaviour of electric circuits. Another goal is to study various uses and applications of electronic circuits.

We will start by revising some basic concepts, such as KVL, KCL and Ohm’s Law. We will then introduce the concept of the electronic amplifier, and then study a device called an operational amplifier (op-amp for short), which has been used as the building block for modern analog electronic circuitry since its invention in the 1960’s.

1.1 Current

Charge in motion represents a *current*. The current present in a discrete path, such as a metallic wire, has both a magnitude and a direction associated with it – it is a measure of the rate at which charge is moving past a given reference point in a specified direction. Current is symbolised by *i* and thus:

Current defined as the rate of change of charge moving past a reference

$$i = \frac{dq}{dt} \tag{1.1}$$

The unit of current is the ampere (A) and is equivalent to Cs^{-1}. In a circuit current is represented by an arrow:

Representation of current in a circuit

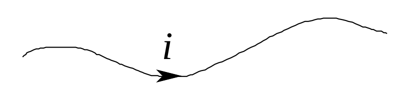


Figure 1.1

The arrow does not indicate the "actual" direction of charge flow, but is simply part of a convention that allows us to talk about the current in an unambiguous manner.

Correct usage of the term "current"

The use of terms such as "a current flows through the resistor" is a tautology and should not be used, since this is saying a "a charge flow flows through the resistor". The correct way to describe such a situation is "there is a current in the resistor".

DC and AC defined

A current which is constant is termed a direct current, or simply DC. Examples of direct currents are those that exist in circuits with a chemical battery as the source. A sinusoidal current is often referred to as alternating current, or AC[1]. Alternating current is found in the normal household electricity supply.

[1] Later we shall also see that a periodic current (e.g. a square wave), with no DC term, can also be referred to as an alternating current.

1.2 Voltage

A *voltage* exists between two points in a circuit when energy is required to move a charge between the two points. The unit of voltage is the volt (V) and is equivalent to JC^{-1}. In a circuit, voltage is represented by a pair of +/- signs:

Representation of voltage in a circuit

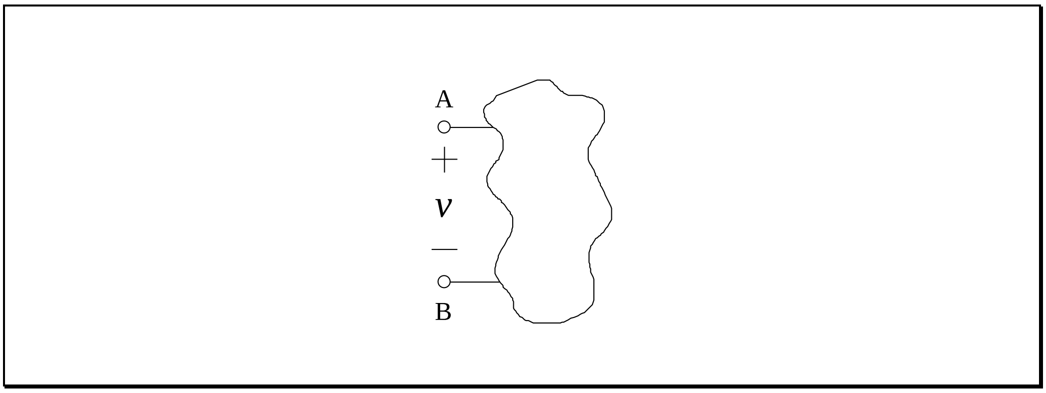
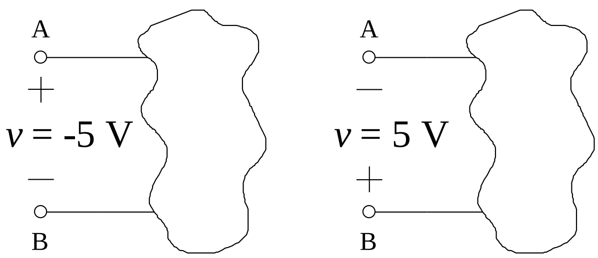


Figure 1.2

Once again, the plus-minus pair does not indicate the "actual" voltage polarity.

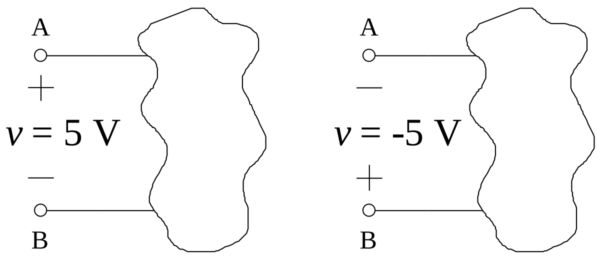
EXAMPLE 1.1 Voltage Polarity

Note the voltages across the circuit elements below:



(a) (b)

In both (a) and (b), terminal B is 5 V positive with respect to terminal A.



(c) (d)

In both (c) and (d), terminal A is 5 V positive with respect to terminal B.

1.3 Circuit Elements and Types of Circuits

Ideal circuit elements are used to model real circuit elements

A circuit element is an *idealised* mathematical model of a two-terminal electrical device that is completely characterised by its voltage-current relationship. Although ideal circuit elements are not "off-the-shelf" circuit components, their importance lies in the fact that they can be interconnected (on paper or on a computer) to approximate actual circuits that are composed of nonideal elements and assorted electrical components – thus allowing for the analysis of such circuits.

Circuit elements can be categorised as either *active* or *passive*.

1.3.1 Active Circuit Elements

Active circuit element defined

Active circuit elements *can* deliver a non-zero average power indefinitely. There are four types of active circuit element, and all of them are termed an *ideal source*. They are:

- the independent voltage source
- the independent current source
- the dependent voltage source
- the dependent current source

1.3.2 Passive Circuit Elements

Passive circuit element defined

Passive circuit elements *cannot* deliver a non-zero average power indefinitely. Some passive elements are capable of storing energy, and therefore delivering power back into a circuit at some later time, but they cannot do so indefinitely.

There are three types of passive circuit element. They are:

- the resistor
- the inductor
- the capacitor

1.3.3 Types of Circuits

Network and circuit defined

Active and passive circuits defined

The interconnection of two or more circuit elements forms an electrical *network*. If the network contains at least one closed path, it is also an electrical *circuit*. A network that contains at least one active element, i.e. an independent or dependent source, is an *active* network. A network that does not contain any active elements is a *passive* network.

1.4 Independent Sources

Independent sources are *ideal* circuit elements that possess a voltage or current value that is independent of the behaviour of the circuits to which they belong.

1.4.1 The Independent Voltage Source

An independent voltage source is characterised by a terminal voltage which is completely independent of the current through it. The representation of an independent voltage source is shown below:

Independent voltage source defined

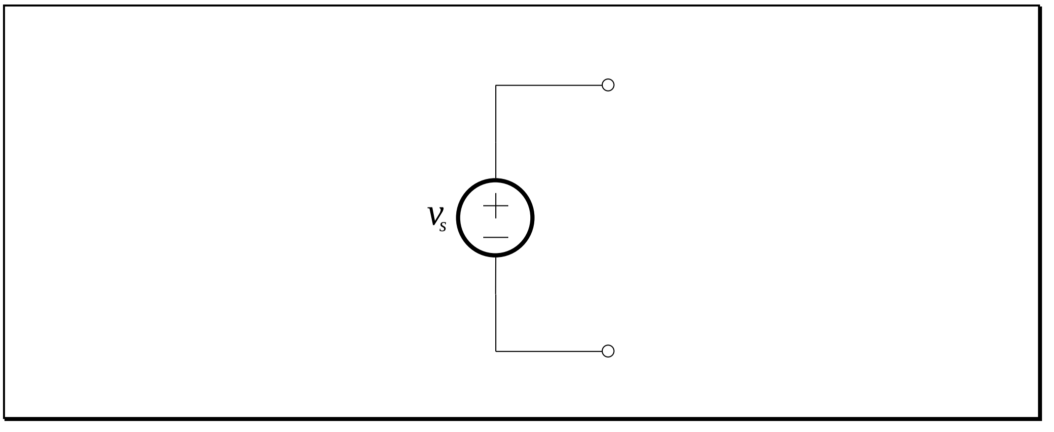


Figure 1.3

If the value of the voltage source is constant, that is, does not change with time, then we can also represent it as an *ideal battery*:

An ideal battery is equivalent to an independent voltage source that has a constant value

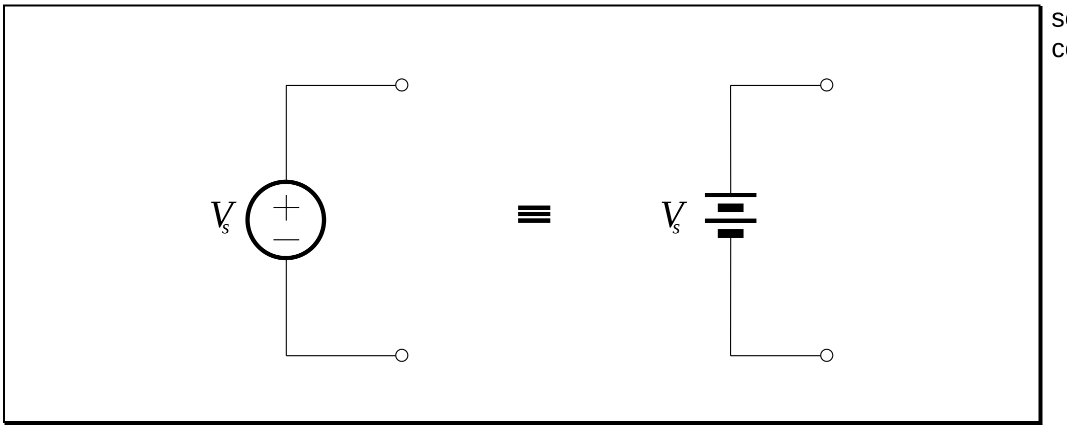


Figure 1.4

Although a "real" battery is not ideal, there are many circumstances under which an ideal battery is a very good approximation.

In general, however, the voltage produced by an ideal voltage source will be a function of time. In this case we represent the voltage symbolically as $v(t)$.

A few typical voltage waveforms are shown below. The waveforms in (a) and (b) are typical-looking amplitude modulation (AM) and frequency modulation (FM) signals, respectively. Both types of signals are used in consumer radio communications. The sinusoid shown in (c) has a wide variety of uses; for example, this is the shape of ordinary household voltage. A "pulse train", such as that in (d), can be used to drive DC motors at a variable speed.

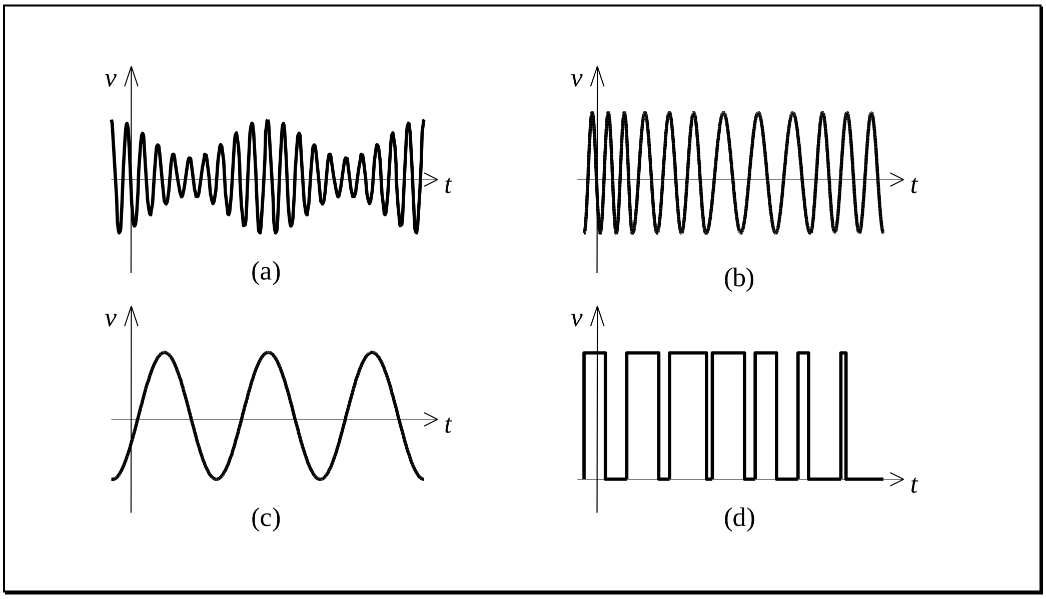


Figure 1.5

Since the voltage produced by a source is in general a function of time, then the most general representation of an ideal voltage source is as shown below:

The most general representation of an ideal independent voltage source

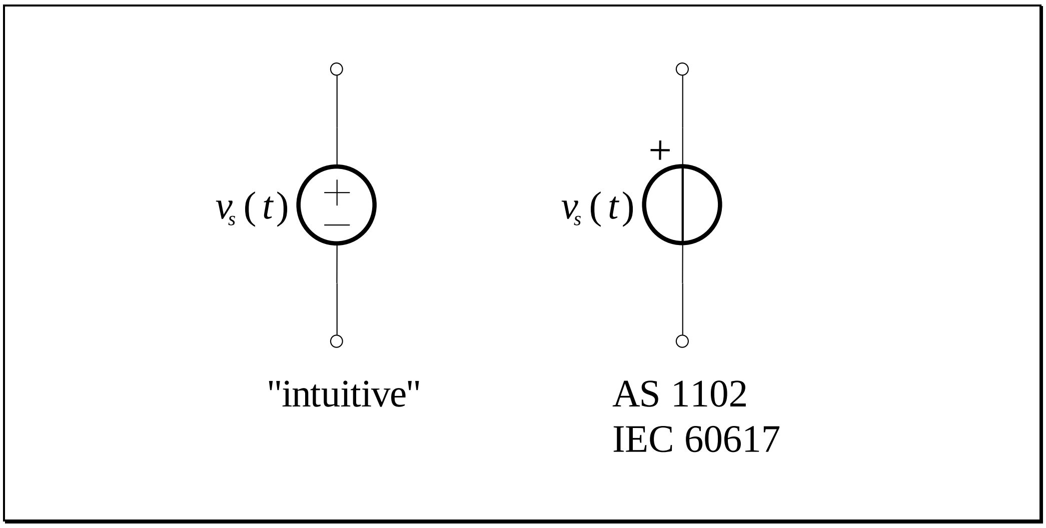


Figure 1.6

1.4.2 The Independent Current Source

An independent current source establishes a current which is independent of the voltage across it. The representation of an independent current source is shown below:

Independent current source defined

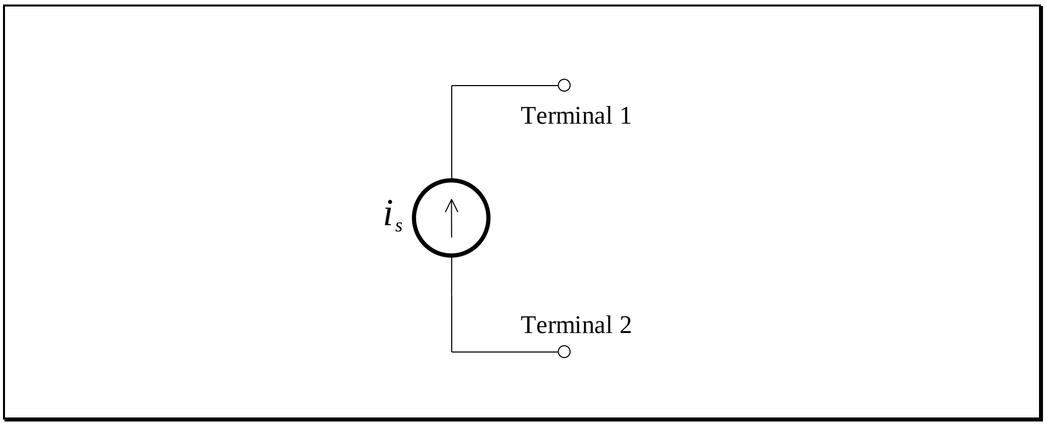


Figure 1.7

In other words, an *ideal current source* is a device that, *when connected to anything*, will always push i_s out of terminal 1 and pull i_s into terminal 2.

Since the current produced by a source is in general a function of time, then the most general representation of an ideal current source is as shown below:

The most general representation of an ideal independent current source

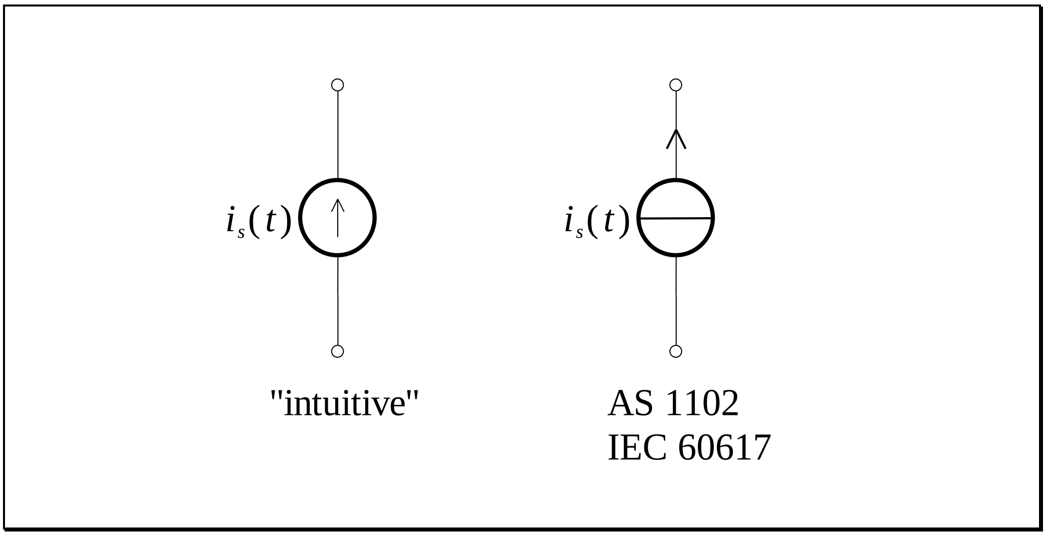


Figure 1.8

1.5 The Resistor and Ohm's Law

In 1827 the German physicist George Ohm published a pamphlet entitled "The Galvanic Circuit Investigated Mathematically". It contained one of the first efforts to measure currents and voltages and to describe and relate them mathematically. One result was a statement of the fundamental relationship we now call Ohm's Law.

Consider a uniform cylinder of conducting material, to which a voltage has been connected. The voltage will cause charge to flow, i.e. a current:

A simple resistor

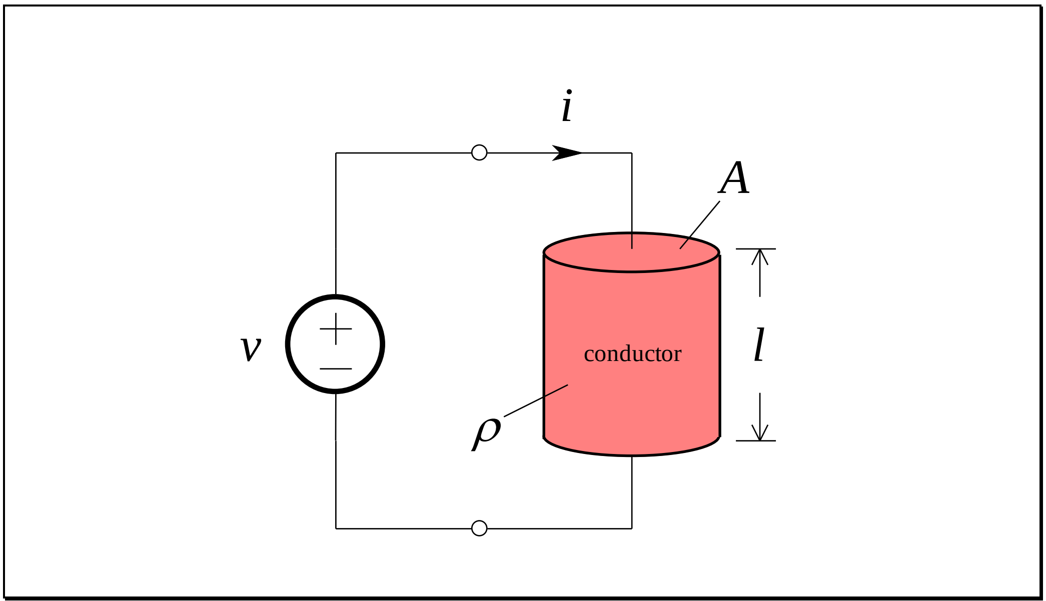


Figure 1.9

Ohm found that in many conducting materials, such as metal, the current is always proportional to the voltage. Since voltage and current are directly proportional, there exists a proportionality constant R, called *resistance*, such that:

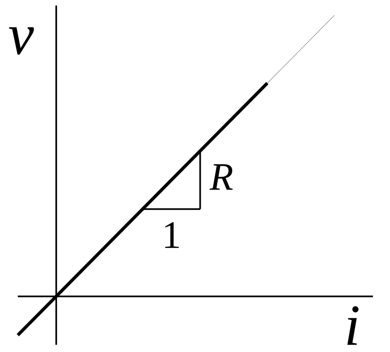
Ohm's Law

$$v = Ri \tag{1.2}$$

This is Ohm's Law. The unit of resistance (volts per ampere) is referred to as the *ohm*, and is denoted by the capital Greek letter omega, Ω.

We refer to a construction in which Ohm's Law is obeyed as a *resistor*.

The ideal resistor relationship is a *straight line through the origin*:



The resistor is a linear circuit element

Figure 1.10

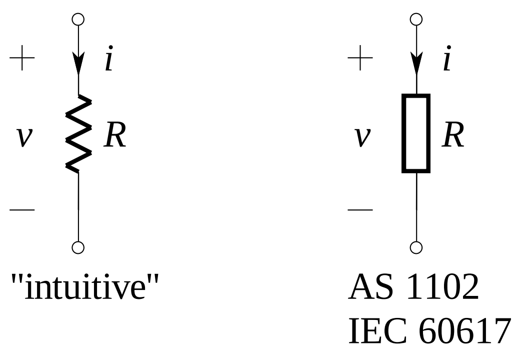
Even though resistance is defined as $R = v/i$, it should be noted that R is a purely geometric property, and depends only on the conductor shape and the material used in the construction. For example, it can be shown for a uniform resistor that the resistance is given by:

$$R = \frac{\rho l}{A} \tag{1.3}$$

The resistance of a uniform resistor

where l is the length of the resistor, and A is the cross-sectional area. The *resistivity*, ρ, is a constant of the conducting material used to make the resistor.

The circuit symbol for the resistor is shown below, together with the direction of current and polarity of voltage that make Ohm's Law algebraically correct:

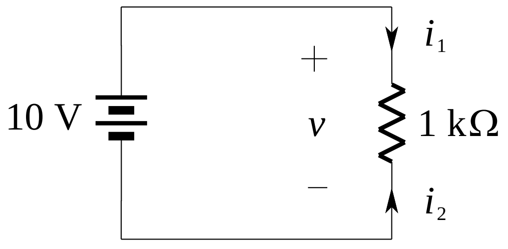


The circuit symbol for the resistor

Figure 1.11

EXAMPLE 1.2 Ohm's Law with a Voltage Source

Consider the circuit shown below.



The voltage across the 1 kΩ resistor is, by definition of an ideal voltage source, $v(t)=10\,\text{V}$. Thus, by Ohm's Law, we get:

$$i_1 = \frac{v}{R} = \frac{10}{1000} = 0.01\ \text{A} = 10\ \text{mA}$$

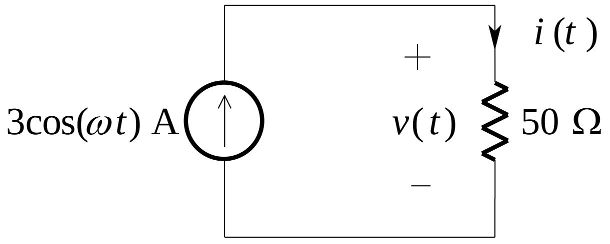
and:

$$i_2 = \frac{-v}{R} = \frac{-10}{1000} = -0.01\ \text{A} = -10\ \text{mA}$$

Note that $i_2 = -i_1$, as expected.

EXAMPLE 1.3 Ohm's Law with a Current Source

Consider the circuit shown below.

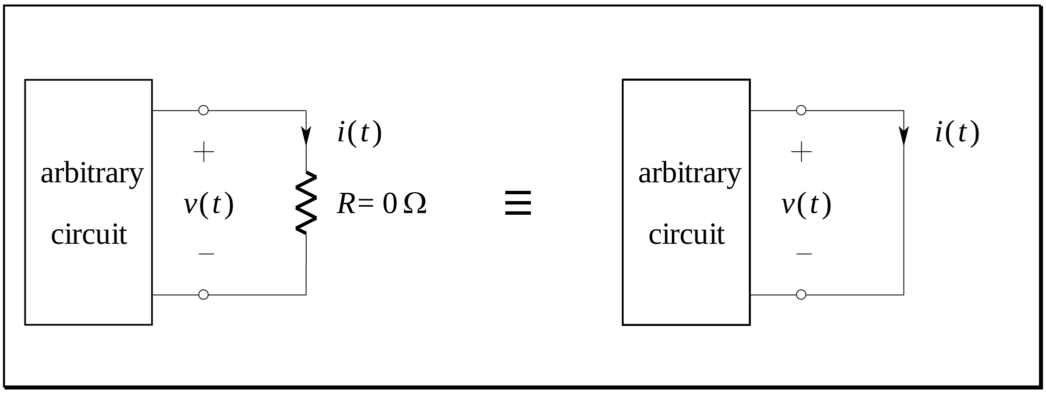


Ohm's Law yields:

$$\begin{aligned} v(t) &= Ri(t) \\ &= 50 \times 3\cos(\omega t) \\ &= 150\cos(\omega t)\ \text{V} \end{aligned}$$

1.5.1 The Short-Circuit

Consider a resistor whose value is zero ohms. An equivalent representation of such a resistance, called a *short-circuit*, is shown below:



The short-circuit

Figure 1.12

By Ohm's Law:

$$\begin{aligned} v &= Ri \\ &= 0i \\ &= 0\text{ V} \end{aligned} \tag{1.4}$$

Thus, no matter what finite value $i(t)$ has, $v(t)$ will be zero. Hence, we see that a *zero-ohm resistor* is equivalent to an *ideal voltage source whose value is zero volts*, provided that the current through it is finite.

1.5.2 The Open-Circuit

Consider a resistor having infinite resistance. An equivalent representation of such a resistance, called an *open-circuit*, is shown below:

The open-circuit

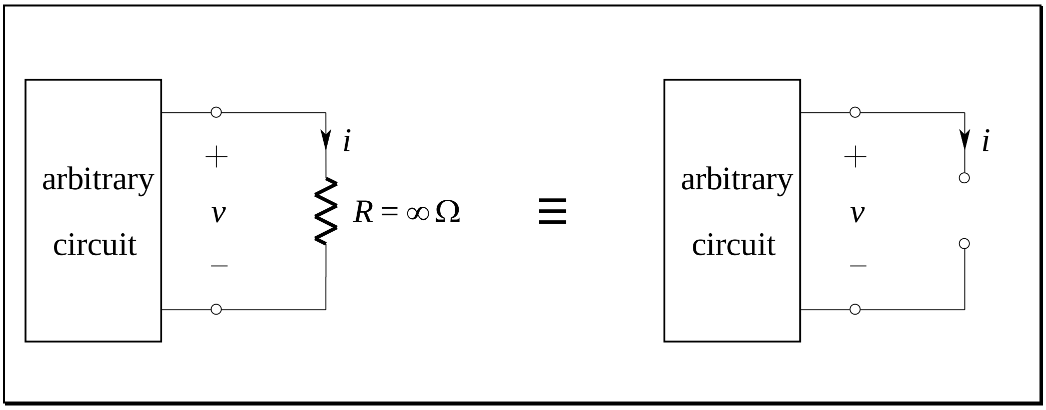


Figure 1.13

By Ohm's Law:

$$\begin{aligned} i &= \frac{v}{R} \\ &= \frac{v}{\infty} \\ &= 0\text{ A} \end{aligned} \tag{1.5}$$

Thus, no matter what finite value $v(t)$ has, $i(t)$ will be zero. Thus, we may conclude that an *infinite resistance* is equivalent to an *ideal current source whose value is zero amperes*, provided that the voltage across it is finite.

1.5.3 Conductance

The reciprocal of resistance is called the *conductance*, *G*:

Conductance defined

$$G = \frac{1}{R} \tag{1.6}$$

The unit of conductance is the *siemen*, and is abbreviated S. The same circuit symbol is used to represent both resistance and conductance.

1.6 Practical Resistors

There are many different types of resistor construction. Some are shown below:

Some types of resistors

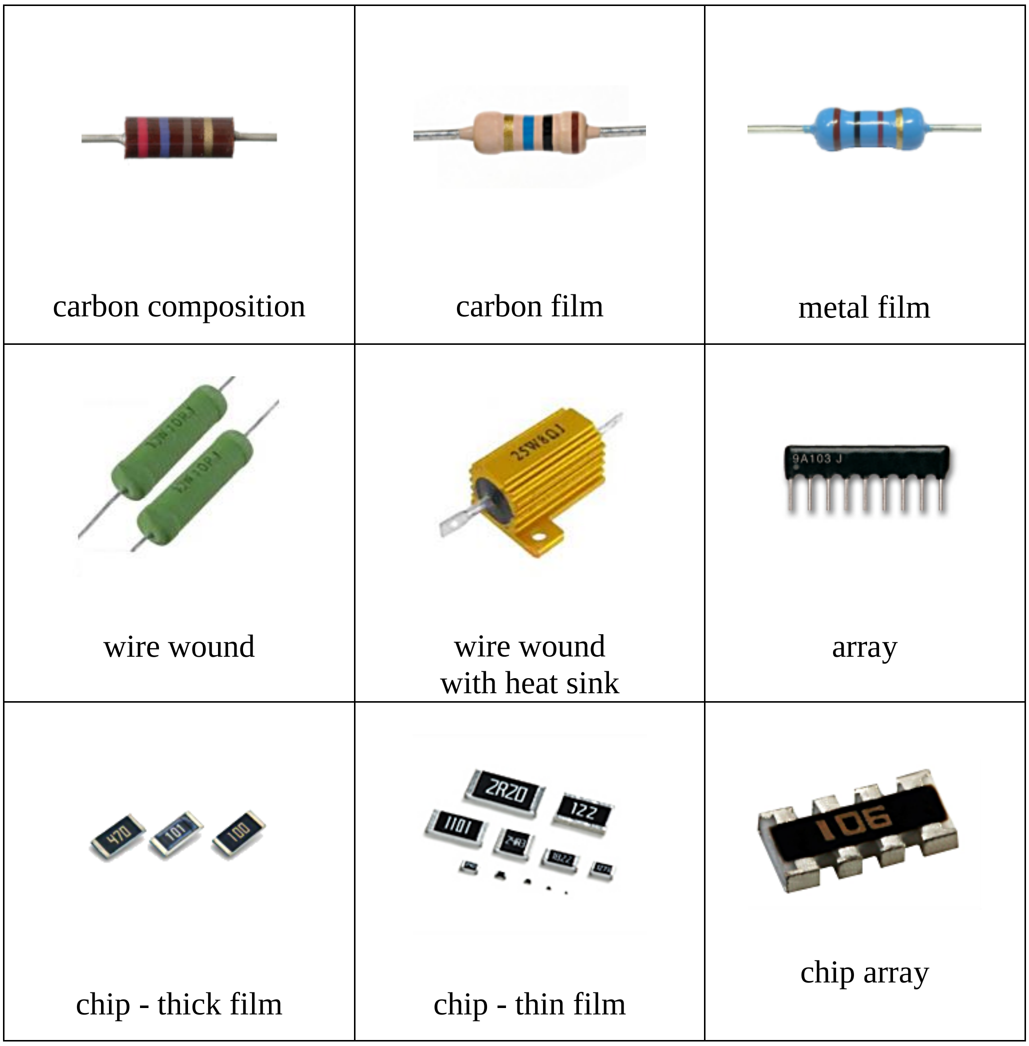


Figure 1.14 – Some types of resistors

The “through-hole” resistors are used by hobbyists and for prototyping real designs. Their material and construction dictate several of their properties, such as accuracy, stability and pulse handling capability.

The wire wound resistors are made for accuracy, stability and high power applications. The array is used where space is a premium and is normally used in digital logic designs where the use of “pull-up” resistors is required.

Modern electronics utilises “surface-mount” components. There are two varieties of surface-mount chip resistor – thick film and thin film.

1.6.1 Preferred Values and the Decade Progression

Fundamental standardization practices require the selection of *preferred* values within the ranges available. Standard values may at first sight seem to be strangely numbered. There is, however, a beautiful logic behind them, dictated by the tolerance ranges available.

Component values have been standardized by the IEC

The *decade progression* of preferred values is based on preferred numbers generated by a geometric progression, repeated in succeeding decades. In 1963, the International Electrotechnical Commission (IEC) standardized the preferred number series for resistors and capacitors (standard IEC 60063). It is based on the fact that we can linearly space values along a logarithmic scale so a percentage change of a value results in a linear change on the logarithmic scale.

Component values are spaced equidistantly on a logarithmic scale

For example, if 6 values per decade are desired, the common ratio is $\sqrt[6]{10} \approx 1.468$. The six rounded-off values become 100, 150, 220, 330, 470, 680.

1.6.2 The 'E' Series Values

The 'E' series values explained

The IEC set the number of values for resistors (and capacitors) per decade based on their tolerance. These tolerances are 0.5%, 1%, 2%, 5%, 10%, 20% and 40% and are respectively known as the E192, E96, E48, E24, E12, E6 and E3 series, the number indicating the quantity of values per decade in that series. For example, if resistors have a tolerance of 5%, a series of 24 values can be assigned to a single decade multiple (e.g. 100 to 999) knowing that the possible extreme values of each resistor overlap the extreme values of adjacent resistors in the same series.

Any of the numbers in a series can be applied to any decade multiple set. Thus, for instance, multiplying 220 by each decade multiple (0.1, 1, 10 100, 1000 etc.) produces values of 22, 220, 2 200, 22 000, 220 000 etc.

The 'E' series of preferred resistor and capacitor values according to IEC 60063 are reproduced in Table 1.1.

0.5% E192	1% E96	2% E48	0.5% E192	1% E96	2% E48	0.5% E192	1% E96	2% E48	0.5% E192	1% E96	2% E48	0.5% E192	1% E96	2% E48	
100	100	100	169	169	169	287	287	287	487	487	487	825	825	825	
101			172			291			493			835			
102	102		174	174		294	294		499	499		845	845		
104			176			298			505			856			
105	105	105	178	178	178	301	301	301	511	511	511	866	866	866	
106			180			305			517			876			
107	107		182	182		309	309		523	523		887	887		
109			184			312			530			898			
110	110	110	187	187	187	316	316	316	536	536	536	909	909	909	
111			189			320			542			920			
113	113		191	191		324	324		549	549		931	931		
114			196			328			556			942			
115	115	115	196	196	196	332	332	332	562	562	562	953	953	953	
117			198			336			569			965			
118	118		200	200		340	340		576	576		976	976		
120			203			344			583			988			
121	121	121	205	205	205	348	348	348	590	590	590				
123			208			352			597			5%	10%	20%	40%
124	124		210	210		357	357		604	604		E24	E12	E6	E3
126			213			361			612						
127	127	127	215	215	215	365	365	365	619	619	619	100	100	100	100
129			218			370			626			110			
130	130		221	221		374	374		634	634		120	120		
132			223			379			642			130			
133	133	133	226	226	226	383	383	383	649	649	649	150	150	150	
135			229			388			657			160			
137	137		232	232		392	392		665	665		180	180		
138			234			397			673			200			
140	140	140	237	237	237	402	402	402	681	681	681	220	220	220	220
142			240			407			690			240			
143	143		243	243		412	412		698	698		270	270		
145			246			417			706			300			
147	147	147	249	249	249	422	422	422	715	715	715	330	330	330	
149			252			427			723			360			
150	150		255	255		432	432		732	732		390	390		
152			259			437			741			430			
154	154	154	261	261	261	442	442	442	750	750	750	470	470	470	470
156			264			448			759			510			
158	158		267	267		453	453		768	768		560	560		
160			271			459			777			620			
162	162	162	274	274	274	464	464	464	787	787	787	680	680	680	
164			277			470			796			750			
165	165		280	280		475	475		806	806		820	820		
167			284			481			816			910			

Table 1.1 – IEC standard 'E' series of values in a decade

1.6.3 Marking Codes

The IEC also defines how manufacturers should mark the values of resistors and capacitors in the standard called IEC 60062. The colours used on fixed leaded resistors are shown below:

IEC labelling for leaded resistors

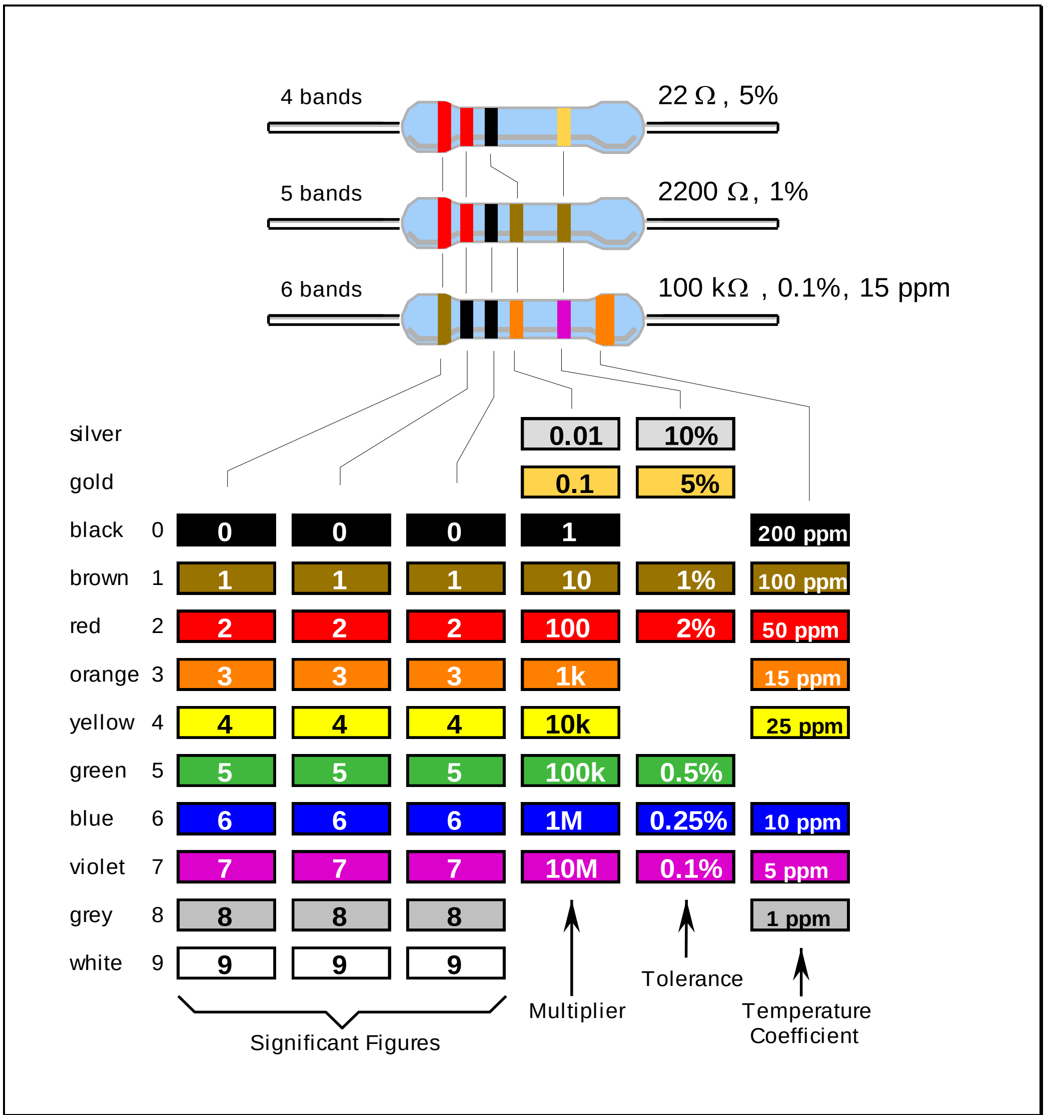


Figure 1.15 – Colour code marking of leaded resistors

The resistance *colour code* consists of three or four colour bands and is followed by a band representing the tolerance. The temperature coefficient band, if provided, is to the right of the tolerance band and is usually a wide band positioned on the end cap.

The resistance colour code includes the first two or three significant figures of the resistance value (in ohms), followed by a multiplier. This is a factor by which the significant-figure value must be multiplied to find the actual resistance value. (i.e. the number of zeros to be added after the significant figures).

Whether two or three significant figures are represented depends on the tolerance: ±5% and wider require two band; ±2% and tighter requires three bands. The significant figures refer to the first two or three digits of the resistance value of the standard series of values in a decade, in accordance with IEC 60063 as indicated in the relevant data sheets and shown in Table 1.1.

The colours used and their basic numerical meanings are recognized internationally for any colour coding used in electronics, not just resistors, but some capacitors, diodes, cabling and other items.

The resistor colour code explained

The colours are easy to remember: Black is the absence of any colour, and therefore represents the absence of any quantity, 0. White (light) is made up of all colours, and so represents the largest number, 9. In between, we have the colours of the rainbow: red, orange, yellow, green, blue and violet. These take up the numbers from 2 to 7. A colour in between black and red would be brown, which has the number 1. A colour intermediate to violet and white is grey, which represents the number 8.

When resistors are labelled in diagrams, such as schematics, IEC 60062 calls for the significant figures to be printed as such, but the decimal point is replaced with the SI prefix of the multiplier. Examples of such labelling are shown below:

IEC labelling for diagrams

Resistor Value	IEC Labelling
0.1 Ω	0R1
1 Ω	1R0
22 Ω	22R
3.3 kΩ	3K3
100 kΩ	100K
4.5 MΩ	4M5

Note how the decimal point is expressed, that the ohm symbol is shown as an R, and that 1000 is shown as a capital K. The use of a letter instead of a decimal point solves a printing problem – the decimal point in a number may not always be printed clearly, and the alternative display method is intended to help misinterpretation of component values in circuit diagrams and parts lists.

We use a letter in place of a decimal point for labelling component values

In circuit diagrams and constructional charts, a resistor’s numerical identity, or *designator*, is usually prefixed by ‘R’. For example, R15 simply means resistor number 15.

A portion of a schematic diagram showing designators and IEC labelling is shown below:

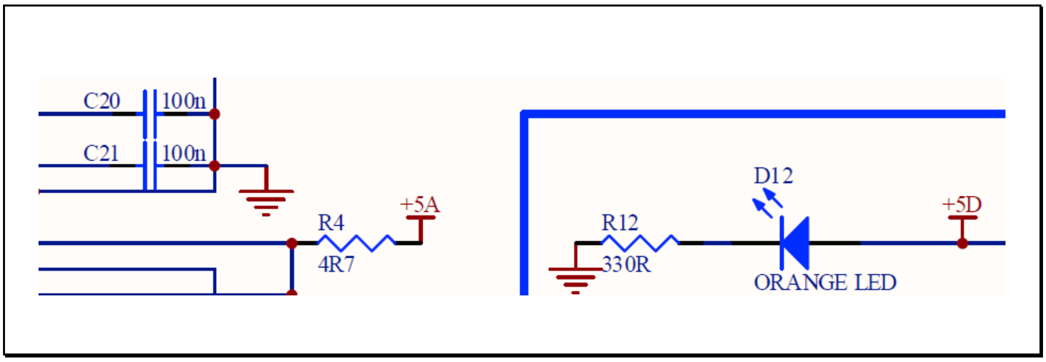


Figure 1.16 – Schematic portion showing IEC labelling

Note that resistor R4 has the value 4.7 Ω and resistor R12 has the value 330 Ω.

1.7 Kirchhoff's Current Law

A connection of two or more elements is called a *node*. An example of a node is depicted in the partial circuit shown below:

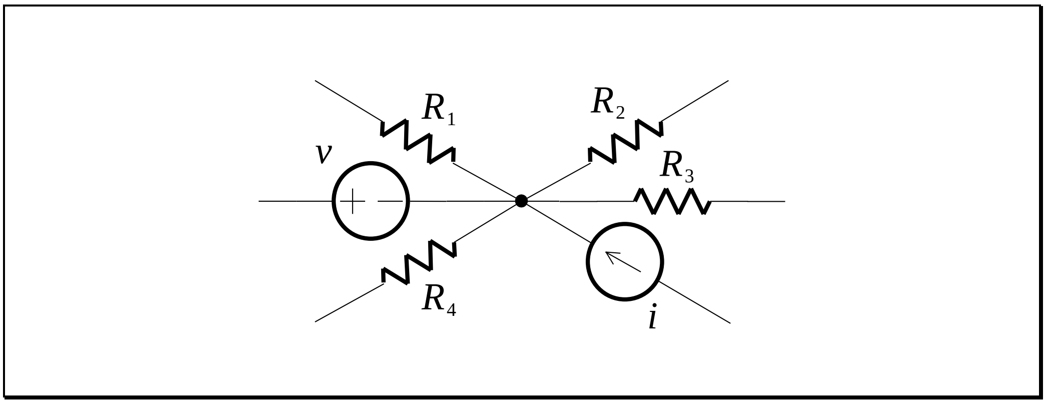


Figure 1.17

Even if the figure is redrawn to make it appear that there may be more than one node, as in the figure below, the connection of the six elements actually constitutes only one node.

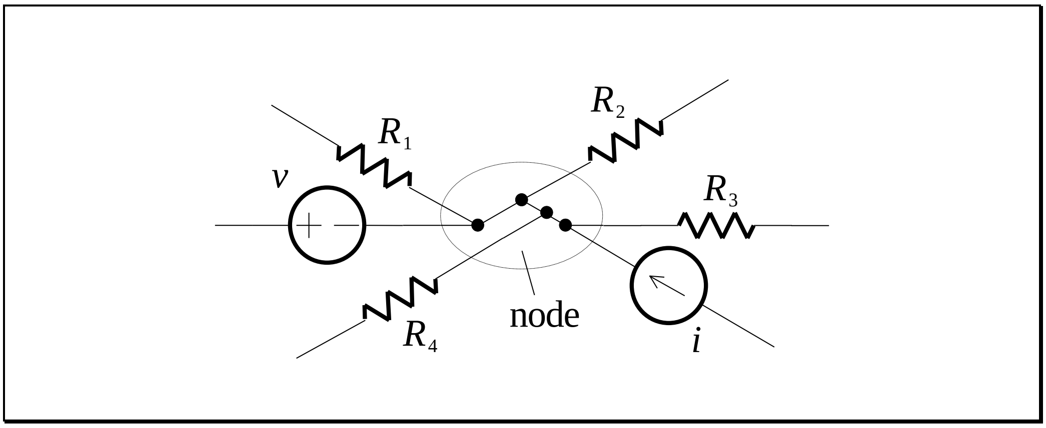


Figure 1.18

Kirchhoff's Current Law (KCL) is essentially the law of conservation of electric charge. If currents directed out of a node are positive in sense, and currents directed into a node are negative in sense (or vice versa), then KCL can be stated as follows:

KCL defined

KCL: At any node of a circuit, the currents algebraically sum to zero. (1.7)

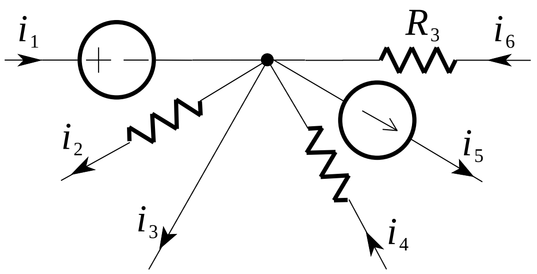
If there are n elements attached to a node then, in symbols, KCL is:

$$\sum_{k=1}^{n} i_k = 0 \tag{1.8}$$

KCL can also be stated as: The sum of the currents entering a node is equal to the sum of the currents leaving a node.

EXAMPLE 1.4 Kirchhoff's Current Law for a Node

As an example of KCL, consider a portion of some circuit, shown below:



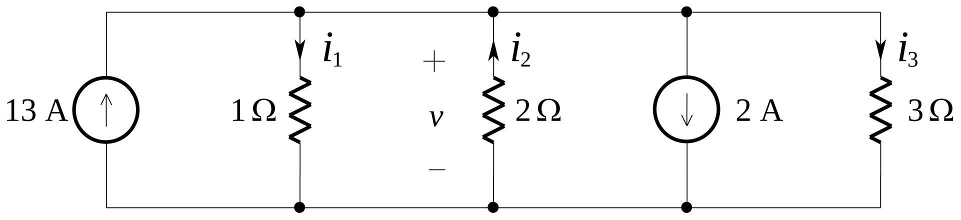
Choosing the positive sense to be leaving, we apply KCL at the node and obtain the equation:

$$-i_1 + i_2 + i_3 - i_4 + i_5 - i_6 = 0$$

Note that even if one of the elements – the one which carries i_3 – is a short-circuit, KCL holds. In other words, KCL applies regardless of the nature of the elements in the circuit.

EXAMPLE 1.5 Kirchhoff's Current Law for a Two-Node Circuit

We want to find the voltage *v*, in the two-node circuit shown below:



The directions of i_1, i_2, i_3 and the polarity of v were chosen arbitrarily (the directions of the 13 A and 2 A sources are given). By KCL (at either of the two nodes), we have:

$$-13 + i_1 - i_2 + 2 + i_3 = 0$$

From this we can write:

$$i_1 - i_2 + i_3 = 11$$

By Ohm's Law:

$$i_1 = \frac{v}{1} \qquad i_2 = \frac{-v}{2} \qquad i_3 = \frac{v}{3}$$

Substituting these into the previous equation yields:

$$\begin{aligned}
\left(\frac{v}{1}\right) - \left(\frac{-v}{2}\right) + \left(\frac{v}{3}\right) &= 11 \\
v + \frac{v}{2} + \frac{v}{3} &= 11 \\
\frac{6v + 3v + 2v}{6} &= 11 \\
\frac{11v}{6} &= 11 \\
v &= 6\text{ V}
\end{aligned}$$

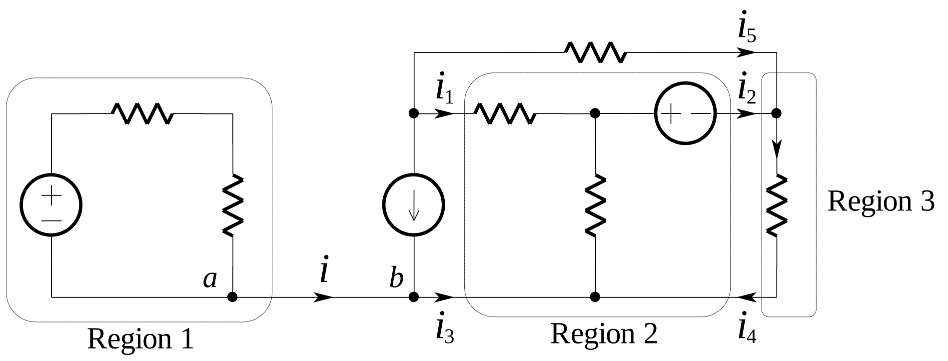
Having solved for *v*, we can now find that:

$$i_1 = \frac{v}{1} = \frac{6}{1} = 6\text{ A} \qquad \text{i}_2 = -\frac{v}{2} = -\frac{6}{2} = -3\text{ A} \qquad \text{i}_3 = \frac{v}{3} = \frac{6}{3} = 2\text{ A}$$

Just as KCL applies to any node of a circuit, so must KCL hold for any closed region, i.e. to satisfy the physical law of conservation of charge, the total current leaving (or entering) a region must be zero.

EXAMPLE 1.6 Kirchhoff's Current Law for a Closed Region

In the circuit shown below, three regions have been identified:



Applying KCL to Region 1, we get:

$$i = 0$$

For Region 2:

$$i_1 + i_3 + i_4 = i_2$$

For Region 3:

$$i_2 + i_5 = i_4$$

You may now ask, "Since there is no current from point *a* to point *b* (or vice versa) why is the connection (a short-circuit) between the points there?" If the connection between the two points is removed, two separate circuits result. The voltages and currents within each individual circuit remain the same as before. Having the connection present constrains points *a* and *b* to be the same node, and hence be at the same voltage. It also indicates that the two separate portions are physically connected (even though there is no current between them).

1.8 Kirchhoff's Voltage Law

Starting at any node in a circuit, we form a *loop* by traversing through elements (open-circuits included!) and returning to the starting node, never encountering any other node more than once.

Loop defined

For example, the paths *fabef* and *fdcef* are loops:

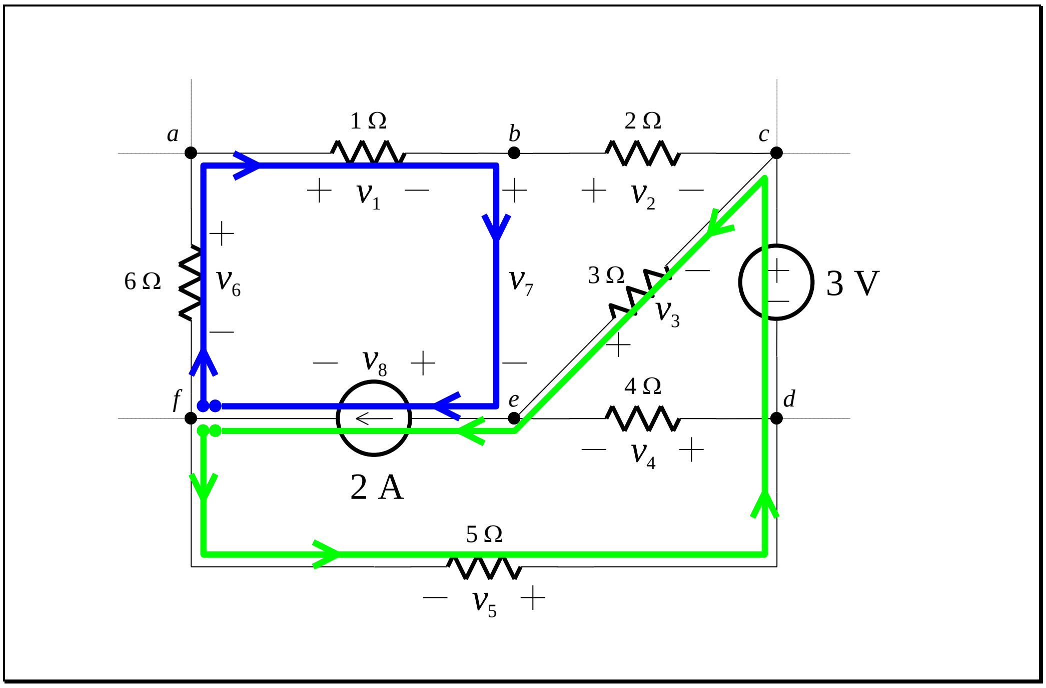


Figure 1.19

whereas the paths *becba* and *fde* are not:

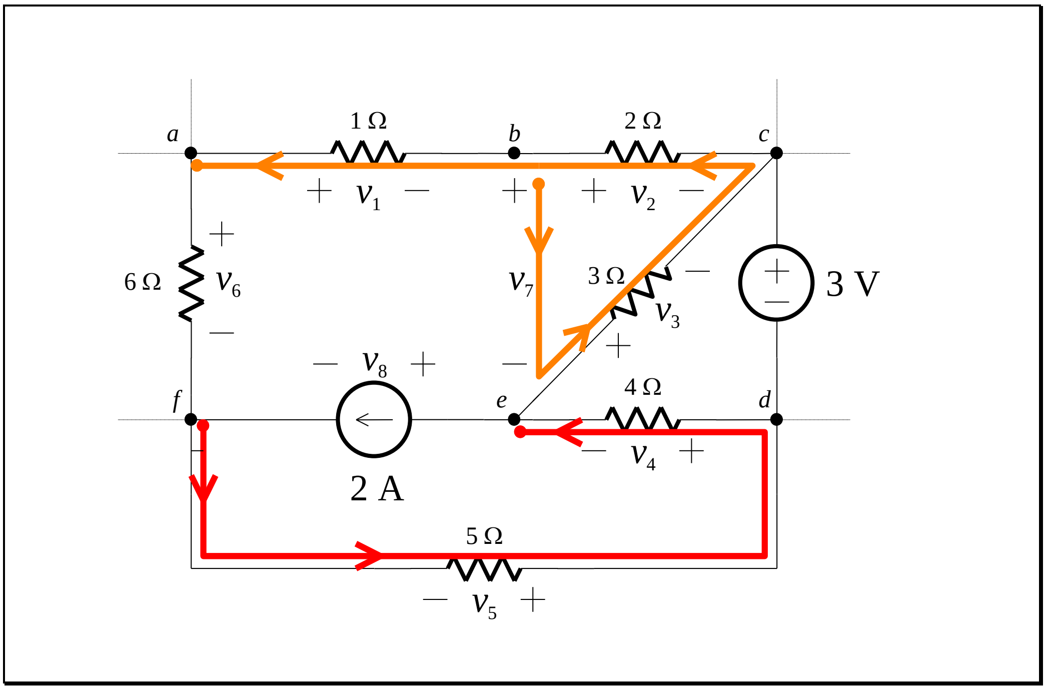


Figure 1.20

Kirchhoff's Voltage Law (KVL) is essentially the law of conservation of energy. If voltages across elements traversed from – to + are positive in sense, and voltages across elements that are traversed from + to – are negative in sense (or vice versa), then KVL can be stated as follows:

KVL defined

KVL: Around any loop in a circuit, the voltages algebraically sum to zero. (1.9)

If there are n elements in the loop then, in symbols, KVL is:

$$\sum_{k=1}^{n} v_k = 0 \tag{1.10}$$

KVL can also be stated as: In traversing a loop, the sum of the voltage rises equals the sum of the voltage drops.

EXAMPLE 1.7 Kirchhoff's Voltage Law Around a Loop

In the circuit shown in Figure 1.19, we select a traversal from – to + to be positive in sense. Then KVL around the loop *abcefa* gives:

$$-v_1 - v_2 + v_3 - v_8 + v_6 = 0$$

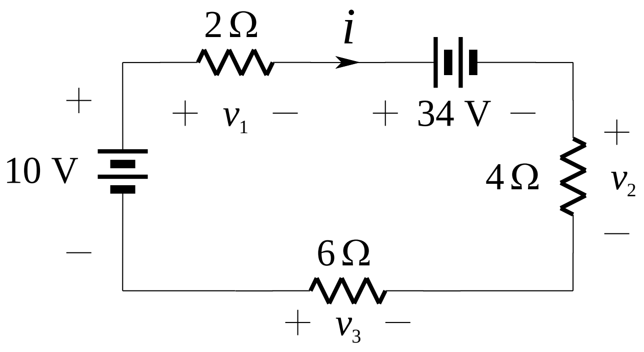
and around loop *bcdeb*, we have:

$$-v_2 - 3 - v_4 + v_7 = 0$$

In this last loop, one of the elements traversed (the element between nodes *b* and *e*) is an open-circuit; however, KVL holds regardless of the nature of the elements in the circuit.

EXAMPLE 1.8 Kirchhoff's Voltage Law Around a Circuit

We want to find the current *i*, in the one-loop circuit shown below:



The polarities of v_1, v_2, v_3 and the direction of *i* were chosen arbitrarily (the polarities of the 10 V and 34 V sources are given). Applying KVL we get:

$$10 - v_1 - 34 - v_2 + v_3 = 0$$

Thus:

$$v_1 + v_2 - v_3 = -24$$

From Ohm's Law:

$$v_1 = 2i \qquad v_2 = 4i \qquad v_3 = -6i$$

Substituting these into the previous equation yields:

$$\begin{aligned} (2i) + (4i) - (-6i) &= -24 \\ 2i + 4i + 6i &= -24 \\ 12i &= -24 \\ i &= -2\text{ A} \end{aligned}$$

Having solved for *i*, we now find that:

$$\begin{aligned} v_1 &= 2i = 2(-2) = -4\text{ V} \\ v_2 &= 4i = 4(-2) = -8\text{ V} \\ v_3 &= -6i = (-6)(-2) = 12\text{ V} \end{aligned}$$

1.9 Combining Resistors

Relatively complicated resistor combinations can be replaced by a single equivalent resistor whenever we are not specifically interested in the current, voltage or power associated with any of the individual resistors.

1.9.1 Series Resistors

Consider the series combination of N resistors shown in (a) below:

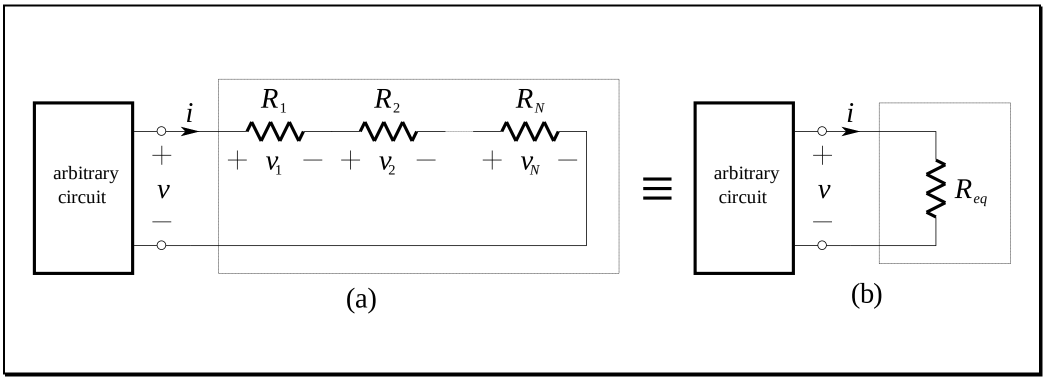


Figure 1.21

We apply KVL:

$$v = v_1 + v_2 + \cdots v_N \tag{1.11}$$

and Ohm's Law:

$$\begin{aligned} v &= R_1 i + R_2 i + \cdots R_N i \\ &= \left(R_1 + R_2 + \cdots R_N\right) i \end{aligned} \tag{1.12}$$

and then compare this result with the simple equation applying to the equivalent circuit shown in Figure 1.21b:

$$v = R_{eq} i \tag{1.13}$$

Thus, the value of the equivalent resistance for N series resistances is:

Combining series resistors

$$R_{eq} = R_1 + R_2 + \cdots + R_N \qquad \text{(series)} \tag{1.14}$$

1.9.2 Parallel Resistors

A similar simplification can be applied to parallel resistors. Consider the parallel combination of N conductances shown in (a) below:

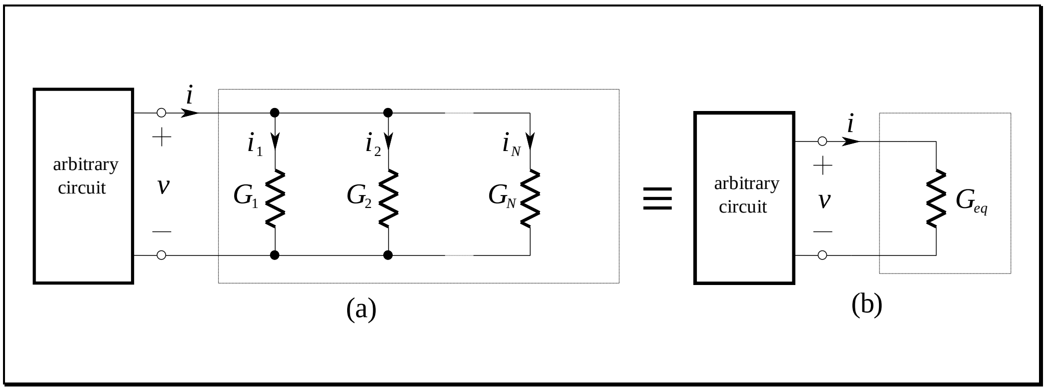


Figure 1.22

We apply KCL:

$$i = i_1 + i_2 + \cdots i_N \tag{1.15}$$

and Ohm's Law:

$$\begin{aligned} i &= G_1 v + G_2 v + \cdots G_N v \\ &= (G_1 + G_2 + \cdots G_N) v \end{aligned} \tag{1.16}$$

whereas the equivalent circuit shown in Figure 1.22b gives:

$$i = G_{eq} v \tag{1.17}$$

and thus the value of the equivalent conductance for N parallel conductances is:

$$G_{eq} = G_1 + G_2 + \cdots + G_N \quad \text{(parallel)} \tag{1.18}$$

Combining parallel conductances

In terms of resistance instead of conductance:

Combining parallel resistors

$$\frac{1}{R_{eq}} = \frac{1}{R_1} + \frac{1}{R_2} + \cdots + \frac{1}{R_N} \quad \text{(parallel)} \tag{1.19}$$

The special case of only two parallel resistors is needed often:

Combining two resistors in parallel…

$$R_{eq} = \frac{R_1 R_2}{R_1 + R_2} \quad \text{(parallel)} \tag{1.20}$$

Note that since $G_{eq} = G_1 + G_2$ then we may deduce that:

$$G_{eq} > G_1 \qquad \text{and} \qquad G_{eq} > G_2 \tag{1.21}$$

Hence:

$$\frac{1}{R_{eq}} > \frac{1}{R_1} \qquad \text{and} \qquad \frac{1}{R_{eq}} > \frac{1}{R_2} \tag{1.22}$$

or:

…results in an equivalent resistance smaller than either resistor

$$R_{eq} < R_1 \qquad \text{and} \qquad R_{eq} < R_2 \tag{1.23}$$

Thus the equivalent resistance of two resistors in parallel is less than the value of either of the two resistors.

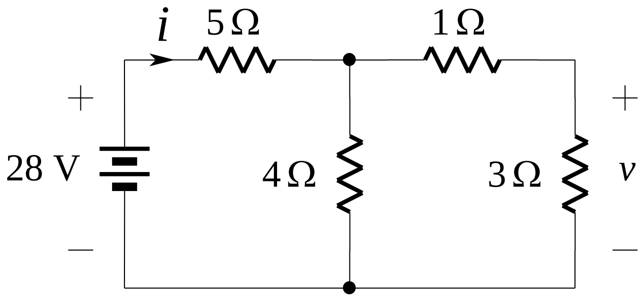
The special case of N resistors of equal value R in parallel is:

Combining the same valued resistors in parallel

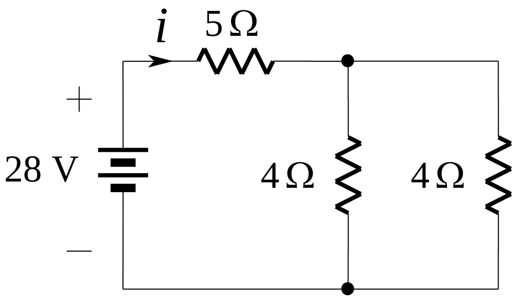
$$R_{eq} = \frac{R}{N} \quad \text{(parallel)} \tag{1.24}$$

EXAMPLE 1.9 Series and Parallel Resistors

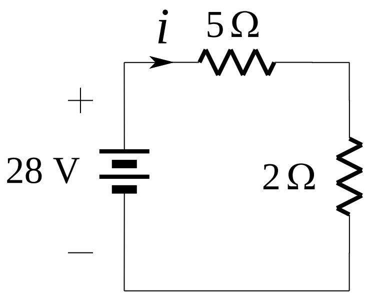
We want to find the current *i* in the circuit below:



In order to find *i*, we can replace series and parallel connections of resistors by their equivalent resistances. We begin by noting that the $1\,\Omega$ and $3\,\Omega$ resistors are in series. Combining them we obtain:



Note that it is not possible to display the original voltage *v* in this figure. Since the two $4\,\Omega$ resistors are connected in parallel, we can further simplify the circuit as shown below:



Here, the $5\,\Omega$ and $2\,\Omega$ resistors are in series, so we may combine them into one $7\,\Omega$ resistor. Then, from Ohm's Law, we have:

$$i = \frac{28}{7} = 4\text{ A}$$

1.10 Combining Independent Sources

An inspection of the KVL equations for a series circuit shows that the order in which elements are placed in a series circuit makes no difference. An inspection of the KCL equations for a parallel circuit shows that the order in which elements are placed in a parallel circuit makes no difference. We can use these facts to simplify voltage sources in series and current sources in parallel.

1.10.1 Combining Independent Voltage Sources in Series

It is not possible to combine independent voltage sources in parallel, since this would violate KVL. However, consider the series connection of two ideal voltage sources shown in (a) below:

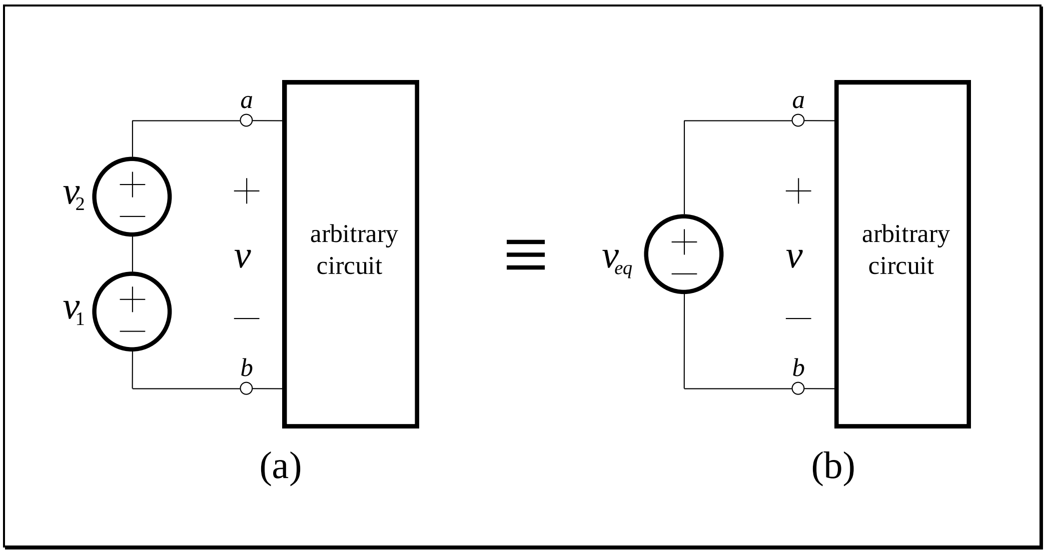


Figure 1.23

From KVL we know that $v = v_1 + v_2$, and by the definition of an ideal voltage source, this must be the voltage between nodes *a* and *b*, regardless of what is connected to them. Thus, the series connection of two ideal voltage sources is equivalent to a single independent voltage source given by:

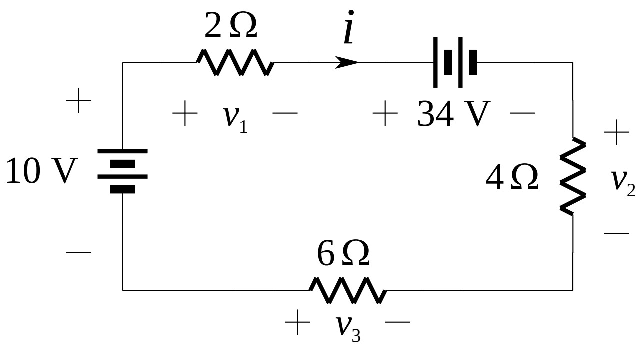
Combining independent voltage sources in series

$$v_{eq} = v_1 + v_2 \qquad \text{(series)} \tag{1.25}$$

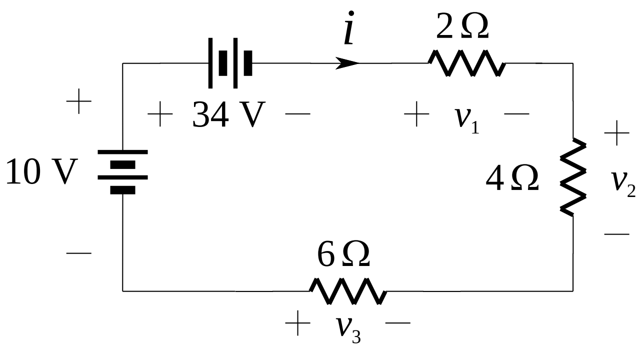
Clearly, the obvious generalization to *N* voltage sources in series holds.

EXAMPLE 1.10 Combining Independent Voltage Sources in Series

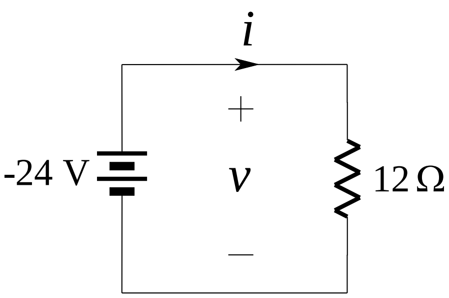
In a previous example we determined the current i in the one-loop circuit shown below:



By rearranging the order in this one loop circuit (of course this does not affect i), we obtain the circuit shown below:



We can now combine the series independent voltage sources and the series resistors into single equivalent elements:



By Ohm's Law:

$$i = \frac{-24}{12} = -2\ \text{A}$$

1.10.2 Combining Independent Current Sources in Parallel

It is not possible to combine independent current sources in series, since this would violate KCL. However, consider the parallel connection of two ideal current sources shown in (a) below:

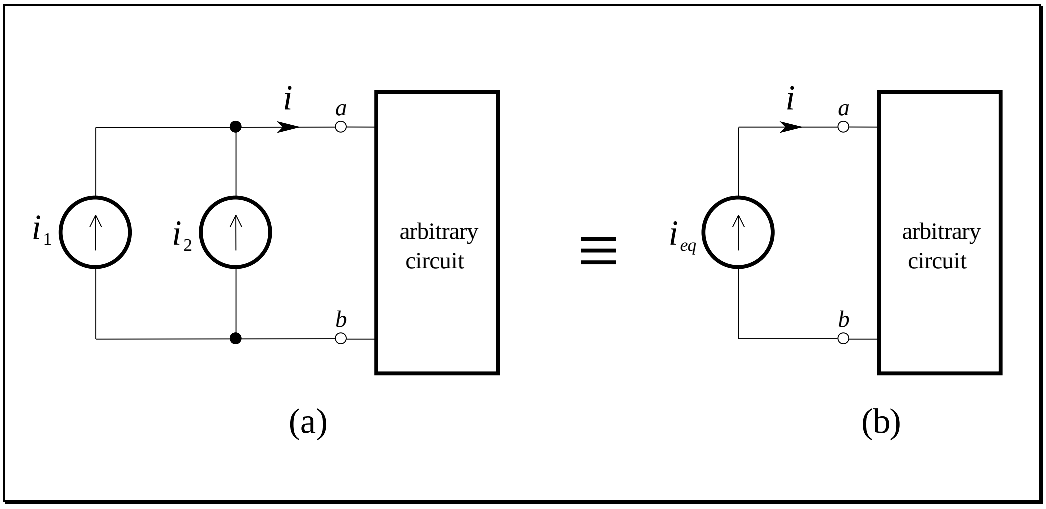


Figure 1.24

From KCL we find that $i = i_1 + i_2$, and by the definition of an ideal current source, this must always be the current into the arbitrary circuit. Thus, the parallel connection of two ideal current sources is equivalent to a single independent current source given by:

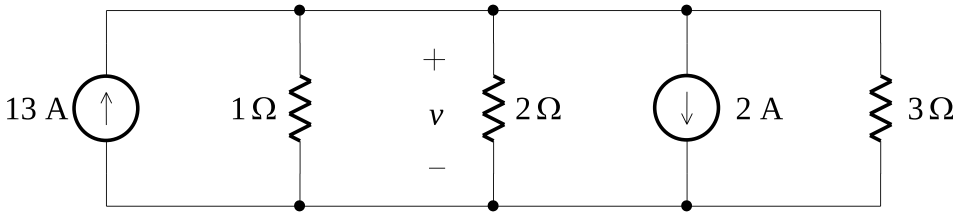
Combining independent current sources in parallel

$$i_{eq} = i_1 + i_2 \qquad \text{(parallel)} \tag{1.26}$$

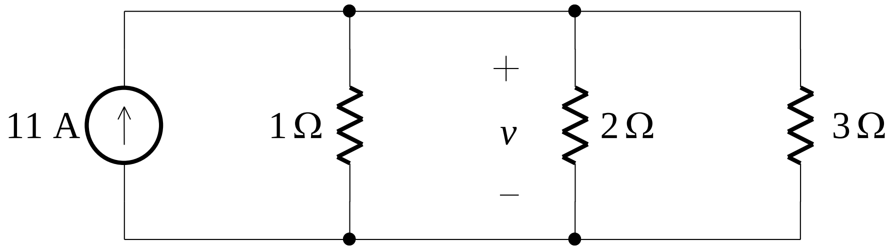
Clearly, the obvious generalization to N current sources in parallel holds.

EXAMPLE 1.11 Combining Independent Current Sources in Parallel

In a previous example, we determined the voltage v in the two-node circuit shown below:



Combining the parallel independent current sources into a single equivalent source, we obtain the circuit:



Since the equivalent resistance of the three resistors in parallel is given by:

$$\frac{1}{R_{eq}} = \frac{1}{1} + \frac{1}{2} + \frac{1}{3} = \frac{6+3+2}{6} = \frac{11}{6}$$

we obtain:

$$R_{eq} = \frac{6}{11}\,\Omega$$

Then, from Ohm's Law:

$$v = \frac{6}{11}(11) = 6\text{ V}$$

1.11 The Voltage Divider Rule

It can be quite useful to determine how a voltage appearing across two series resistors "divides" between them. Consider the circuit shown below:

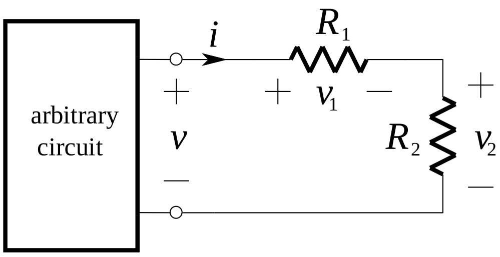


Figure 1.25

By Ohm's Law, the current in the resistors is:

$$i = \frac{v}{R_1 + R_2} \tag{1.27}$$

By application of Ohm's Law again, the voltage across R_1 is:

$$v_1 = R_1 i \tag{1.28}$$

and therefore:

Voltage divider rule defined

$$v_1 = \frac{R_1}{R_1 + R_2} v \tag{1.29}$$

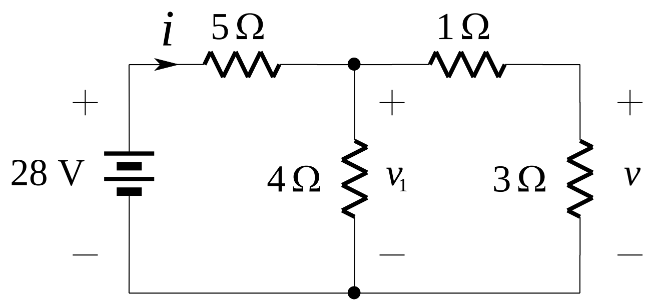
Similarly, the voltage across R_2 is:

$$v_2 = \frac{R_2}{R_1 + R_2} v \tag{1.30}$$

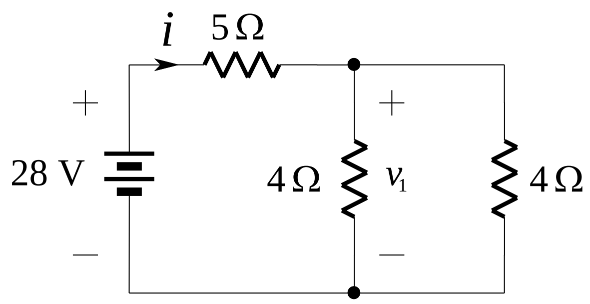
These equations describe how the voltage is divided between the resistors. Because of this, a pair of resistors in series is often called a *voltage divider*.

EXAMPLE 1.12 Voltage Divider Rule

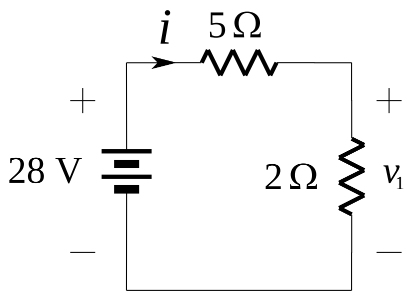
We want to find the voltage v in the circuit below:



Combining the series connection of the 1Ω and 3Ω resistors, we obtain the circuit below:



Now the pair of 4Ω resistors in parallel can be combined as shown below:



By voltage division:

$$v_1 = \frac{2}{2+5} \times 28 = \frac{56}{7} = 8\,\text{V}$$

Returning to the original circuit and applying voltage division again yields:

$$v = \frac{3}{3+1} v_1 = \frac{3}{4} \times 8 = 6\,\text{V}$$

1.12 The Current Divider Rule

It can be quite useful to determine how a current entering two parallel resistors "divides" between them. Consider the circuit shown below:

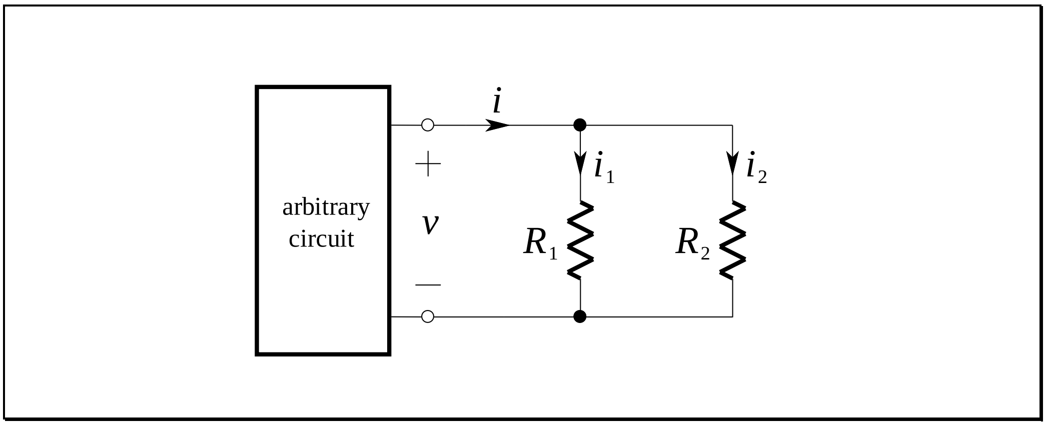


Figure 1.26

We replace the parallel connection of R_1 and R_2 by its equivalent resistance. Thus, Ohm's Law gives:

$$v = R_{eq} i = \frac{R_1 R_2}{R_1 + R_2} i \tag{1.31}$$

By application of Ohm's Law again, the current in R_1 is $i_1 = v / R_1$ and thus:

Current divider rule defined

$$i_1 = \frac{R_2}{R_1 + R_2} i \tag{1.32}$$

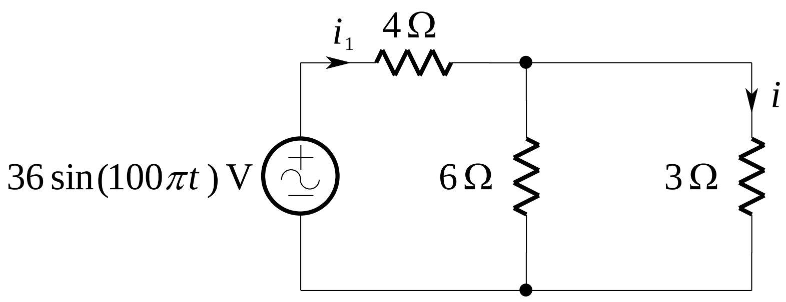
Similarly, the current in R_2 is:

$$i_2 = \frac{R_1}{R_1 + R_2} i \tag{1.33}$$

These equations describe how the current is divided between the resistors. Because of this, a pair of resistors in parallel is often called a *current divider*. Note that a larger amount of current will exist in the smaller resistor – thus current tends to take the path of least resistance!

EXAMPLE 1.13 Current Divider Rule

We want to find the current *i* in the circuit below:



The total current delivered by the source is:

$$\begin{aligned} i_1 &= \frac{36\sin(100\pi t)}{4+(6)(3)/(6+3)} \\ &= 6\sin(100\pi t)\,\text{A} \end{aligned}$$

Therefore the desired current is:

$$i = \frac{6}{6+3} i_1 = \frac{2}{3} \times 6\sin(100\pi t) = 4\sin(100\pi t)\,\text{A}$$

The current divider rule can also be derived using conductances. Referring to Figure 1.26, the voltage across the parallel resistors is:

$$v = R_{eq} i = \frac{i}{G_{eq}} = \frac{i}{G_1 + G_2} \tag{1.34}$$

The current in resistor R_1 is $i_1 = G_1 v$ and thus:

$$i_1 = \frac{G_1}{G_1 + G_2} i \tag{1.35}$$

A similar result obviously holds for current i_2. The advantage of this form of the current divider rule is that it is the *dual* of the voltage divider rule – we replace voltages with currents, and resistors with conductances.

1.13 Dependent Sources

An ideal source, either voltage or current, whose value depends upon some parameter (usually a voltage or current) in the circuit to which the source belongs is known as a *dependent* or *controlled* source.

1.13.1 The Dependent Voltage Source

Dependent voltage sources defined

A dependent voltage source establishes a voltage across its terminals, independent of the current through it, which is determined by the voltage or current at some other location in the electrical system. There are two types of dependent voltage source – the voltage-controlled voltage source (VCVS) and the current-controlled voltage source (CCVS).

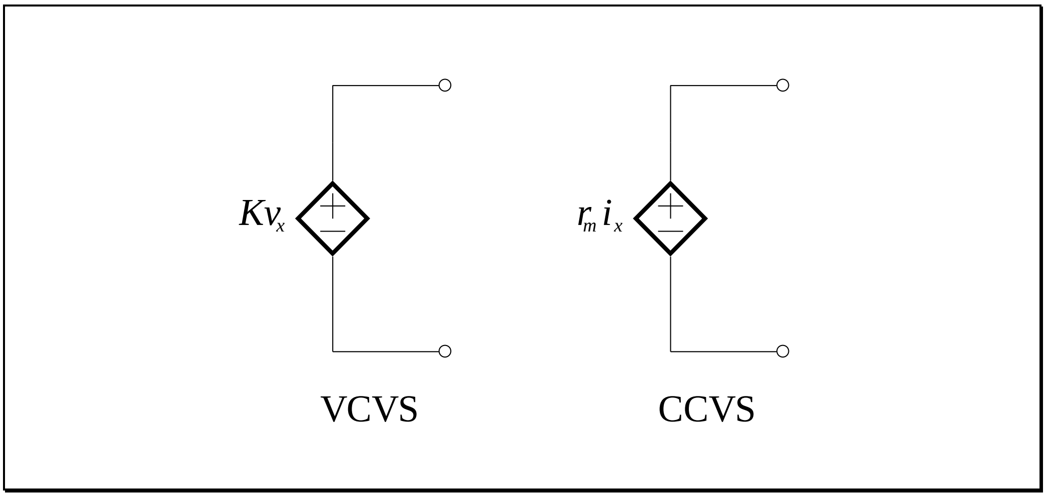


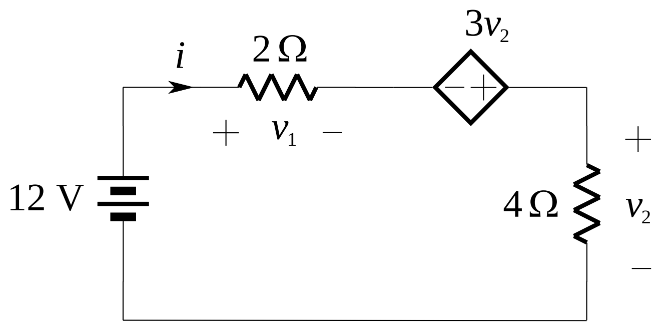
Figure 1.27

Note that the dependent source is represented by a diamond-shaped symbol so as not to confuse it with an independent source.

These sources are mathematical models that are useful in modelling real circuits and systems, e.g. they are used in modelling operational amplifiers.

EXAMPLE 1.14 Circuit with a Dependent Voltage Source

Consider the circuit shown below. This circuit contains a dependent source whose value in this case depends on the voltage across the 4 Ω resistor – it is a VCVS.



To analyse the circuit, we apply KVL and obtain:

$$v_1 - 3v_2 + v_2 = 12$$

or:

$$v_1 - 2v_2 = 12$$

By Ohm's Law:

$$v_1 = 2i \qquad \text{and} \qquad v_2 = 4i$$

Therefore:

$$\begin{aligned} 2i - 2(4i) &= 12 \\ 2i - 8i &= 12 \\ -6i &= 12 \\ i &= -2\text{ A} \end{aligned}$$

Hence:

$$v_2 = 4i = -8\text{ V}$$

and the value of the dependent voltage source is:

$$3v_2 = -24\text{ V}$$

1.13.2 The Dependent Current Source

Dependent current sources defined

A dependent current source establishes a current, which is independent of the voltage across it, that is determined by the voltage or current at some other location in the electrical system. There are two types of dependent current source – the voltage-controlled current source (VCCS) and the current-controlled current source (CCCS).

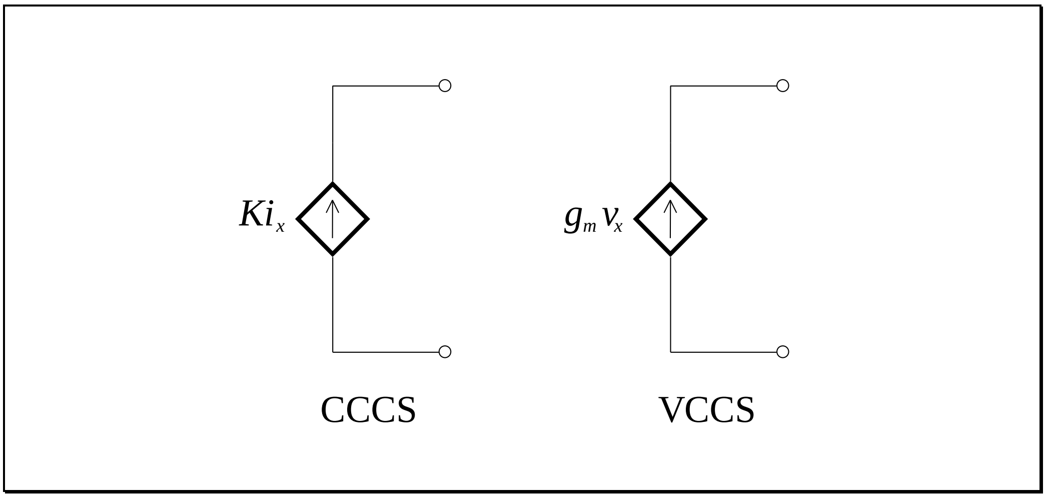
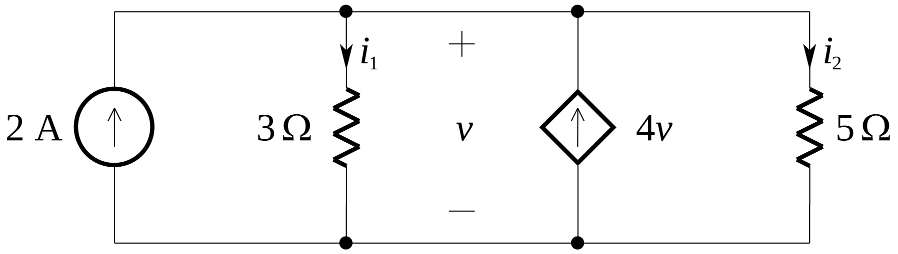


Figure 1.28

These sources are mathematical models that are useful in modelling real circuits and systems, e.g. they are used in modelling transistors.

EXAMPLE 1.15 Circuit with a Dependent Current Source

Consider the circuit shown below. In this circuit the value of the dependent current source is specified by a voltage – it is a VCCS.



To solve for *v*, we apply KCL and obtain:

$$i_1 + i_2 = 4v + 2$$

Thus:

$$\begin{aligned} \frac{v}{3} + \frac{v}{5} &= 4v + 2 \\ -4v + \frac{8v}{15} &= 2 \\ \frac{-52v}{15} &= 2 \\ v &= \frac{-30}{52} = -\frac{15}{26}\text{ V} \end{aligned}$$

Consequently:

$$4v = -\frac{30}{13}\text{ A}$$

and this is the value of the dependent current source, in amperes. The other variables in the circuit are:

$$i_1 = \frac{v}{3} = -\frac{5}{26}\text{ A} \quad \text{and} \quad i_2 = \frac{v}{5} = -\frac{3}{26}\text{ A}$$

1.14 Power

Power is the rate at which work is done or energy is expended. Taking the product of voltage (energy per unit charge) and current (charge per unit time) we get a quantity that measures energy per unit time. It's for this reason that we define p, the *instantaneous power absorbed* by an electrical circuit element, to be the product of voltage and current:

Instantaneous power defined

$$p = vi \tag{1.36}$$

The fundamental unit of power is the watt (W) and is equivalent to Js^{-1}. In using the formula for instantaneous power, we need to be careful in establishing the correct voltage polarity and current direction. Consider the circuit element:

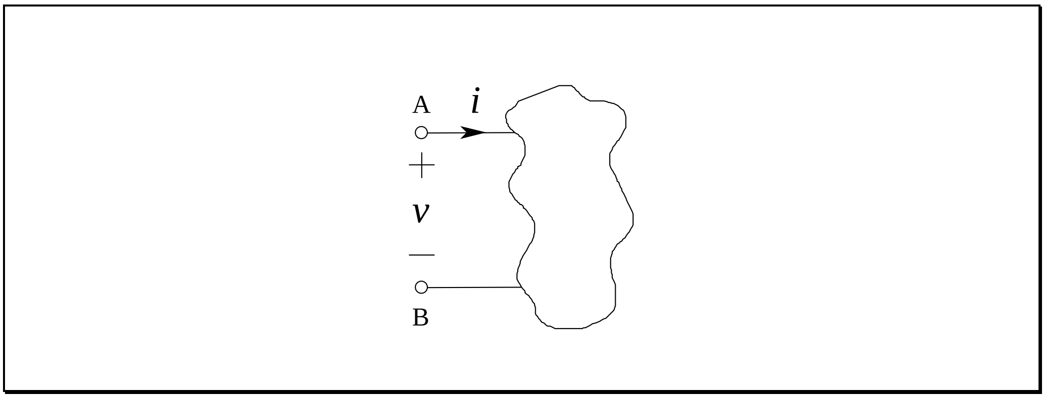


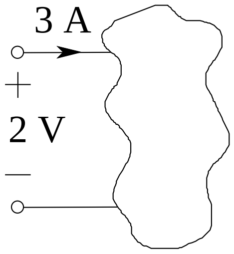
Figure 1.29

Passive sign convention defined – it gives power absorbed by a circuit element

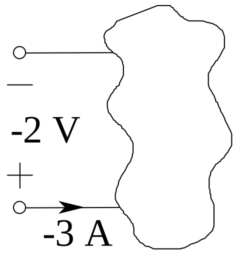
If one terminal of the element (A) is v volts positive with respect to the other terminal (B), and if a current i is entering the element through terminal A, then a power $p = vi$ is being *absorbed* or *delivered* to the element. When the current arrow is directed into the element at the plus-marked terminal, we satisfy the *passive sign convention*. If the numerical value of the power using this convention is negative, then we say that the element is *generating* or *delivering* power.

EXAMPLE 1.16 Power Absorbed

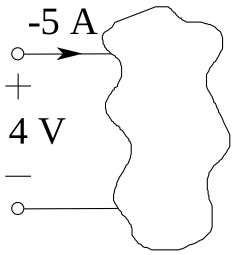
Note the power in the circuit elements below:



$p = (2)(3) = 6\text{ W}$ absorbed



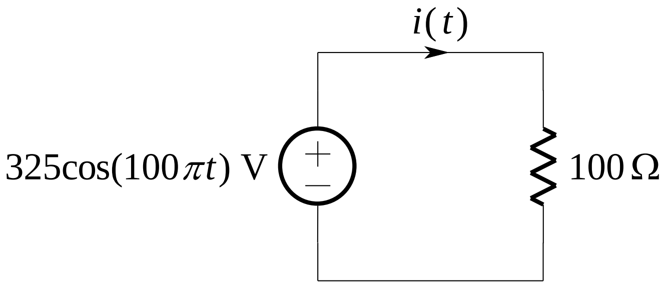
$p = (-2)(-3) = 6\text{ W}$ absorbed



$p = (4)(-5) = -20\text{ W}$ absorbed
(20 W generated)

EXAMPLE 1.17 Power Absorbed by a Resistor

Consider the circuit shown below:



By Ohm's Law:

$$i(t) = \frac{v(t)}{R} = \frac{325}{100}\cos(100\pi t) \text{ A}$$

By definition, the power absorbed by the resistor is:

$$\begin{aligned} p_R(t) &= v(t)i(t) \\ &= \frac{325^2}{100}\cos^2(100\pi t) \\ &= 1056\cos^2(100\pi t) \text{ W} \end{aligned}$$

In particular, at time $t = 0$ the power absorbed by the resistor is:

$$\begin{aligned} p_R(0) &= 1056\cos^2(0) \\ &= 1056 \text{ W} \end{aligned}$$

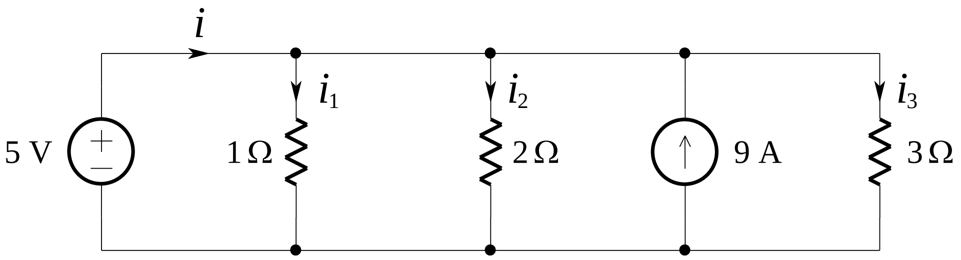
At time $t = 5\,\text{ms}$, however, since:

$$\cos\left(100\pi \times 5 \times 10^{-3}\right) = \cos\left(\frac{\pi}{2}\right) = 0$$

then the resistor absorbs 0 watts.

EXAMPLE 1.18 Power Absorbed by Circuit Elements

Consider the circuit shown below:



We shall determine the power absorbed in each of the elements.

Note that the voltage across each of the elements is 5 V since all the elements are in parallel. Therefore, by Ohm's Law:

$$i_1 = \frac{5}{1} = 5\text{ A} \qquad i_2 = \frac{5}{2}\text{ A} \qquad i_3 = \frac{5}{3}\text{ A}$$

and the powers absorbed in the $1\,\Omega$, $2\,\Omega$ and $3\,\Omega$ resistors are:

$$p_1 = 5i_1 = 5(5) = 25\text{ W}$$

$$p_2 = 5i_2 = 5\left(\frac{5}{2}\right) = \frac{25}{2}\text{ W}$$

$$p_3 = 5i_3 = 5\left(\frac{5}{3}\right) = \frac{25}{3}\text{ W}$$

respectively, for a total of:

$$25 + \frac{25}{2} + \frac{25}{3} = \frac{150 + 75 + 50}{6} = \frac{275}{6}\text{ W}$$

absorbed by the resistors.

1.48

By KCL:

$$i+9=i_1+i_2+i_3=\frac{5}{1}+\frac{5}{2}+\frac{5}{3}$$

or:

$$i=\frac{30+15+10}{6}-9=\frac{1}{6}\text{ A}$$

Thus the power delivered by the voltage source is:

$$p_v=5i=\frac{5}{6}\text{ W}$$

Also, the power delivered by the current source is:

$$p_i=5(9)=45\text{ W}$$

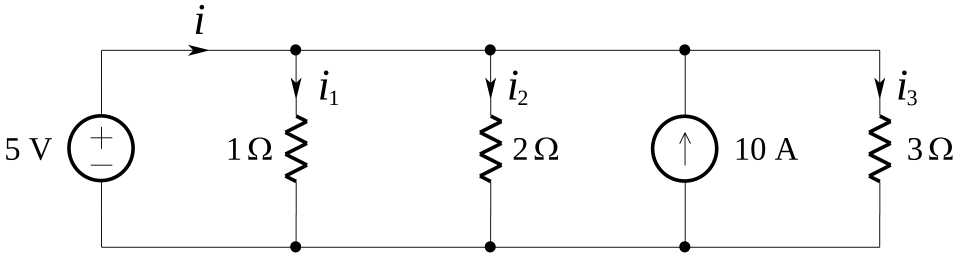
Hence the total power delivered by the sources is:

$$45+\frac{5}{6}=\frac{270+5}{6}=\frac{275}{6}\text{ W}$$

We see that the total power delivered by the sources is equal to the total power absorbed by the resistors. Since power delivered by a circuit element is equal to the negative of the power absorbed, this is equivalent to saying that the total power absorbed by all circuit elements is zero. Thus, the principle of conservation of energy (and therefore power) is satisfied in this circuit (as it is in any circuit).

EXAMPLE 1.19 Power Conservation

Consider the circuit shown below, which is identical to the previous example except for the value of the current source:



In this case:

$$p_1 = 25\,\text{W} \qquad p_2 = \frac{25}{2}\,\text{W} \qquad p_3 = \frac{25}{3}\,\text{W}$$

as before. By KCL, however:

$$i + 10 = i_1 + i_2 + i_3$$

and thus:

$$i = \frac{55}{6} - 10 = -\frac{5}{6}\,\text{A}$$

Therefore, the powers delivered by the sources are:

$$p_v = 5\left(-\frac{5}{6}\right) = -\frac{25}{6}\,\text{W} \qquad p_i = 5(10) = 50\,\text{W}$$

Hence the total power absorbed is:

$$p_1 + p_2 + p_3 + p_v + p_i = \frac{275}{6} + \frac{25}{6} - 50 = 0\,\text{W}$$

and again energy (power) is conserved. However, in this case not only do the resistors absorb power, but so does the voltage source. It is the current source that supplies all the power absorbed in the rest of the circuit.

1.14.1 Power Absorbed in a Resistor

The power absorbed in every resistor is always a nonnegative number. Consider the resistor shown below:

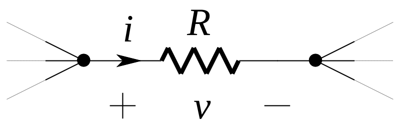


Figure 1.30

By definition, the power absorbed in the resistor is $p = vi$. But by Ohm's Law, $v = Ri$. Thus $p = (Ri)i$, or:

$$p = Ri^2 \tag{1.37}$$

Also, $i = v/R$, so that $p = v(v/R)$, or:

$$p = \frac{v^2}{R} \tag{1.38}$$

A real resistor always absorbs power

Both formulas for calculating power absorbed in a resistor R demonstrate that p is always a nonnegative number when R is positive. Therefore a resistor always absorbs power.

In a physical resistor, this power is dissipated as radiation (light and/or heat). In some types of resistors (such as an incandescent bulb, a toaster, or an electric heater), this property is desirable in that the net result may be light or warmth. In other types of resistors, such as those found in electronic circuits, the heat dissipated in a resistor may be a problem that cannot be ignored.

Real resistors have a power rating that must not be exceeded

The physical size of a resistor determines the amount of power it can safely dissipate. A power dissipation that exceeds the rating of a resistor can physically damage the resistor. In many electronic applications, resistors need dissipate only small amounts of power, allowing their use in integrated circuits.

1.15 Summary

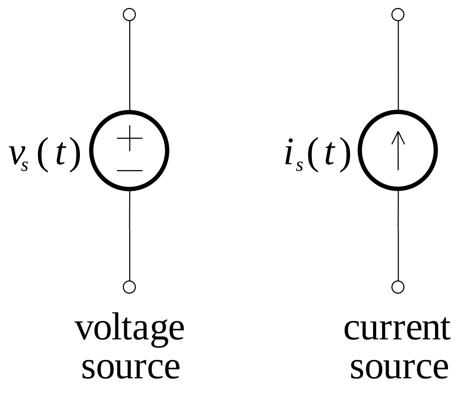
- Current is defined as the rate of flow of charge past a certain cross-sectional area:

$$i = \frac{dq}{dt}$$

- Voltage is defined as the work done per unit charge in moving it from one point to another in a circuit.

- A circuit element is an *idealised* mathematical model of a two-terminal electrical device that is completely characterised by its voltage-current relationship. Active circuit elements *can* deliver a non-zero average power indefinitely, whilst passive circuit elements *cannot*. A connection of circuit elements is called a *network*. If the network contains at least one closed path, it is also an electrical *circuit*.

- Independent sources are *ideal* circuit elements that possess a voltage or current value that is independent of the behaviour of the circuits to which they belong. There are two types, voltage and current:



- The resistor is a linear passive circuit element that obeys Ohm's Law:

$$v = Ri$$

A resistance of $0\,\Omega$ is known as a *short-circuit*.

A resistance of $\infty\,\Omega$ is known as an *open-circuit*.

The reciprocal of resistance is called the *conductance*:

$$G = \frac{1}{R}$$

- Practical resistors come in a large variety of shapes, materials and construction which dictate several of their properties, such as accuracy, stability, pulse handling capability, resistor value, size and cost.

- Kirchhoff's Current Law (KCL) states: "At any node of a circuit, the currents algebraically sum to zero":

$$\sum_{k=1}^{n} i_k = 0$$

- Kirchhoff's Voltage Law (KVL) states: "Around any loop in a circuit, the voltages algebraically sum to zero":

$$\sum_{k=1}^{n} v_k = 0$$

- Resistors in *series* can be combined into one equivalent resistor:

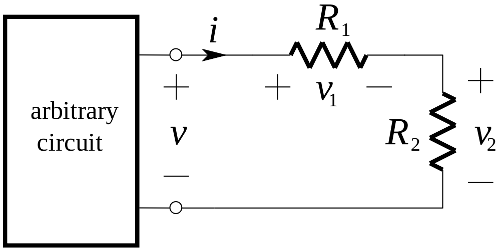
$$R_{eq} = R_1 + R_2 + \cdots + R_N$$

- Resistors in *parallel* can be combined into one equivalent resistor:

$$\frac{1}{R_{eq}} = \frac{1}{R_1} + \frac{1}{R_2} + \cdots + \frac{1}{R_N}$$

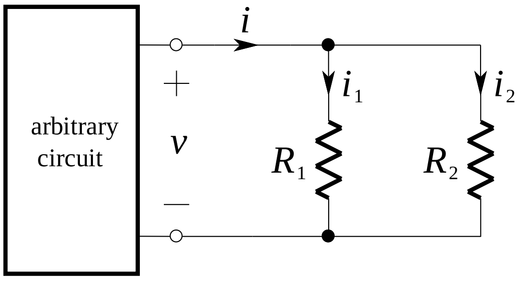
- Independent voltage sources in *series* can be added. Independent current sources in *parallel* can be added.

- Two resistors in series form a voltage divider:



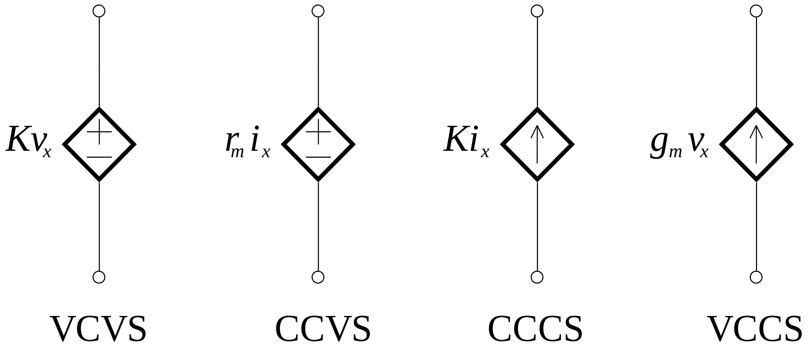
The voltage divider rule is: $$v_2 = \frac{R_2}{R_1 + R_2} v$$

- Two resistors in parallel form a current divider:



The current divider rule is: $$i_2 = \frac{R_1}{R_1 + R_2} i$$

- An ideal source, either voltage or current, whose value depends upon some parameter (usually a voltage or current) in the circuit to which the source belongs is known as a *dependent* or *controlled* source. There are four types:



- The *instantaneous power absorbed* by an electrical circuit element is the product of voltage and current:

$$p = vi$$

The power absorbed in a resistor is:

$$p = Ri^2 = \frac{v^2}{R}$$

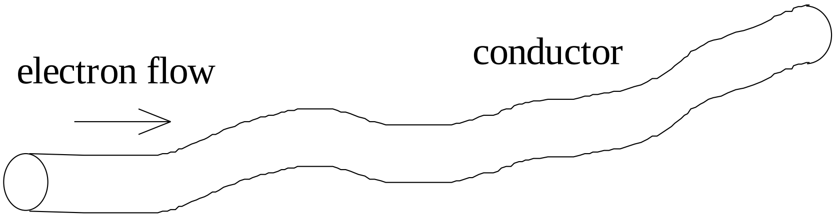
1.16 References

Hayt, W. & Kemmerly, J.: *Engineering Circuit Analysis*, 3rd *Ed.*, McGraw-Hill, 1984.

Exercises

1.

A large number of electrons are moving through a conductor:



The number varies with time *t* seconds.

(a) What is the direction of current?

(b) If the total charge to pass a certain point on the conductor varies according to the equation:

$$q(t) = 3\left(1 - e^{-100t}\right) \quad \text{mC}$$

then find the current in amperes as a function of time.

(c) When will the current be 200 mA?

(d) If the conductor has a uniform diameter of 1 mm throughout its length, find the current density as a function of time (express as A/mm^2).

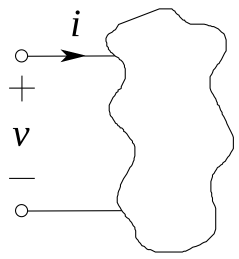
(e) Sketch charge and current as functions of time.

(f) How many electrons are moving through the conductor at time t = 50 ms?

1.56

2.

The total charge that has entered the upper terminal of the element below is given by $5\sin 1000\pi t$ μC.



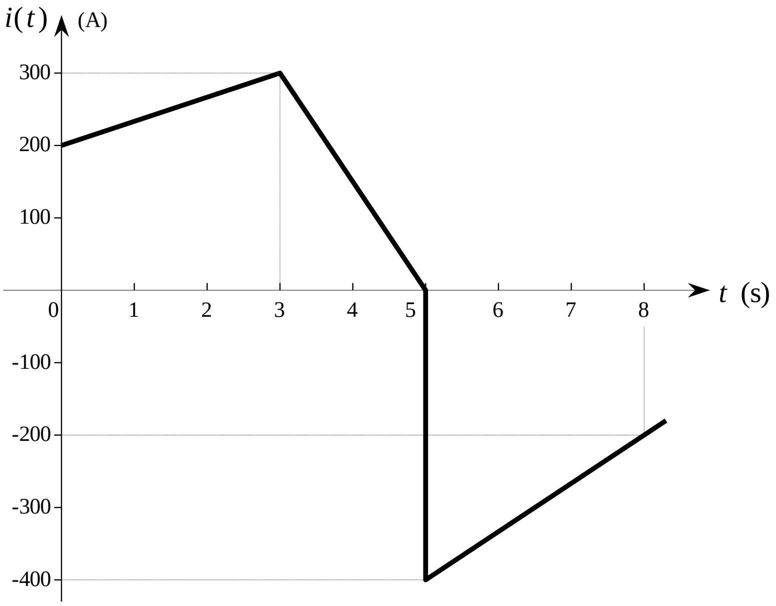
(a) How much charge enters that terminal between $t = -0.5\,\text{ms}$ and $t = 0.5\,\text{ms}$?

(b) How much charge leaves the lower terminal in the same time interval?

(c) Find i at $t = 0.2\,\text{ms}$.

3.

For the current waveform shown below:



determine the total charge transferred between $t = 0$ and $t =$:

(a) 4 s (b) 7 s

4.

The charging current supplied to a 12 V automotive battery enters its positive terminal. It is given as a function of time by:

$$i = \begin{cases} 0 & t < 0 \\ 4e^{-t/10000}\ \text{A} & 0 \le t \le 15000\ \text{s} \\ 0 & t > 15000\ \text{s} \end{cases}$$

(a) What is the total charge delivered to the battery in the 15000 s charging interval?

(b) What is the maximum power absorbed by the battery?

(c) What is the total energy supplied?

(d) What is the average power delivered in the 15000 s interval?

5.

The voltage v has its positive reference at terminal A of a certain circuit element. The power absorbed by the circuit element is $4(t-1)^2$ W for $t > 0$. If $v = (2t-2)$ V for $t > 0$, how much charge enters terminal A between $t = 0$ and $t = 2$ s?

6.

The resistance of 10 mm^2 copper wire is 1.725 Ω/km, and, with a certain type of insulation, it can safely carry 70 A without overheating. With a one kilometre length of wire operating at maximum current:

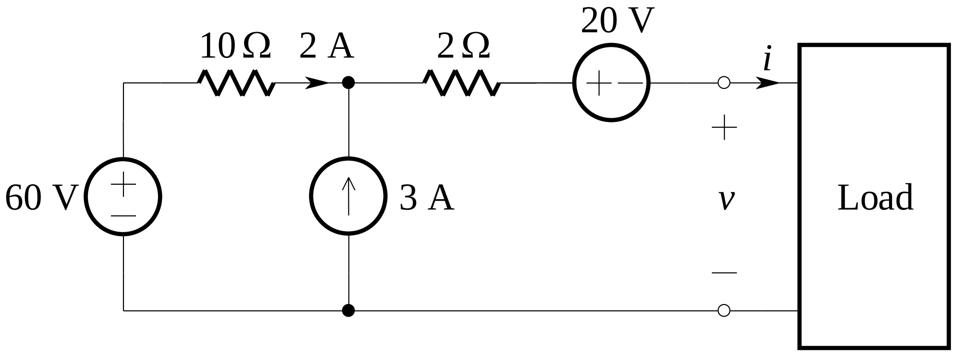
(a) What voltage exists between the ends of the wire?

(b) How much power is dissipated in the conductor?

(c) What is the power dissipation per square mm of surface area?

7.

For the circuit shown below:

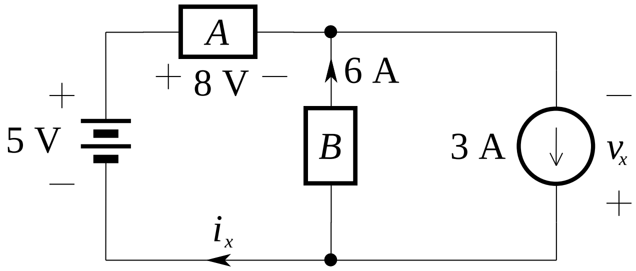


find:

(a) v (b) i (c) the power absorbed by the load.

8.

With reference to the network shown below:

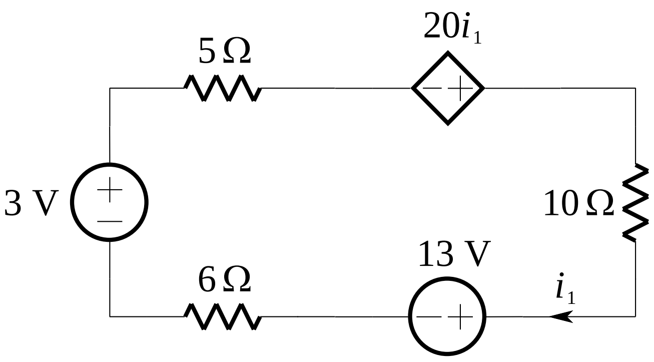


find:

(a) i_x (b) v_x (c) the power absorbed by the battery.

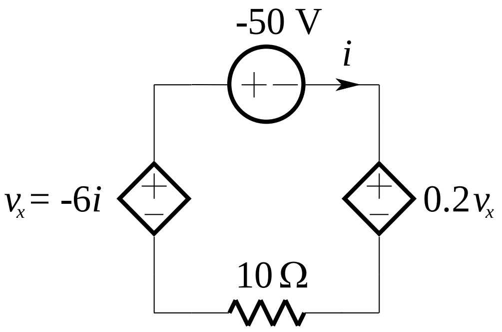
9.

Find the power supplied by the 3 V source in the circuit below:



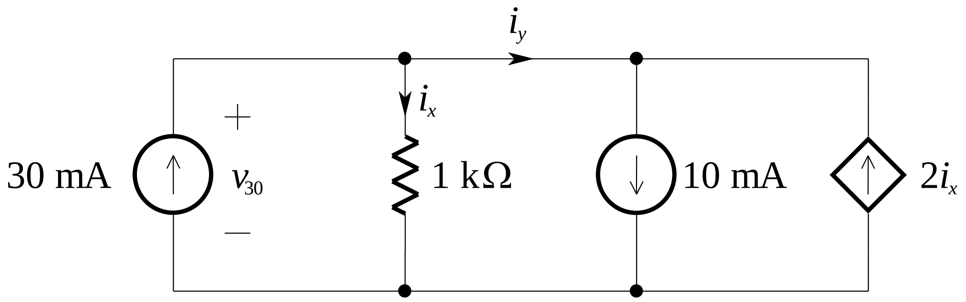
10.

Find the power absorbed by each element in the circuit below:



11.

Consider the circuit shown below:

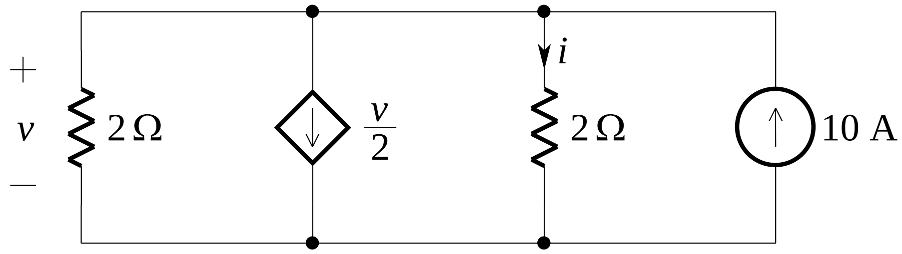


(a) Find v_{30}, i_x, and i_y.

(b) Change the control on the dependent source from $2i_x$ to $2i_y$ and then find v_{30}, i_x, and i_y.

12.

Consider the circuit shown below:



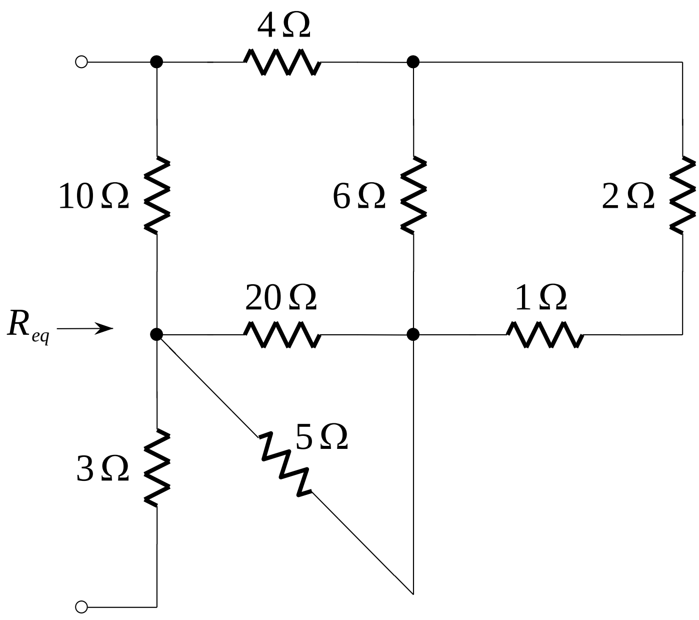
(a) Find v, i, and the power delivered by the independent source.

(b) Repeat if the arrow of the dependent source is reversed.

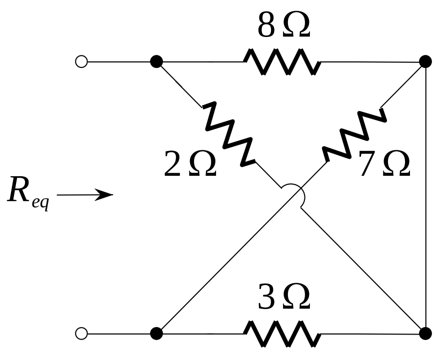
13.

Find R_{eq} for each of the networks shown below:

(a)



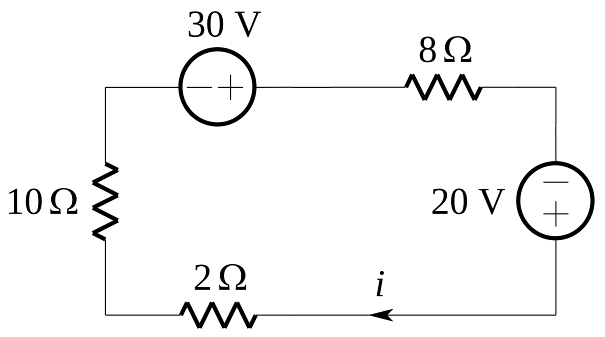
(b)



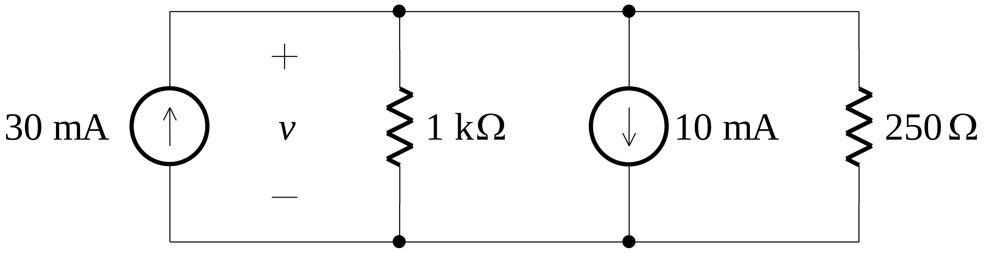
14.

By combining independent sources and resistances as appropriate, find:

(a) The current i in the circuit below.

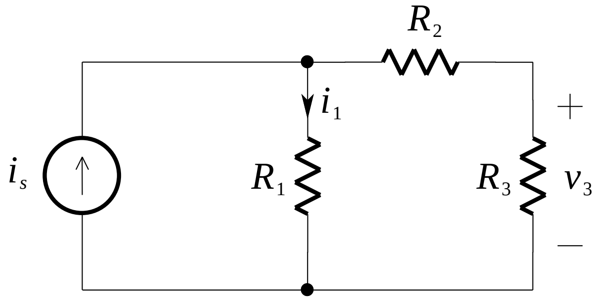


(b) The voltage v in the circuit below.



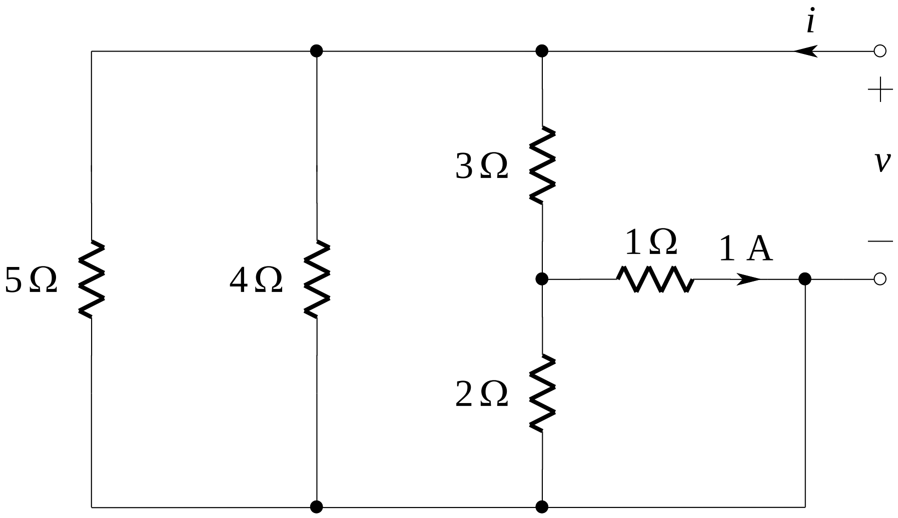
15.

Use the concepts of current division, voltage division, and resistance combination to write expressions (by inspection) for v_3 and i_1 in the circuit shown below:



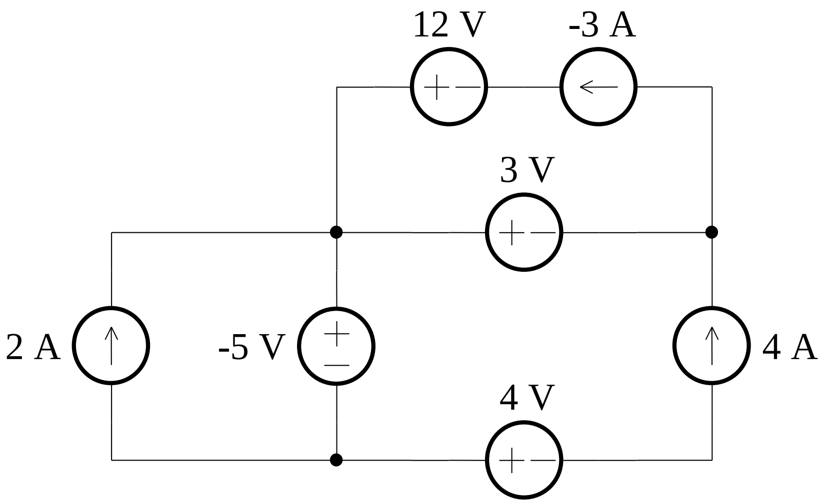
16.

Determine the necessary values of *v* and *i* in the circuit shown below:



17.

The circuit shown below exhibits several examples of independent current and voltage sources in series and in parallel.



(a) Find the power supplied by the -5 V source.

(b) To what value should the 4 A source be changed to reduce the power supplied by the -5 V source to zero?

2 Amplifiers

Contents

2.2

Introduction

Amplifiers play a fundamental role in analog electronic circuits. For example, an electronic amplifier can play a salient role in amplifying an ultrasound wave from deep within a human body, or measure and help reduce the error of a feedback system, or help translate signals rapidly and accurately between the world of tangible physics and the world of abstract digits.

We will look at some basic amplifier characteristics (such as “gain”) and see how they are used at a “block diagram” level. We will also look at some real limitations of amplifiers, such as the fact they need a power supply to operate.

A device called an operational amplifier (op-amp for short), which has been used as the building block for modern analog electronic circuitry since its invention and widespread adoption in the 1960’s, is the principal piece of technology underlying analog electronic circuits.

We will examine the characteristics of the op-amp, how it works, and how it is used within a circuit. In most cases, we will use an ideal model of the op-amp which is suitable for hand-analysis and design work. We will see that the op-amp is the versatile building block of a whole range of anlaog circuits.

2.1 Amplifiers

A linear amplifier is a device that increases the amplitude of a signal (a voltage or a current) whilst preserving waveform shape. The circuit symbol for an amplifier is a triangle which clearly shows the direction of signal travel.

An example of a voltage amplifier is shown below:

The amplifier's circuit symbol and its transfer characteristic

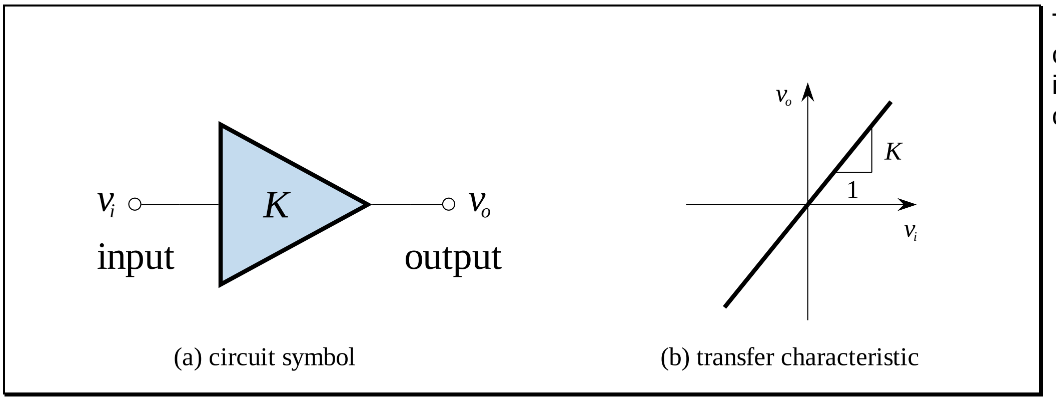


Figure 2.1

Note the use of subscript notation: *i* for input, and *o* for output. The relationship between the input and output for the voltage amplifier is:

The input-output relationship for an ideal linear voltage amplifier

$$v_o = Kv_i \tag{2.1}$$

The quantity K is referred to as the *gain*. If the gain is a positive number, then the amplifier is said to be *non-inverting*. If the gain is a negative number, then the amplifier is said to be *inverting*. Note that a negative number does not imply a decrease in the signal – it implies an inversion.

Since the purpose of an amplifier is to increase signal amplitude, we normally have $|K| > 1$. Circuits with $|K| < 1$ are said to *attenuate*, but as we shall see, they may still be implemented with an "amplifier".

Amplifiers are used in numerous places and form one of the basic building-blocks of electronic circuits. For example, signals in telecommunications that come from antennas are particularly "weak" and could be in the microvolt or millivolt range. Reliable processing of these small signals is made easier if the signal magnitude is much larger.

2.1.1 Units of Gain

The gain K of a voltage amplifier can be expressed in two ways. The first way is as a straight voltage ratio, with units "volts per volt":

Amplifier gain expressed in volts per volt

$$K = \frac{v_o}{v_i} \text{ V/V} \tag{2.2}$$

Note that K must be a dimensionless quantity.

The second way comes from the historical development of amplifiers which were first used extensively throughout telecommunication systems. In these applications, since signals were audio in nature, it became common to compare signal amplitudes in terms of the audio power they could deliver. Thus, we can express the voltage gain with units of decibels:[1]

Amplifier gain expressed in decibels

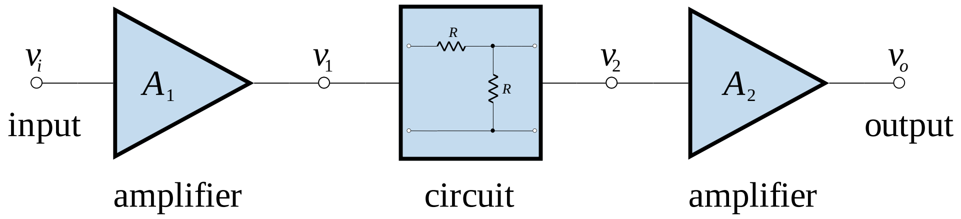
$$|K| = 20\log_{10}\left|\frac{v_o}{v_i}\right| \text{ dB} \tag{2.3}$$

The dB unit of voltage gain is useful when circuits are *cascaded* – a cascade occurs when the output of one circuit is fed into the input of another (and it has been ensured, through careful design, that one circuit does not "load" the next, i.e. each individual circuit's behaviour is independent of the load placed on it). For cascaded circuits you can add the voltage gains in dB instead of multiplying the standard voltage gains.

[1] Historically the Bel (named after Alexander Graham Bell – the inventor of the first commercially viable and practical telephone) was used to define ratios of audio loudness i.e. ratios of power. In the metric system, a convenient unit to use is the decibel (dB): $1\,\text{decibel} = 10\log_{10}(P_o/P_i)$. If electrical power is assumed to be dissipated across equal resistors, then since $P = V^2/R$, the power ratio is $20\log_{10}|V_o/V_i|$. This power ratio became a way to express the voltage gain of amplifiers. Note that the decibel is dimensionless, so it can be applied to any dimensionless ratio, if one wished.

EXAMPLE 2.1 Cascaded Amplifiers and Circuits

A cascade of amplifiers and circuits is shown below, with the gain (or attenuation) expressed in V/V and dB.



$$\frac{v_1}{v_i} = A_1 = -10 \text{ V/V} \qquad \frac{v_2}{v_1} = K = \tfrac{1}{2} \text{ V/V} \qquad \frac{v_o}{v_2} = A_2 = 20 \text{ V/V}$$

$$|A_1| = 20 \text{ dB} \qquad |K| = -6 \text{ dB} \qquad |A_2| = 26 \text{ dB}$$

The overall gain can be expressed as:

$$\frac{v_o}{v_i} = \frac{v_1}{v_i}\frac{v_2}{v_1}\frac{v_o}{v_2} = A_1 K A_2 = -10 \cdot \tfrac{1}{2} \cdot 20 = -100 \text{ V/V}$$

Note that when the gain in V/V is negative, then the signal is inverted.

The gain expressed in decibels is:

$$\begin{aligned}\left|\frac{v_o}{v_i}\right| &= 20\log_{10}|A_1 K A_2| \\ &= 20\log_{10}|A_1| + 20\log_{10}|K| + 20\log_{10}|A_2| \\ &= 20 - 6 + 26 \\ &= 40 \text{ dB}\end{aligned}$$

As a check, we can perform the dB calculation directly on the overall V/V gain:

$$\left|\frac{v_o}{v_i}\right| = 20\log_{10}|-100| = 20 \times 2 = 40 \text{ dB}$$

When the gain is expressed in dB, it refers to the *magnitude* of the gain only – it conveys no phase information. When the gain in dB is a negative number, then we have attenuation.

2.1.2 Amplifier Power Supplies

Amplifiers require a separate and independent power supply to operate. Some amplifiers are *bipolar*, meaning that they are designed to amplify both positive and negative signals. These amplifiers require a bipolar DC power supply. For example, it is common to power bipolar amplifiers with +15 V and -15 V. A lot of amplifier power supply labelling is used to reflect the technology it is implemented with, or is based on historical precedent. For example, it is common to see bipolar power supplies represented on amplifiers as:

A bipolar amplifier showing the power supply connections

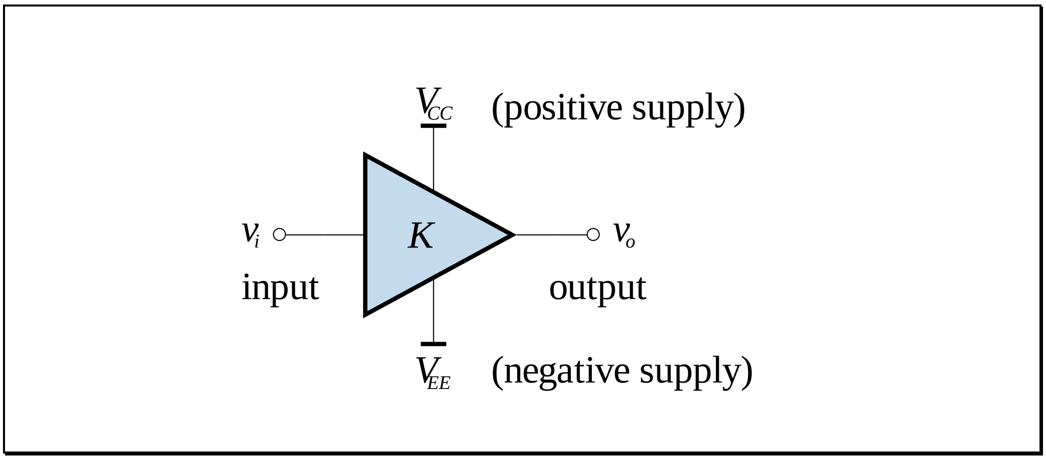


Figure 2.2

where $V_{CC} = 15\text{ V}$ and $V_{EE} = -15\text{ V}$, with respect to a circuit "common". The "*CC*" subscript in this case refers to the voltage at a "collector" of a transistor inside the amplifier package, and the "*EE*" to an "emitter" of a transistor.

Some amplifiers are *unipolar*, which means they are designed to amplify signals that are of one polarity only. These amplifiers only require a single DC power supply, such as +5V.

A unipolar amplifier showing the power supply connections

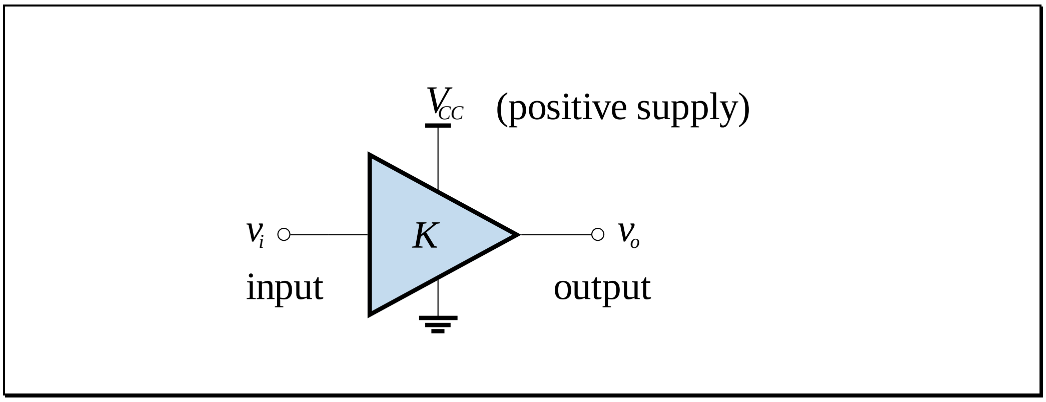
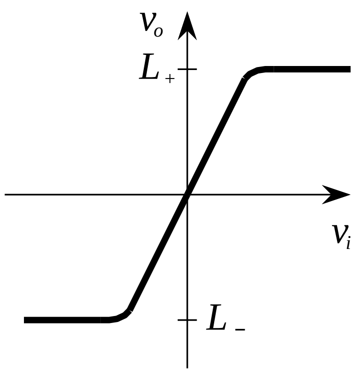


Figure 2.3

2.1.3 Saturation

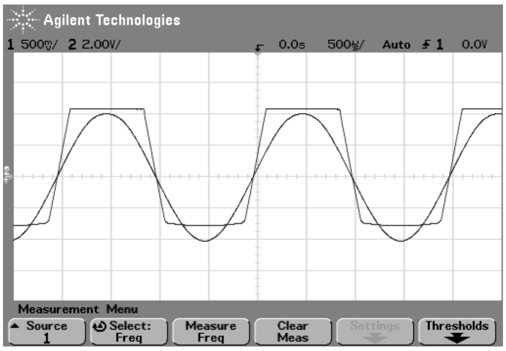
Real amplifiers can only output a voltage signal that is within the capabilities of the internal circuitry and the external DC power supplies. When amplifier outputs approach their output limitation, they are said to *saturate* – they cannot provide the output that is required by a linear characteristic. The resulting transfer characteristic, with the positive and negative saturation levels denoted L_+ and L_- respectively, is shown below:



The transfer characteristic of a real amplifier, showing that it saturates eventually

Figure 2.4

Each of the two saturation levels is usually within a volt or so of the voltage of the corresponding power supply. Obviously, in order to avoid distorting the output signal waveform, the input signal swing must be kept within the linear range of operation. If we don't, then the output waveform becomes distorted and eventually gets *clipped* at the output saturation levels.



The input signal and the output signal of a saturated amplifier showing clipping

Figure 2.5

2.1.4 Circuit Model

For an ideal voltage amplifier, the output voltage is independent of both the source resistance and the load resistance. Thus, to model an ideal voltage amplifier, we would use a voltage-controlled voltage source:

The model of an ideal voltage amplifier

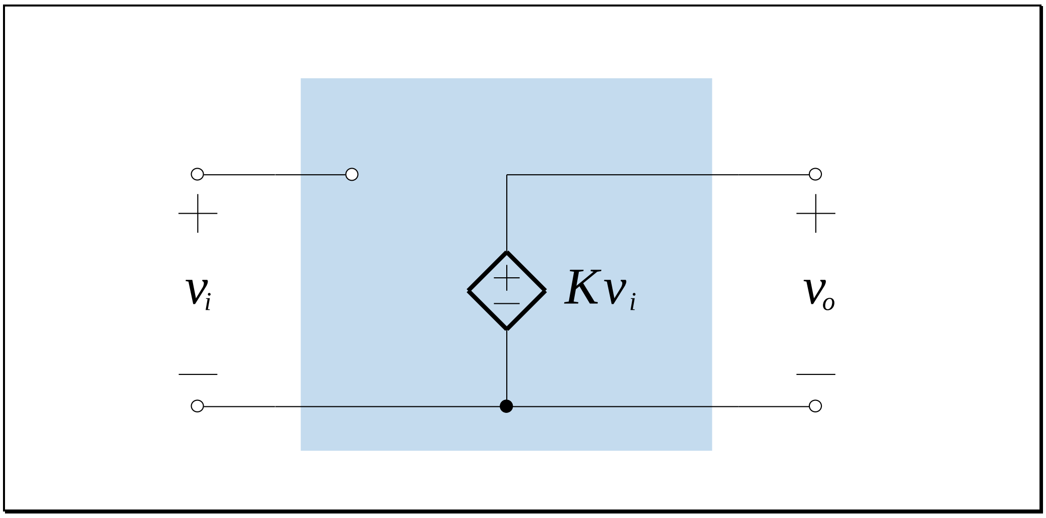


Figure 2.6

Real voltage amplifiers have a finite input resistance as well as a finite output resistance. Thus, a model of a real amplifier is:

A linear model of a real voltage amplifier

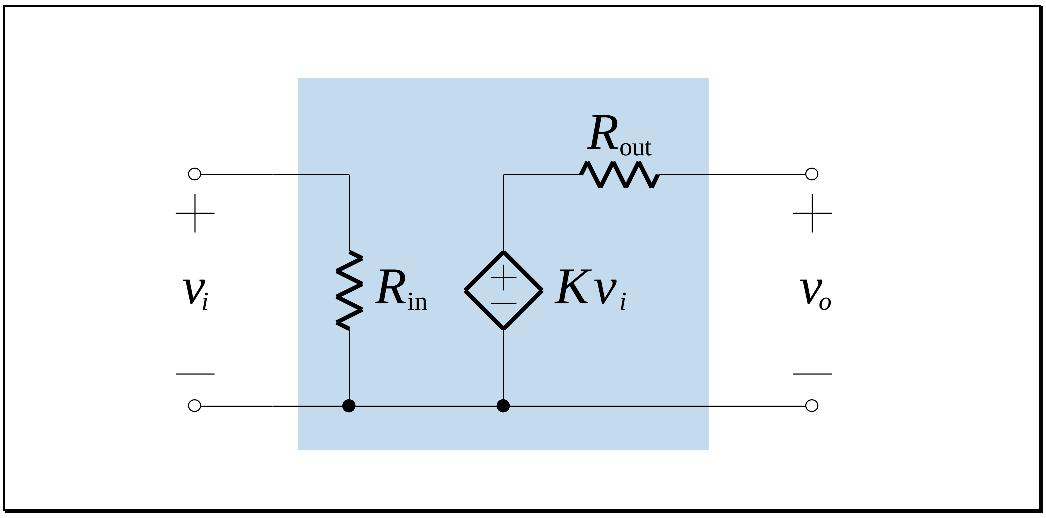


Figure 2.7

This model is only valid in its *linear region of operation*. Also note that the amplifier is *unilateral* – there is no "path" for a voltage at the output to appear in some way at the input of the amplifier. Thus, the use of the voltage-controlled voltage source creates a "one way path" for the voltage from the input to the output.

2.2 The Operational Amplifier

An operational amplifier (op-amp) is an integrated circuit amplifier consisting of dozens of transistors. An op-amp amplifies the voltage *difference* between its two input terminals, and produces a *single-ended* output voltage, i.e. the output voltage is with respect to the power supply "common". The circuit symbol for the op-amp is shown below:

Op-amps defined

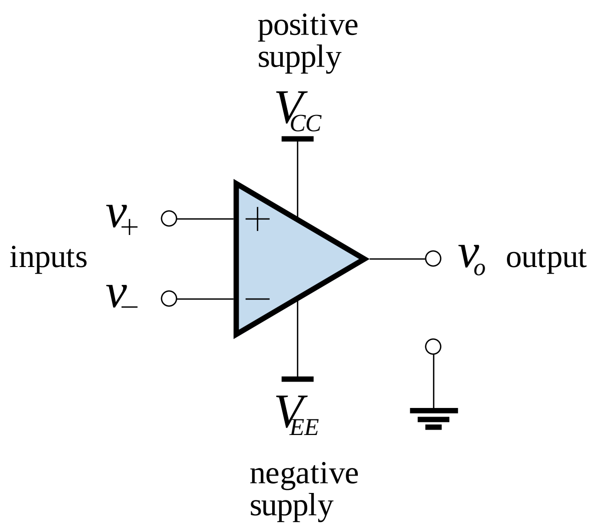


Figure 2.8

An op-amp circuit symbol showing all connections

In many circuit diagrams it is customary to omit the power supply and common connections (the output being understood to be taken with respect to the circuit common), and so we normally draw:

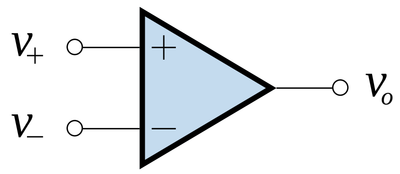


Figure 2.9

A simplified op-amp circuit symbol

The input labelled v_+ is termed the *noninverting* terminal, and the input labelled v_- is termed the *inverting* terminal. This naming is a result of the op-amps ability to amplify the *difference* between these two voltages.

2.10

2.2.1 Feedback

The op-amp is an amplifier intended for use with *external feedback elements*, where these elements determine the resultant function, or *operation*[2]. As we shall see, op-amp circuits can perform a variety of mathematical operations, such as addition, subtraction, integration and differentiation of voltage signals.

The feedback elements are connected between the op-amp's output and its inverting terminal, thus providing what is known as *negative feedback*.

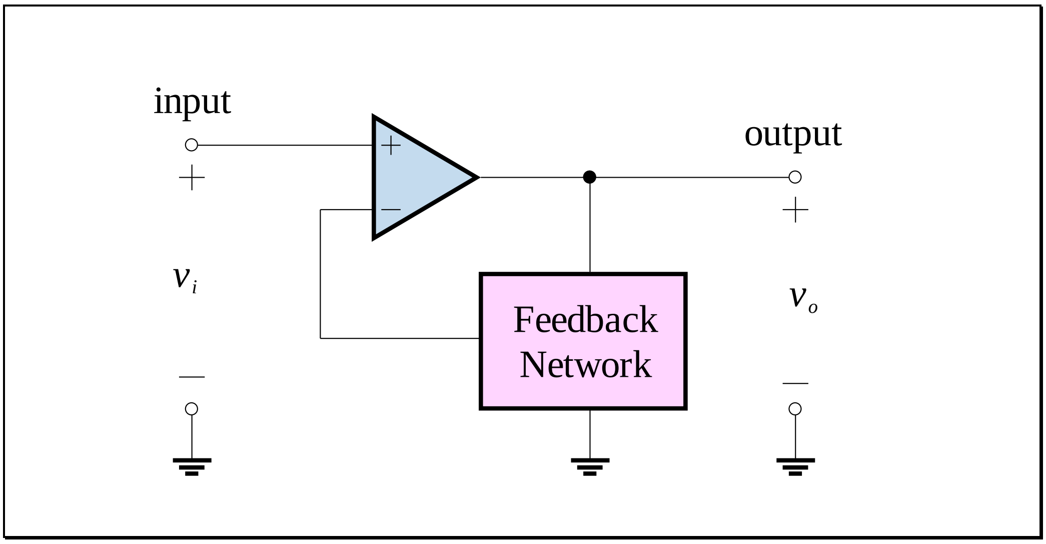


Figure 2.10

In the figure, the input is applied between the op-amp (+) input and a *common*, or *reference point*, as denoted by the "ground" symbol. This reference point is also common to the output, the feedback network and the power supply.

The feedback network can be resistive or reactive, linear or nonlinear, or any combination of these. More detailed analysis, shown later, shows that the circuit's overall gain characteristic is predominantly determined by the feedback network. It is critically important to note that op-amps are *never* intended for use without a feedback network.

---

[2] The naming of the operational amplifier occurred in the classic paper by John R. Ragazzini, Robert H. Randall and Frederick A. Russell, "Analysis of Problems in Dynamics by Electronic Circuits," *Proceedings of the IRE*, Vol. 35, May 1947, pp. 444-452. This paper references the op-amp circuits (feedback amplifiers) used by Bell Labs in the development of the "M9 gun director", a weapon system which was instrumental in winning WWII.

2.2.2 Circuit Model

Op-amps are voltage amplifiers. They are designed to have an extremely large input resistance, a very low output resistance, and a very large gain. Even though they are comprised of dozens of transistors, a simple linear *macro-model* of a real op-amp – valid over a certain range of operating conditions – is shown below:

A simple linear model of a real op-amp

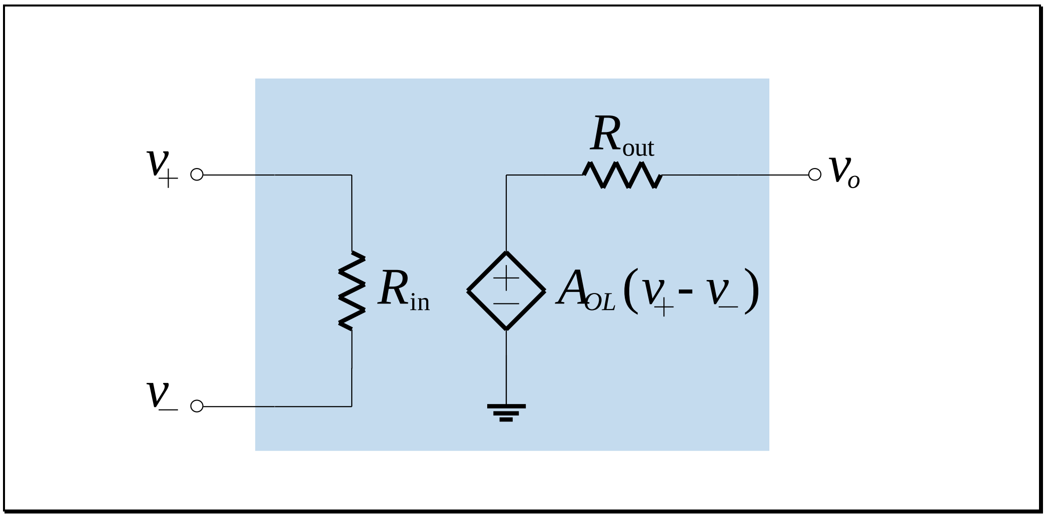


Figure 2.11

Note that, under open-circuit conditions on the output (i.e. no load is attached to the op-amp output terminal), the op-amp's output voltage is given by:

The open-loop output voltage of an op-amp, under no-load conditions

$$v_o = A_{OL}(v_+ - v_-) \tag{2.4}$$

The gain of the amplifier under these conditions, A_{OL}, is termed the *open-loop* gain, hence the "*OL*" subscript. The reason for this name will become apparent shortly.

The model parameters for a *general purpose* op-amp, such as the TL071, are tabulated below:

Parameter	**Symbol**	**Typical Value**
Open-loop voltage gain	A_{OL}	200 000 V/V
Input resistance	R_{in}	1 TΩ
Output resistance	R_{out}	200 Ω

2.2.3 The Ideal Op-Amp

The "ideal op-amp" is a theoretical device that pushes the typical op-amp parameters to their ideal values:

Parameter	**Symbol**	**Ideal Value**
Open-loop voltage gain	A_{OL}	∞ V/V
Input resistance	R_{in}	∞ Ω
Output resistance	R_{out}	0 Ω

Thus, the *ideal* op-amp has the circuit model:

Model of an ideal op-amp

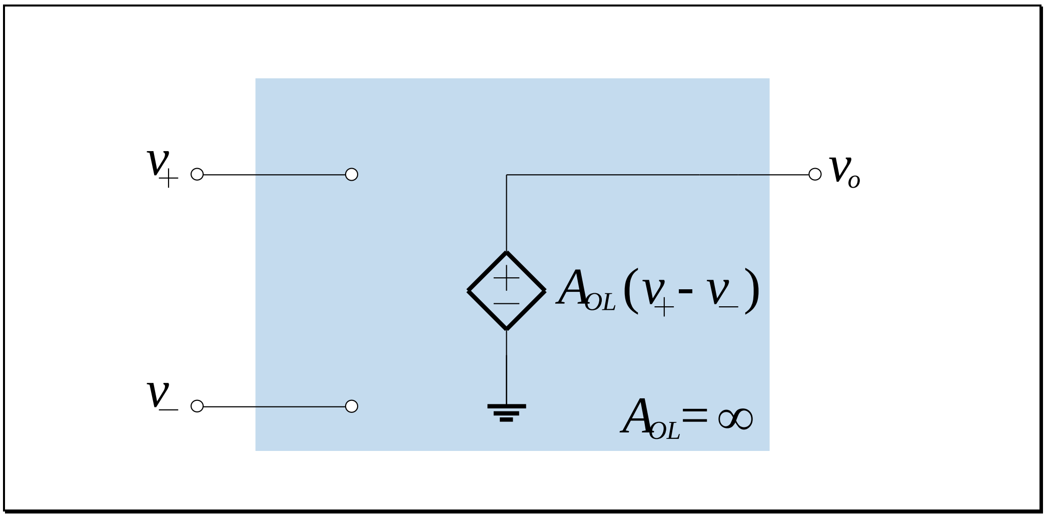


Figure 2.12

There are several interesting characteristics of this model that will be useful when analysing (and designing) circuits with ideal op-amps.

The first characteristic is due to the infinite input resistance:

An ideal op-amp draws no input current	(2.5)

An ideal op-amp draws no input current

The second characteristic of the model is that the output voltage is constrained by a dependent voltage source (there is no output resistance), and thus:

An ideal op-amp has an output voltage that is independent of its load	(2.6)

An ideal op-amp has an output voltage that is independent of the load placed on it

The last and most important characteristic is due to the infinite open-loop gain. At first glance the idealisation that $A_{OL}=\infty$ appears problematic from a circuit analysis viewpoint, since for a *finite* input voltage difference the output will be *infinite*. However, the ideal op-amp *can* produce a *finite* output voltage, but *only* so long as the input voltage is *zero*. Thus, for an ideal op-amp to produce a finite output voltage v_o, the input voltage difference must be:

$$v_+ - v_- = \frac{v_o}{A_{OL}} = \frac{\text{finite}}{\infty} = 0 \tag{2.7}$$

and therefore:

$$\boxed{v_+ = v_-} \tag{2.8}$$

An ideal op-amp has equal input voltages if it has a finite output voltage

Thus:

An ideal op-amp has equal input voltages	(2.9)

Since the ideal op-amp has equal input voltages (like a short-circuit), but draws no input current (like an open-circuit), we say there is a *virtual short-circuit* across its input terminals. We will use the concept of the virtual short-circuit as the fundamental basis for the analysis and design of circuits containing ideal op-amps.

The virtual short-circuit defined

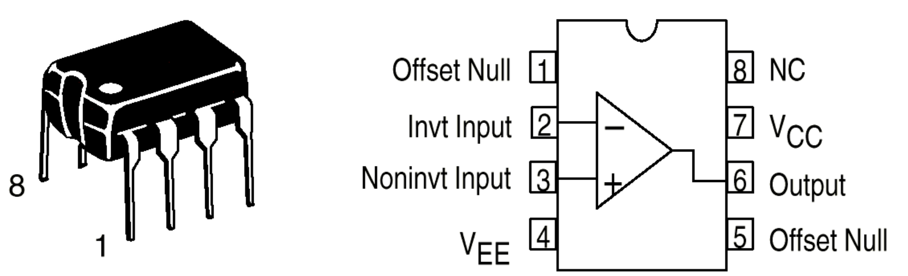
2.2.4 Op-Amp Fabrication and Packaging

There are many designs for the internal circuit of an op-amp, with each design optimising a particular parameter (or parameters) of interest to the designer. Such parameters may be the *open-loop gain* (how much the input voltage difference is amplified), the *bandwidth* (the highest frequency it can amplify), or the *bias current* (how much DC current it draws from the input terminals). You will become familiar with these terms later when we look more closely at real op-amp limitations (as opposed to the ideal op-amp).

There are several device fabrication technologies that are used to construct an op-amp. For *general-purpose* op-amps, bipolar junction transistors (BJTs) are mostly used at the input because they are easy to match and are capable of carrying large currents. However, some operational amplifiers have a field effect transistor (FET) input, with the rest of the circuit being made from BJTs. Complementary metal-oxide-semiconductor (CMOS) transistors are used in op-amps that find application in the design of analog and mixed-signal very large scale integrated (VLSI) circuits.

It is important to be able to recognise the standard pin-outs of an op-amp IC. All ICs conform to a standard pin numbering scheme. There is usually a notch or mark on one end of the chip. With the notch oriented to the left, pin 1 is the first pin on the bottom of the package. The pins are then numbered in a counter-clockwise direction. An example is shown below for the TL071 op-amp, which is a single op-amp housed in an 8-pin package:

Single op-amp IC package details



TL071 package details

Figure 2.13

2.3 Negative Feedback

The concept of negative feedback is fundamental to life. A simple experiment will illustrate this point: close your eyes and then bring your index fingers together so that they touch at the tips. You will probably miss. By closing your eyes you have broken a feedback loop which is vital to most human actions; in order to perform an operation accurately we must be able to see what we are doing and thus apply any small corrections as and when necessary. In effect, we are taking the output (the action) and feeding it back to the input (the mental 'instruction' or intention) in such a way that the output is made equal to the input. In other words, the action is forced to correspond exactly with the intention.

An everyday example of negative feedback

Examples of negative feedback can also be found in the field of mechanical engineering. One of the clearest examples is the governor which is used to control the speed of rotating machinery. The most historically significant form of governor was that developed by James Watt in 1788, following the suggestion of his business partner, Matthew Boulton. The *conical pendulum* governor was one of the final series of innovations that Watt made to the steam engine that ushered in the industrial revolution and became the prime source of motive power in the 19th century.

The conical pendulum governor is shown in basic form below:

Watt's conical pendulum governor

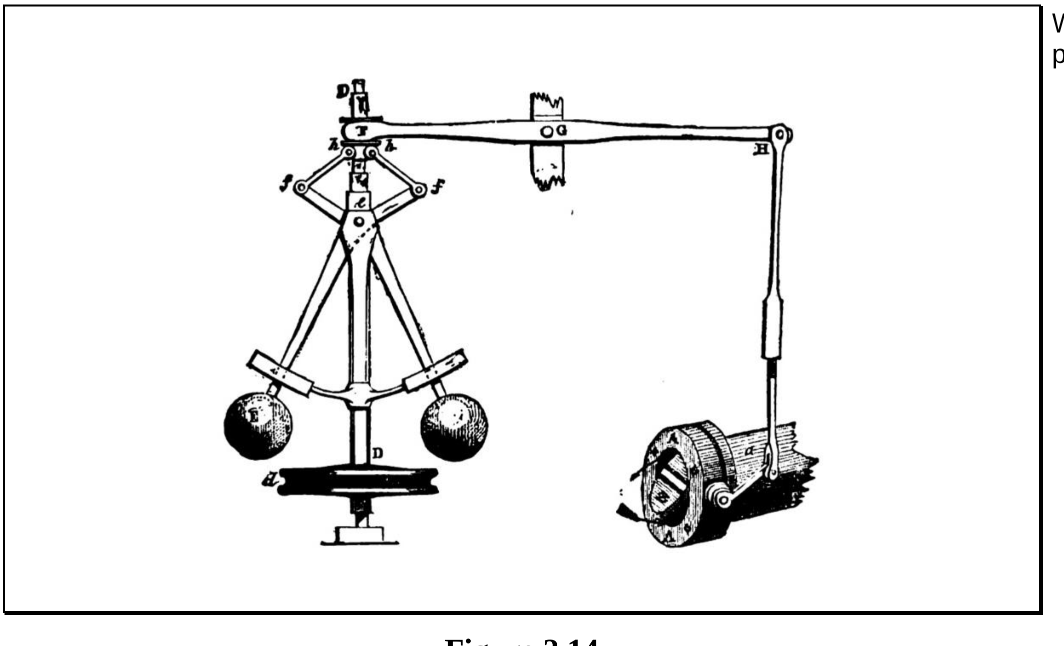


Figure 2.14

As the speed of the engine, and hence of the governor shaft, increases, the centrifugal force causes the weights to fly outwards on their linkage. The linkage is connected via a system of levers to the main steam valve so that, if the speed increases, the movement of the governor weights throttles back the main steam supply to the engine. Conversely, a tendency to slow down, perhaps due to increased load, will allow more steam in to boost the speed back to normal. The speed will thus settle down to a happy medium and be largely independent of variations in the load on the engine.

In this example, as in the physiological illustration discussed first, the system is kept under control by feeding a measure of its output back to the input. Mechanical control systems such as the governor are often known as servo-systems (literally, slave systems) and are fundamental to industrial automation.

2.3.1 Negative Feedback in Electronics

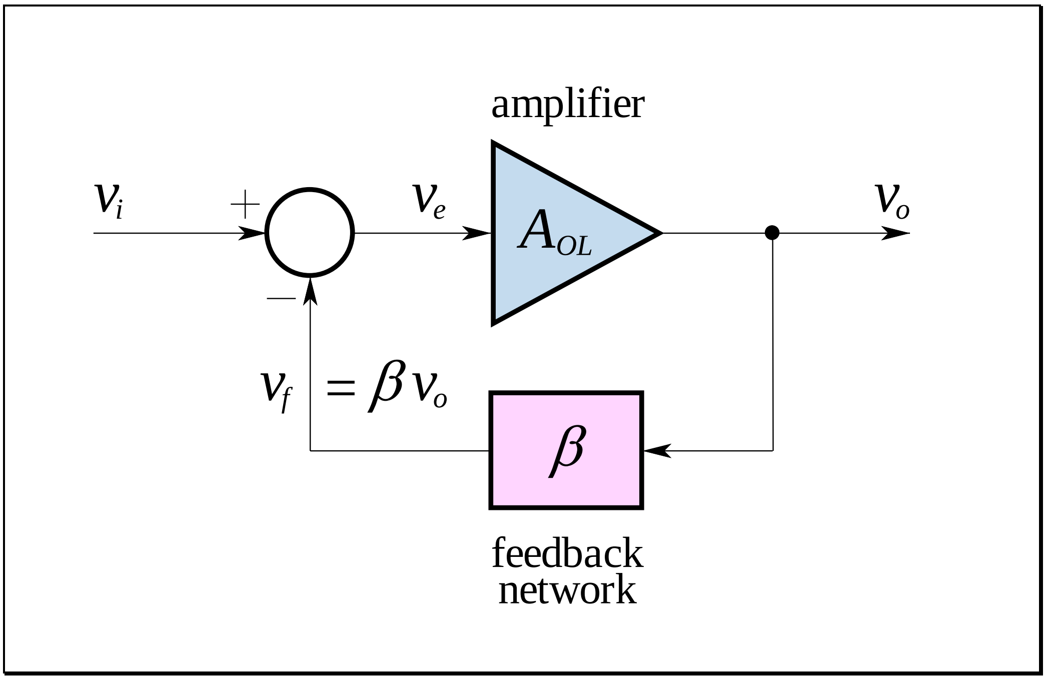
The gain of an amplifier, such as an op-amp, will vary from device to device due to the many manufacturing variations in the transistors and resistors that comprise it. Such component variations result in considerable uncertainty in the overall voltage gain. For example, an op-amp datasheet may specify the typical value for the open-loop gain as 200 000, but some specimens may achieve a gain as low as 25 000. The open-loop gain also changes with temperature, power supply voltage, signal frequency and signal amplitude.

Negative feedback is used to precisely set the gain of an amplifier

Just as the steam engine needs the controlling influence of the governor, so most electronic amplifiers require electrical negative feedback if their gain is to be accurately predictable and remain constant with varying environmental conditions.

2.3.2 An Amplifier with Negative Feedback

The figure below shows a *block diagram* of an amplifier with negative feedback:



Block diagram of an amplifier with negative feedback

Figure 2.15

The amplifier has a voltage gain A_{OL} and the *feedback network* is an attenuator which feeds a fixed fraction, β, of the output back to the input. The feedback signal, $v_f = \beta v_o$, is *subtracted* from the input signal (we thus have *negative* feedback – if the signal were *added*, we would have *positive* feedback).

We can now determine the effective voltage gain, A_{CL}, of the amplifier with feedback. This is given simply by the ratio of the output voltage to input voltage:

$$A_{CL} = \frac{v_o}{v_i} \tag{2.10}$$

The signal at the input to the basic amplifier is:

$$v_e = v_i - v_f = v_i - \beta v_o \tag{2.11}$$

Also, the signal at the output of the basic amplifier is given by:

$$v_o = A_{OL} v_e \tag{2.12}$$

Therefore:

$$v_o = A_{OL}\left(v_i - \beta v_o\right) \tag{2.13}$$

Rearranging:

$$\begin{aligned} v_o\left(1 + A_{OL}\beta\right) &= A_{OL} v_i \\ \frac{v_o}{v_i} &= \frac{A_{OL}}{1 + A_{OL}\beta} \end{aligned} \tag{2.14}$$

Hence:

The closed-loop gain of an amplifier with negative feedback

$$A_{CL} = \frac{A_{OL}}{1 + A_{OL}\beta} \tag{2.15}$$

This is the general equation for an amplifier with negative feedback. The basic gain of the amplifier, A_{OL}, is known as the *open-loop* gain and the gain with feedback, A_{CL}, as the *closed-loop* gain.

Engineers design the circuit by starting with a basic amplifier with a very large open-loop gain (e.g., the open-loop gain of an op-amp is $A_{OL} \approx 100{,}000$) and then ensure by design that the feedback network provides an attenuation β so that:

$$A_{OL}\beta >> 1 \tag{2.16}$$

When this is the case, we can neglect the '1' in the denominator of Eq. (2.15) so that:

$$A_{CL} \approx \frac{1}{\beta} \tag{2.17}$$

The closed-loop gain of an amplifier with negative feedback, if the open-loop gain is very large

This is a very significant equation because we have 'designed' an amplifier with a precisely determined voltage gain. As long as the open-loop gain is much larger than the closed-loop gain (e.g. a hundred times greater) then the closed-loop gain is independent of the amplifier characteristics and dependent only on the feedback network β. The feedback network usually depends upon just a handful of passive elements, such as resistors and capacitors. These elements are the most stable components in electronics; their values can be precisely specified to very high levels of accuracy (better than 0.1% for resistors, and 1% for capacitors), and their value does not vary greatly with environmental changes such as temperature and aging. Negative feedback extends these attributes of accuracy and stability to the gain of the entire amplifier.

Later, we will also see that negative feedback increases the frequency of signals that we can apply to the amplifier, reduces nonlinear distortion, increases input resistance and decreases output resistance. The price for these benefits is a reduction in the amplifier gain – a trade-off that is well worth making.

2.4 The Noninverting Amplifier

We seek a circuit implementation of the amplifier with negative feedback presented in Figure 2.15. We have already seen that an op-amp amplifies the difference between its two input voltages. Thus, an op-amp implements the subtracter and the amplifier in one device:

An op-amp implements the subtracter and the amplifier in one device

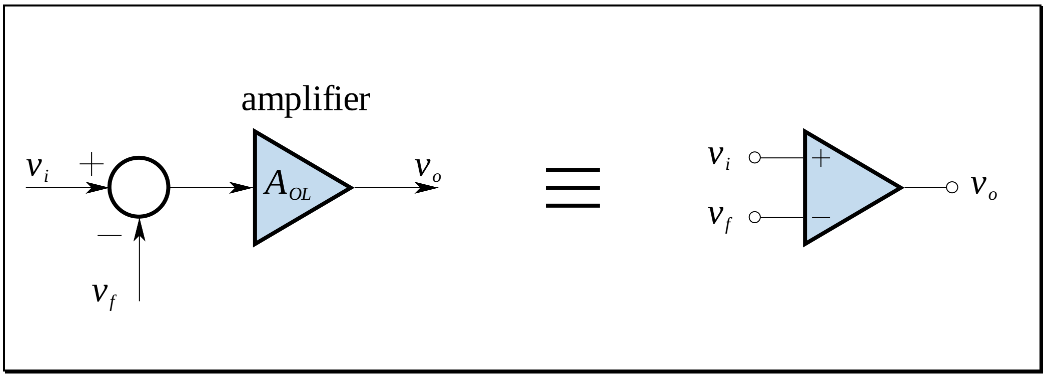


Figure 2.16

The feedback network can consist of any combination of passive or active elements. The simplest feedback network provides attenuation, using a resistive voltage divider:

A simple feedback network that provides a fixed attenuation

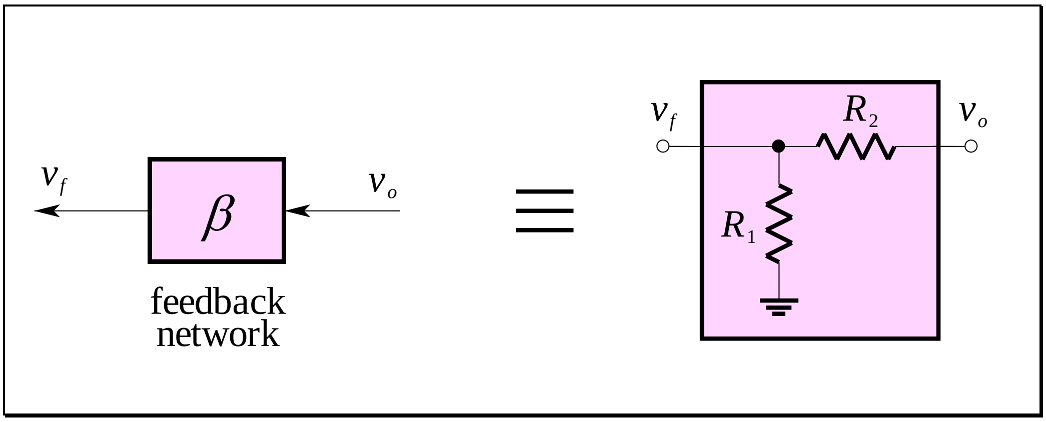


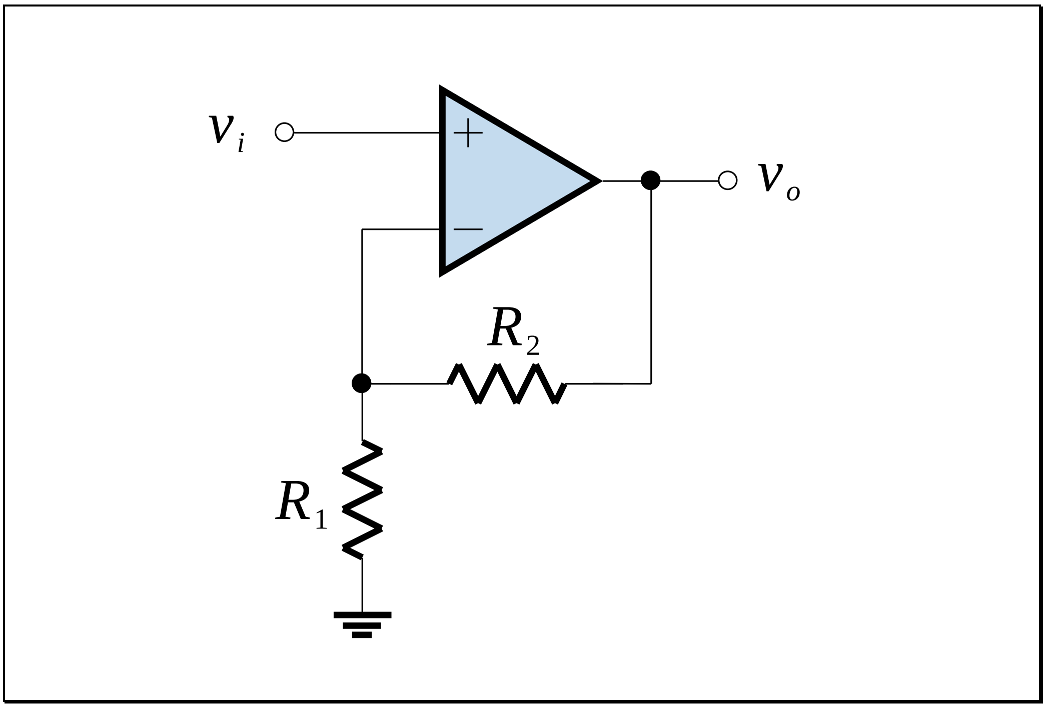
Figure 2.17

Note that for this circuit the input is on the right and the output is on the left, as we are providing a *feedback* path from the output of the op-amp, and back to its inverting input terminal.

The fixed fraction, β, of the output which is fed back to the input is given by the voltage divider rule:

$$\beta = \frac{v_f}{v_o} = \frac{R_1}{R_1 + R_2} \tag{2.18}$$

Thus, a circuit that implements the amplifier with negative feedback is:



The noninverting amplifier

Figure 2.18

To reiterate, the closed-loop gain of this amplifier is:

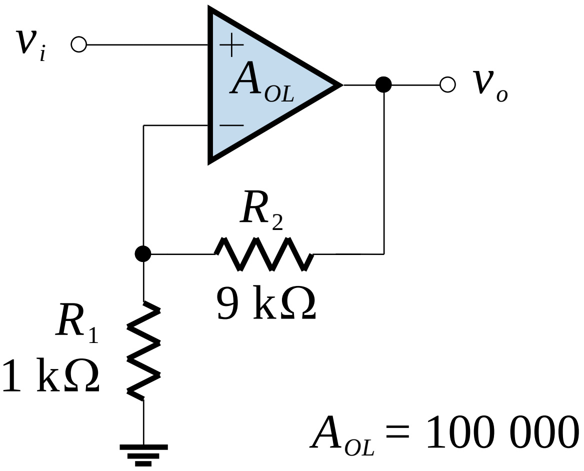
$$\frac{v_o}{v_i} = A_{CL} = \frac{A_{OL}}{1 + A_{OL}\beta} \tag{2.19}$$

and if we have $A_{OL}\beta >> 1$, then the closed-loop gain is approximately:

$$A_{CL} \approx \frac{1}{\beta} = 1 + \frac{R_2}{R_1} \tag{2.20}$$

EXAMPLE 2.2 A Real Noninverting Amplifier

An op-amp with an open-loop gain of 100 000 is connected in a noninverting amplifier configuration to give a nominal gain of 10, as shown below:



The feedback factor is:

$$\beta = \frac{R_1}{R_1 + R_2} = \frac{1\text{k}}{1\text{k} + 9\text{k}} = \frac{1}{10}$$

Therefore the closed-loop gain is:

$$A_{CL} = \frac{A_{OL}}{1 + A_{OL}\beta} = \frac{10^5}{1 + 10^5 \cdot 10^{-1}} = \frac{10^5}{10001} = 9.9990 \approx 10$$

If the op-amp open-loop gain is changed to 200 000 (e.g. a different op-amp is used) then the closed-loop gain changes to:

$$A_{CL} = \frac{A_{OL}}{1 + A_{OL}\beta} = \frac{2 \cdot 10^5}{1 + 2 \cdot 10^5 \cdot 10^{-1}} = \frac{2 \cdot 10^5}{20001} = 9.9995 \approx 10$$

Thus, the closed-loop gain changes by only 0.005%, even though the open-loop gain changed by 100%. This is because $A_{OL}\beta >> 1$, and therefore, by Eq. (2.19), $A_{CL} \approx 1/\beta = 10$. Thus, so long as $A_{OL}\beta >> 1$ is satisfied, the closed-loop amplifier maintains a nominal gain of 10 to a very high accuracy.

2.4.1 The Noninverting Amplifier with an Ideal Op-Amp

Assuming an ideal op-amp with infinite open-loop gain ($A_{OL} = \infty$), then the overall closed-loop gain of the amplifier is given by Eq. (2.19):

$$\begin{aligned} A_{CL} &= \frac{A_{OL}}{1 + A_{OL}\beta} \\ &= \frac{1}{1/A_{OL} + \beta} \\ &= \frac{1}{1/\infty + \beta} \\ &= \frac{1}{\beta} \end{aligned} \tag{2.21}$$

Thus, the overall closed-loop gain is:

$$A_{CL} = 1 + \frac{R_2}{R_1} \quad \text{(noninverting)} \tag{2.22}$$

The gain of an ideal noninverting amplifier

Thus, the approximation for the closed-loop gain that we used for a real op-amp, $A_{CL} \approx 1/\beta$, now turns into an exact equation, $A_{CL} = 1/\beta$. Thus, we will find it expedient to analyse op-amp circuits by assuming that ideal op-amps are used, with an understanding that the real circuits will differ in performance by only a tiny amount.

A key point to note in this formula is that the *ratio* of the resistors determines the gain. In practice this means that a range of actual R_1 and R_2 values can be used, so long as they provide the same ratio.

The amplifier in this configuration provides a gain which is always greater than or equal to 1. The output is also "in phase" with the input, since the gain is positive. Hence, this configuration is referred to as a *noninverting amplifier*.

The closed-loop gain of the amplifier can also be derived using circuit analysis and the concept of the virtual short-circuit. For such a simple circuit, the results of the analysis can be written directly on the circuit diagram:

Analysis steps for the ideal noninverting amplifier

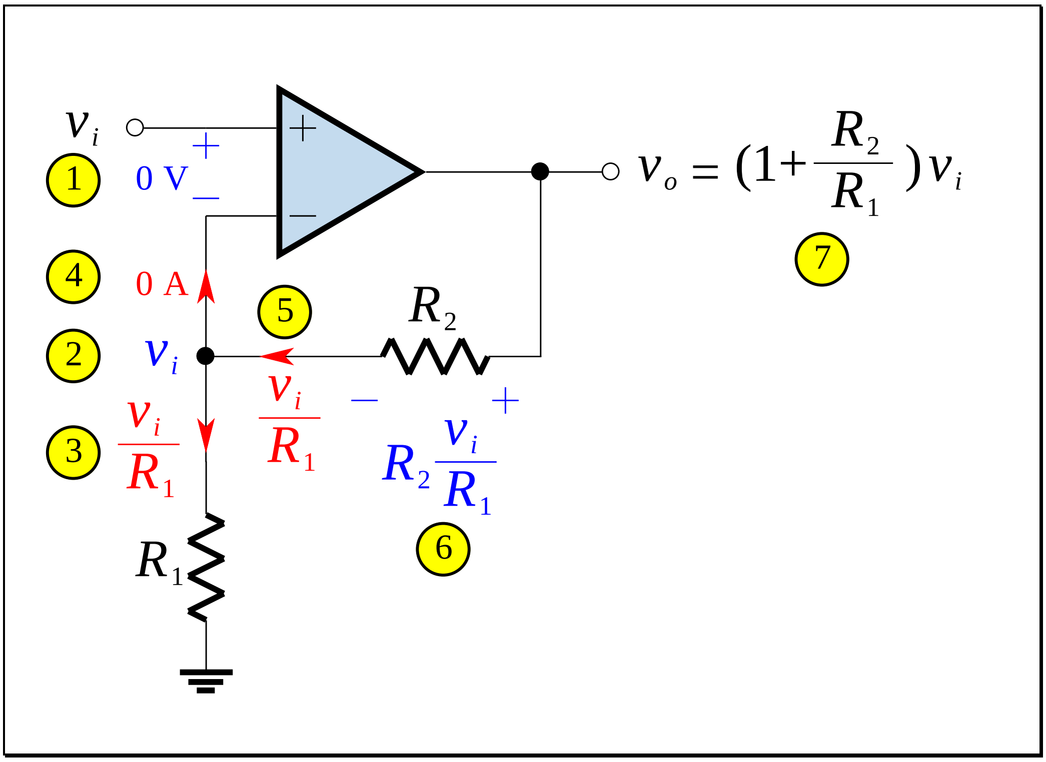


Figure 2.19

The analysis steps are:

1. We assume an ideal op-amp, and also assume that since there is a negative feedback path around the op-amp, then it is producing a finite output voltage (i.e., the overall amplifier is "working"). Thus, the ideal op-amp must have a virtual short-circuit (VSC) at its input terminals. We label the voltage across the input terminals as 0 V.

2. Since there is no difference in the voltages across the VSC, the voltage at the inverting terminal is $v_- = v_i$.

3. The current through resistor R_1 is given by Ohm's Law, $i_1 = v_i / R_1$.

4. Due to the infinite input resistance of the ideal op-amp, the current entering the inverting terminal is 0 A.

5. KCL at the inverting terminal now gives $i_2 = i_1 = v_i / R_1$.

6. The voltage drop across the resistor R_2 is given by Ohm's Law, $v_{R_2} = R_2 i_2 = R_2 \dfrac{v_i}{R_1}$, with the polarity shown.

7. KVL, from the common, across R_1, across R_2 and to the output terminal gives $v_o = v_i + R_2 \dfrac{v_i}{R_1} = \left(1 + \dfrac{R_2}{R_1}\right) v_i$.

Once again, we see that the overall closed-loop gain is given by $1 + R_2 / R_1$.

2.4.2 Input Resistance of the Noninverting Amplifier

The input voltage in a noninverting amplifier is connected directly to the noninverting terminal of the op-amp, which is effectively an open-circuit as far as current is concerned. Thus, the "input resistance" of the circuit is ideally infinite. This is an important property of the noninverting amplifier that is used in a variety of situations.

The input resistance of the ideal noninverting amplifier is infinite

2.4.3 Equivalent Circuit of the Noninverting Amplifier

The equivalent circuit of the noninverting amplifier is:

Equivalent circuit of the ideal noninverting amplifier

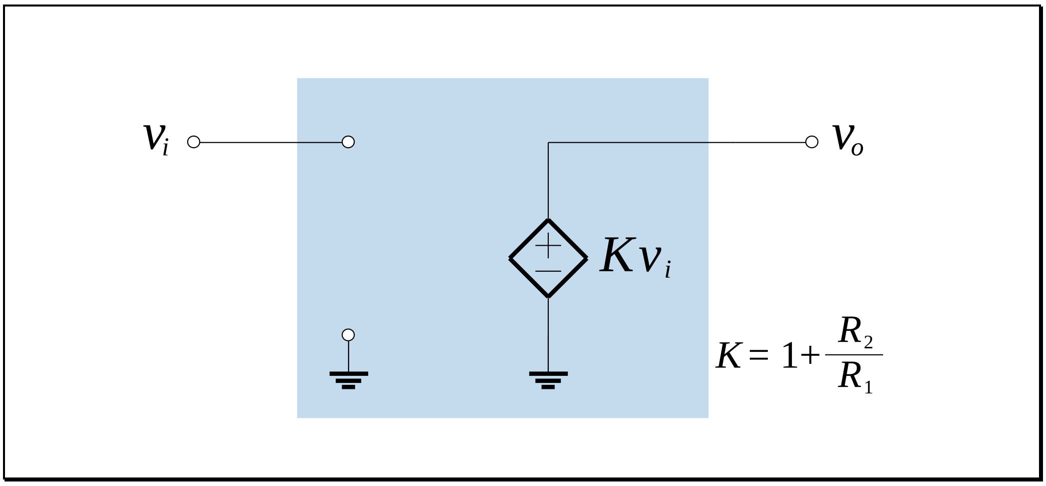


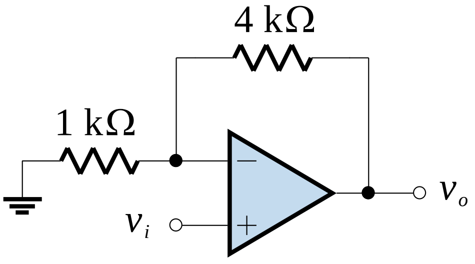
Figure 2.20

In our analysis of op-amp circuits from now on, we will assume ideal op-amps and make frequent use of the virtual short-circuit concept.

EXAMPLE 2.3 Design of a Noninverting Amplifier

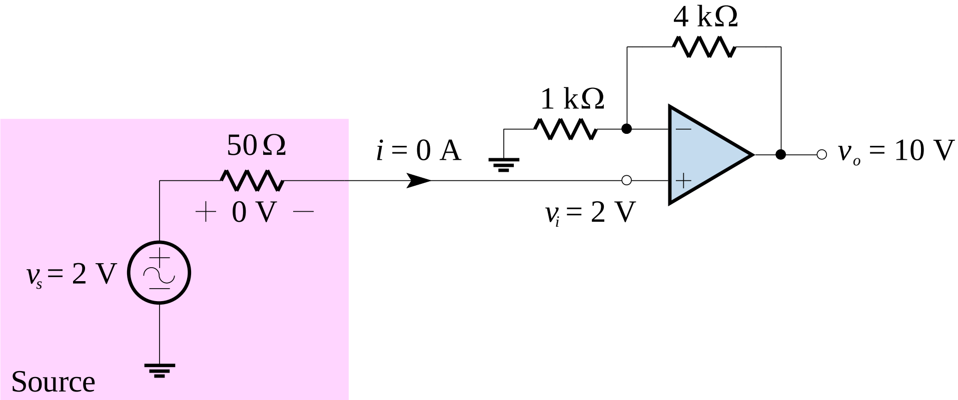
It is required to design an amplifier with a gain of 5 V/V that presents an almost infinite input resistance to any attached voltage source.

In this case, we decide to choose feedback resistors in the kΩ range:



Notice that we have decided to draw the inverting terminal of the op-amp at the top, and the feedback resistors pass over the top of the op-amp. The circuit is still the same as before. Either representation can be used, and will depend on such factors as space or clarity in the circuit schematic.

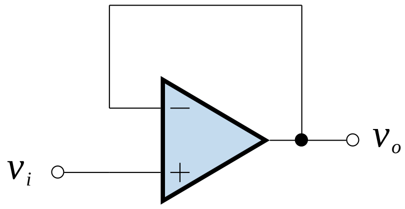
When we attach any type of source to the input of this circuit, no current will be drawn. For example:



We have therefore met the design specifications.

EXAMPLE 2.4 The Buffer

An ideal op-amp is connected in a noninverting amplifier configuration as shown below:

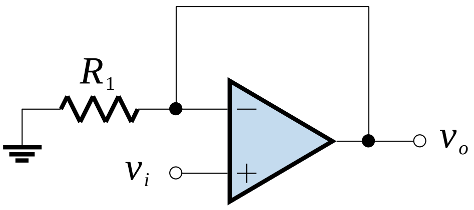


We identify $R_1 = \infty$ and $R_2 = 0$. Thus, the overall closed-loop gain is:

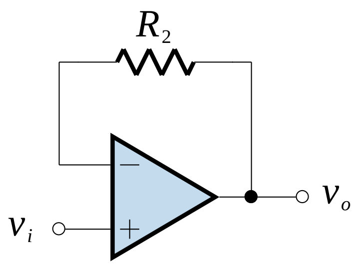
$$A_{CL} = 1 + \frac{R_2}{R_1} = 1 + \frac{0}{\infty} = 1$$

Thus, the circuit provides a "gain" of 1, and is called a *buffer* or a *follower*.

You can also see that a buffer can be created by:



or by:



Both of these circuits will operate as buffers, but the circuit presented first uses one less component.

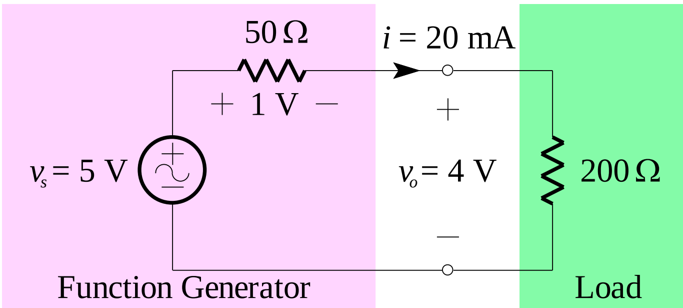
2.4.4 The Buffer

The buffer is used to "couple" one circuit to another

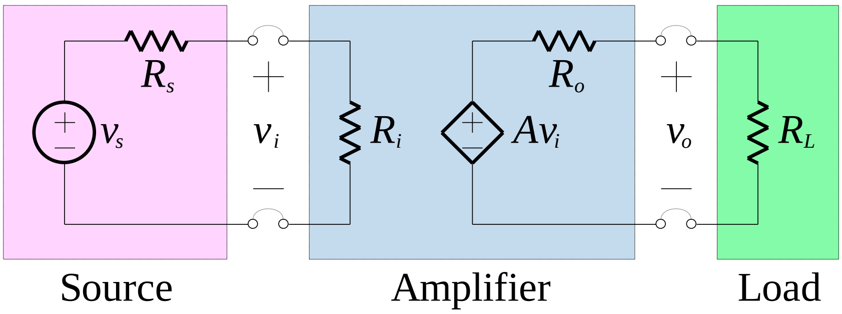
You may wonder "What is the point of a buffer if it only provides a gain of 1?" The answer lies in the other properties of the circuit – its infinite input resistance and zero output resistance.

EXAMPLE 2.5 Buffering a Source

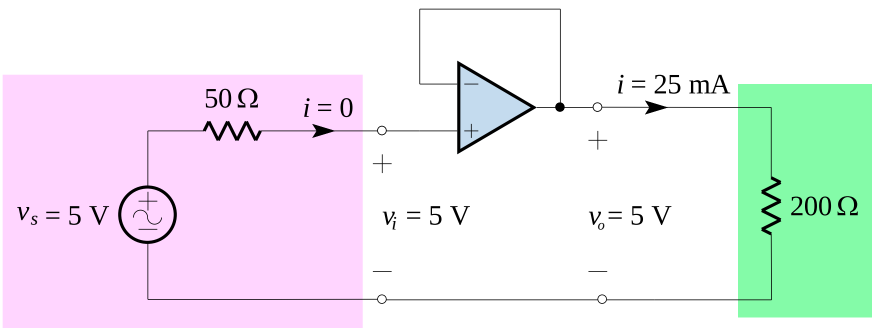
A certain function generator has an "output resistance" of 50 Ω and so its output will experience a significant internal *Ri* voltage drop when the attached load draws a "large" current, resulting in a drop in the output terminal voltage:



Therefore, we need to "buffer" the function generator with an amplifier which presents a high input resistance to the source and which also provides a low output resistance to the load:

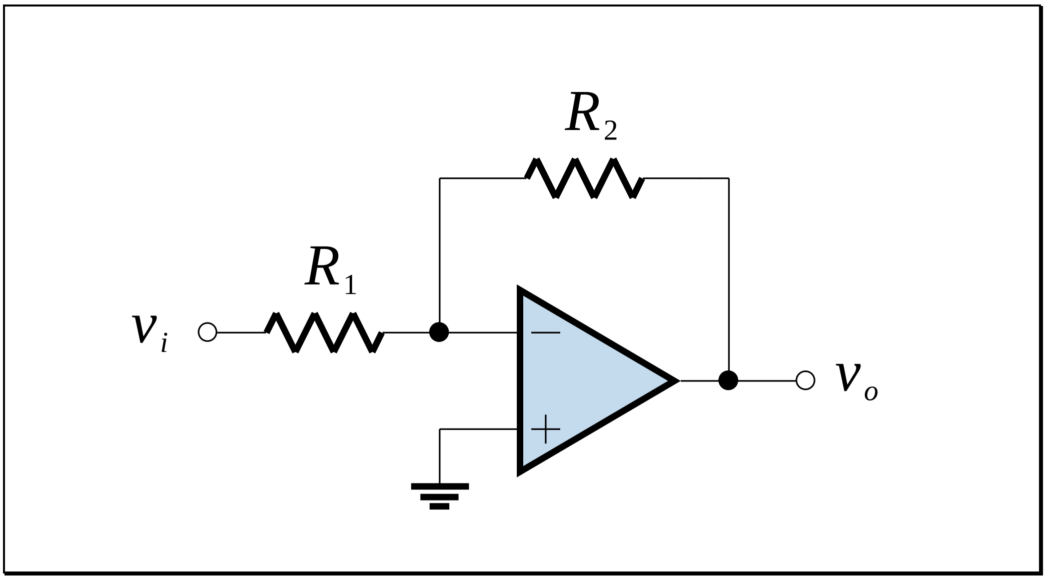


An ideal buffer amplifier with a gain of 1, when placed in between the function generator and the load, delivers the full source voltage to the load:



2.5 The Inverting Amplifier

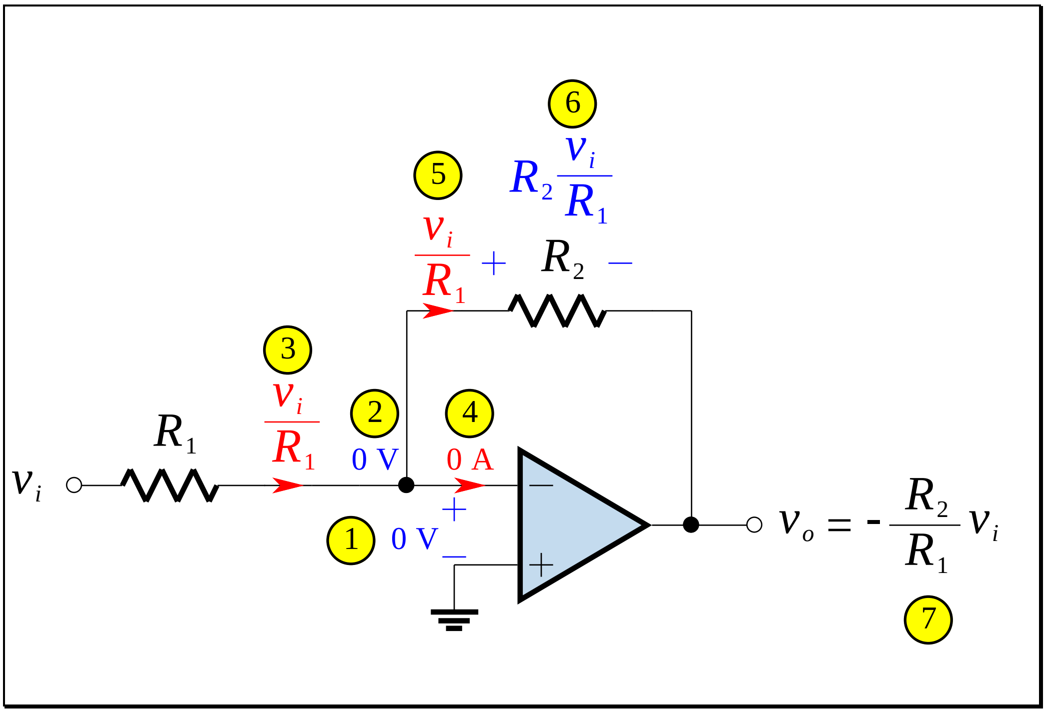
An inverting amplifier is created by swapping the input and common connections of the noninverting amplifier. The result is:



The inverting amplifier

Figure 2.21

We will analyse this circuit using the concept of the virtual short-circuit. This can be done on the circuit schematic:



Analysis steps for the ideal inverting amplifier

Figure 2.22

The analysis steps are:

1. We assume an ideal op-amp, and also assume that since there is a negative feedback path around the op-amp, then it is producing a finite output voltage (i.e., the overall amplifier is “working”). Thus, the ideal op-amp must have a virtual short-circuit (VSC) at its input terminals. We label the voltage across the input terminals as 0 V.

2. Since there is no difference in the voltages across the VSC, the voltage at the inverting terminal is $v_{-} = 0$.

3. The current through resistor R_1 is given by Ohm’s Law, $i_1 = v_i / R_1$.

4. Due to the infinite input resistance of the ideal op-amp, the current entering the inverting terminal is 0 A.

5. KCL at the inverting terminal now gives $i_2 = i_1 = v_i / R_1$.

6. The voltage drop across the resistor R_2 is given by Ohm’s Law, $v_{R_2} = R_2 i_2 = R_2 \frac{v_i}{R_1}$, with the polarity shown.

7. KVL, from the common, across the VSC, across R_2 and to the output terminal gives $v_o = 0 - R_2 \frac{v_i}{R_1} = -\frac{R_2}{R_1} v_i$.

Thus, the overall closed-loop gain is:

The gain of an ideal inverting amplifier

$$A_{CL} = -\frac{R_2}{R_1} \qquad \text{(inverting)} \tag{2.23}$$

The negative sign indicates that there is an inversion of the signal (i.e. a 180° phase change), so that a waveform will appear amplified, but “upside down”.

2.5.1 Input Resistance of the Inverting Amplifier

The input resistance of the inverting amplifier (i.e. the resistance "seen" by the input voltage source), is, by the definition of input resistance:

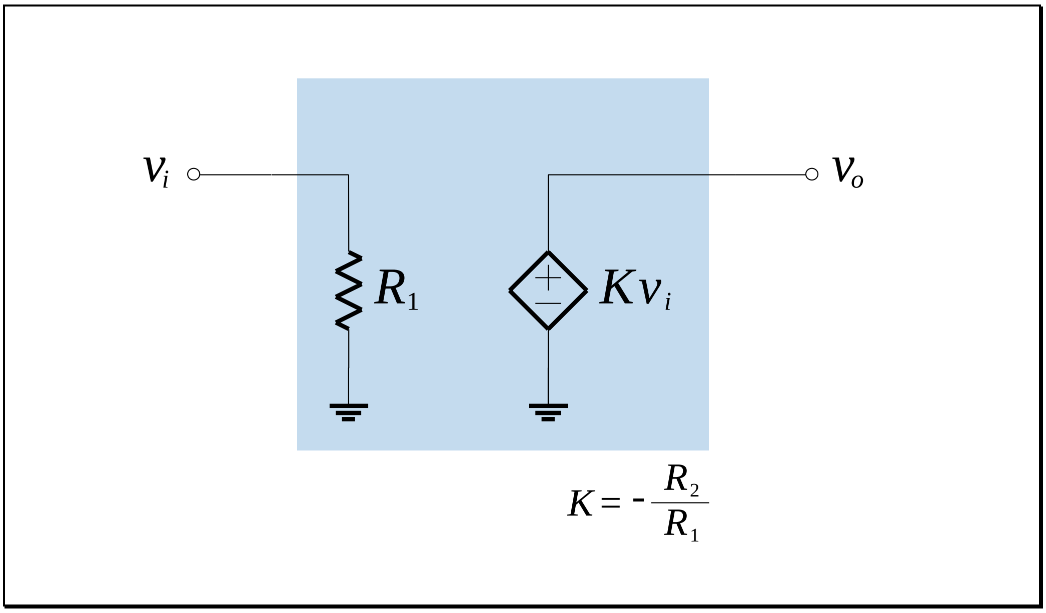
$$R_{\text{in}} = v_i / i_i = v_i / i_1 = \frac{v_i}{v_i / R_1} = R_1 \tag{2.24}$$

The input resistance of the ideal inverting amplifier is finite

Thus, the "input resistance" of the circuit is equal to R_1. This is a disadvantage compared to the noninverting amplifier, and careful use of the inverting amplifier is required.

2.5.2 Equivalent Circuit of the Inverting Amplifier

An equivalent circuit of the inverting amplifier is:



Equivalent circuit of the ideal inverting amplifier

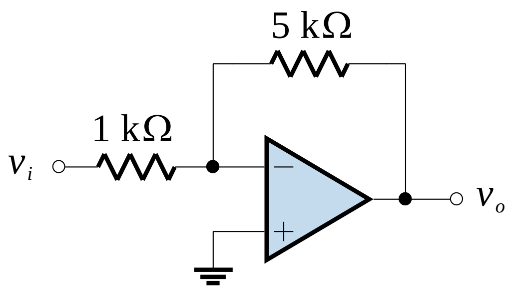
Figure 2.23

One advantage of the inverting amplifier over the noninverting amplifier is that you can achieve gain magnitudes less than one, i.e. build circuits that can *attenuate*, as well as amplify.

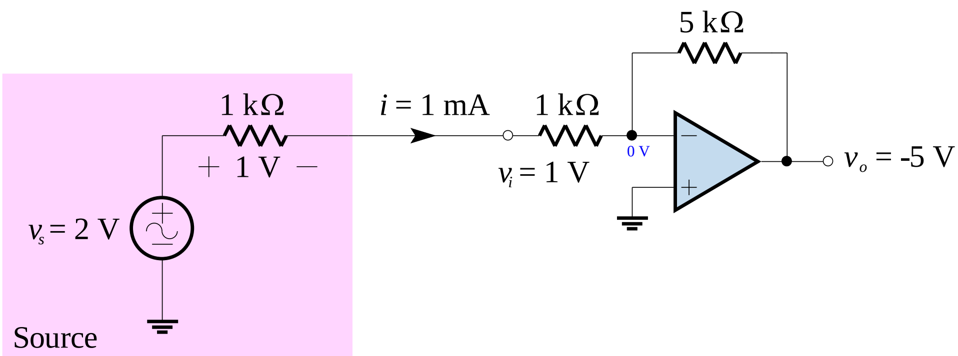
The inverting amplifier is also the basis for many other useful circuits that we will encounter later, such as the summer, integrator and differentiator.

EXAMPLE 2.6 Design of an Inverting Amplifier

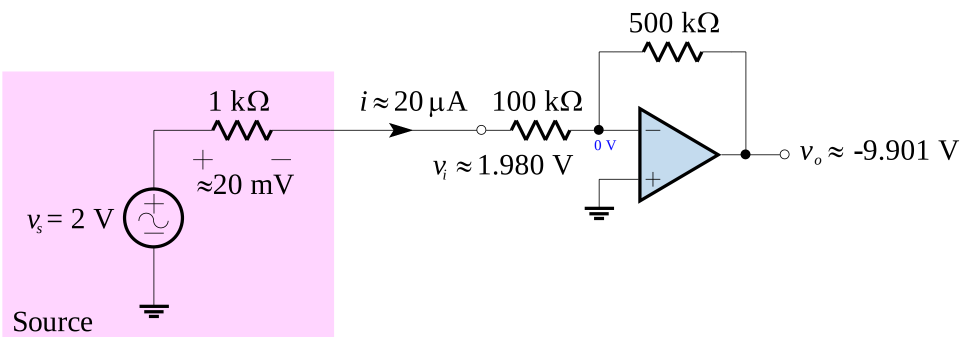
It is required to design an amplifier with a gain of -5 V/V. In this case, we decide to choose feedback resistors in the kΩ range:



Suppose we now attach a source, which has an internal resistance of 1 kΩ, to this amplifier:



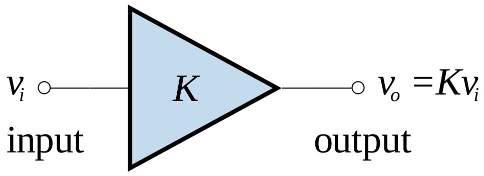
We see that the 1 kΩ input resistance of the inverting amplifier has caused the source to deliver current, and therefore there is a significant voltage drop across its internal resistance. A better design to suit this particular source would use resistors in the 100's of kΩ:



Now there is less than 1% error in the gain that it provides to the source.

2.6 Summary

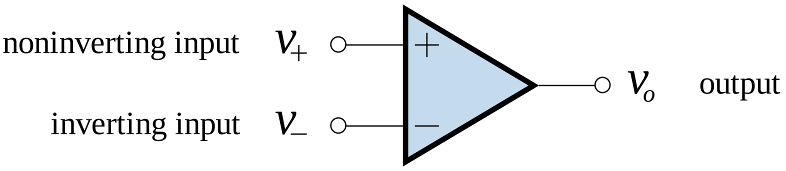
- A linear amplifier is a device that increases the amplitude of a signal (a voltage or a current) whilst preserving waveform shape. The most common is a voltage amplifier:



The gain of a voltage amplifier is expressed in either V/V or in dB:

$$K = \frac{v_o}{v_i} \text{ V/V} \qquad \text{or} \qquad |K| = 20\log_{10}\left|\frac{v_o}{v_i}\right| \text{ dB}$$

- The operational amplifier (op-amp) is an integrated circuit amplifier consisting of dozens of transistors. An op-amp amplifies the voltage *difference* between its two input terminals, and produces a *single-ended* output voltage. It is the "building block" for a vast array of useful electronic circuits.



The op-amp is always operated with a negative feedback network.

The characteristics of the ideal op-amp are:

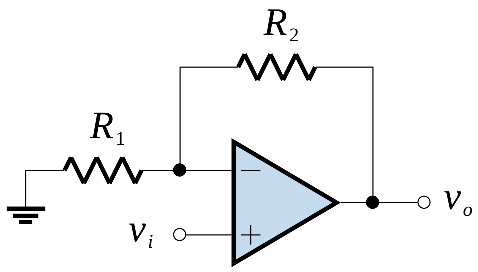
Parameter	**Symbol**	**Ideal Value**
Open-loop voltage gain	A_{OL}	∞ V/V
Input resistance	R_{in}	$\infty\ \Omega$
Output resistance	R_{out}	$0\ \Omega$

If an ideal op-amp has negative feedback, then a *virtual short-circuit* appears across its input terminals. This is the key to analysing and designing op-amp circuits.

- Negative feedback is a universal concept that arises in fields as diverse as biology, mechanics and electronics. If a proportion of the output of a device is fed back to the input and subtracted (so-called negative feedback), then the device will exhibit certain desirable overall qualities, such as accuracy and stability. In particular, the gain of an electronic amplifier with an open-loop gain of A_{OL} and a feedback factor of β will have a closed-loop gain given by:

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

- The noninverting amplifier is one of the most fundamental op-amp arrangements:

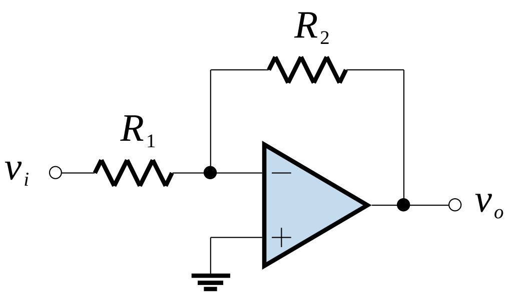


The gain of the noninverting amplifier is:

$$A_{CL} = 1 + \frac{R_2}{R_1}$$

An important special case of the noninverting amplifier is the buffer, for which $R_1 = \infty\,\Omega$, $R_2 = 0\,\Omega$ and therefore $A_{CL} = 1$.

- The inverting amplifier is one of the most fundamental op-amp arrangements:



The gain of the inverting amplifier is:

$$A_{CL} = -\frac{R_2}{R_1}$$

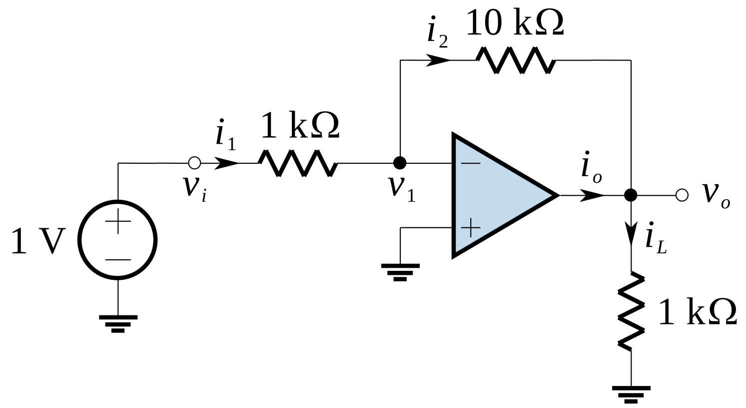
2.7 References

Hayt, W. & Kemmerly, J.: *Engineering Circuit Analysis*, 3rd *Ed*., McGraw-Hill, 1984.

Exercises

1.

For the ideal op-amp circuit below:



Determine the values of:

(a) v_1, i_1, i_2, v_o, i_L and i_o

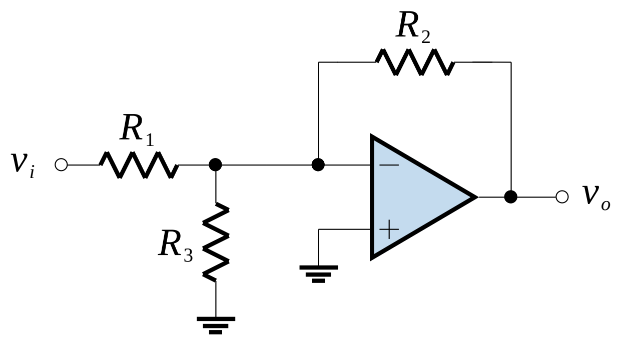
(b) the voltage gain v_o/v_i

(a) the current gain i_L/i_i

(d) the power gain P_L/P_i

2.

Given the ideal op-amp circuit below:



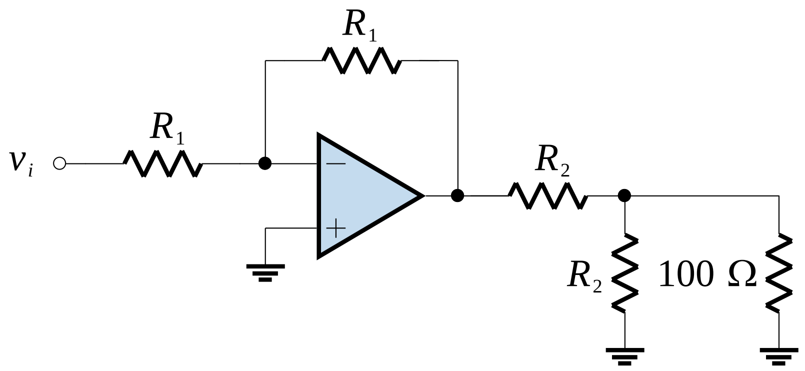
(a) Find the voltage gain v_o/v_i .

(b) Find the resistance "seen" by the voltage source v_i .

(c) How do the results of (a) and (b) differ from the case when $R_3 = \infty\,\Omega$? Why?

3.

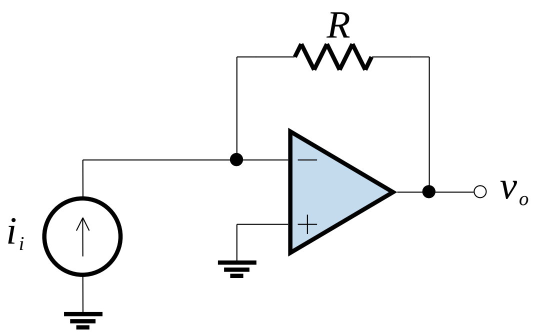
Consider the ideal op-amp circuit below:



Choose values for resistors R_1 and R_2 such that the $100\ \Omega$ resistor absorbs 10 mW when $v_i = 4\text{ V}$.

4.

Consider the ideal op-amp circuit below:

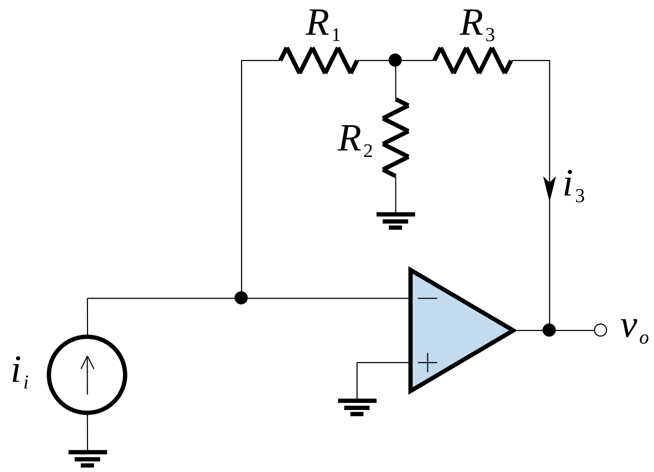


(a) Find the ratio v_o/i_i.

(b) Find the resistance "seen" by the current source i_i.

5.

Consider the ideal op-amp circuit below:



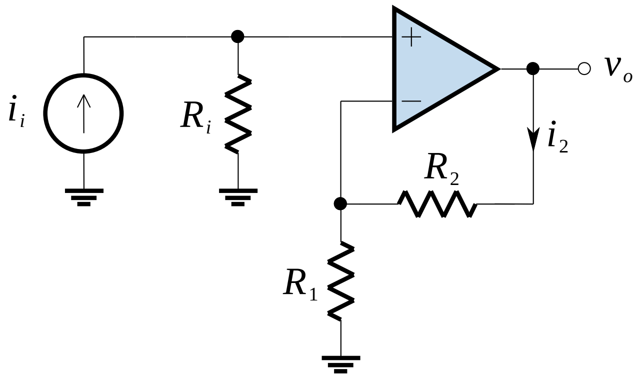
(a) Find the ratio v_o/i_i .

(b) Find the resistance "seen" by the current source i_i .

(c) Find the current gain i_3/i_i .

6.

Given the ideal op-amp circuit below:



(a) Find the ratio v_o/i_i .

(b) Find the resistance "seen" by the current source i_i .

(c) Find the current gain i_2/i_i .

3 Nodal and Mesh Analysis

Contents

3.2

Introduction

After becoming familiar with Ohm's Law and Kirchhoff's Laws and their application in the analysis of simple series and parallel resistive circuits, we must begin to analyse more complicated and practical circuits.

Physical systems that we want to analyse and design include electronic control circuits, communication systems, energy converters such as motors and generators, power distribution systems, mobile devices and embedded systems. We will also be confronted with allied problems involving heat flow, fluid flow, and the behaviour of various mechanical systems.

To cope with large and complex circuits, we need powerful and general methods of circuit analysis. Nodal analysis is a method which can be applied to any circuit, and mesh analysis is a method that can be applied to any *planar* circuit (i.e. to circuits that are able to be laid out on a 2D surface without crossing elements). Both of these methods are widely used in hand design and computer simulation. A third technique, known as *loop analysis*, generalises mesh analysis and can be applied to any circuit – it is effectively the "dual" of nodal analysis.

We will find that the judicious selection of an analysis technique can lead to a drastic reduction in the number of equations to solve, and we should therefore try to develop an ability to select the most convenient analysis method for a particular circuit.

3.1 Nodal Analysis

In general terms, nodal analysis for a circuit with N nodes proceeds as follows:

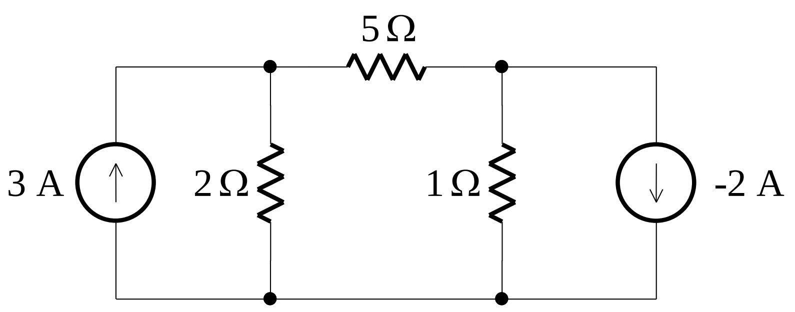
The general principle of nodal analysis

1. Select one node as the reference node, or *common* (all nodal voltages are defined with respect to this node in a positive sense).
2. Assign a voltage to each of the remaining $(N-1)$ nodes.
3. Write KCL at each node, in terms of the nodal voltages.
4. Solve the resulting set of simultaneous equations.

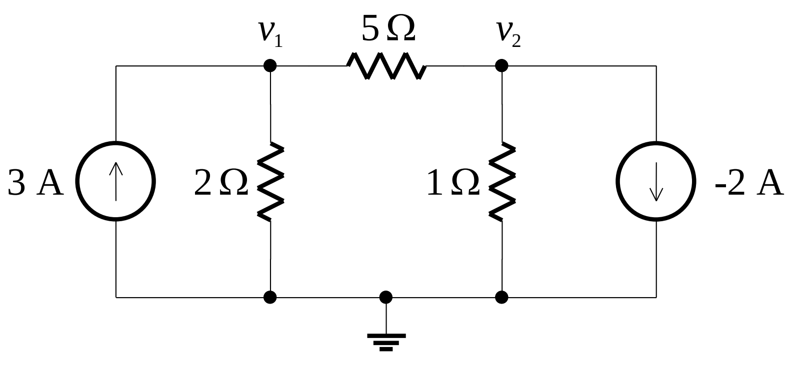
As will be seen, the method outlined above becomes a little complicated if the circuit contains voltage sources and / or controlled sources, but the principle remains the same.

EXAMPLE 3.1 Nodal Analysis with Independent Sources

We apply nodal analysis to the following 3-node circuit:



Following the steps above, we assign a reference node and then assign nodal voltages:



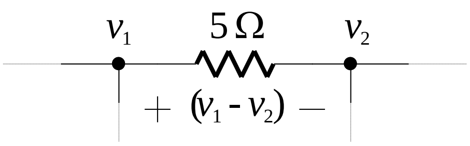
We chose the bottom node as the reference node, but either of the other two nodes could have been selected. A little simplification in the resultant equations is obtained if the node to which the greatest number of branches is connected is identified as the reference node.

In many practical circuits the reference node is one end of a power supply which is generally connected to a metallic case or chassis in which the circuit resides; the chassis is often connected through a good conductor to the Earth. Thus, the metallic case may be called "ground", or "earth", and this node becomes the most convenient reference node.

The distinction between "common" and "earth"

To avoid confusion, the reference node will be called the "common" unless it has been specifically connected to the Earth (such as the outside conductor on a digital storage oscilloscope, function generator, etc).

Note that the voltage across any branch in a circuit may be expressed in terms of nodal voltages. For example, in our circuit the voltage across the $5\,\Omega$ resistor is $(v_1 - v_2)$ with the positive polarity reference on the left:



We must now apply KCL to nodes 1 and 2. We do this be equating the total current *leaving* a node to zero. Thus:

$$\frac{v_1}{2} + \frac{v_1 - v_2}{5} - 3 = 0$$

$$\frac{v_2 - v_1}{5} + \frac{v_1}{1} + (-2) = 0$$

Simplifying, the equations can be written:

$$0.7v_1 - 0.2v_2 = 3$$

$$-0.2v_1 + 1.2v_2 = 2$$

Rewriting in matrix notation, we have:

$$\begin{bmatrix} 0.7 & -0.2 \\ -0.2 & 1.2 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$$

These equations may be solved by a simple process of elimination of variables, or by Cramer's rule and determinants. Using the latter method we have:

$$v_1 = \frac{\begin{vmatrix} 3 & -0.2 \\ 2 & 1.2 \end{vmatrix}}{\begin{vmatrix} 0.7 & -0.2 \\ -0.2 & 1.2 \end{vmatrix}} = \frac{3.6 + 0.4}{0.84 - 0.04} = \frac{4}{0.8} = 5\,\mathrm{V}$$

$$v_2 = \frac{\begin{vmatrix} 0.7 & 3 \\ -0.2 & 2 \end{vmatrix}}{0.8} = \frac{1.4 + 0.6}{0.8} = \frac{2}{0.8} = 2.5\,\mathrm{V}$$

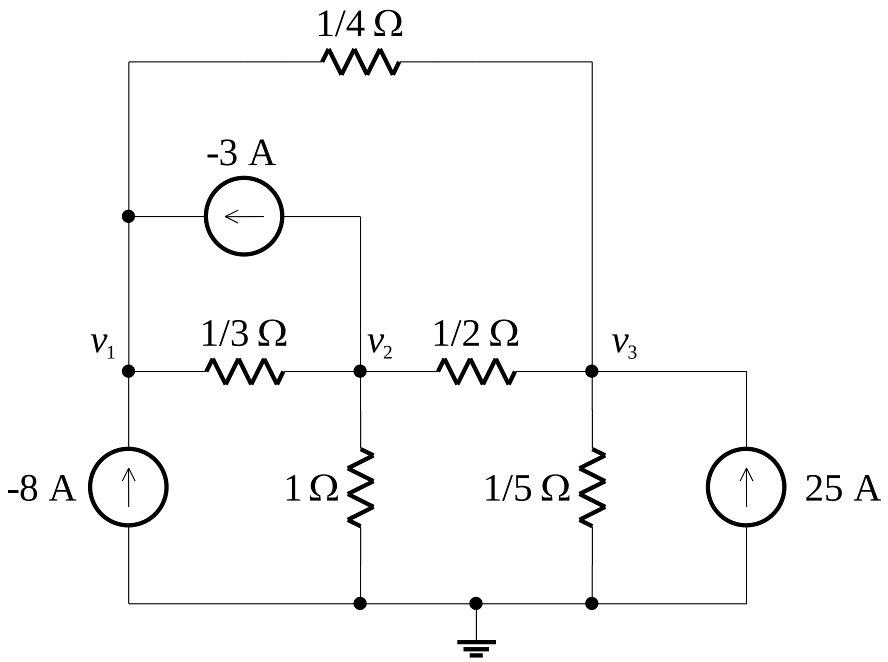
Everything is now known about the circuit – any voltage, current or power in the circuit may be found in one step. For example, the voltage at node 1 with respect to node 2 is $(v_1 - v_2) = 2.5\,\mathrm{V}$, and the current directed downward through the $2\,\Omega$ resistor is $v_1/2 = 2.5\,\mathrm{A}$.

3.1.1 Circuits with Resistors and Independent Current Sources Only

A further example will reveal some interesting mathematical features of nodal analysis, at least for the case of circuits containing only resistors and independent current sources. We will find it is much easier to consider conductance, *G*, rather than resistance, *R*, in the formulation of the equations.

EXAMPLE 3.2 Nodal Analysis with Independent Sources Only

A circuit is shown below with a convenient reference node and nodal voltages specified.



We sum the currents leaving node 1:

$$3(v_1 - v_2) + 4(v_1 - v_3) - (-8) - (-3) = 0$$
$$7v_1 - 3v_2 - 4v_3 = -11$$

At node 2:

$$3(v_2 - v_1) + 1v_2 + 2(v_2 - v_3) - 3 = 0$$
$$-3v_1 + 6v_2 - 2v_3 = 3$$

At node 3:

$$4(v_3 - v_1) + 2(v_3 - v_2) + 5v_3 - 25 = 0$$
$$-4v_1 - 2v_2 + 11v_3 = 25$$

Rewriting in matrix notation, we have:

$$\begin{bmatrix} 7 & -3 & -4 \\ -3 & 6 & -2 \\ -4 & -2 & 11 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} -11 \\ 3 \\ 25 \end{bmatrix}$$

For circuits that contain only resistors and independent current sources, we define the *conductance matrix* of the circuit as:

$$\mathbf{G} = \begin{bmatrix} 7 & -3 & -4 \\ -3 & 6 & -2 \\ -4 & -2 & 11 \end{bmatrix}$$

It should be noted that the nine elements of the matrix are the ordered array of the coefficients of the KCL equations, each of which is a conductance value. Thr first row is composed of the coefficients of the Kirchhoff current law equation at the first node, the coefficients being given in the order of v_1, v_2 and v_3. The second row applies to the second node, and so on.

The conductance matrix defined

The major diagonal (upper left to lower right) has elements that are positive. The conductance matrix is symmetrical about the major diagonal, and all elements not on this diagonal are negative. This is a general consequence of the systematic way in which we ordered the equations, and in circuits consisting of only resistors and independent current sources it provides a check against errors committed in writing the circuit equations.

We also define the voltage and current source vectors as:

$$\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} \qquad \mathbf{i} = \begin{bmatrix} -11 \\ 3 \\ 25 \end{bmatrix}$$

Nodal analysis expressed in matrix notation

Our KCL equations can therefore be written succinctly in matrix notation as:

$$\mathbf{G}\mathbf{v} = \mathbf{i}$$

The solution of the matrix equation is just:

$$\mathbf{v} = \mathbf{G}^{-1}\mathbf{i}$$

Computer programs that do nodal analysis use sophisticated numerical methods to efficiently invert the $\mathbf{G}$ matrix and solve for $\mathbf{v}$. When solving the equations by hand we resort to matrix reduction techniques, or use Cramer's rule (up to 3 x 3). Thus:

$$v_1 = \frac{\begin{vmatrix} -11 & -3 & -4 \\ 3 & 6 & -2 \\ 25 & -2 & 11 \end{vmatrix}}{\begin{vmatrix} 7 & -3 & -4 \\ -3 & 6 & -2 \\ -4 & -2 & 11 \end{vmatrix}}$$

To reduce work, we expand the numerator and denominator determinants by minors along their first columns to get:

$$\begin{aligned} v_1 &= \frac{-11\begin{vmatrix} 6 & -2 \\ -2 & 11 \end{vmatrix} - 3\begin{vmatrix} -3 & -4 \\ -2 & 11 \end{vmatrix} + 25\begin{vmatrix} -3 & -4 \\ 6 & -2 \end{vmatrix}}{7\begin{vmatrix} 6 & -2 \\ -2 & 11 \end{vmatrix} - (-3)\begin{vmatrix} -3 & -4 \\ -2 & 11 \end{vmatrix} + (-4)\begin{vmatrix} -3 & -4 \\ 6 & -2 \end{vmatrix}} \\ &= \frac{-11(62) - 3(-41) + 25(30)}{7(62) + 3(-41) - 4(30)} = \frac{-682 + 123 + 750}{434 - 123 - 120} \\ &= \frac{191}{191} = 1\ \text{V} \end{aligned}$$

Similarly:

$$v_2 = \frac{\begin{vmatrix} 7 & -11 & -4 \\ -3 & 3 & -2 \\ -4 & 25 & 11 \end{vmatrix}}{|191|} = 2\ \text{V} \qquad v_3 = \frac{\begin{vmatrix} 7 & -3 & -11 \\ -3 & 6 & 3 \\ -4 & -2 & 25 \end{vmatrix}}{|191|} = 3\ \text{V}$$

3.1.2 Nodal Analysis Using Branch Element Stamps

The previous example shows that nodal analysis leads to the equation $\mathbf{Gv} = \mathbf{i}$. We will now develop a method whereby the equation $\mathbf{Gv} = \mathbf{i}$ can be built up on an element-by-element basis by inspection of each branch in the circuit.

Consider a resistive element connected between nodes *i* and *j*:

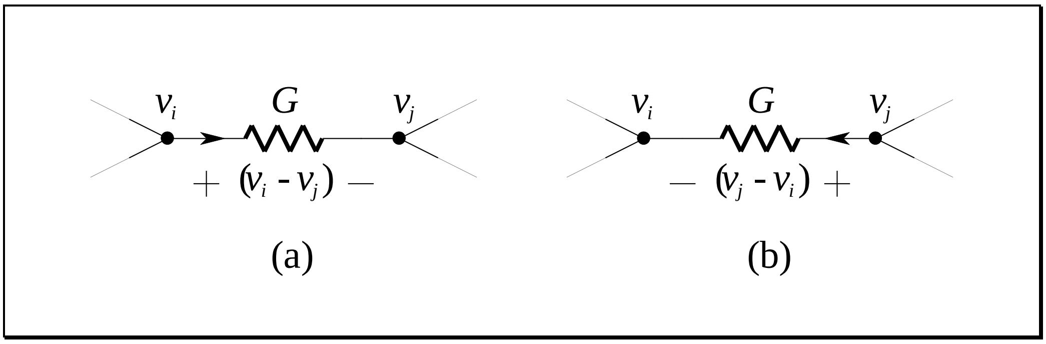


Figure 3.1

Suppose that we are writing the i^{th} KCL equation because we are considering the current leaving node *i* (see Figure 3.1a). The term that we would write in this equation to take into account the branch connecting nodes *i* and *j* is:

$$\cdots + G\left(v_i - v_j\right) + \cdots = 0 \tag{3.1}$$

This term appears in the i^{th} row when writing out the matrix equation.

If we are dealing with the j^{th} KCL equation because we are considering the current leaving node *j* (see Figure 3.1b) then the term that we would write in this equation to take into account the branch connecting nodes *j* and *i* is:

$$\cdots + G\left(v_j - v_i\right) + \cdots = 0 \tag{3.2}$$

This term appears in the j^{th} row when writing out the matrix equation.

Thus, the branch between nodes i and j contributes the following *element stamp* to the conductance matrix, $\mathbf{G}$:

The element stamp for a conductance

$$\begin{array}{c c} & \begin{array}{cc} i & j \end{array} \\ \begin{array}{c} i \\ j \end{array} & \begin{bmatrix} G & -G \\ -G & G \end{bmatrix} \end{array} \tag{3.3}$$

If node i or node j is the reference node, then the corresponding row and column are eliminated from the element stamp shown above.

For any circuit containing only resistors and independent current sources, the conductance matrix can now be built up by inspection. The result will be a $\mathbf{G}$ matrix where each diagonal element g_{ii} is the sum of conductances connected to node i, and each off-diagonal element g_{ij} is the total conductance between nodes i and j but with a negative sign.

Now consider a current source connected between nodes i and j:

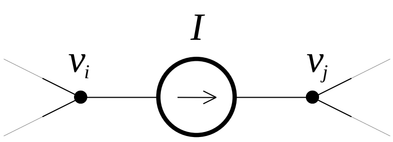


Figure 3.2

In writing out the i^{th} KCL equation we would introduce the term:

$$\cdots + I + \cdots = 0 \tag{3.4}$$

In writing out the j^{th} KCL equation we would introduce the term:

$$\cdots - I + \cdots = 0 \tag{3.5}$$

Thus, a current source contributes to the right-hand side (rhs) of the matrix equation the terms:

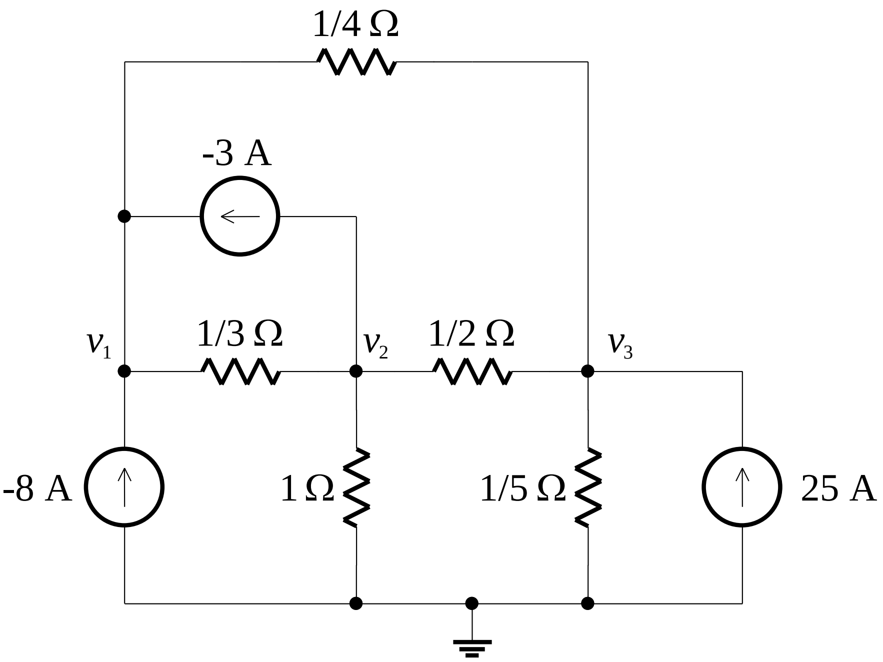
$$\begin{matrix} i \\ j \end{matrix}\begin{bmatrix} -I \\ I \end{bmatrix} \tag{3.6}$$

The element stamp for an independent current source

Thus, the $\mathbf{i}$ vector can also be built up by inspection – each row is the addition of all current sources *entering* a particular node. This makes sense since $\mathbf{Gv} = \mathbf{i}$ is the mathematical expression for KCL in the form of "current leaving a node = current entering a node".

EXAMPLE 3.3 Nodal Analysis Using the "Formal" Approach

We will analyse the previous circuit but use the "formal" approach to nodal analysis.



By inspection of each branch, we build the matrix equation:

$$\begin{matrix} \\ 1 \\ 2 \\ 3 \end{matrix}\begin{matrix} 1 & 2 & 3 \\ \left[\begin{matrix} 3+4 \\ -3 \\ -4 \end{matrix}\right. & \begin{matrix} -3 \\ 1+3+2 \\ -2 \end{matrix} & \left.\begin{matrix} -4 \\ -2 \\ 5+4+2 \end{matrix}\right] \end{matrix}\begin{matrix} \\ \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} \end{matrix} \begin{matrix} \\ = \end{matrix} \begin{matrix} \\ \begin{bmatrix} (-8)+(-3) \\ -(-3) \\ 25 \end{bmatrix} \end{matrix}$$

This is the same equation as derived previously.

3.1.3 Circuits with Voltage Sources

Voltage sources present a problem in undertaking nodal analysis, since by definition the voltage across a voltage source is *independent* of the current through it. Thus, when we consider a branch with a voltage source when writing a nodal equation, there is no way by which we can express the current through the branch as a function of the nodal voltages across the branch.

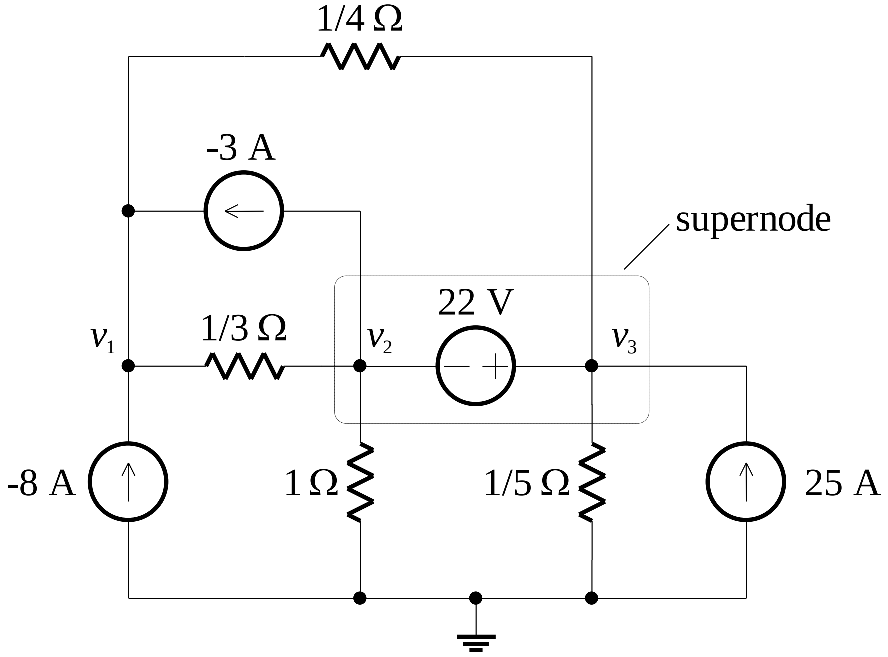
There are two ways around this problem. The more difficult is to assign an unknown current to each branch with a voltage source, proceed to apply KCL at each node, and then apply KVL across each branch with a voltage source. The result is a set of equations with an increased number of unknown variables.

The concept of a supernode

The easier method is to introduce the concept of a *supernode*. A supernode encapsulates the voltage source, and we apply KCL to both end nodes at the same time. The result is that the number of nodes at which we must apply KCL is reduced by the number of voltage sources in the circuit.

EXAMPLE 3.4 Nodal Analysis with Voltage Sources

Consider the circuit shown below, which is the same as the previous circuit except the $1/2\,\Omega$ resistor between nodes 2 and 3 has been replaced by a 22 V voltage source:



KCL at node 1 remains unchanged:

$$3(v_1 - v_2) + 4(v_1 - v_3) - (-8) - (-3) = 0$$
$$7v_1 - 3v_2 - 4v_3 = -11$$

We find six branches connected to the supernode around the 22 V source (suggested by a broken line in the figure). Beginning with the $1/3\,\Omega$ resistor branch and working clockwise, we sum the six currents leaving this supernode:

$$3(v_2 - v_1) + (-3) + 4(v_3 - v_1) + (-25) + 5v_3 + 1v_2 = 0$$
$$-7v_1 + 4v_2 + 9v_3 = 28$$

We need one additional equation since we have three unknowns, and this is provided by KVL between nodes 2 and 3 inside the supernode:

$$v_3 - v_2 = 22$$

Rewriting these last three equations in matrix form, we have:

$$\begin{bmatrix} 7 & -3 & -4 \\ -7 & 4 & 9 \\ 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} -11 \\ 28 \\ 22 \end{bmatrix}$$

The solution turns out to be $v_1 = -4.5\,\text{V}$, $v_2 = -15.5\,\text{V}$ and $v_3 = 6.5\,\text{V}$.

Note the lack of symmetry about the major diagonal in the $\mathbf{G}$ matrix as well as the fact that not all of the off-diagonal elements are negative. This is the result of the presence of the voltage source. Note also that it does not make sense to call the $\mathbf{G}$ matrix the *conductance* matrix, for the bottom row comes from the equation $-v_2 + v_3 = 22$, and this equation does not have any terms that are related to a conductance.

The presence of a dependent source destroys the symmetry in the **G** matrix

A supernode can contain any number of independent or dependent voltage sources. In general, the analysis procedure is the same as the example above – *one* KCL equation is written for currents leaving the supernode, and then a KVL equation is written for *each* voltage source inside the supernode.

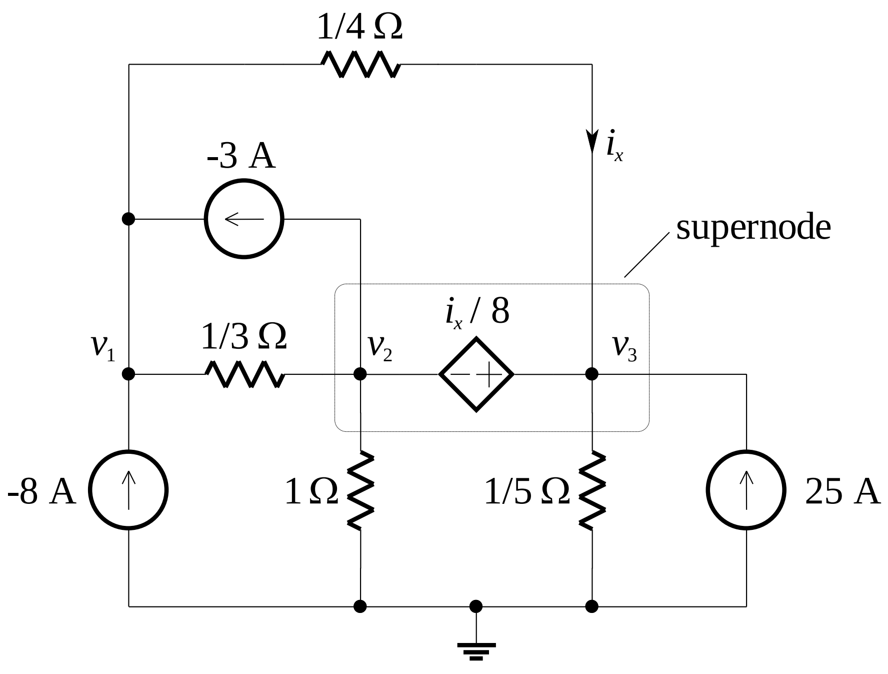
3.1.4 Circuits with Dependent Sources

Dependent current sources are fairly easy to include into nodal analysis – we just need to express the dependent current in terms of nodal voltages.

Dependent voltages sources are dealt with using the concept of the supernode. Of the two types of dependent voltage source, the current controlled voltage source (CCVS) requires the most effort to incorporate into nodal analysis. We will analyse this case before summarizing the method of nodal analysis for any resistive circuit.

EXAMPLE 3.5 Nodal Analysis with Dependent Voltage Sources

Consider the circuit shown below, which is the same as the previous circuit except now the 22 V voltage source has been replaced by a current controlled voltage source:



KCL at node 1 remains unchanged:

$$3(v_1 - v_2) + 4(v_1 - v_3) - (-8) - (-3) = 0$$
$$7v_1 - 3v_2 - 4v_3 = -11$$

KCL at the supernode, formed by considering nodes 2 and 3 together, remains unchanged:

$$\begin{aligned} 3(v_2 - v_1) + (-3) + 4(v_3 - v_1) + (-25) + 5v_3 + 1v_2 &= 0 \\ -7v_1 + 4v_2 + 9v_3 &= 28 \end{aligned}$$

Finally, we turn our attention to the dependent source. We rewrite the dependent current in terms of nodal voltages:

$$i_x = 4(v_1 - v_3)$$

and then we write KVL for the dependent source as:

$$\begin{aligned} v_3 - v_2 &= \frac{i_x}{8} = \frac{4(v_1 - v_3)}{8} \\ v_1 + 2v_2 - 3v_3 &= 0 \end{aligned}$$

In matrix form we now have:

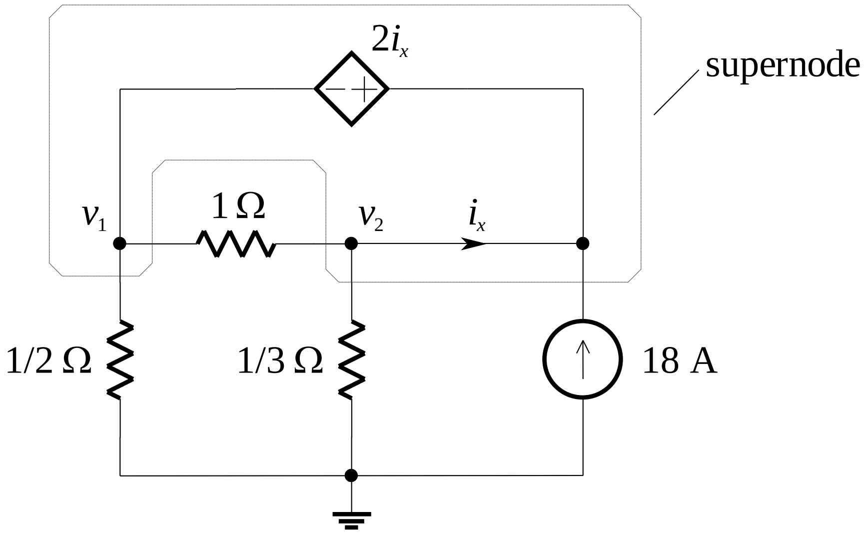
$$\begin{bmatrix} 7 & -3 & -4 \\ -7 & 4 & 9 \\ 1 & 2 & -3 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} -11 \\ 28 \\ 0 \end{bmatrix}$$

The nodal voltages turn out to be $v_1 = 1.431\text{V}$, $v_2 = 3.373\text{V}$ and $v_3 = 2.726\ \text{V}$.

In the previous example, the source dependency (i_x) was easily expressed using Ohm's Law across the $1/4\,\Omega$ resistor. In other cases, the dependency may need to be found using KCL.

EXAMPLE 3.6 Nodal Analysis with a Dependent Voltage Source

Consider the circuit shown below:



KCL at the supernode, formed by considering nodes 1 and 2 together, gives:

$$2v_1 + 1(v_1 - v_2) + +1(v_2 - v_1) + 3v_2 - 18 = 0$$
$$2v_1 + 3v_2 = 18$$

Note how the 1Ω resistor contributes nothing to the KCL equation. Next, we turn our attention to the dependent source inside the supernode. We rewrite the dependent current in terms of nodal voltages using KCL at node 2:

$$\frac{v_2 - v_1}{1} + 3v_2 + i_x = 0$$
$$i_x = v_1 - 4v_2$$

and then we write KVL for the dependent source as:

$$v_2 - v_1 = 2i_x = 2(v_1 - 4v_2)$$
$$3v_1 - 9v_2 = 0$$

In matrix form we now have:

$$\begin{bmatrix} 2 & 3 \\ 3 & -9 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} 18 \\ 0 \end{bmatrix}$$

The nodal voltages turn out to be $v_1 = 6\text{ V}$, $v_2 = 2\text{ V}$.

3.1.5 Summary of Nodal Analysis

We perform nodal analysis for any resistive circuit with N nodes by the following method:

The general procedure to follow when undertaking nodal analysis

1. Make a neat, simple, circuit diagram. Indicate all element and source values. Each source should have its reference symbol.

2. Select one node as the reference node, or *common*. Then write the node voltages v_1, v_2, …, v_{N-1} at their respective nodes, remembering that each node voltage is understood to be measured with respect to the chosen reference.

3. If the circuit contains dependent sources, express those sources in terms of the variables v_1, v_2, …, v_{N-1}, if they are not already in that form.

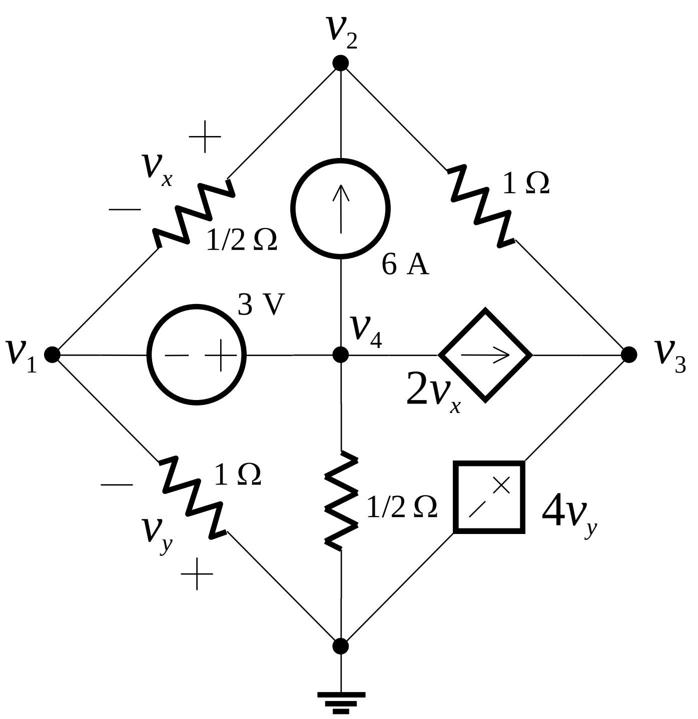
4. If the circuit contains voltage sources, temporarily modify the given circuit by replacing each voltage source by a short-circuit to form *supernodes*, thus reducing the number of nodes by one for each voltage source that is present. The assigned nodal voltages should not be changed. Relate each supernode's source voltage to the nodal voltages.

5. Apply KCL at each of the nodes or supernodes. If the circuit has only resistors and independent current sources, then the equations may be built using the "element stamp" approach.

6. Solve the resulting set of simultaneous equations.

EXAMPLE 3.7 Nodal Analysis with all Four Types of Sources

Consider the circuit shown below, which contains all four types of sources and has five nodes.



We select the bottom node as the reference and assign v_1 to v_4 in a clockwise direction starting from the left node.

We next relate the controlled sources to the nodal voltages:

$$2v_x = 2(v_2 - v_1)$$
$$4v_y = -4v_1$$

We form supernodes around the two voltage sources, and write relations for them in terms of the nodal voltages:

$$v_4 - v_1 = 3$$
$$v_3 = 4v_y = -4v_1$$

Thanks to the supernodes, we see that we only need to write KCL equations at node 2 and the supernode containing both nodes 1 and 4. At node 2:

$$2(v_2 - v_1) - 6 + 1(v_2 - v_3) = 0$$
$$-2v_1 + 3v_2 - v_3 = 6$$

while at the supernode:

$$2(v_1 - v_2) + 6 + \underbrace{2(v_2 - v_1)}_{2v_x} + 2v_4 + 1v_1 = 0$$

$$v_1 + 2v_4 = -6$$

Thus, we obtain four equations in the four node voltages:

$$\begin{bmatrix} -1 & 0 & 0 & 1 \\ 4 & 0 & 1 & 0 \\ -2 & 3 & -1 & 0 \\ 1 & 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \end{bmatrix} = \begin{bmatrix} 3 \\ 0 \\ 6 \\ -6 \end{bmatrix}$$

The solution is:

$$\begin{aligned} v_1 &= -4 \text{ V} \\ v_2 &= 14/3 \text{ V} \\ v_3 &= 16 \text{ V} \\ v_4 &= -1 \text{ V} \end{aligned}$$

The technique of nodal analysis described here is completely general and can always be applied to any electrical circuit.

3.2 Mesh Analysis

Before we embark on mesh analysis, we need to define the concept of: a planar circuit, a path through a circuit, a loop and a mesh. We can then outline the analysis strategy using these terms.

3.2.1 Planar Circuits

Planar circuits defined

A *planar* circuit is one where it is possible to draw the circuit on a plane surface in such a way that no branch passes over or under any other branch.

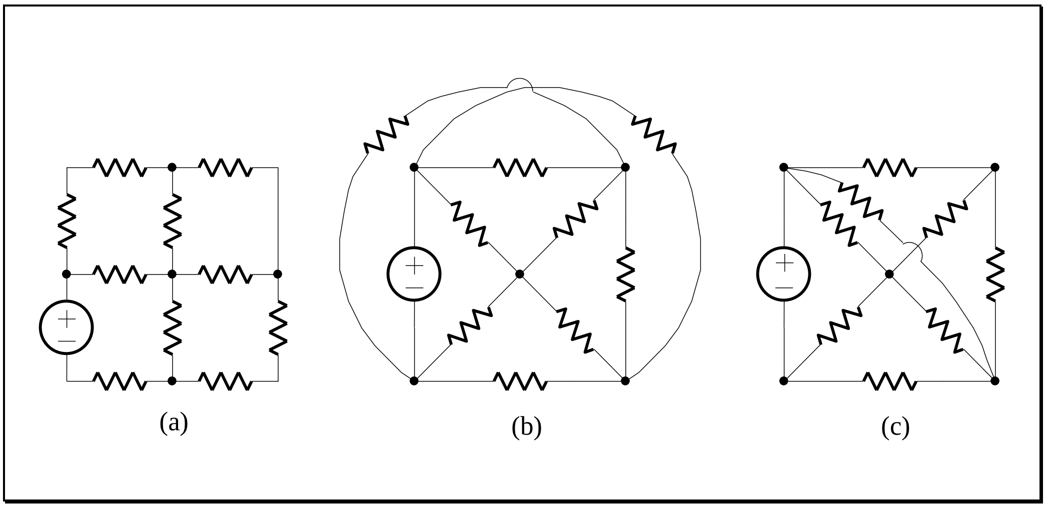


Figure 3.3

In the figure above, circuit (a) is planar, circuit (b) is nonplanar and circuit (c) is planar, but drawn so that it appears nonplanar.

3.2.2 Paths, Loops and Meshes

A *path* is made through a circuit when we start on one node and traverse elements and nodes without encountering any nodes previously visited. A *loop* is any *closed path* – i.e. the last node visited is the same as the starting node. A *mesh* is a loop which does not contain any other loops within it. A mesh is a property of a planar circuit and is not defined for a nonplanar circuit.

Examples of loops and meshes

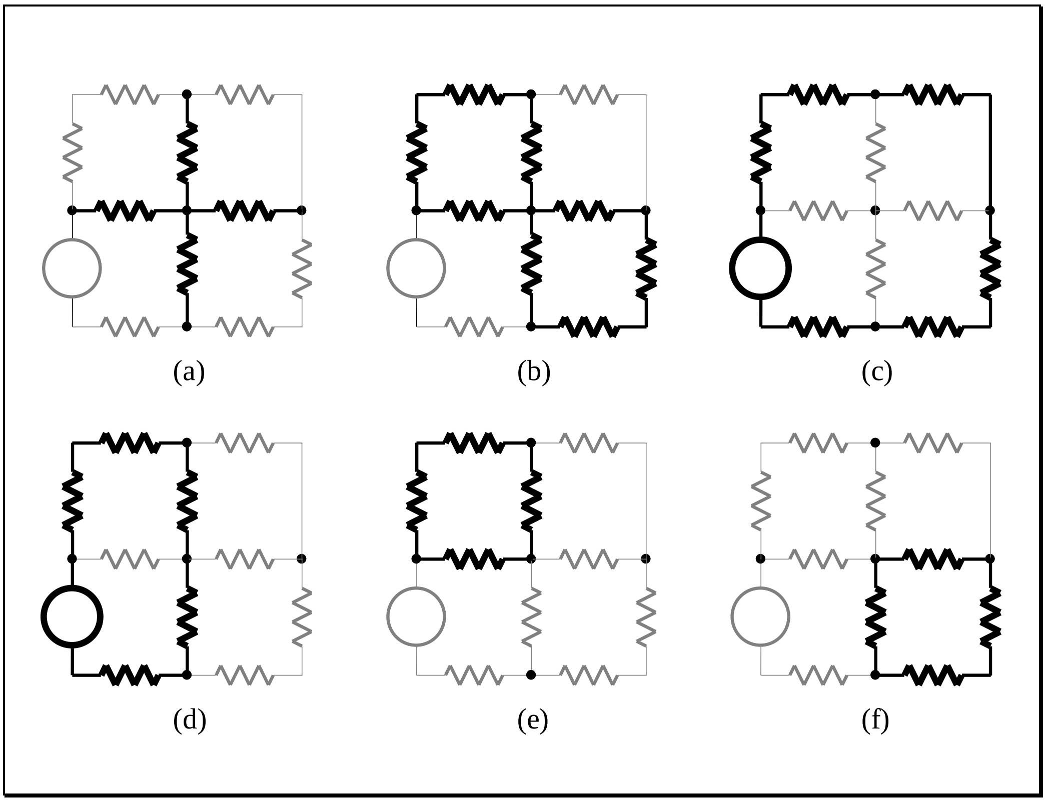


Figure 3.4

In the figure above, the set of branches in (a) identified by the heavy lines is neither a path nor a loop. In (b) the set of branches is not a path since it can be traversed only by passing through the central node twice. In (c) the closed path is a loop but not a mesh. In (d) the closed path is also a loop but not a mesh. In (e) and (f) each of the closed paths is both a loop and a mesh. This circuit contains four meshes.

3.2.3 Mesh Current

We define a *mesh current*[1] as a "mathematical" (or imaginary) current in which charge flows only around the perimeter of a mesh. A mesh current is indicated by a curved arrow that almost closes on itself and is drawn inside the appropriate mesh.

Mesh currents defined

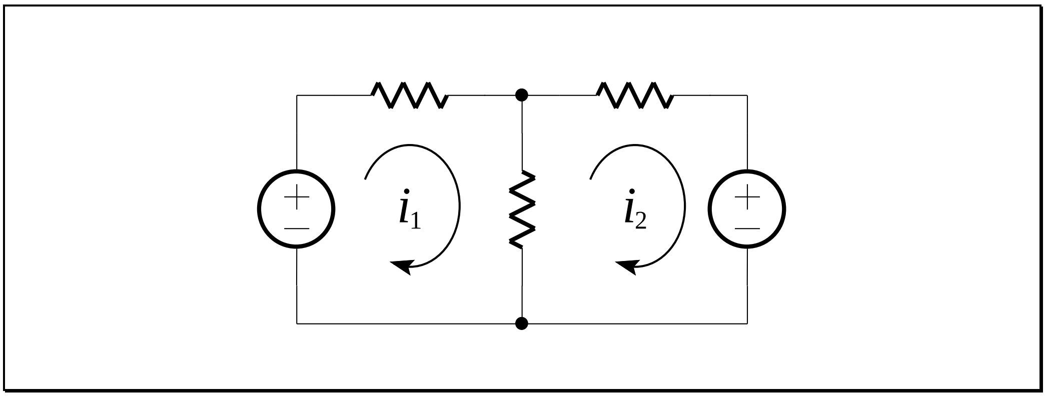


Figure 3.5

Although the direction of mesh currents is arbitrary, we draw the mesh currents in a clockwise direction so that a symmetry in the equations results when performing mesh analysis. One of the great advantages of mesh currents is that KCL is automatically satisfied, and no branch can appear in more than two meshes.

Branch currents can be expressed in terms of mesh currents

We no longer have a current or current arrow shown on each branch in the circuit. The current through any branch may be determined by superimposing each mesh current that exists in it. For example, the branch current heading *down* in the middle resistor in the circuit above is given by $(i_1 - i_2)$.

[1] It was the famous Scottish mathematical physicist James Clark Maxwell who invented the concept of a mesh current, and the associated methodology of formulating the "mesh equations". The analysis of planar circuits using mesh currents was thus reduced to solving a set of linear equations, in the same manner as nodal analysis.

3.2.4 Mesh Analysis Methodology

In general terms, mesh analysis for a planar circuit with M meshes proceeds as follows:

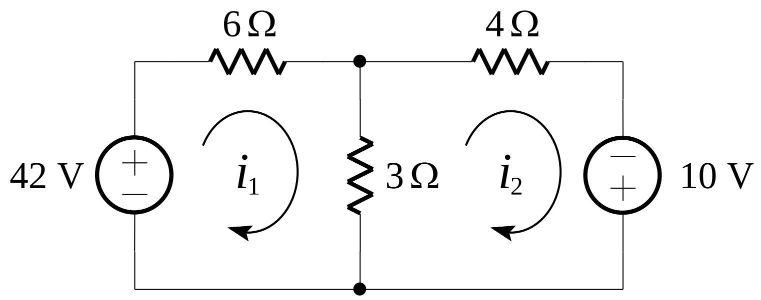
1. Assign a clockwise mesh current in each of the M meshes.
2. Write KVL around each mesh, in terms of the mesh currents.
3. Solve the resulting set of simultaneous equations.

The general principle of mesh analysis

As will be seen, the method outlined above becomes a little complicated if the circuit contains current sources and / or controlled sources, but the principle remains the same.

EXAMPLE 3.8 Mesh Analysis with Independent Sources

A two-mesh circuit is shown below.



We apply KVL to each mesh. For the left-hand mesh:

$$\begin{aligned} 42 - 6i_1 - 3(i_1 - i_2) &= 0 \\ 9i_1 - 3i_2 &= 42 \end{aligned}$$

For the right-hand mesh:

$$\begin{aligned} -3(i_2 - i_1) - 4i_2 + 10 &= 0 \\ -3i_1 + 7i_2 &= 10 \end{aligned}$$

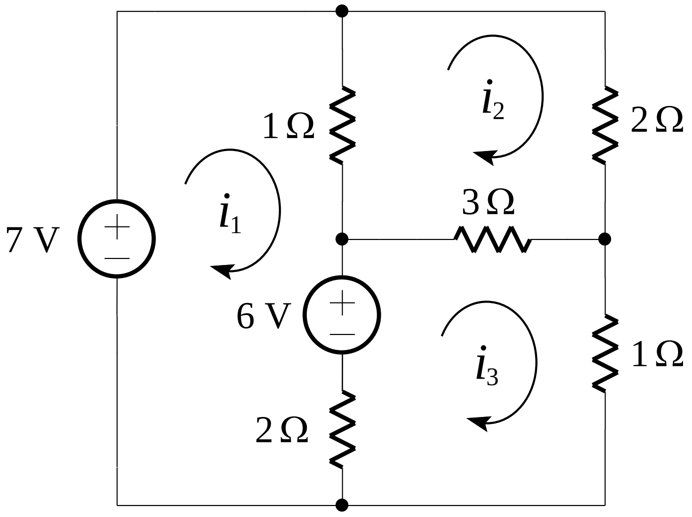
The solution is obtained by solving simultaneously: $i_1 = 6\text{ A}$ and $i_2 = 4\text{ A}$.

3.2.5 Circuits with Resistors and Independent Voltage Sources Only

When the circuit contains only resistors and voltages sources, the KVL equations have a certain symmetrical form and we can define a *resistance matrix* with the circuit. We will find again that the matrix equation can be formulated by inspection of the circuit.

EXAMPLE 3.9 Mesh Analysis with Independent Sources Only

Consider the five-node, three-mesh circuit shown below.



The three required mesh currents are assigned as indicated, and we methodically apply KVL about each mesh:

$$
\begin{aligned}
7-1(i_1-i_2)-6-2(i_1-i_3)&=0\\
-1(i_2-i_1)-2i_2-3(i_2-i_3)&=0\\
-2(i_3-i_1)+6-3(i_3-i_2)-1i_3&=0
\end{aligned}
$$

Simplifying and writing as a matrix equation:

$$
\begin{bmatrix} 3 & -1 & -2 \\ -1 & 6 & -3 \\ -2 & -3 & 6 \end{bmatrix}
\begin{bmatrix} i_1 \\ i_2 \\ i_3 \end{bmatrix}
=
\begin{bmatrix} 1 \\ 0 \\ 6 \end{bmatrix}
$$

For circuits that contain only resistors and independent voltage sources, we define the *resistance matrix* of the circuit as:

The resistance matrix defined

$$\mathbf{R} = \begin{bmatrix} 3 & -1 & -2 \\ -1 & 6 & -3 \\ -2 & -3 & 6 \end{bmatrix}$$

Once again we note the symmetry about the major diagonal. This occurs *only* for circuits with resistors and independent voltage sources when we order the equations correctly (rows correspond to meshes).

We also define the current and voltage source vectors as:

$$\mathbf{i} = \begin{bmatrix} i_1 \\ i_2 \\ i_3 \end{bmatrix} \qquad \mathbf{v} = \begin{bmatrix} 1 \\ 0 \\ 6 \end{bmatrix}$$

Our KVL equations can therefore be written succinctly in matrix notation as:

$$\mathbf{Ri} = \mathbf{v}$$

Mesh analysis expressed in matrix notation

Applying Cramer's rule to the formulation for i_1 gives:

$$i_1 = \frac{\begin{vmatrix} 1 & -1 & -2 \\ 0 & 6 & -3 \\ 6 & -3 & 6 \end{vmatrix}}{\begin{vmatrix} 3 & -1 & -2 \\ -1 & 6 & -3 \\ -2 & -3 & 6 \end{vmatrix}} = \frac{27 - 0 + 90}{81 - 12 - 30} = \frac{117}{39} = 3\,\text{A}$$

The other mesh currents are: $i_2 = 2\,\text{A}$ and $i_3 = 3\,\text{A}$.

3.2.6 Circuits with Current Sources

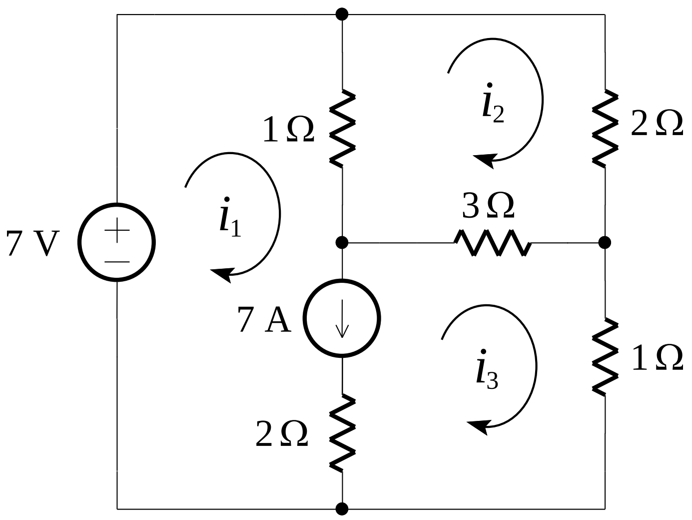
When a mesh has a current source in it, we must modify the procedure for forming the circuit equations. There are two possible methods. In the first method, we can relate the source current to the assigned mesh currents, assign an arbitrary voltage across it (thereby increasing the number of variables by one) and write KVL equations using this voltage. Alternately, a better method is to take a lead from nodal analysis and formulate the *dual* of a supernode - a *supermesh*.

The concept of a supermesh

To create a supermesh, we open-circuit or remove current sources, thereby reducing the total number of meshes. We apply KVL only to those meshes in the modified circuit.

EXAMPLE 3.10 Mesh Analysis with Current Sources

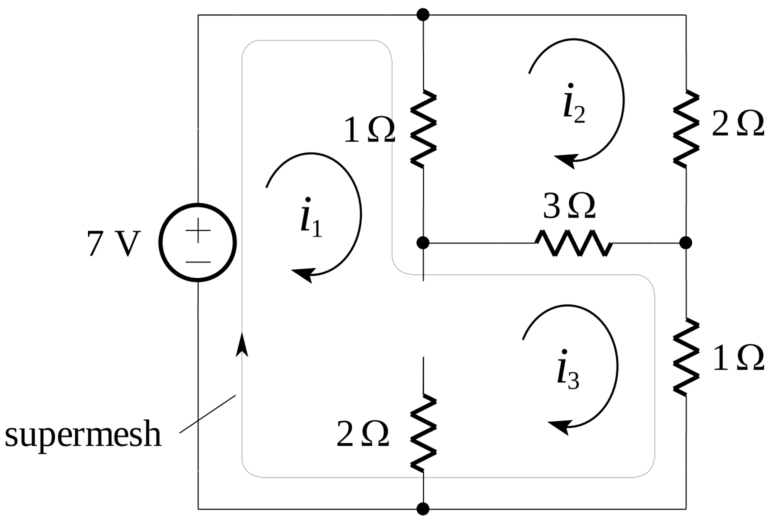
Consider the circuit shown below in which a 7 A independent current source is in the common boundary of two meshes.



For the independent current source, we relate the source current to the mesh currents:

$$i_1 - i_3 = 7$$

We then mentally open-circuit the current source, and form a supermesh whose interior is that of meshes 1 and 3:



Applying KVL about the supermesh:

$$7 - 1(i_1 - i_2) - 3(i_3 - i_2) - 1i_3 = 0$$

Around mesh 2 we have:

$$-1(i_2 - i_1) - 2i_2 - 3(i_2 - i_3) = 0$$

Rewriting these last three equations in matrix form, we have:

$$\begin{bmatrix} 1 & 0 & -1 \\ 1 & -4 & 4 \\ -1 & 6 & -3 \end{bmatrix} \begin{bmatrix} i_1 \\ i_2 \\ i_3 \end{bmatrix} = \begin{bmatrix} 7 \\ 7 \\ 0 \end{bmatrix}$$

Notice that we have lost all symmetry in the matrix equation $\mathbf{Ri} = \mathbf{v}$, and we can no longer call $\mathbf{R}$ the resistance matrix. Applying Cramer's rule for i_1:

$$i_1 = \frac{\begin{vmatrix} 7 & 0 & -1 \\ 7 & -4 & 4 \\ 0 & 6 & -3 \end{vmatrix}}{\begin{vmatrix} 1 & 0 & -1 \\ 1 & -4 & 4 \\ -1 & 6 & -3 \end{vmatrix}} = \frac{-84 - 42}{-12 - 2} = \frac{-126}{-14} = 9\text{ A}$$

The other mesh currents are: $i_2 = 2.5\text{ A}$ and $i_3 = 2\text{ A}$.

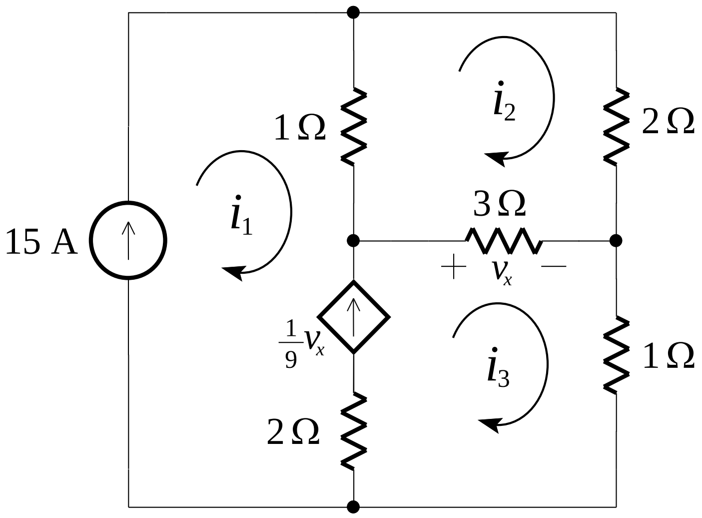
3.2.7 Circuits with Dependent Sources

Dependent voltage sources are fairly easy to include into mesh analysis – we just need to express the dependent voltage in terms of mesh currents.

Dependent current sources are dealt with using the concept of the supermesh. Of the two types of dependent current source, the voltage controlled current source (VCCS) requires the most effort to incorporate into mesh analysis. We will analyse this case before summarizing the method of mesh analysis for any resistive circuit.

EXAMPLE 3.11 Mesh Analysis with Dependent Current Sources

In the circuit shown below we have both a dependent and an independent current source:



For the independent current source, we relate the source current to the mesh currents:

$$i_1 = 15$$

Turning our attention to the dependent source, we describe the dependency in terms of the mesh currents:

$$i_3 - i_1 = \frac{v_x}{9} = \frac{3(i_3 - i_2)}{9}$$
$$-3i_1 + i_2 + 2i_3 = 0$$

Since the current sources appear in meshes 1 and 3, when they are open-circuited, only mesh 2 remains. Around mesh 2 we have:

$$-1(i_2 - i_1) - 2i_2 - 3(i_2 - i_3) = 0$$

In matrix form we now have:

$$\begin{bmatrix} 1 & 0 & 0 \\ -3 & 1 & 2 \\ -1 & 6 & -3 \end{bmatrix} \begin{bmatrix} i_1 \\ i_2 \\ i_3 \end{bmatrix} = \begin{bmatrix} 15 \\ 0 \\ 0 \end{bmatrix}$$

The mesh currents turn out to be $i_1 = 15\text{ A}$, $i_2 = 11\text{ A}$ and $i_3 = 17\text{ A}$. We note that we wasted a little time in assigning a mesh current i_1 to the left mesh – we should simply have indicated a mesh current and labelled it 15 A.

3.2.8 Summary of Mesh Analysis

The general procedure to follow when undertaking mesh analysis

We perform mesh analysis for any resistive circuit with M meshes by the following method:

1. Make certain that the circuit is a planar circuit. If it is nonplanar, then mesh analysis is not applicable.

2. Make a neat, simple, circuit diagram. Indicate all element and source values. Each source should have its reference symbol.

3. Assuming that the circuit has M meshes, assign a clockwise mesh current in each mesh, $i_1, i_2, \ldots, i_M$.

4. If the circuit contains dependent sources, express those sources in terms of the variables $i_1, i_2, \ldots, i_M$, if they are not already in that form.

5. If the circuit contains current sources, temporarily modify the given circuit by replacing each current source by an open-circuit to form *supermeshes*, thus reducing the number of meshes by one for each current source that is present. The assigned mesh currents should not be changed. Relate each source current to the mesh currents.

6. Apply KVL around each of the meshes or supermeshes. If the circuit has only resistors and independent voltage sources, then the equations may be formed by inspection.

7. Solve the resulting set of simultaneous equations.

3.3 Summary

- Nodal analysis can be applied to any circuit. Apart from relating source voltages to nodal voltages, the equations of nodal analysis are formed from application of Kirchhoff's Current Law.

- In nodal analysis, a supernode is formed by short-circuiting a voltage source and treating the two ends as a single node.

- Mesh analysis can only be applied to planar circuits. Apart from relating source currents to mesh currents, the equations of mesh analysis are formed from application of Kirchhoff's Voltage Law.

- In mesh analysis, a supermesh is formed by open-circuiting a current source and treating the perimeter of the original two meshes as a single mesh.

3.4 References

Hayt, W. & Kemmerly, J.: *Engineering Circuit Analysis*, 3rd *Ed.*, McGraw-Hill, 1984.

Exercises

1.

(a) Find the value of the determinant:

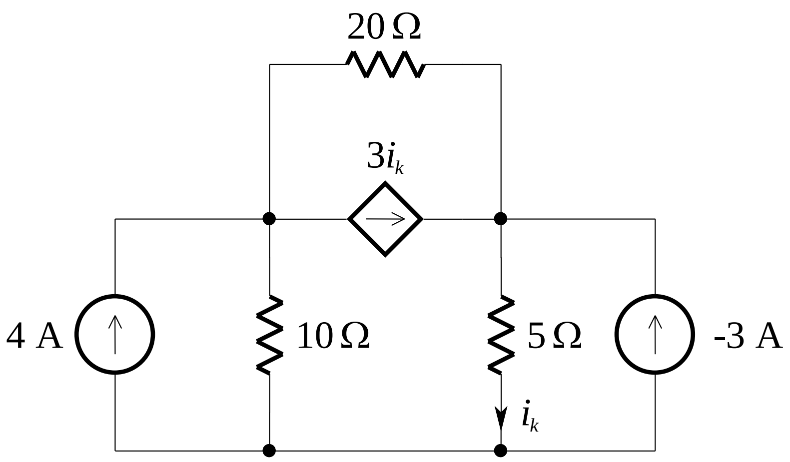
$$\Delta = \begin{vmatrix} 2 & -1 & 0 & -3 \\ -1 & 1 & 0 & -1 \\ 4 & 0 & 3 & -2 \\ -3 & 0 & 0 & 1 \end{vmatrix}$$

(b) Use Cramer's rule to find v_1, v_2 and v_3 if:

$$\begin{aligned} 2v_1 - 35 - v_2 + 3v_3 &= 0 \\ -2v_3 - 3v_2 - 4v_1 &= 56 \\ v_2 + 3v_1 - 28 - v_3 &= 0 \end{aligned}$$

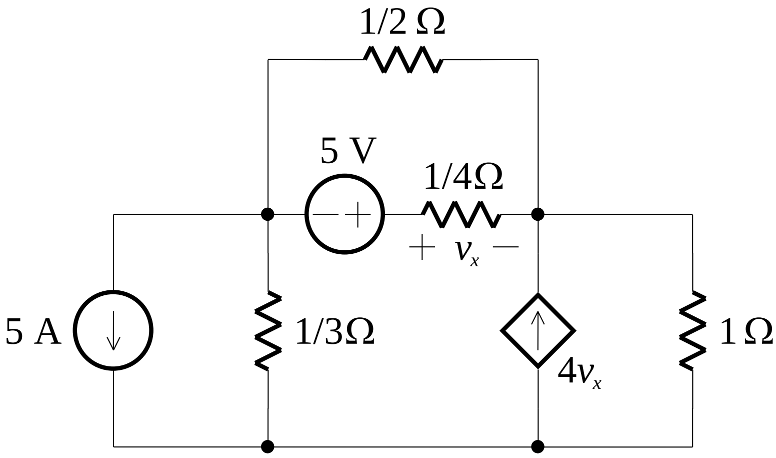
2.

Use nodal techniques to determine i_k in the circuit shown below:



3.

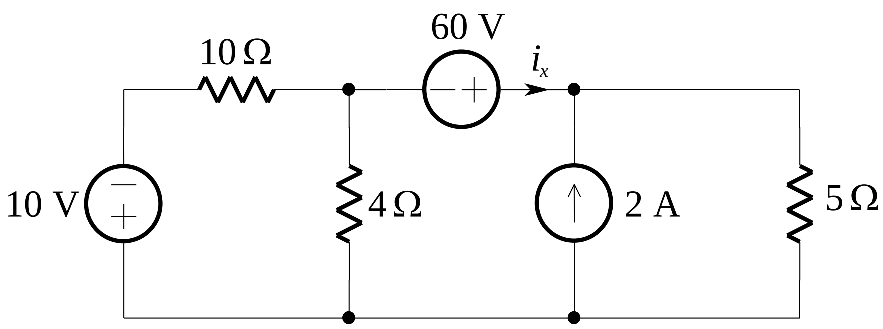
Set up nodal equations for the circuit shown below and then find the power furnished by the 5 V source.



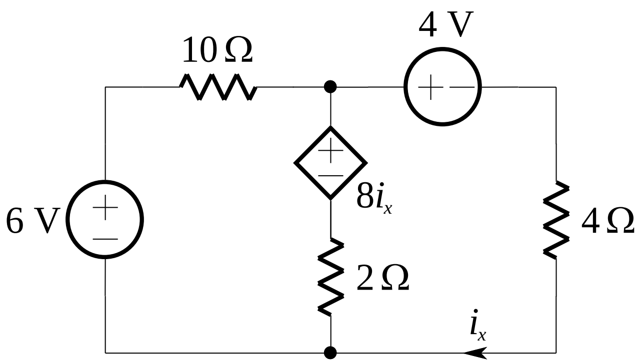
4.

Write mesh equations and then determine i_x in each of the circuits shown below:

(a)

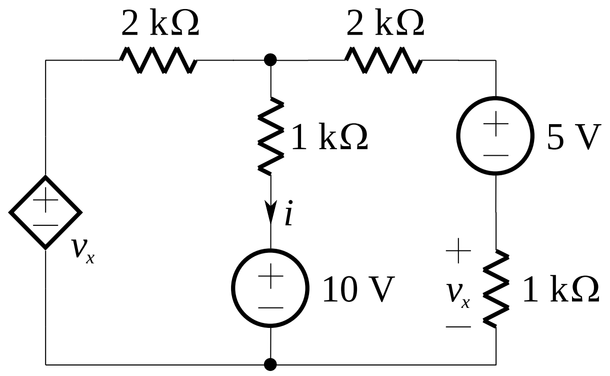


(b)



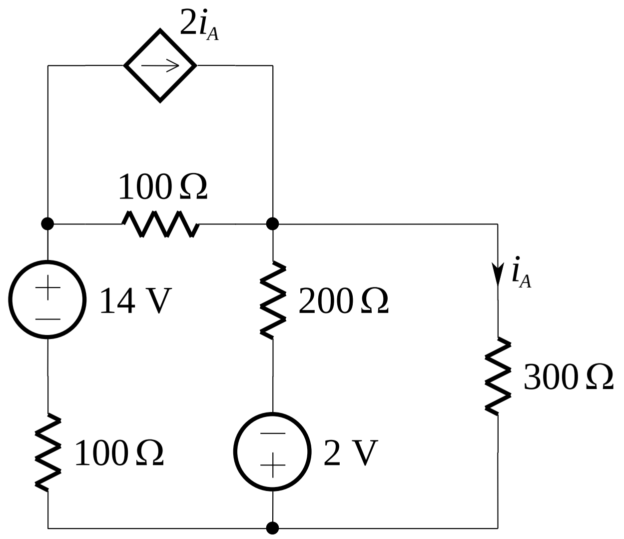
5.

Assign mesh currents in the circuit below, write a set of mesh equations, and determine *i*.

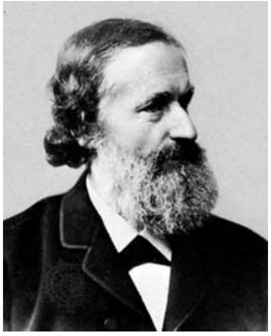


6.

With reference to the circuit shown below, use mesh equations to find i_A and the power supplied by the dependent source.



Gustav Robert Kirchhoff (1824-1887)



Kirchhoff was born in Russia, and showed an early interest in mathematics. He studied at the University of Königsberg, and in 1845, while still a student, he pronounced Kirchhoff's Laws, which allow the calculation of current and voltage for *any* circuit. They are the Laws electrical engineers apply on a routine basis – they even apply to non-linear circuits such as those containing semiconductors, or distributed parameter circuits such as microwave striplines.

He graduated from university in 1847 and received a scholarship to study in Paris, but the revolutions of 1848 intervened. Instead, he moved to Berlin where he met and formed a close friendship with Robert Bunsen, the inorganic chemist and physicist who popularized use of the "Bunsen burner".

In 1857 Kirchhoff extended the work done by the German physicist Georg Simon Ohm, by describing charge flow in three dimensions. He also analysed circuits using topology. In further studies, he offered a general theory of how electricity is conducted. He based his calculations on experimental results which determine a constant for the speed of the propagation of electric charge. Kirchhoff noted that this constant is approximately the speed of light – but the greater implications of this fact escaped him. It remained for James Clerk Maxwell to propose that light belongs to the electromagnetic spectrum.

Kirchhoff's most significant work, from 1859 to 1862, involved his close collaboration with Bunsen. Bunsen was in his laboratory, analysing various salts that impart specific colours to a flame when burned. Bunsen was using coloured glasses to view the flame. When Kirchhoff visited the laboratory, he suggested that a better analysis might be achieved by passing the light from the flame through a prism. The value of spectroscopy became immediately clear. Each element and compound showed a spectrum as unique as any fingerprint, which could be viewed, measured, recorded and compared.

Spectral analysis, Kirchhoff and Bunsen wrote not long afterward, promises "the chemical exploration of a domain which up till now has been completely closed." They not only analysed the known elements, they discovered new ones. Analyzing salts from evaporated mineral water, Kirchhoff and Bunsen detected a blue spectral line – it belonged to an element they christened *caesium* (from the Latin *caesius*, sky blue). Studying lepidolite (a lithium-based mica) in 1861, Bunsen found an alkali metal he called rubidium (from the Latin *rubidius*, deepest red). Both of these elements are used today in atomic clocks. Using spectroscopy, ten more new elements were discovered before the end of the century, and the field had expanded enormously – between 1900 and 1912 a "handbook" of spectroscopy was published by Kayser in six volumes comprising five thousand pages!

"[Kirchhoff is] a perfect example of the true German investigator. To search after truth in its purest shape and to give utterance with almost an abstract self-forgetfulness, was the religion and purpose of his life." – Robert von Helmholtz, 1890.

Kirchhoff's work on spectrum analysis led on to a study of the composition of light from the Sun. He was the first to explain the dark lines (Fraunhofer lines) in the Sun's spectrum as caused by absorption of particular wavelengths as the light passes through a gas. Kirchhoff wrote "It is plausible that spectroscopy is also applicable to the solar atmosphere and the brighter fixed stars." We can now analyse the collective light of a hundred billion stars in a remote galaxy billions of light-years away – we can tell its composition, its age, and even how fast the galaxy is receding from us – simply by looking at its spectrum!

As a consequence of his work with Fraunhofer's lines, Kirchhoff developed a general theory of emission and radiation in terms of thermodynamics. It stated that a substance's capacity to emit light is equivalent to its ability to absorb it at the same temperature. One of the problems that this new theory created was the "blackbody" problem, which was to plague physics for forty years. This fundamental quandary arose because heating a black body – such as a metal bar – causes it to give off heat and light. The spectral radiation, which depends only on the temperature and not on the material, could not be predicted by classical physics. In 1900 Max Planck solved the problem by discovering quanta, which had enormous implications for twentieth-century science.

In 1875 he was appointed to the chair of mathematical physics at Berlin and he ceased his experimental work. An accident-related disability meant he had to spend much of his life on crutches or in a wheelchair. He remained at the University of Berlin until he retired in 1886, shortly before his death in 1887.

4 Circuit Analysis Techniques

Contents

4.2

Introduction

Many of the circuits that we analyse and design are *linear* circuits. Linear circuits possess the property that "outputs are proportional to inputs", and that "a sum of inputs leads to a sum of corresponding outputs". This is the principle of superposition and is a very important consequence of linearity. As will be seen later, this principle will enable us to analyse circuits with multiple sources in an easy way.

Nonlinear circuits can be analysed and designed with graphical methods or numerical methods (with a computer) – the mathematics that describe them can only be performed by hand in the simplest of cases. Examples of nonlinear circuits are those that contain diodes, transistors, and ferromagnetic material.

In reality all circuits are nonlinear, since there must be physical limits to the linear operation of devices, e.g. voltages will eventually break down across insulation, resistors will burn because they can't dissipate heat to their surroundings, etc. Therefore, when we draw, analyse and design a linear circuit, we keep in mind that it is a *model* of the real physical circuit, and it is only valid under a defined range of operating conditions.

In modelling real physical circuit elements, we need to consider *practical* sources as opposed to *ideal* sources. A practical source gives a more realistic representation of a physical device. We will study methods whereby practical current and voltage sources may be interchanged without affecting the remainder of the circuit. Such sources will be called *equivalent* sources.

Finally, through the use of Thévenin's theorem and Norton's theorem, we will see that we can replace a large portion of a complex circuit (often a complicated and uninteresting part) with a very simple equivalent circuit, thus enabling analysis and focus on one particular element of the circuit.

4.1 Linearity

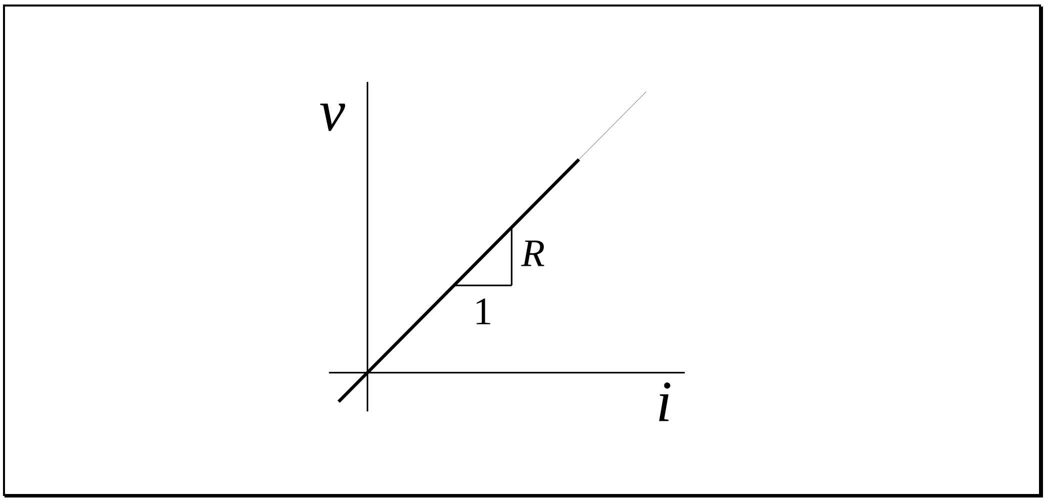
A linear circuit is one that contains linear elements, independent sources, and linear dependent sources.

A linear circuit defined

A *linear element* is one that possesses a linear relationship between a cause and an effect. For example, when a voltage is impressed across a resistor, a current results, and the amount of current (the effect) is proportional to the voltage (the cause). This is expressed by Ohm's Law, $v = Ri$. Notice that a *linear element* means simply that if the cause is increased by some multiplicative constant *K*, then the effect is also increased by the same constant *K*.

A linear element defined

If a linear element's relationship is graphed as "cause" vs. "effect", the result is a *straight line through the origin*. For example, the resistor relationship is:



A linear relationship is defined by a straight line through the origin

Figure 4.1

A *linear dependent source* is one whose output voltage or current is proportional only to the first power of some current or voltage variable in the circuit (or a sum of such quantities). For example, a dependent voltage source given by $v_s = 0.6i_1 - 14v_2$ is linear, but $v_s = 0.6i_1^2$ and $v_s = 0.6i_1v_2$ are not.

A linear dependent source defined

From the definition of a linear circuit, it is possible to show that "the response is proportional to the source", or that multiplication of all *independent* sources by a constant *K* increases all the current and voltage responses by the same factor *K* (including the dependent source outputs).

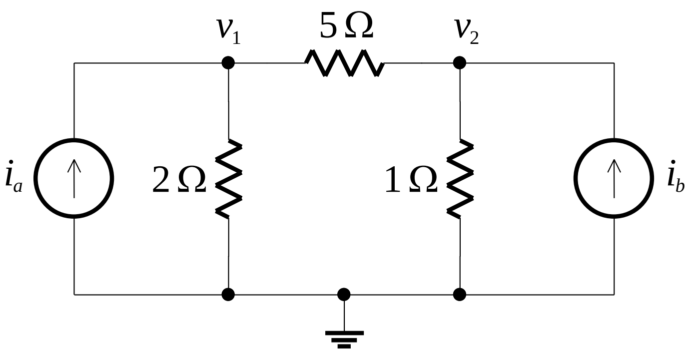
Output is proportional to input for a linear circuit

4.2 Superposition

The linearity property of a circuit leads directly to the principle of superposition. To develop the idea, consider the following example:

EXAMPLE 4.1 Superposition

We have a 3-node circuit:



There are two independent current sources which force the currents i_a and i_b into the circuit. Sources are often called *forcing functions* for this reason, and the voltages they produce at each node in this circuit may be termed *response functions*, or simply *responses*.

The two nodal equations for this circuit are:

$$\begin{aligned} 0.7v_1 - 0.2v_2 &= i_a \\ -0.2v_1 + 1.2v_2 &= i_b \end{aligned}$$

Now we perform experiment x. We change the two current sources to i_{ax} and i_{bx}; the two unknown node voltages will now be different, and we let them be v_{1x} and v_{2x}. Thus:

$$\begin{aligned} 0.7v_{1x} - 0.2v_{2x} &= i_{ax} \\ -0.2v_{1x} + 1.2v_{2x} &= i_{bx} \end{aligned}$$

If we now perform experiment *y* by changing the current sources again, we get:

$$\begin{aligned} 0.7v_{1y} - 0.2v_{2y} &= i_{ay} \\ -0.2v_{1y} + 1.2v_{2y} &= i_{by} \end{aligned}$$

We now add or *superpose* the two results of the experiments:

$$\begin{aligned} 0.7(v_{1x} + v_{1y}) - 0.2(v_{2x} + v_{2y}) &= (i_{ax} + i_{ay}) \\ -0.2(v_{1x} + v_{1y}) + 1.2(v_{2x} + v_{2y}) &= (i_{bx} + i_{by}) \end{aligned}$$

Compare this with the original set of equations:

$$\begin{aligned} 0.7v_1 - 0.2v_2 &= i_a \\ -0.2v_1 + 1.2v_2 &= i_b \end{aligned}$$

We can draw an interesting conclusion. If we let $i_{ax} + i_{ay} = i_a$, $i_{bx} + i_{by} = i_b$, then the desired responses are given by $v_1 = v_{1x} + v_{1y}$ and $v_2 = v_{2x} + v_{2y}$. That is, we may perform experiment *x* and note the responses, perform experiment *y* and note the responses, and finally add the corresponding responses. These are the responses of the original circuit to independent sources which are the sums of the independent sources used in experiments *x* and *y*.

This is the fundamental concept involved in the superposition principle. It is evident that we may break an independent source into as many pieces as we wish, so long as the algebraic sum of the pieces is equal to the original source.

Superposition allows us to treat inputs separately, then combine individual responses to obtain the total response

In practical applications of the superposition principle, we usually set each independent source to zero, so that we can analyse the circuit one source at a time.

4.2.1 Superposition Theorem

We can now state the superposition theorem as it is mostly applied to circuits:

The superposition theorem

> In any linear network containing several sources, we can calculate any response by adding algebraically all the individual responses caused by each independent source acting alone, with all other independent sources set to zero. (4.1)

When we set the value of an independent voltage source to zero, we create a *short-circuit* by definition. When we set the value of an independent current source to zero, we create an *open-circuit* by definition.

Setting a voltage source to zero creates a short-circuit. Setting a current source to zero creates an open-circuit.

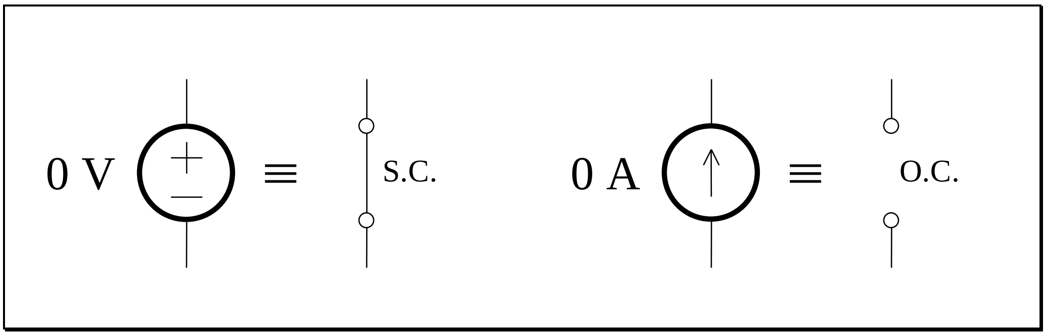


Figure 4.2

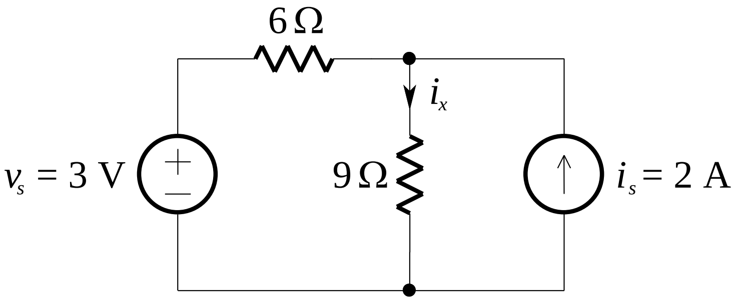
Note that *dependent* sources **cannot** be arbitrarily set to zero, and are generally active when considering *every* individual independent source.

The theorem as stated above can be made much stronger – a group of independent sources may be made active and inactive collectively. For example, sometimes it is handy to consider all voltage sources together, so that mesh analysis can be applied easily, and then all current sources together so that nodal analysis may be applied easily.

There is also no reason that an independent source must assume only its given value or zero – it is only necessary that the sum of the several values be equal to the original value. However, an inactive source almost always leads to the simplest circuit.

EXAMPLE 4.2 Superposition with Independent Sources

We use superposition in the following circuit to write an expression for the unknown branch current i_x.



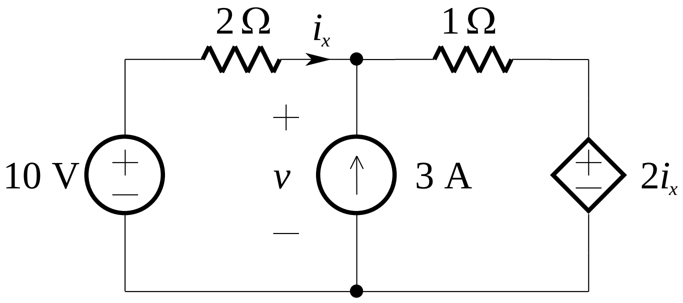
We first set the current source equal to zero (an open-circuit) and obtain the portion of i_x due to the voltage source as 0.2 A. Then if we let the voltage source be zero (a short-circuit) and apply the current divider rule, the remaining portion of i_x is seen to be 0.8 A.

We may write the answer in detail as:

$$i_x = i_x\big|_{i_s=0} + i_x\big|_{v_s=0} = \frac{3}{6+9} + \frac{6}{6+9}2 = 0.2 + 0.8 = 1\,\text{A}$$

EXAMPLE 4.3 Superposition with Dependent Sources

The circuit below contains a dependent source.



We seek i_x, and we first open-circuit the 3 A source. The single mesh equation is:

$$10 - 2i_x' - 1i_x' - 2i_x' = 0$$

so that:

$$i_x' = 2$$

Next, we short-circuit the 10 V source and write the single node equation:

$$\frac{v''}{2} + \frac{v'' - 2i_x''}{1} - 3 = 0$$

and relate the dependent-source-controlling quantity to v'':

$$v'' = -2i_x''$$

We find:

$$i_x'' = -0.6$$

and thus:

$$i_x = i_x' + i_x'' = 2 - 0.6 = 1.4\text{ A}$$

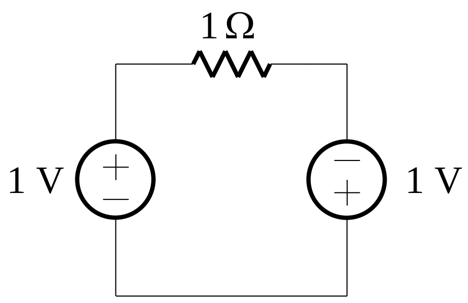
It usually turns out that little, if any, time is saved in analysing a circuit containing dependent sources by use of the superposition principle, because there are at least two sources in operation: one independent source and all the dependent sources.

We must be aware of the limitations of superposition. It is applicable only to linear responses, and thus the most common nonlinear response – power – is not subject to superposition.

Superposition is often *misapplied* to power in a circuit element

EXAMPLE 4.4 Superposition Cannot be Applied to Power

The circuit below contains two 1 V batteries in series.



If we apply superposition, then each voltage source alone delivers 1 A and furnishes 1 W. We might then mistakenly calculate the total power delivered to the resistor as 2 W. This is incorrect.

Each source provides 1 A, making the total current in the resistor 2 A. The power delivered to the resistor is therefore 4 W.

4.3 Source Transformations

4.3.1 Practical Voltage Sources

The ideal voltage source is defined as a device whose terminal voltage is independent of the current through it. Graphically, it's characteristic is:

An ideal voltage source, and its terminal characteristic

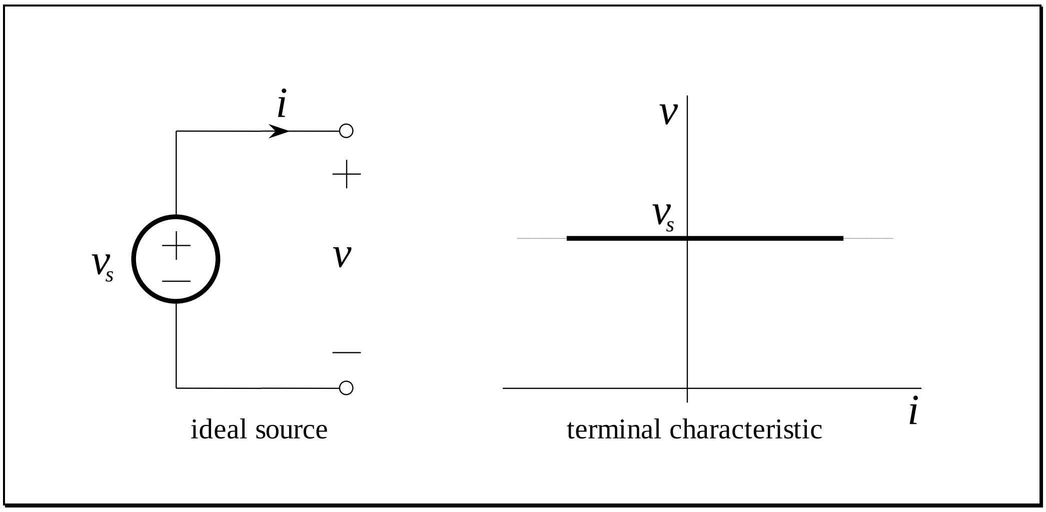


Figure 4.3

The ideal voltage source can provide any amount of current, and an unlimited amount of power. No such device exists practically. All practical voltage sources suffer from a voltage drop when they deliver current – the larger the current, the larger the voltage drop. Such behaviour can be *modelled* by the inclusion of a resistor in *series* with an ideal voltage source:

A practical voltage source, and its terminal characteristic

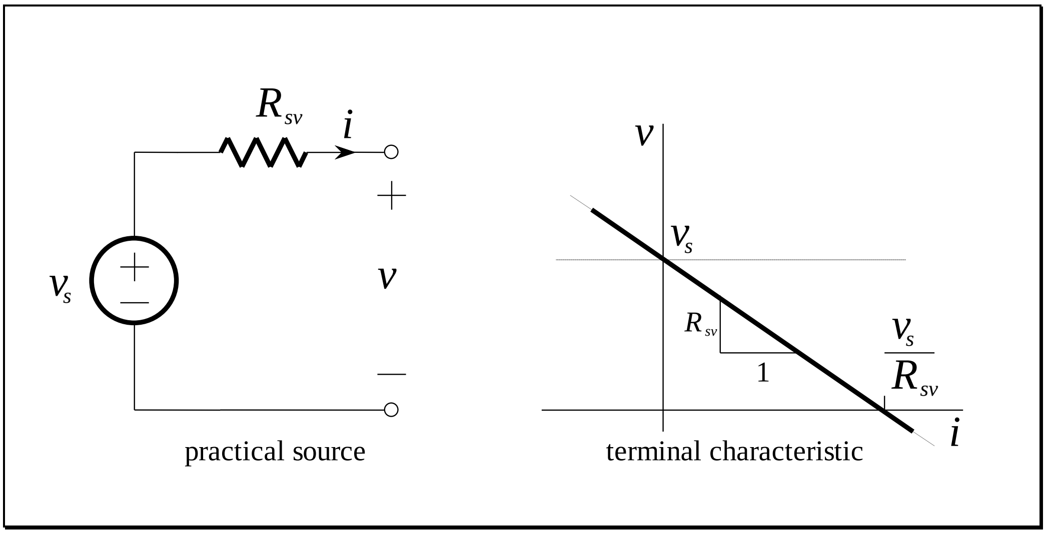


Figure 4.4

The terminal characteristic of the practical voltage source is given by KVL:

$$v = v_s - R_{sv} i \tag{4.2}$$

The terminal characteristic of a practical voltage source…

The resistance R_{sv} is known as the *internal resistance* or *output resistance*. This resistor (in most cases) is not a real physical resistor that is connected in series with a voltage source – it merely serves to account for a terminal voltage which decreases as the load current increases.

…shows the effect of the internal resistance

The applicability of this model to a practical source depends on the device and the operating conditions. For example, a DC power supply such as found in a laboratory will maintain a linear relationship in its terminal characteristic over a larger range of currents than a chemical battery.

When we attach a load to a practical voltage source:

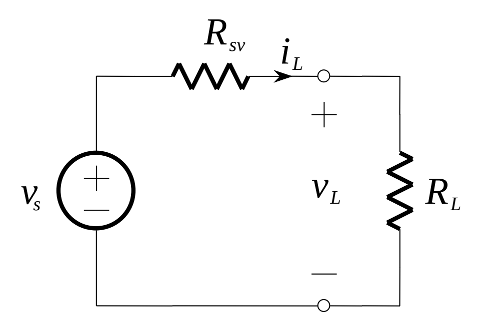


Figure 4.5

A load attached to a practical voltage source will always exhibit less voltage and current than the ideal case

we get a load voltage which is always less than the open-circuit voltage, and given by the voltage divider rule:

$$v_L = \frac{R_L}{R_{sv} + R_L} v_s < v_s \tag{4.3}$$

The load current will also be less than we expect from an ideal source:

$$i_L = \frac{v_s}{R_{sv} + R_L} < \frac{v_s}{R_L} \tag{4.4}$$

4.3.2 Practical Current Sources

The ideal current source is defined as a device whose current is independent of the voltage across it. Graphically, it's characteristic is:

An ideal current source, and its terminal characteristic

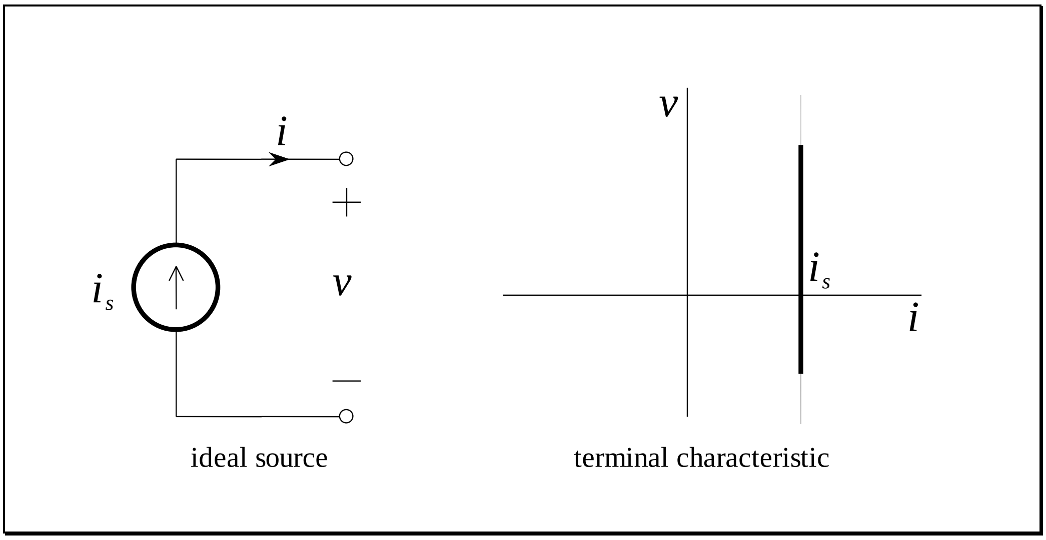


Figure 4.6

The ideal current source can support any terminal voltage regardless of the load resistance to which it is connected, and an unlimited amount of power. An ideal current source is nonexistent in the real world. For example, transistor circuits and op-amp circuits can deliver a constant current to a wide range of load resistances, but the load resistance can always be made sufficiently large so that the current through it becomes very small. Such behaviour can be *modelled* by the inclusion of a resistor in *parallel* with an ideal current source:

A practical current source, and its terminal characteristic

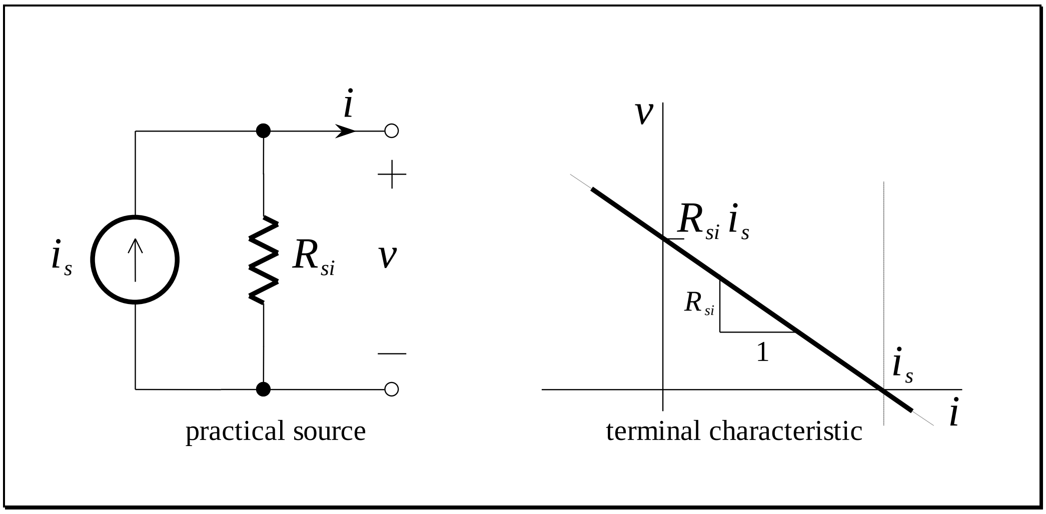


Figure 4.7

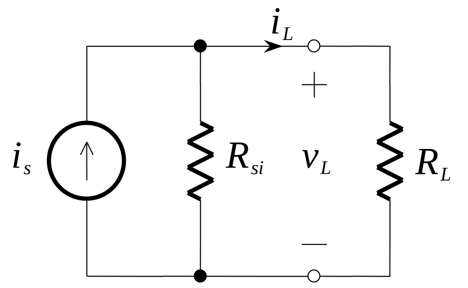
The terminal characteristic of the practical current source is given by KCL:

$$i = i_s - \frac{v}{R_{si}}$$

$$\text{or} \quad v = R_{si} i_s - R_{si} i \tag{4.5}$$

The terminal characteristic of a practical current source

When we attach a load to a practical current source:



A load attached to a practical current source will always exhibit less voltage and current than the ideal case

Figure 4.8

we get a load current which is always less than the short-circuit current, and given by the current divider rule:

$$i_L = \frac{R_{si}}{R_{si} + R_L} i_s < i_s \tag{4.6}$$

The load voltage will also be less than we expect from an ideal source:

$$v_L = \frac{R_{si} R_L i_s}{R_{si} + R_L} < R_L i_s \tag{4.7}$$

4.3.3 Practical Source Equivalence

We define two sources as being *equivalent* if each produces identical current and identical voltage for *any* load which is placed across it terminals. With reference to the practical voltage and current source terminal characteristics:

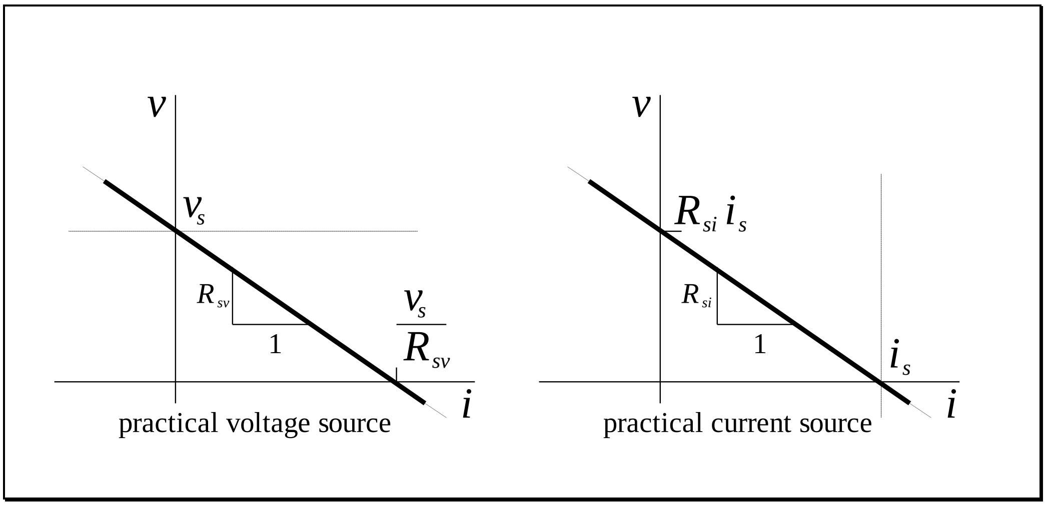


Figure 4.9

we can easily establish the conditions for equivalence. We must have:

If practical sources are equivalent then they have the same internal resistance

$$R_{sv} = R_{si} = R_s \tag{4.8}$$

so that the slopes of the two terminal characteristics are equal. We now let R_s represent the internal resistance of either practical source. To achieve the same voltage and current axes intercepts, we must have, respectively:

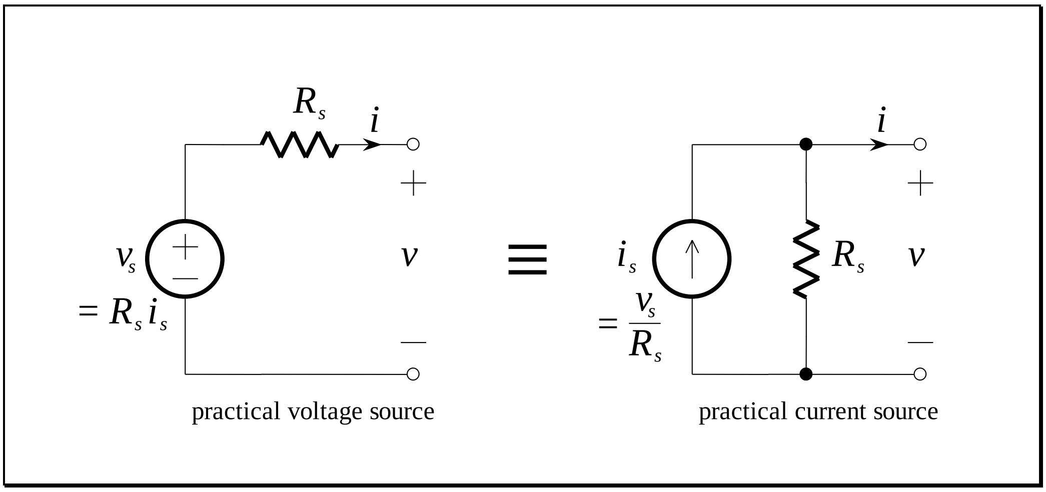
$$v_s = R_{si} i_s \quad \text{and} \quad \frac{v_s}{R_{sv}} = i_s \tag{4.9}$$

But since $R_{sv} = R_{si} = R_s$, these two relations turn into just one requirement:

The relationship between a practical voltage source and a practical current source

$$v_s = R_s i_s \tag{4.10}$$

We can now transform between practical voltage and current sources:

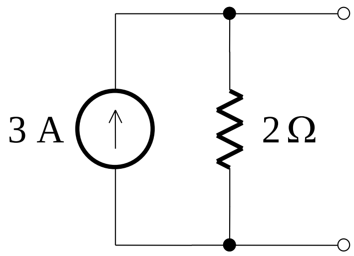


The equivalence of practical sources

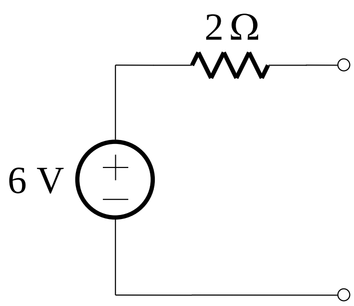
Figure 4.10

EXAMPLE 4.5 Equivalent Practical Sources

Consider the practical current source shown below:



Since its internal resistance is $2\,\Omega$, the internal resistance of the equivalent practical voltage source is also $2\,\Omega$. The voltage of the ideal voltage source contained within the practical voltage source is $v_s = R_s i_s = 2 \times 3 = 6\,\text{V}$. The equivalent practical voltage source is shown below:



4.3.4 Maximum Power Transfer Theorem

Consider a practical DC voltage source:

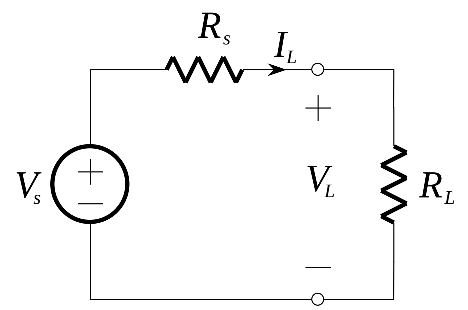


Figure 4.11

The power delivered to the load R_L is:

$$P_L = R_L I_L^2 = \frac{R_L V_s^2}{(R_s + R_L)^2} \tag{4.11}$$

Assume that V_S and R_S are known and fixed, and that R_L is allowed to vary.

A graph of the load power P_L versus load resistance R_L is shown below:

A graph of the power delivered to a load versus the load resistance shows clearly that a peak occurs at a certain resistance

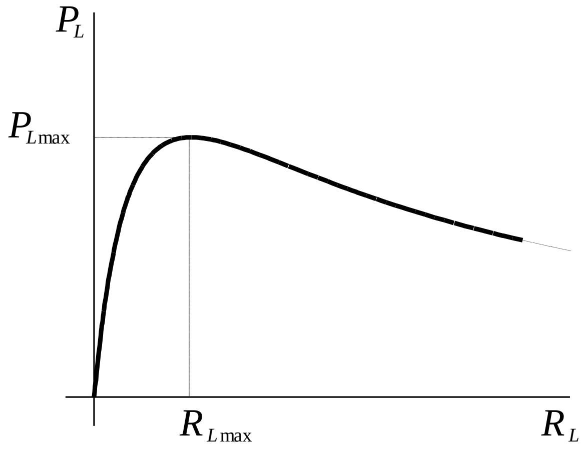


Figure 4.12

To find the value of R_L that absorbs maximum power from the practical source, we differentiate with respect to R_L (using the quotient rule):

$$\frac{\partial P_L}{\partial R_L} = \frac{\left(R_s + R_L\right)^2 V_s^2 - V_s^2 R_L \left(2\right)\left(R_s + R_L\right)}{\left(R_s + R_L\right)^4} \tag{4.12}$$

and equate the derivative to zero to obtain the relative maximum:

$$\left(R_s + R_L\right)^2 - 2R_L\left(R_s + R_L\right) = 0 \tag{4.13}$$

or:

$$R_L = R_s \tag{4.14}$$

The load resistance which maximizes power delivered from a practical source

Since the values $R_L = 0$ and $R_L = \infty$ both give a minimum ($P_L = 0$), then this value is the absolute maximum (and not just a relative maximum).

Since we have already proved the equivalence between practical voltage and current sources, we have proved the following *maximum power transfer* theorem:

The maximum power transfer theorem…

An independent voltage source in series with a resistance R_s, or an independent current source in parallel with a resistance R_s, delivers a maximum power to that load resistance R_L when $R_L = R_s$.

(4.15)

We can only apply the maximum power transfer theorem when we have control over the load resistance, i.e. if we know the source resistance, then we can choose $R_L = R_s$ to maximize power transfer. On the other hand, if we are given a load resistance and we are free to design or choose a source resistance, we do **not** choose $R_s = R_L$ to maximize power transfer – by examining Eq. (4.11), we see that for a voltage source we should choose $R_s = 0$ (and for a current source we should choose $R_s = \infty$).

…only applies to a choice of load resistor

If we choose $R_L = R_s$ to obtain maximum power transfer to a load, then by Eq. (4.11) that maximum power is:

The maximum power delivered from a practical source

$$P_{L\max} = \frac{V_s^2}{4R_L} = \frac{V_s^2}{4R_s} \qquad (4.16)$$

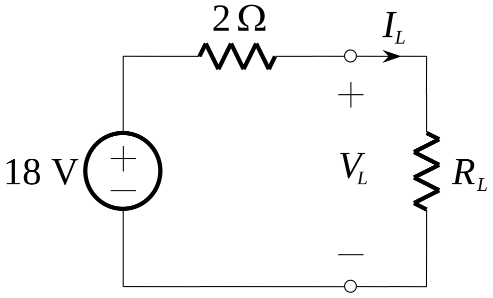
There is a distinct difference between *drawing* maximum power from a source and *delivering* maximum power to a load. If the load is sized such that $R_L = R_s$, it will receive maximum power from that source. However, considering just the practical source itself, we draw maximum possible power from the ideal voltage source by drawing the maximum possible current – which is achieved by shorting the source's terminals. However, in this extreme case, we *deliver* zero power to the "load" (a 0 Ω resistor).

Power matching is used in three situations:

- where the signal levels are very small, so any power lost gives a worse signal to noise ratio. e.g. in antenna to receiver connections in television, radio and radar.
- high frequency electronics
- where the signal levels are very large, where the maximum efficiency is desirable on economic grounds. e.g. a broadcast antenna, audio amplifier.

EXAMPLE 4.6 Power Transfer

Consider the circuit shown below:



We want to determine the values of the load resistor that draw *half* the maximum power deliverable by the practical source. The maximum power deliverable by the source is:

$$P_{L\max} = \frac{V_s^2}{4R_s} = \frac{18^2}{4 \times 2} = 40.5\ \text{W}$$

Half the maximum power deliverable is therefore 20.25 W. The power dissipated by the load resistor is:

$$P_L = R_L I_L^2 = R_L \left(\frac{V_s}{R_s + R_L} \right)^2$$

Substituting values gives:

$$\begin{aligned} 20.25 &= R_L \left(\frac{18}{2 + R_L} \right)^2 = \frac{324 R_L}{(2 + R_L)^2} \\ (2 + R_L)^2 &= 16 R_L \\ R_L^2 - 12 R_L + 4 &= 0 \end{aligned}$$

Solving this quadratic gives:

$$\begin{aligned} R_L &= \frac{12 \pm \sqrt{12^2 - 4 \times 4}}{2} \\ &= 6 \pm \sqrt{32} \\ &= 11.66 \text{ or } 0.3431\ \Omega \end{aligned}$$

4.4 Thévenin's and Norton's Theorem

Thévenin's and Norton's theorems greatly simplify the analysis of many linear circuits. Léon Thévenin was a French engineer working in telegraphy who first published a statement of the theorem in 1883. Edward Norton was a scientist with the Bell Telephone Laboratories who mentioned his theorem in a technical memorandum in 1926, but never published it.

Thévenin's theorem tells us that it is possible to replace a large portion of a linear circuit (often a complicated and / or uninteresting portion) by an equivalent circuit containing only an independent voltage source in series with a resistor:

Thévenin's theorem allows us to replace part of a circuit with a practical voltage source

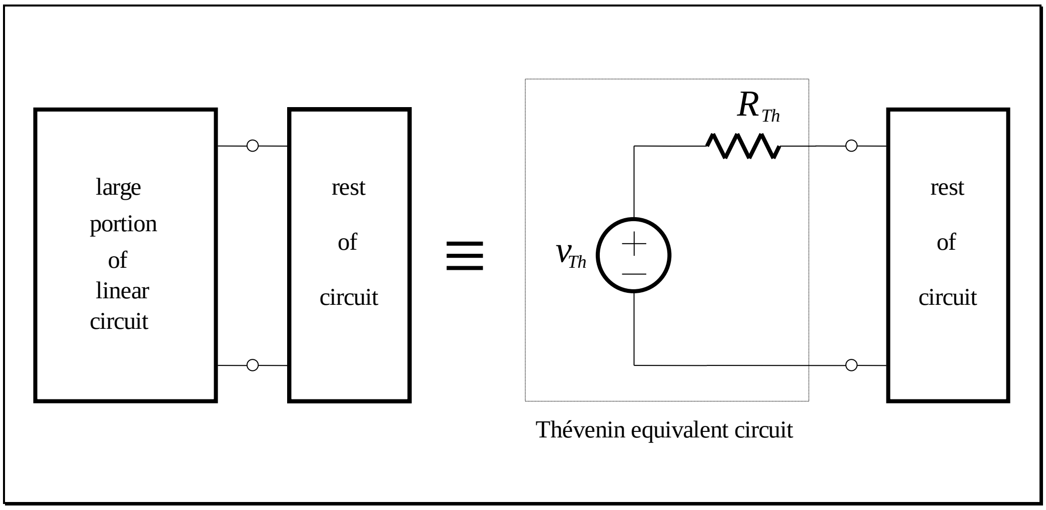


Figure 4.13

Norton's theorem is the dual of Thévenin's theorem, and uses a current source:

Norton's theorem allows us to replace part of a circuit with a practical current source

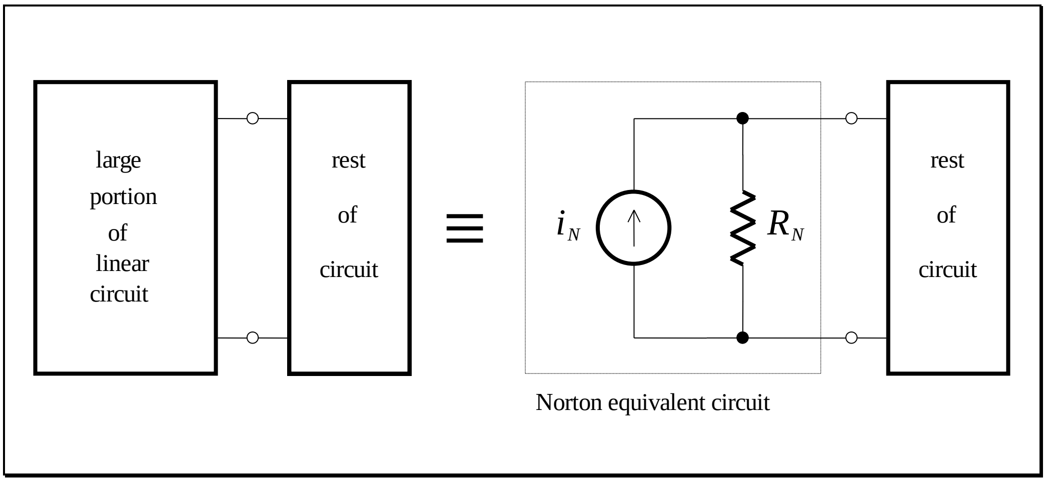


Figure 4.14

We can state Thévenin's theorem formally as:

Thévenin's theorem

> Given any linear circuit, rearrange it in the form of two circuits *A* and *B* that are connected together at two terminals.
> If either circuit contains a dependent source, its control variable must be in the same circuit.
> Define a voltage v_{oc} as the open-circuit voltage which would appear across the terminals of *A* if *B* were disconnected so that no current is drawn from *A*. Then all the currents and voltages in *B* will remain unchanged if *A* has all its independent sources set to zero and an independent *voltage source* v_{oc} is connected in *series* with the inactive *A* network.

(4.17)

The inactive circuit *A* will always reduce to a single resistor, which we call the Thévenin resistance, R_{Th}. Also, since v_{oc} appears as an independent voltage source in the Thévenin equivalent circuit, it is also denoted as v_{Th}.

We can state Norton's theorem formally as:

Norton's theorem

> Given any linear circuit, rearrange it in the form of two circuits *A* and *B* that are connected together at two terminals.
> If either circuit contains a dependent source, its control variable must be in the same circuit.
> Define a current i_{sc} as the short-circuit current which would appear across the terminals of *A* if *B* were short-circuited so that no voltage is provided by *A*. Then all the voltages and currents in *B* will remain unchanged if *A* has all its independent sources set to zero and an independent *current source* i_{sc} is connected in *parallel* with the inactive *A* network.

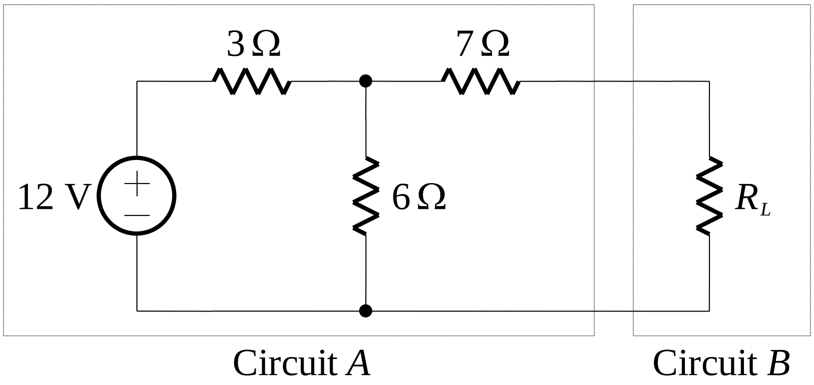
(4.18)

The inactive circuit *A* will always reduce to a single resistor, which we call the Norton resistance, R_N. Also, since i_{sc} appears as an independent current source in the Norton equivalent circuit, it is also denoted as i_N.

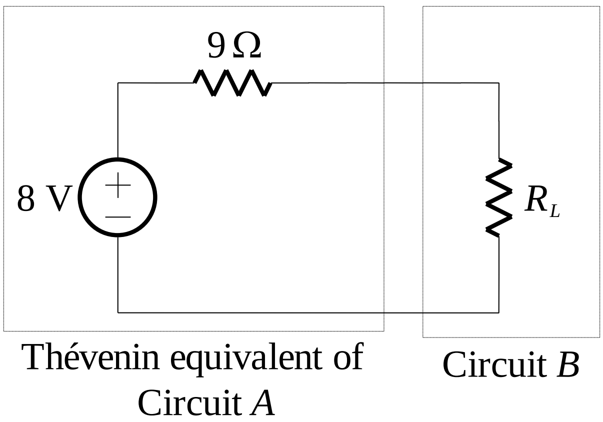
The Thévenin or Norton equivalent circuit allows us to draw a new simpler circuit and make rapid calculations of the voltage, current, and power. It also allows us to easily choose a “load” resistance for maximum power transfer.

EXAMPLE 4.7 Thévenin and Norton Equivalent Circuits

Consider the circuit shown below:



The broken lines separate the original circuit into circuits *A* and *B*. We shall assume that our main interest is in circuit *B*, which consists only of a “load” resistor R_L. To form the Thévenin equivalent circuit, we disconnect circuit *B* and use voltage division to determine that $v_{oc} = 8\text{ V}$. When we set all independent sources in circuit *A* to zero, we replace the 12 V source with a short-circuit. “Looking back” into the inactive *A* circuit, we “see” a $7\,\Omega$ resistor connected in series with the parallel combination of $6\,\Omega$ and $3\,\Omega$. Thus, the inactive *A* circuit can be represented by a $9\,\Omega$ resistor. If we now replace circuit *A* by its Thévenin equivalent circuit, we have:



Note that the Thévenin equivalent circuit we have obtained for circuit *A* is completely independent of circuit *B* – an equivalent for *A* may be obtained *no matter what arrangement of elements is connected to the* A *circuit, even if circuit* B *is nonlinear!*

From the viewpoint of the load resistor R_L, the Thévenin equivalent circuit is identical to the original; from our viewpoint, the circuit is much simpler and we can now easily compute various quantities. For example, the power delivered to the load is:

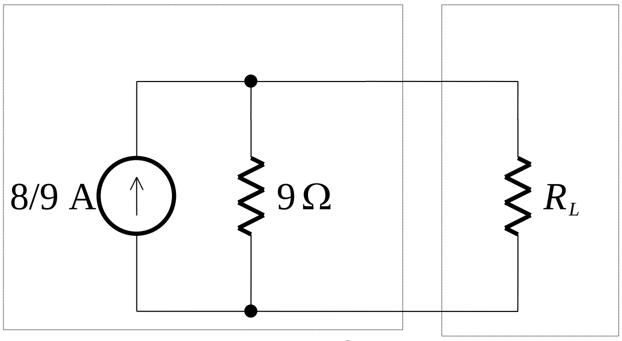
$$P_L = \left(\frac{8}{9 + R_L} \right)^2 R_L$$

Furthermore, we can now easily see that the maximum voltage which can be obtained across R_L is 8 V when $R_L = \infty$. A quick transformation of the Thévenin equivalent circuit to a practical current source (the Norton equivalent) indicates that the maximum current which may be delivered to the load is $8/9\text{ A}$ for $R_L = 0$. The maximum power transfer theorem shows that a maximum power is delivered to R_L when $R_L = 9\,\Omega$. None of these facts is readily apparent from the original circuit.

To form the Norton equivalent circuit, we short-circuit the *B* circuit and use the current divider rule to discover:

$$i_{sc} = \frac{6}{6+7}\left(\frac{12}{3+\frac{6\cdot 7}{6+7}}\right) = \frac{72}{39+42} = \frac{72}{81} = \frac{8}{9}\ \text{A}$$

When we set all independent sources in circuit *A* to zero, we get the same results as for the Thévenin circuit, and so $R_N = 9\,\Omega$. The Norton equivalent circuit is therefore:



Norton equivalent of Circuit *A*

Circuit *B*

It should be apparent from the previous example that we can easily find the Norton equivalent circuit from the Thévenin equivalent circuit, and vice versa, by a simple source transformation. Using our previous results, we must have:

The Thévenin and Norton equivalent resistances are the same

$$R_{Th} = R_N \tag{4.19}$$

Because of this result, we usually just refer to the resistor in either equivalent circuit as the Thévenin resistance, R_{Th}.

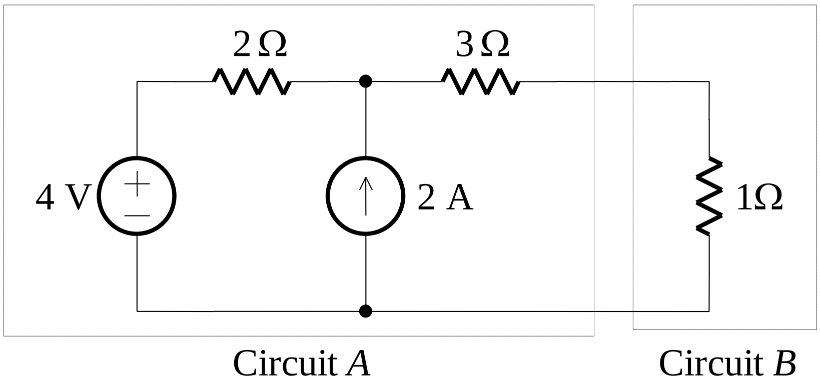
We also have:

The relationship between the Thévenin and Norton equivalent circuits

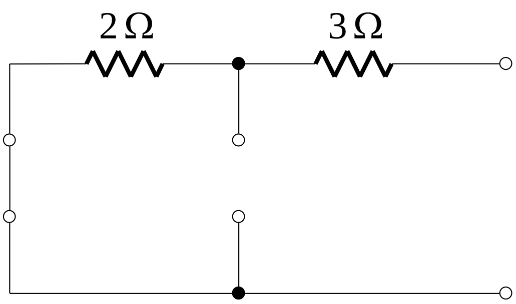
$$v_{oc} = R_{Th} i_{sc} \tag{4.20}$$

EXAMPLE 4.8 Thévenin and Norton Equivalent Circuits

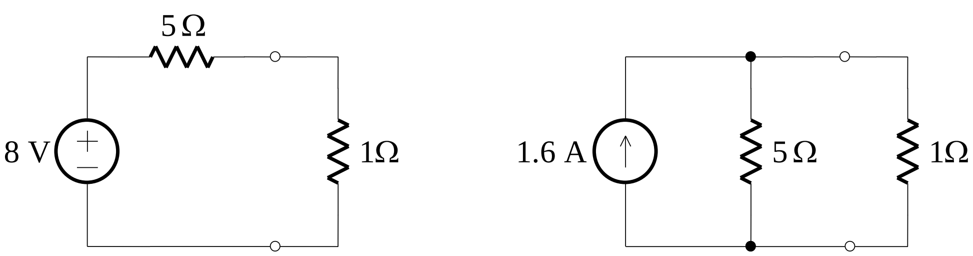
Consider the circuit shown below:



The Thévenin and Norton equivalent circuits are desired from the perspective of the 1Ω resistor. We determine R_{Th} for the inactive network, and then find either v_{oc} or i_{sc}. Making the independent sources inactive, we have:



We thus have $R_{Th} = 5\,\Omega$. The open-circuit voltage can be determined using superposition. With only the 4 V source operating , the open-circuit voltage is 4 V. When only the 2 A source is on, the open-circuit voltage is also 4 V. Thus, with both sources operating, we have $v_{oc} = 8\,\text{V}$. This determines the Thévenin equivalent, and from it the Norton equivalent, shown below:



We will often find it convenient to determine either the Thévenin or Norton equivalent by finding both the open-circuit voltage and the short-circuit current, and then determining the Thévenin resistance by:

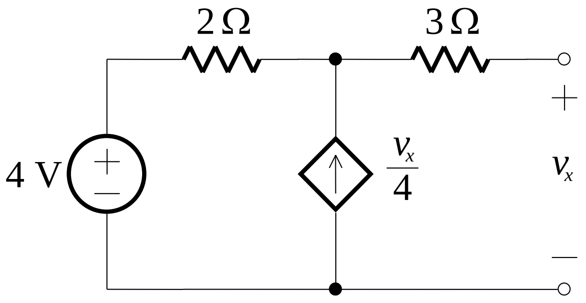
The Thévenin equivalent resistance can be obtained from the open-circuit voltage and short-circuit current

$$R_{Th} = \frac{v_{oc}}{i_{sc}} \tag{4.21}$$

This is especially true for circuits that contain *dependent* sources.

EXAMPLE 4.9 Thévenin Equivalent with Mixed Sources

Consider the circuit shown below:

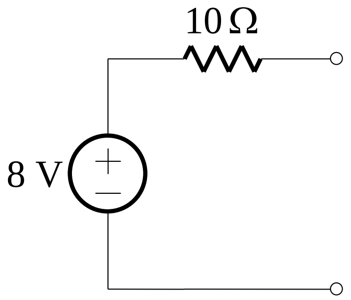


To find v_{oc} we note that $v_x = v_{oc}$, and that the dependent source current must pass through the $2\,\Omega$ resistor since there is an open circuit to the right. KCL at the top of the dependent source gives:

$$\frac{v_{oc} - 4}{2} - \frac{v_{oc}}{4} = 0$$

$$v_{oc} = 8\text{ V}$$

Upon short-circuiting the output terminals, it is apparent that $v_x = 0$ and the dependent current source is zero. Hence $i_{sc} = 4/(2+3) = 0.8\text{ A}$. Thus $R_{Th} = v_{oc}/i_{sc} = 8/0.8 = 10\,\Omega$ and the Thévenin equivalent is:

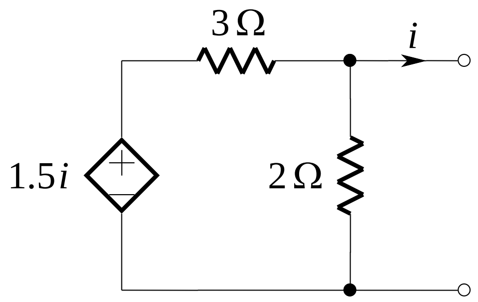


If a circuit contains only dependent sources and no independent sources, then the circuit qualifies as the inactive *A* circuit and $v_{oc} = 0$ and $i_{sc} = 0$. We seek the value of R_{Th} represented by the circuit, but in this case $R_{Th} = v_{oc}/i_{sc}$ is undefined, and we must use a different approach. We apply either an independent voltage source or an independent current source *externally* to the *A* circuit and "measure" the resultant current or voltage.

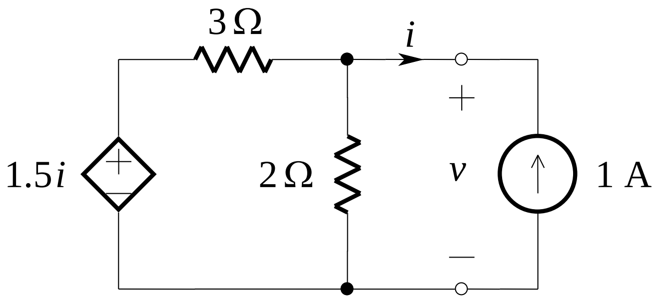
If a circuit does not contain independent sources, then one must be applied to obtain the Thévenin equivalent resistance

EXAMPLE 4.10 Thévenin Equivalent with Dependent Source Only

Consider the circuit shown below:



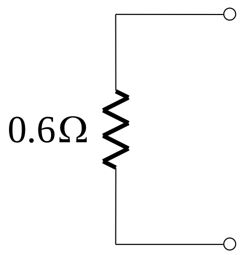
If we apply a 1 A source *externally*, and "measure" the resultant voltage, then $R_{Th} = v/1$:



We see that $i = -1\text{ A}$ and KCL gives:

$$\frac{v - 1.5(-1)}{3} + \frac{v}{2} - 1 = 0$$

so that $v = 0.6\text{ V}$ and $R_{Th} = 0.6\,\Omega$. The Thévenin equivalent is:



4.4.1 Summary of Finding Thévenin Equivalent Circuits

We have seen three approaches to finding the Thévenin equivalent circuit. The first example contained only independent sources and resistors, and we could use several different methods on it. One involved finding v_{oc} for the active circuit, and then R_{Th} for the inactive circuit. We could also have found i_{sc} and R_{Th}, or v_{oc} and i_{sc}. In the second example, both independent and dependent sources were present, and the method we used required us to find v_{oc} and i_{sc}. The last example did not contain any independent sources, and we found R_{Th} by applying a 1 A source and finding $R_{Th} = v/1$. We could also have applied a 1 V source and determined $R_{Th} = 1/i$.

These important techniques and the types of circuits to which they may be applied most readily are indicated in the table below:

Suitable methods to obtain the Thévenin equivalent circuit

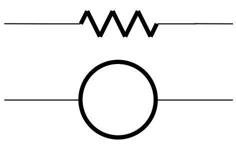
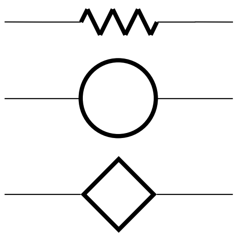
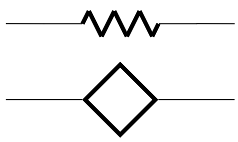
	Circuit contains		
Methods			
R_{Th} and v_{oc} or i_{sc}	✓	–	–
v_{oc} and i_{sc}	Possible	✓	–
$i = 1\text{A}$ or $v = 1\text{V}$	–	–	✓

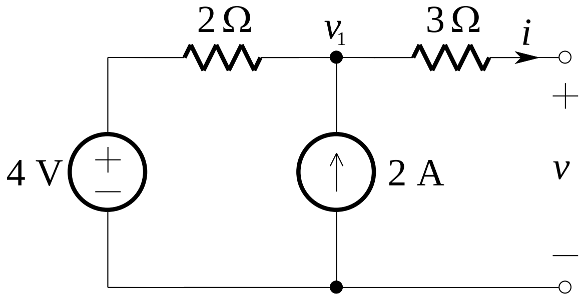
Table 4.1 – Suitable methods to obtain the Thévenin equivalent circuit

All possible methods do not appear in the table. Another method has a certain appeal because it can be used for any of the three types of circuit tabulated. Simply label the terminals of the *A* circuit as v, define the current leaving the positive polarity as i, then analyse the *A* circuit to obtain an equation in the form $v = v_{oc} - R_{Th} i$.

A method to obtain the Thévenin equivalent circuit that works for *all* circuits

EXAMPLE 4.11 Thévenin Equivalent Using a Linear Equation

Consider the circuit shown below:



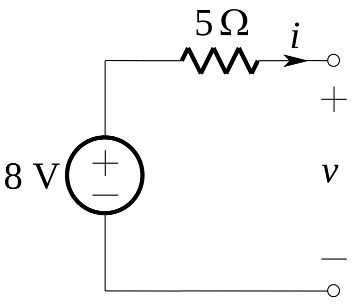
KCL at the middle node gives:

$$\frac{v_1 - 4}{2} - 2 + i = 0$$
$$v_1 = 8 - 2i$$

KVL at the output gives:

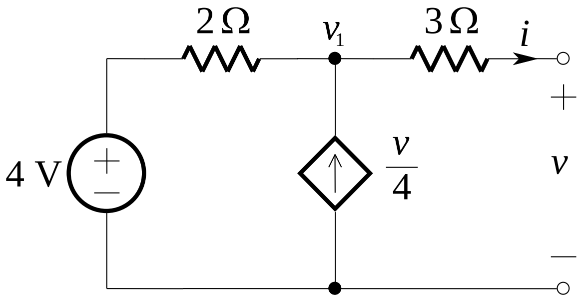
$$\begin{aligned} v &= v_1 - 3i \\ &= 8 - 5i \\ &= v_{oc} - R_{Th} i \end{aligned}$$

from which $v_{oc} = 8\text{ V}$ and $R_{Th} = 5\,\Omega$, as before:



EXAMPLE 4.12 Thévenin Equivalent Using a Linear Equation

Consider the circuit shown below:



KCL at the middle node gives:

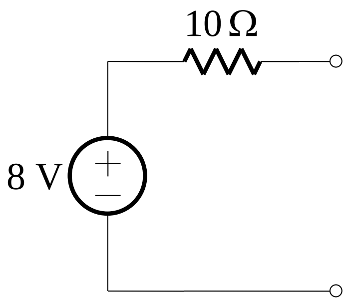
$$\frac{v_1 - 4}{2} - \frac{v}{4} + i = 0$$

$$v_1 = 4 + \frac{v}{2} - 2i$$

KVL at the output gives:

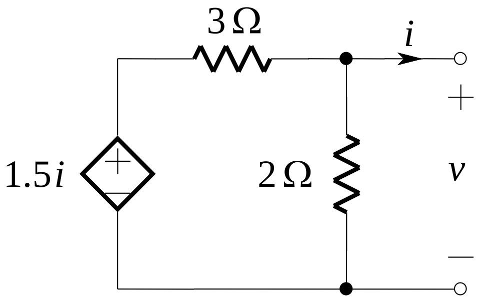
$$\begin{aligned} v &= v_1 - 3i \\ &= 4 + \frac{v}{2} - 5i \\ &= 8 - 10i \\ &= v_{oc} - R_{Th} i \end{aligned}$$

from which $v_{oc} = 8\text{ V}$ and $R_{Th} = 10\ \Omega$, as before:



EXAMPLE 4.13 Thévenin Equivalent Using a Linear Equation

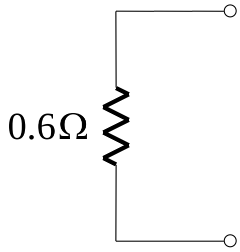
Consider the circuit shown below:



KCL at the top right node gives:

$$\begin{aligned} \frac{v - 1.5i}{3} + \frac{v}{2} + i &= 0 \\ 5v + 3i &= 0 \\ v &= 0 - 0.6i \\ &= v_{oc} - R_{Th} i \end{aligned}$$

from which $v_{oc} = 0\text{ V}$ and $R_{Th} = 0.6\,\Omega$, as before:



This procedure is universally applicable, but one of the other methods is usually easier and quicker to perform.

4.5 Summary

- A linear circuit is one that contains linear elements, independent sources, and linear dependent sources. For a linear circuit, it is possible to show that "the response is proportional to the source".

- The superposition theorem states that, in evaluating the "response" in a linear circuit due to several sources, we are free to treat each independent source separately, collectively, or in any number of parts, and then superpose the response caused by each part.

- A practical voltage source consists of an ideal voltage source v_s in series with a resistance R_{sv}. A practical current source consists of an ideal current source i_s in parallel with a resistance R_{si}. The practical sources can be made equivalent by setting $R_{sv} = R_{si} = R_s$ and $v_s = R_s i_s$.

- The maximum power transfer theorem states that if we know the source resistance R_s of a practical source, then to maximize power transfer to a load R_L, we set $R_L = R_s$.

- Thévenin's theorem tells us that it is possible to replace a large portion of a linear circuit by an equivalent circuit containing only an independent voltage source in series with a resistor.

- Norton's theorem tells us that it is possible to replace a large portion of a linear circuit by an equivalent circuit containing only an independent current source in parallel with a resistor.

- There are several methods that can be applied to determine a Thévenin or Norton equivalent circuit. Some methods are only applicable to certain circuits, and the most convenient analysis method should be chosen.

4.6 References

Hayt, W. & Kemmerly, J.: *Engineering Circuit Analysis*, 3rd *Ed*., McGraw-Hill, 1984.

Exercises

1.

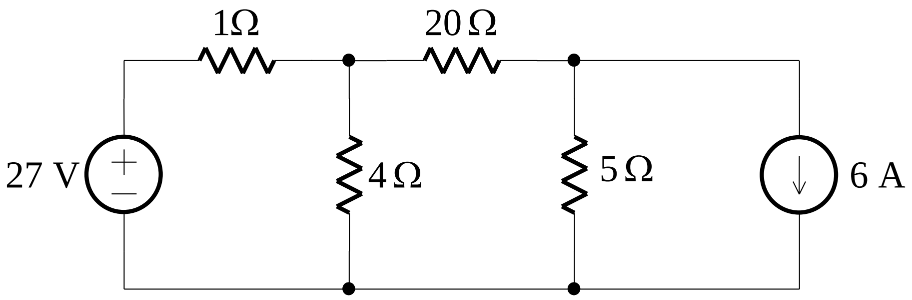
Find the power dissipated in the $20\,\Omega$ resistor of the circuit shown below by each of the following methods:

(a) nodal equations

(b) mesh equations

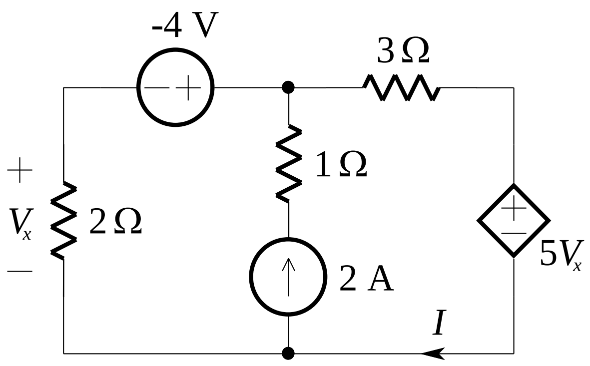
(c) source transformations to eliminate all current sources, then a method of your choice

(d) source transformations to eliminate all voltage sources, followed by any method you wish.



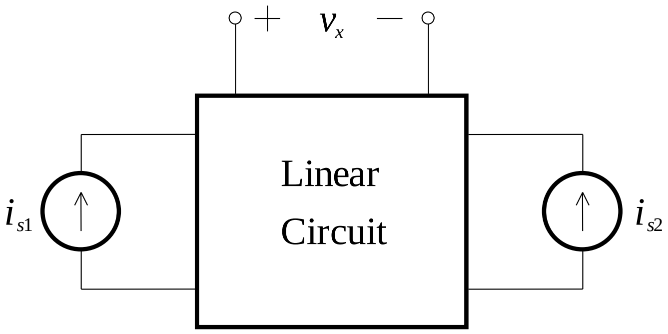
2.

The circuit shown below contains a dependent source. Use superposition to find I.



3.

Consider the linear circuit shown below.

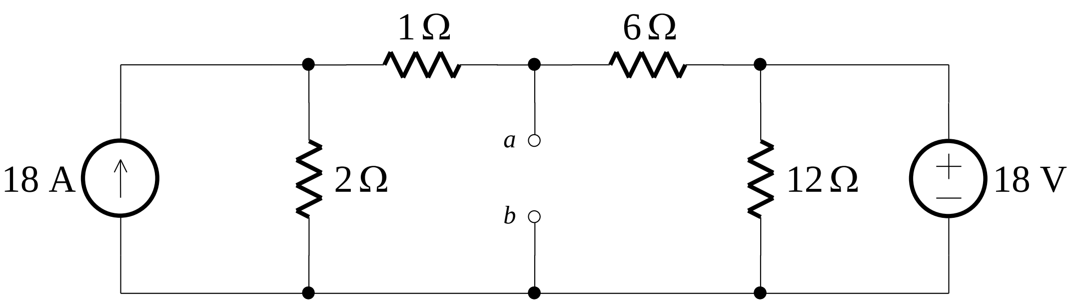


(a) The circuit contains only resistors. If $i_{s1} = 8\text{ A}$ and $i_{s2} = 12\text{ A}$, v_x is found to be 80 V. However, if $i_{s1} = -8\text{ A}$ and $i_{s2} = 4\text{ A}$, then $v_x = 0\text{ V}$. Find v_x when $i_{s1} = i_{s2} = 20\text{ A}$.

(b) The circuit now contains a source such that $v_x = -40\text{ V}$ when $i_{s1} = i_{s2} = 0\text{ A}$. All data in part (a) are still correct. Find v_x when $i_{s1} = i_{s2} = 20\text{ A}$.

4.

Consider the circuit shown below:



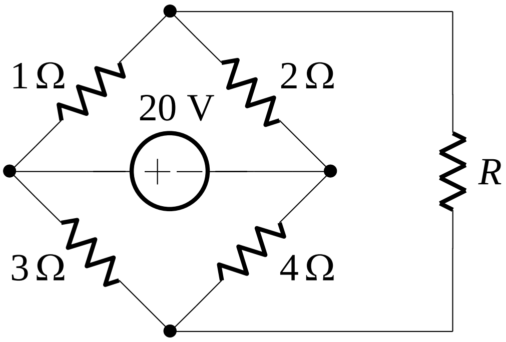
(a) Find the Norton equivalent of the circuit.

(b) If a variable resistor *R* were placed between terminals *a* and *b*, what value would result in maximum power being drawn from the terminals?

(c) Find the maximum power that could be drawn from terminals *a* and *b*.

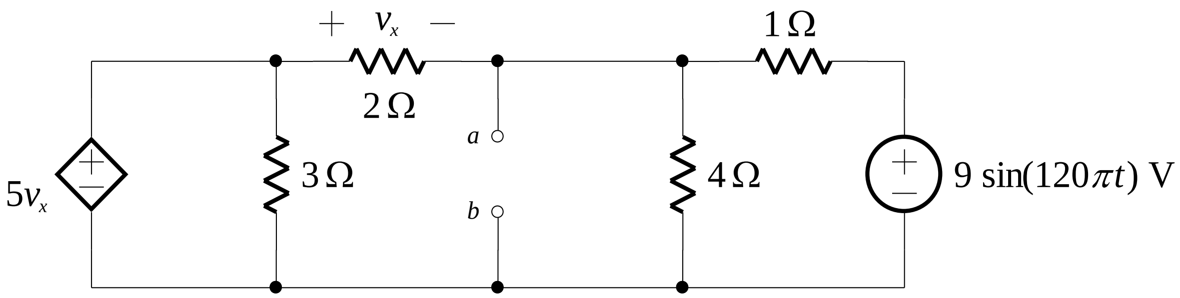
5.

Find the maximum power that can be delivered to a variable R in the circuit below:



6.

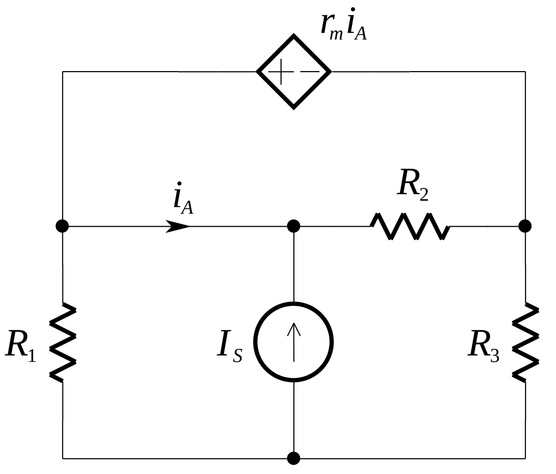
In the circuit below, what value of resistance should be connected between terminals *a*-*b* to draw maximum power?



4.36

7.

Consider the circuit below:



(a) Determine an expression for the current i_A.

(b) If $r_m = R_2$, what is the current i_A?

(c) What would happen if you tried to build such a circuit with $r_m = R_2$?

5 Linear Op-Amp Applications

Contents

5.2

Introduction

One of the reasons for the popularity of the op-amp is its versatility. As we shall see shortly, you can do almost anything with op-amps! More importantly, the IC op-amp has characteristics that closely approach the assumed ideal. This implies that it is quite easy to design circuits using the IC op-amp. It also means that a real op-amp circuit will work in a manner that is very close to the predicted theoretical performance.

Op-amp circuits which exhibit a linear characteristic (i.e. the response is proportional to the source) are known as linear op-amp circuits. They are chiefly, but not necessarily, composed of linear circuit elements, such as resistors and capacitors, Many important op-amp circuits can be created with linear circuit elements. For example, the weighted summer is able to output a voltage which is a weighted sum of the input signals. The difference amplifier is able to subtract two signals. Integrators, of both the inverting and noninverting variety, are used in many control and signal processing applications.

Also, by exploiting the active nature of op-amp circuits, we can develop circuits which have no passive equivalent, such as the negative impedance converter, and the voltage-to-current converter.

5.1 Summing Amplifier

Consider the circuit shown below:

An inverting summing amplifier

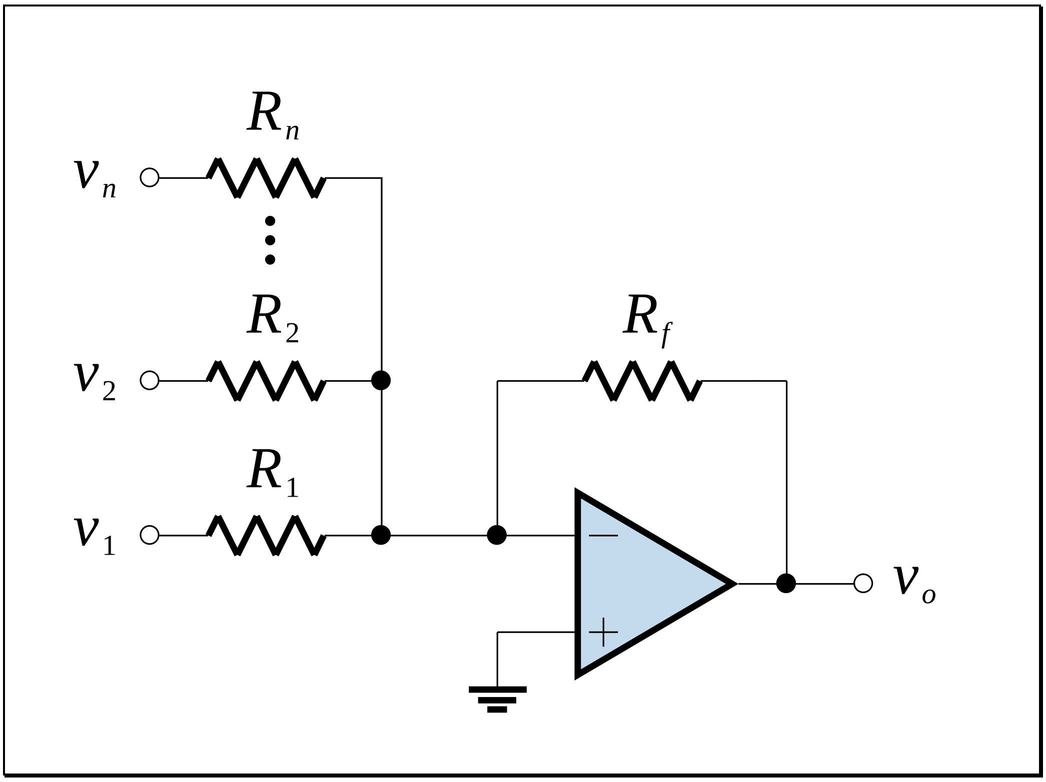


Figure 5.1 – An Inverting Summing Amplifier

The basic configuration is that of an inverting amplifier. We have a resistance R_f in the negative feedback path, but there are a number of input signals, v_1, v_2, …, v_n each applied to a corresponding resistor R_1, R_2, …, R_n which are connected to the inverting terminal of the op-amp. Just like in the inverting amplifier configuration, the op-amp and the negative feedback will maintain a virtual short-circuit across the op-amp input terminals, and therefore maintain the inverting terminal at 0 V. Ohm's Law then tells us that the currents i_1, i_2, …, i_n are given by:

$$i_1 = \frac{v_1}{R_1}, \qquad i_2 = \frac{v_2}{R_2}, \qquad \ldots, \qquad i_n = \frac{v_n}{R_n} \tag{5.1}$$

All these currents sum together at the inverting terminal, also known as the *summing junction*, to produce the current *i*:

$$i = i_1 + i_2 + \cdots + i_n \tag{5.2}$$

Remembering that the ideal op-amp input terminals are open-circuits, application of KCL at the inverting terminal shows that this current is forced through R_f. The output voltage v_o may now be determined by an application of Ohm's Law and KVL:

$$v_o = 0 - R_f i = -R_f i \tag{5.3}$$

Thus:

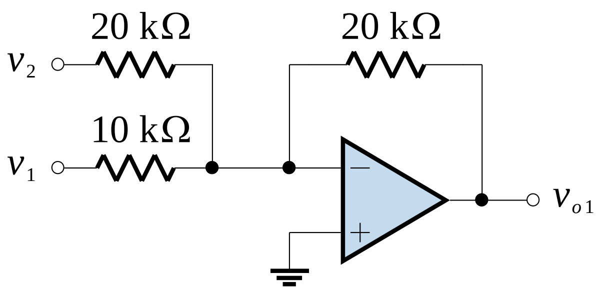
$$v_o = -\left(\frac{R_f}{R_1} v_1 + \frac{R_f}{R_2} v_2 + \cdots + \frac{R_f}{R_n} v_n \right) \tag{5.4}$$

The output voltage is a weighted sum of the input signals v_1, v_2, …, v_n. This circuit is therefore called a *weighted summer*. Note that each summing coefficient may be independently adjusted by varying the corresponding "feed-in" resistor (R_1 to R_n). This nice property is a consequence of the "virtual common" that exists at the inverting op-amp terminal.

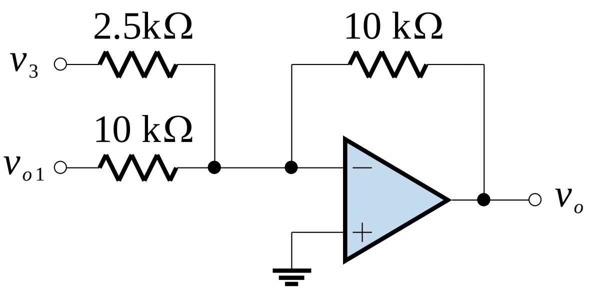
EXAMPLE 5.1 Summing Amplifier

We want to design a weighted summer such that $v_o = 2v_1 + v_2 - 4v_3$.

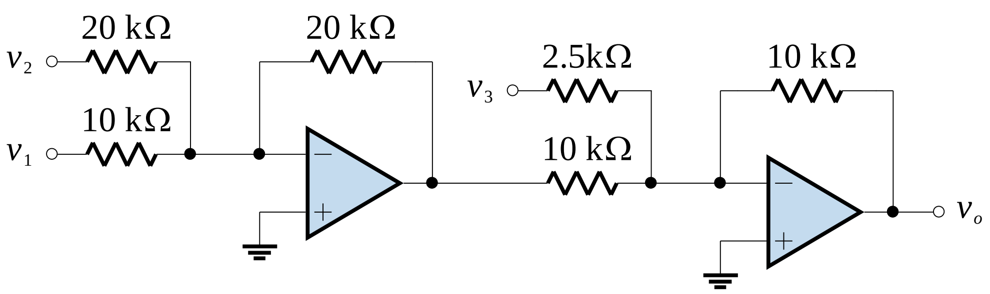
We can firstly form the sum $v_{o1} = -(2v_1 + v_2)$ by using the circuit:



Then we can form $v_o = -(v_{o1} + 4v_3)$ using:



Finally, we can put the two circuits together, in a cascade, to create:



5.2 Difference Amplifier

A simple difference amplifier can be constructed with four resistors and an op-amp, as shown below:

The difference amplifier

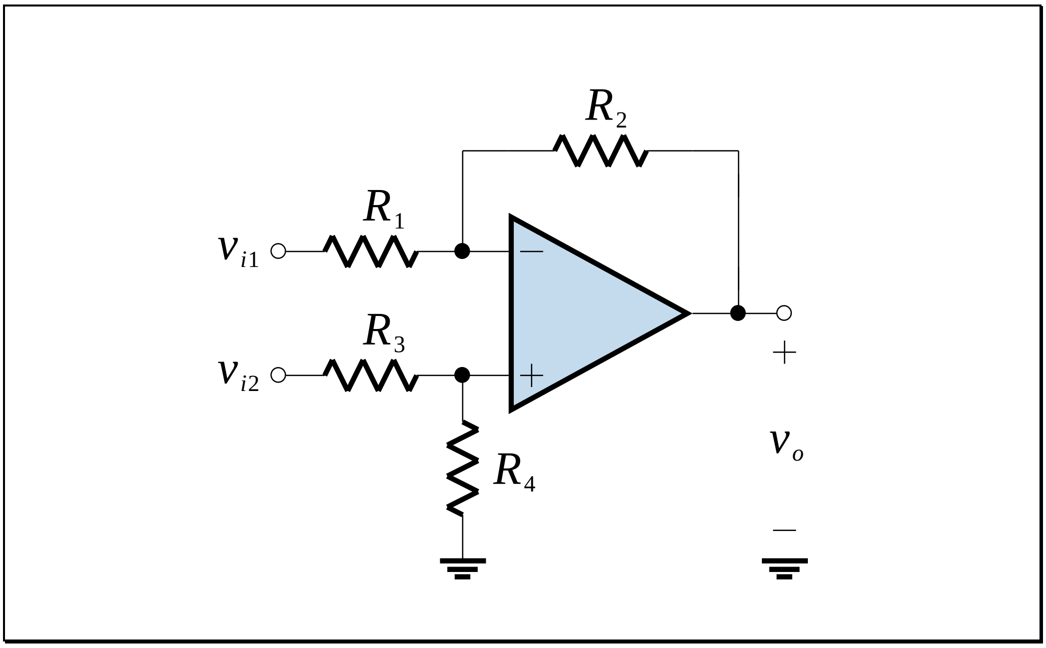


Figure 5.2 – A Difference Amplifier

There are a number of ways to find the output voltage, but the easiest uses the principle of superposition (since the circuit is linear). To apply superposition we first reduce v_{i2} to zero – that is, connect the terminal to which v_{i2} is applied to the common – and then find the corresponding output voltage, which will be due entirely to v_{i1}. We denote this output v_{o1}, as shown in (a) below:

Analyzing a difference amplifier using superposition

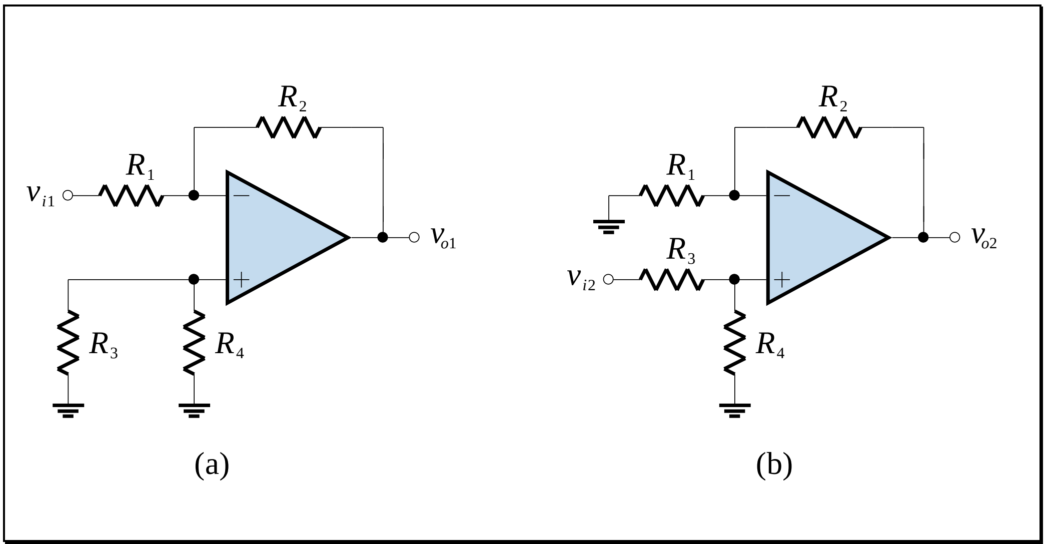


Figure 5.3 – Analyzing a Difference Amplifier

We recognize the (a) circuit as that of an inverting amplifier. The existence of R_3 and R_4 does not affect the gain expression, since there is no current in either of them. Thus:

$$v_{o1} = -\frac{R_2}{R_1} v_{i1} \tag{5.5}$$

Next, we reduce v_{i1} to zero and evaluate the corresponding output voltage v_{o2}. The circuit will now take the form shown in Figure 5.3(b), which we recognize as the noninverting configuration with an additional voltage divider, made up of R_3 and R_4, connected across the input v_{i2}. The output voltage v_{o2} is therefore given by:

$$v_{o2} = \left(1 + \frac{R_2}{R_1}\right) \frac{R_4}{R_3 + R_4} v_{i2} \tag{5.6}$$

The superposition principle tells us that the output voltage v_o is equal to the sum of v_{o1} and v_{o2}. Thus we have:

$$v_o = -\frac{R_2}{R_1} v_{i1} + \frac{1 + R_2/R_1}{1 + R_3/R_4} v_{i2} \tag{5.7}$$

To act as a difference amplifier, the output must produce a signal which is proportional to $v_{i2} - v_{i1}$. That is, the coefficients of v_{i1} and v_{i2} in Eq. (5.7) must be equal in magnitude but opposite in sign. This requirement leads to the condition:

$$\frac{R_2}{R_1} = \frac{1 + R_2/R_1}{1 + R_3/R_4} \tag{5.8}$$

Simplifying we get:

$$\frac{R_2}{R_1} = \frac{R_4}{R_3} \tag{5.9}$$

Substituting in Eq. (5.7) results in the output voltage:

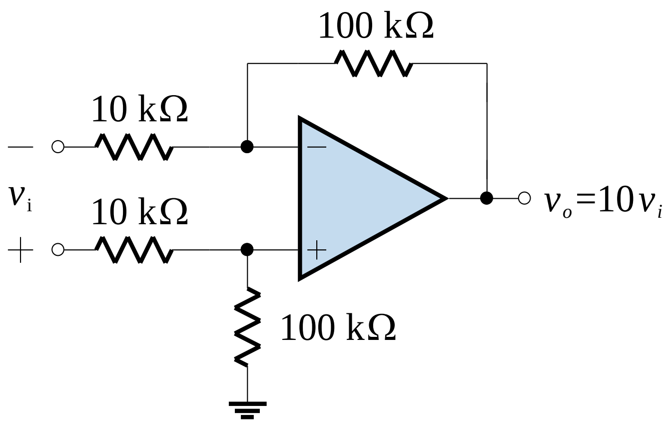
The output of a difference amplifier

$$v_o = \frac{R_2}{R_1}\left(v_{i2} - v_{i1}\right) \tag{4.1}$$

Thus, if we choose $R_4/R_3 = R_2/R_1$ then we have produced a difference amplifier with a gain of R_2/R_1.

EXAMPLE 5.2 Difference Amplifier

A difference amplifier with an input resistance of $20\ \text{k}\Omega$ and a gain of 10 is shown below:



Note that "input resistance" is defined as the resistance "seen" between the two input terminals. Thanks to the virtual short-circuit at the op-amp input terminals, KVL around the input resistors gives $R_{in} = 20\ \text{k}\Omega$.

5.3 Inverting Integrator

An inverting integrator, also known as the Miller integrator after its inventor, is shown below:

The inverting integrator

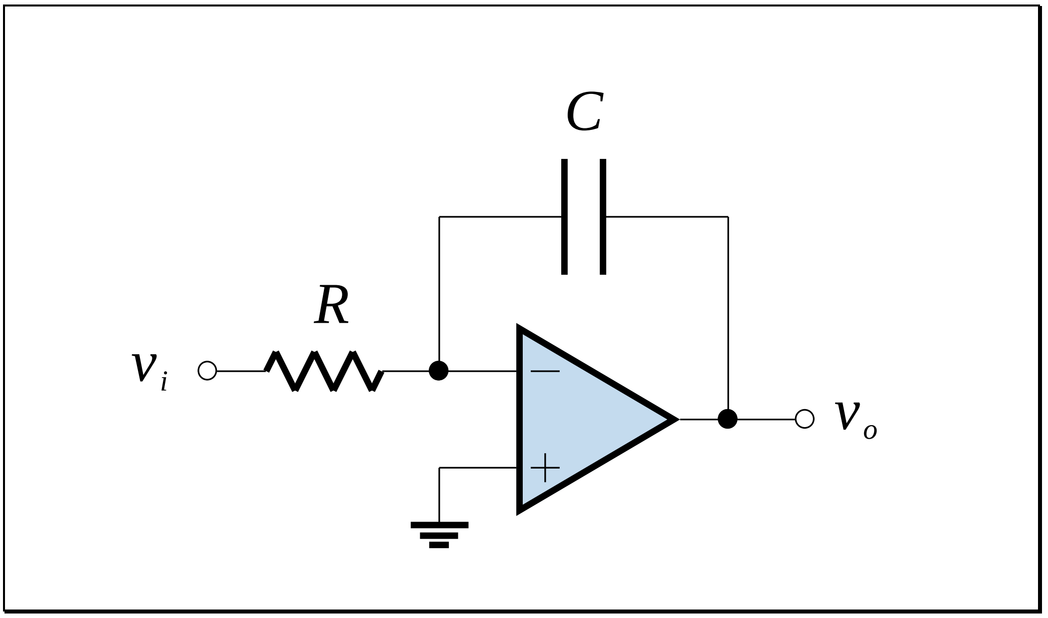


Figure 5.4 – The Miller Integrator

We use the virtual short-circuit concept to analyse the circuit (it is essentially the same analysis as for the inverting amplifier). KCL at the inverting terminal, which is held at 0 V by the op-amp and negative feedback, gives:

$$\frac{v_i}{R} = -C\frac{dv_o}{dt} \tag{5.10}$$

and therefore:

$$v_o(t) = -\frac{1}{RC}\int_0^t v_i(t)dt + v_o(0) \tag{4.2}$$

Thus, $v_o(t)$ is the time integral of $v_i(t)$, and $v_o(0)$ is the initial condition of this integration process. The product RC has units of time and is called the *integration time constant*. This integrator circuit is inverting because of the minus sign in front of the integral.

Unfortunately, for DC input voltages, the op-amp circuit is operating as an *open-loop*, which is easily seen when we recall that capacitors behave as open circuits with DC. Thus, the above analysis does not apply, since there is no virtual short circuit. With an ideal op-amp and a DC input voltage, there is no current and v_i appears at the inverting terminal – the resulting output of the ideal op-amp will be infinite. In practice the output of a real op-amp will saturate close to one of the supply rail voltages (depending on the sign of the DC input voltage).

A practical circuit that alleviates this problem is shown below:

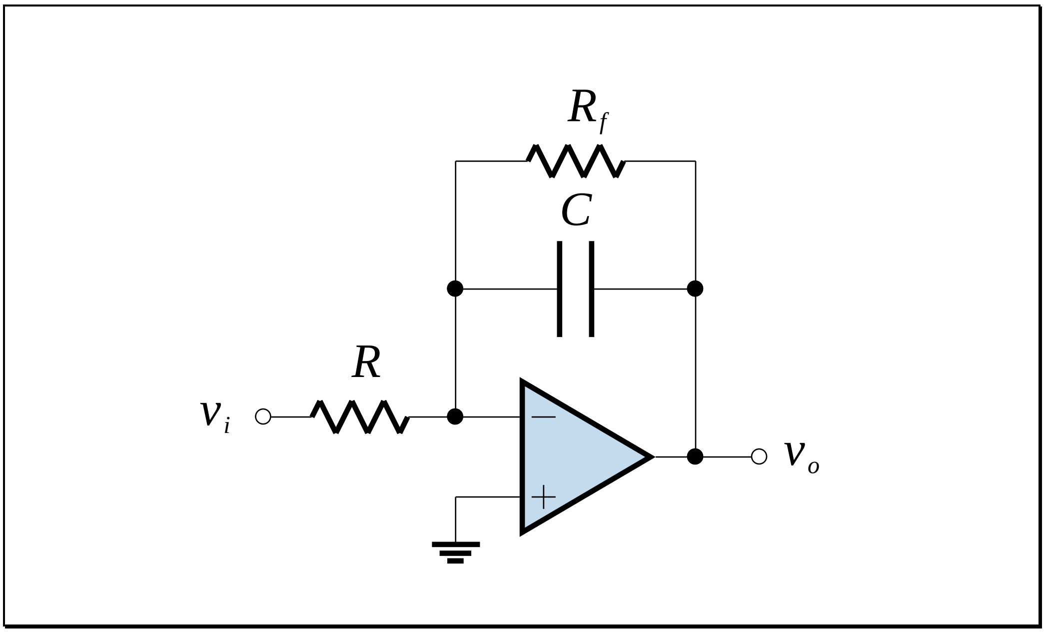


Figure 5.5 – A Practical Approximation to the Miller Integrator

This circuit provides a feedback path for DC voltages (i.e. the op-amp circuit is operating in a *closed-loop*) and prevents the output from saturating. To keep the DC offset at the output of the integrator low, we should select a small R_f. Unfortunately, however, the lower the value of R_f, the less ideal the integrator becomes. Thus selecting a value for R_f is a trade-off between DC performance and integrator performance.

5.4 Differentiator

An ideal differentiator is shown below:

The differentiator

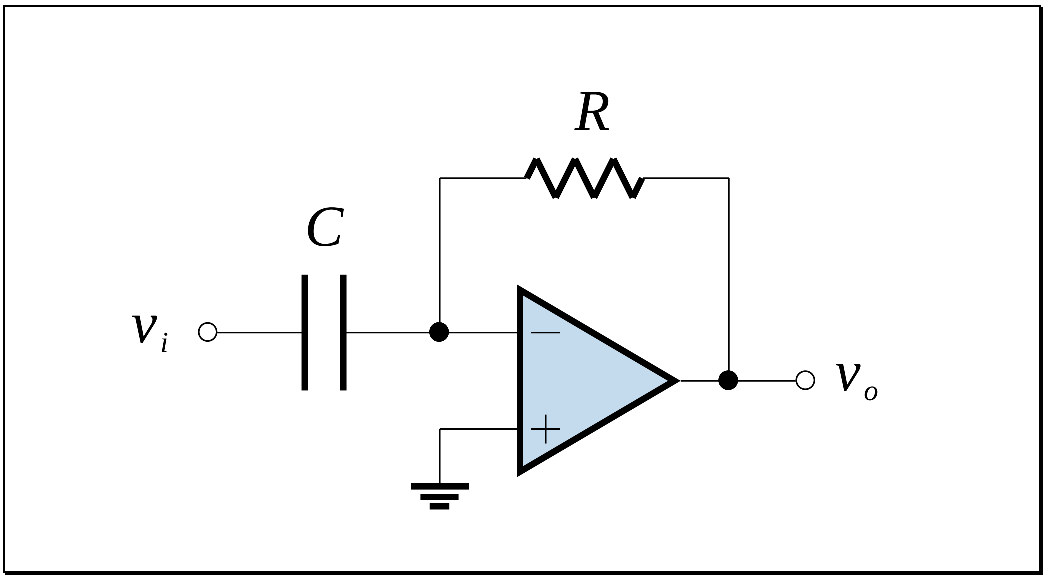


Figure 5.6 – The Differentiator

We use the virtual short circuit concept again to analyze the circuit. KCL at the inverting terminal, which is held at 0 V by the op-amp and negative feedback, gives:

$$C\frac{dv_i}{dt} = -\frac{v_o}{R} \tag{5.11}$$

and therefore:

$$v_o(t) = -RC\frac{dv_i}{dt} \tag{5.12}$$

Thus, $v_o(t)$ is the time derivative of $v_i(t)$.

The very nature of a differentiator circuit causes it to be a "noise magnifier". This is due to spikes introduced at the output every time there is a sharp change in the input. For this reason they are rarely used in practice.

5.5 Negative Impedance Converter

A negative impedance[1] converter (NIC) is shown below:

The negative impedance converter

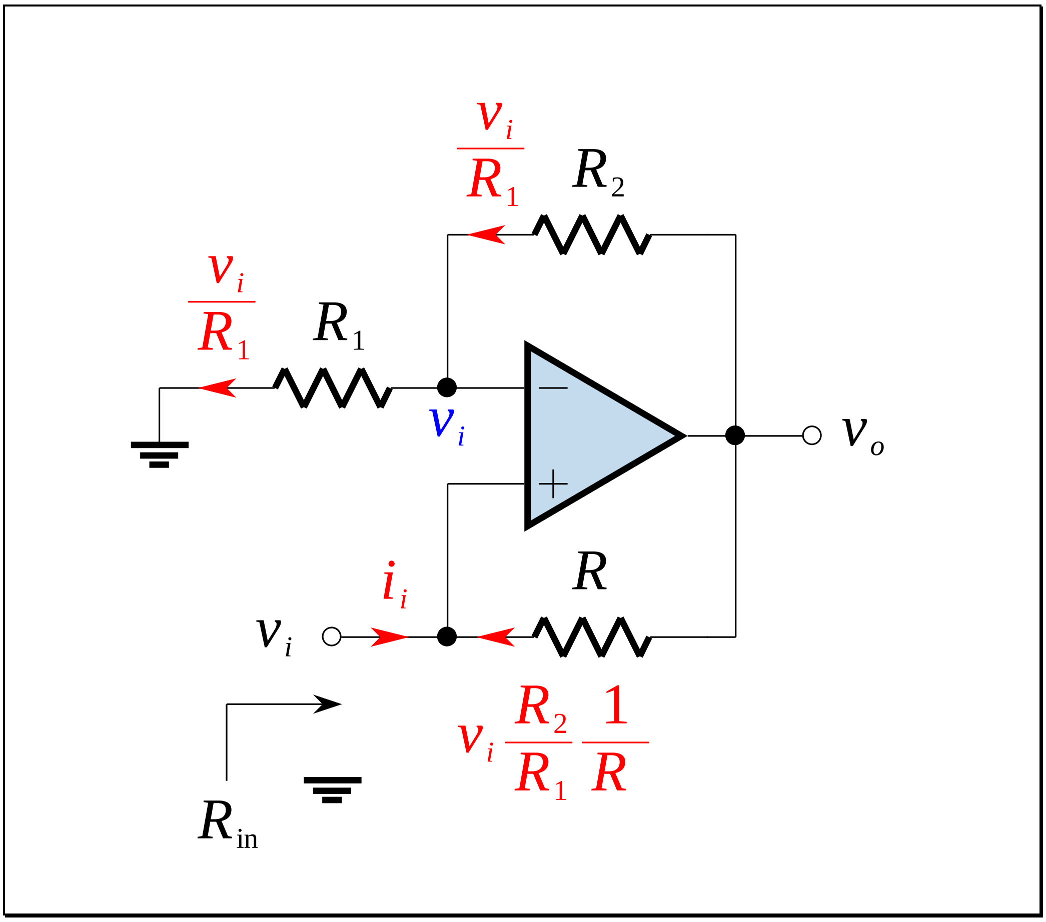


Figure 5.7 – The Negative Impedance Converter

To investigate the operation of this circuit, we will evaluate the input resistance R_{in} of the circuit. To find R_{in} we apply an input voltage v_i and evaluate the input current i_i. Then, by definition, $R_{in} = v_i / i_i$.

Owing to the virtual short circuit between the op-amp input terminals, the voltage at the inverting terminal will be equal to v_i. The current through R_1 will therefore be v_i / R_1. Since the input resistance of the ideal op-amp is

[1] Impedance generalises the concept of resistance – as will be seen later with the introduction of sinusoidal steady-state analysis, phasors, and reactance.

infinite, the current through R_2 will also be v_i/R_1. Thus, the voltage at the op-amp output will be:

$$v_o = v_i + R_2 \frac{v_i}{R_1} = \left(1 + \frac{R_2}{R_1}\right) v_i \tag{5.13}$$

which we recognise as the output of a normal noninverting amplifier. We now apply Ohm's Law to *R* and obtain the current through it as:

$$\frac{v\left(1 + R_2/R_1\right) - v}{R} = v \frac{R_2}{R_1} \frac{1}{R} \tag{5.14}$$

Since there is no current into the positive input terminal of the op-amp, KCL gives:

$$i_i = -\frac{v}{R} \frac{R_2}{R_1} \tag{5.15}$$

Thus:

$$\boxed{R_{\text{in}} = -R \frac{R_1}{R_2}} \tag{5.16}$$

That is, the input resistance is negative with a magnitude equal to *R*, the resistance in the positive-feedback path, multiplied by the ratio R_1/R_2. We now see why the circuit is called a *negative impedance converter* (NIC), where *R* may in general be replaced by an arbitrary circuit element, such as a capacitor or inductor.

5.6 Voltage-to-Current Converter

To investigate the NIC further, consider the case $R_1 = R_2 = r$, where r is an arbitrary value. It follows that $R_{\text{in}} = -R$. Let the input be fed with a voltage source V_s having a source resistance equal to R, as shown below:

A voltage-to-current converter, invented by Prof. Bradford Howland, MIT, around 1962

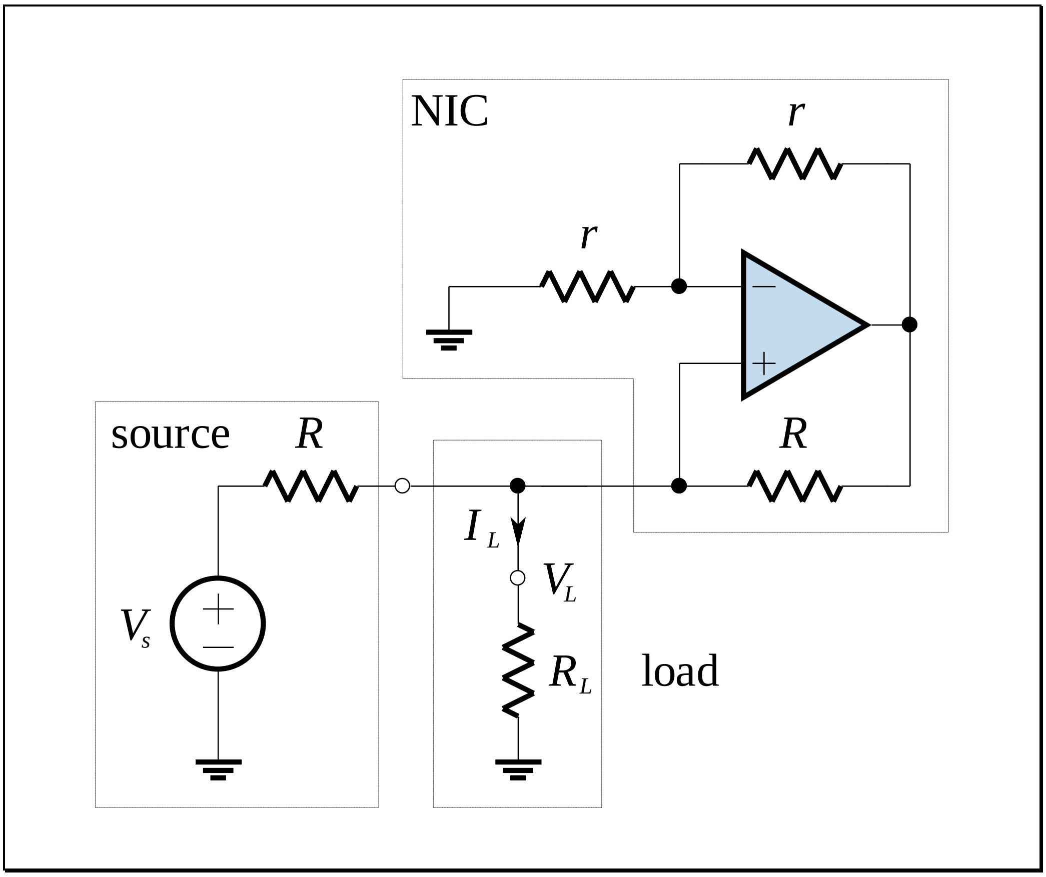


Figure 5.8 – Application of the Negative Impedance Converter

Utilizing the information gained previously, we can replace the NIC by a resistance equal to $-R$:

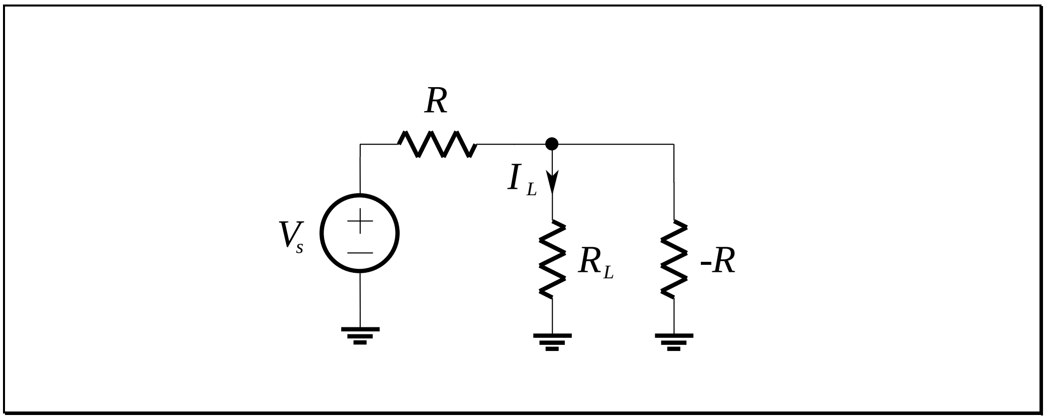


Figure 5.9

The figure below illustrates the conversion of the voltage source to its Norton equivalent:

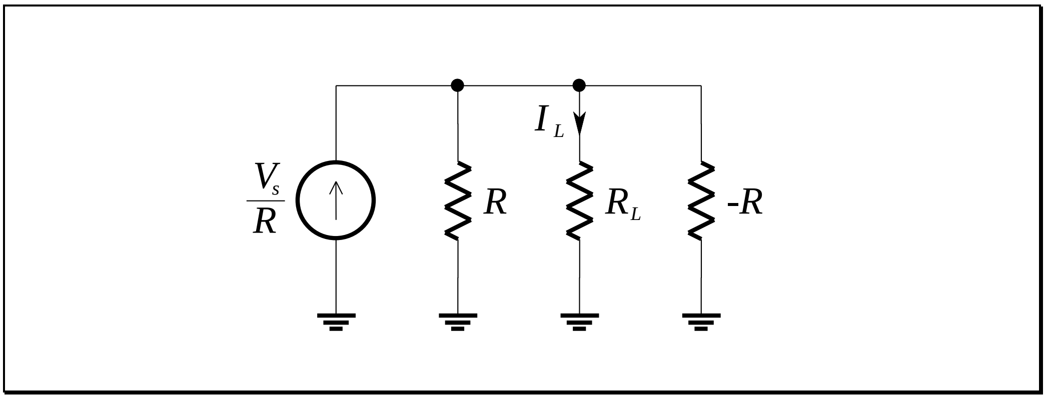


Figure 5.10

Finally, the two parallel resistances R and $-R$ are combined to produce an infinite resistance, resulting in the equivalent circuit:

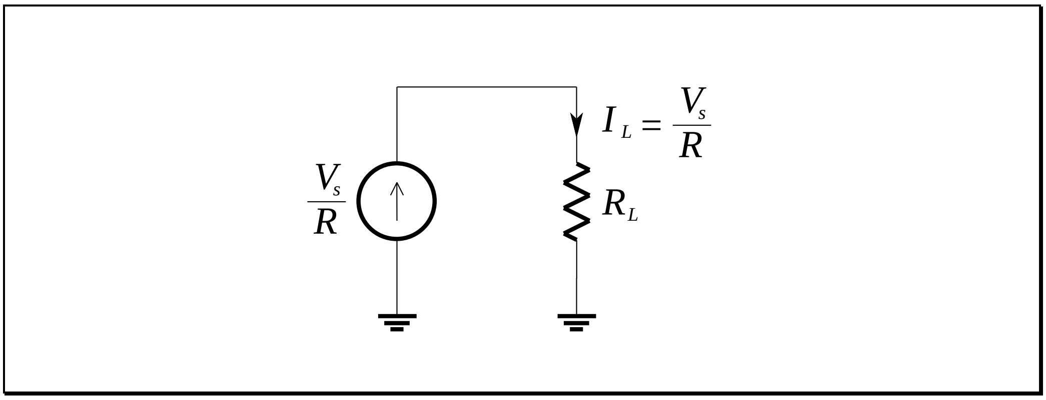


Figure 5.11

We see that the load current is given by:

$$I_L = \frac{V_s}{R} \tag{5.17}$$

independent of the value of R_L!. This is an interesting result; it tells us that the circuit of Figure 5.8 acts as a *voltage-to-current converter*, providing a current I_L that is directly proportional to V_s and is independent of the value of the load resistance. That is, the output terminal acts as a current-source output, with the impedance looking back into the output terminal equal to infinity. Note that this infinite resistance is obtained via the cancellation of the positive source resistance R with the negative input resistance $-R$.

5.7 Noninverting Integrator

A noninverting integrator, invented by Gordon Deboo in 1966 whilst working at the NASA Ames Research Centre

A specific application of the voltage-to-current converter, where a capacitor C is used as a load, is illustrated below:

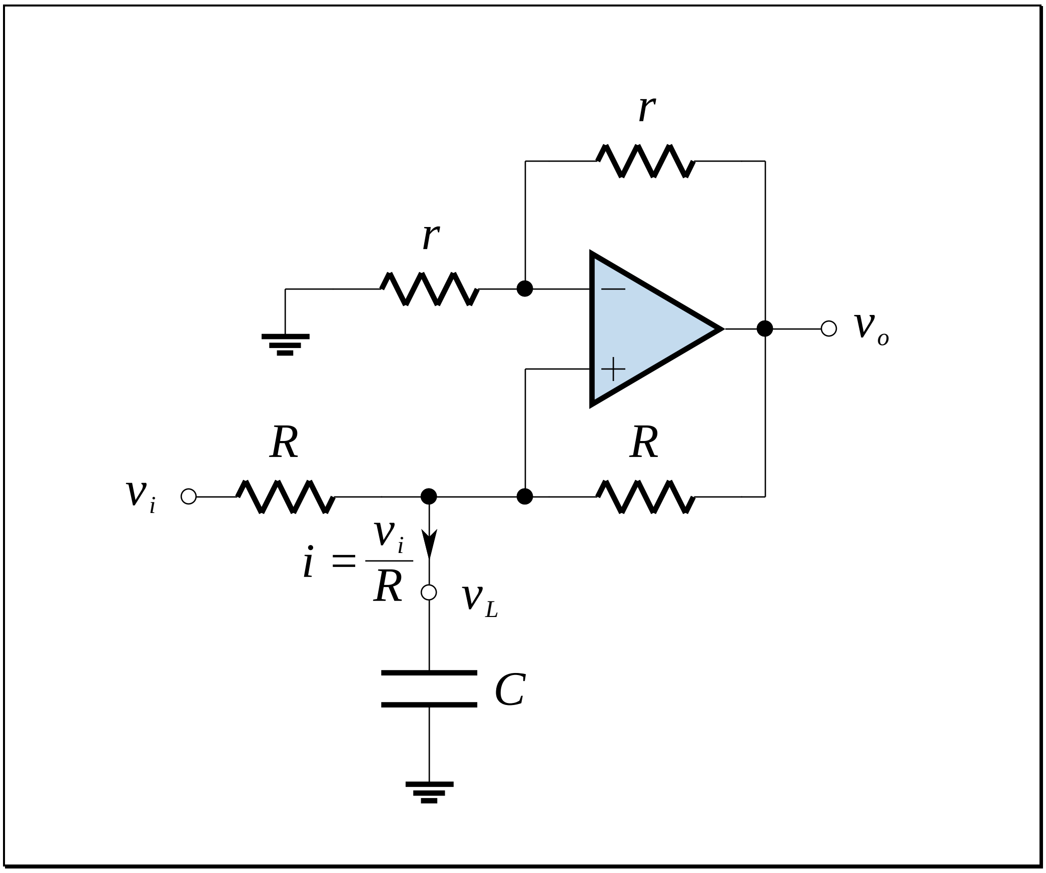


Figure 5.12

From the previous analysis of the voltage-to-current converter we conclude that capacitor C will be supplied by a current $i = v_i / R$. From the branch relationship for a capacitor we conclude:

$$C \frac{dv_L}{dt} = \frac{v_i}{R} \tag{5.18}$$

and thus:

$$v_L(t) = \frac{1}{RC} \int_0^t v_i(t) dt + v_L(0) \tag{5.19}$$

Thus, $v_L(t)$ is the time integral of $v_i(t)$.

This circuit has some interesting properties. The integral relationship does not have an associated negative sign, as is the case with the Miller integrator. Noninverting, or positive, integrators are required in many applications. Another useful property is the fact that one terminal of the capacitor is connected to common. This would simplify the initial charging of the capacitor to $v_L(0)$, as may be necessary to set an initial condition.

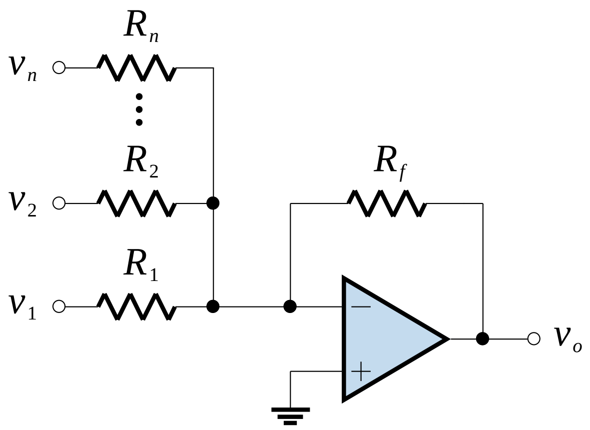
The output of the circuit cannot be taken at the terminal labelled v_L since the connection of a load there will change the preceding analysis. Fortunately, a voltage source output is available that is proportional to v_L – at the output of the op-amp where you can easily verify that $v_o = 2v_L$. Thus the output of the circuit is:

$$v_o(t) = \frac{2}{RC}\int_0^t v_i(t)dt + v_o(0) \tag{5.20}$$

5.8 Summary

- The op-amp is a versatile electronic building block. Real op-amps perform close to the ideal, making circuit design and verification relatively easy.

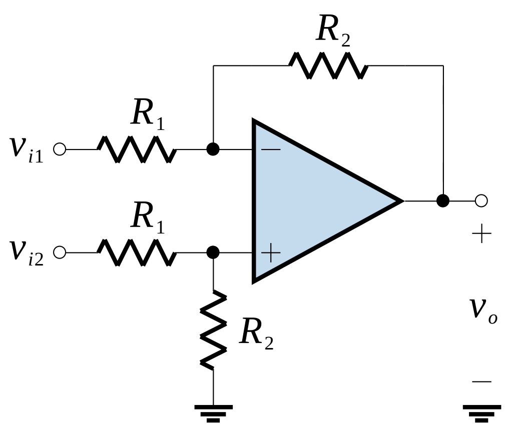
- The inverting summing amplifier:



has an output equal to:

$$v_o = -\left(\frac{R_f}{R_1} v_1 + \frac{R_f}{R_2} v_2 + \cdots + \frac{R_f}{R_n} v_n \right)$$

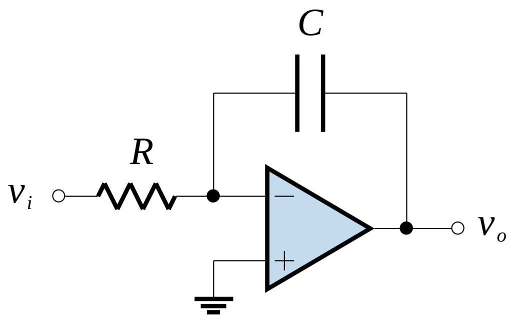
- The difference amplifier:



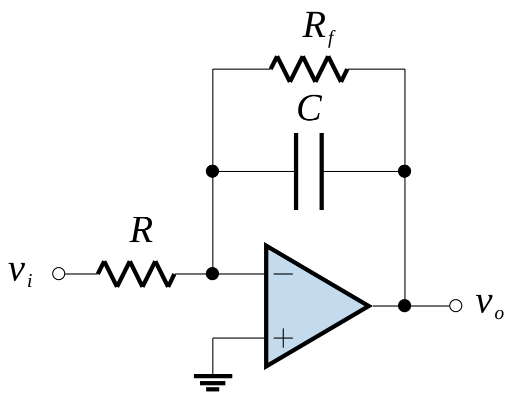
has an output equal to:

$$v_o = \frac{R_2}{R_1}\left(v_{i2} - v_{i1} \right)$$

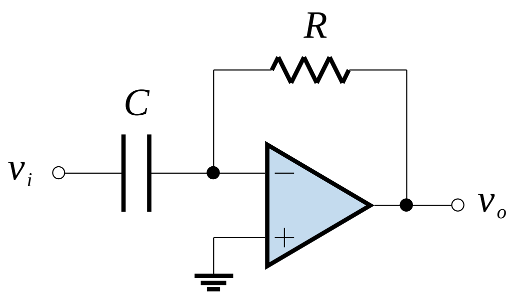
- The inverting integrator:



is a poor implementation of an integrator since the op-amp is operating in an "open-loop" for DC input voltages, causing the output to saturate. A practical *approximation* that alleviates this problem is:



- The differentiator:



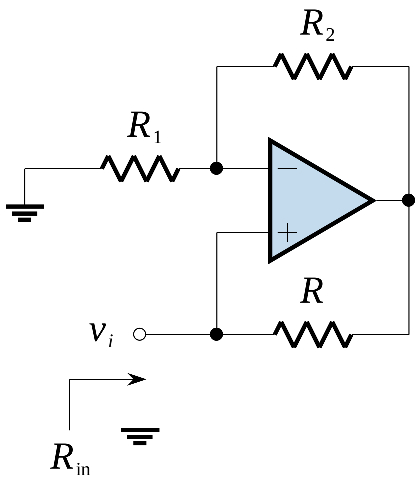
has an output equal to:

$$v_o(t) = -RC\frac{dv_i}{dt}$$

It is rarely used in practice, because it tends to act as a "noise magnifier".

5.20

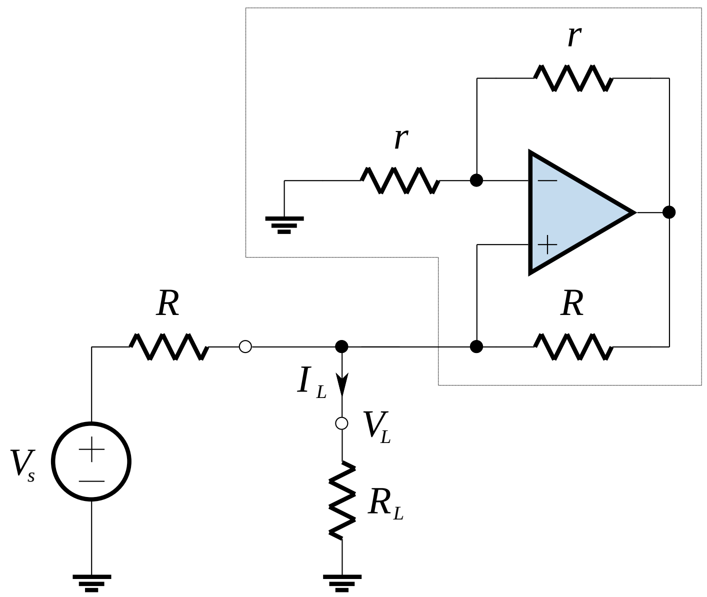
- A negative impedance converter (NIC):



has an input resistance given by:

$$R_{in} = -R\frac{R_1}{R_2}$$

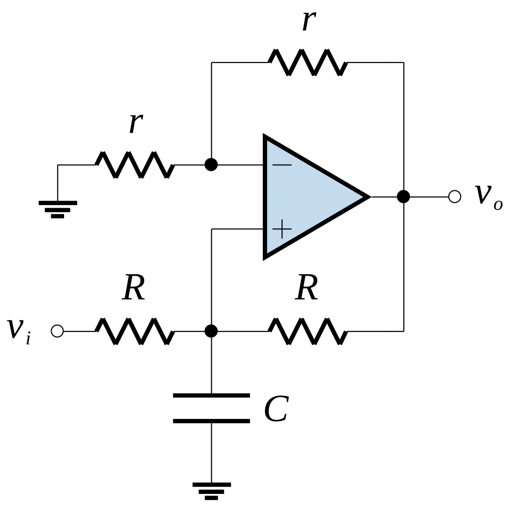
- The Howland voltage-to-current converter:



creates a voltage-controlled current source where:

$$I_L = \frac{V_s}{R}$$

- The Deboo noninverting integrator:



has an output:

$$v_o(t) = \frac{2}{RC}\int_0^t v_i(t)dt + v_o(0)$$

It is a very practical and useful integrator circuit.

5.9 References

Deboo, Gordon: "A Novel Integrator", NASA-TM-X-57906, NASA Ames Research Center, 1966.

Sedra, A. and Smith, K.: *Microelectronic Circuits*, Saunders College Publishing, New York, 1991.

Sheingold, D.: "Impedance & Admittance Transformations", *The Lightning Empiricist*, vol. 12, no. 1, Philbrick Researches, Inc., Boston, 1964.

5.22

Exercises

1.

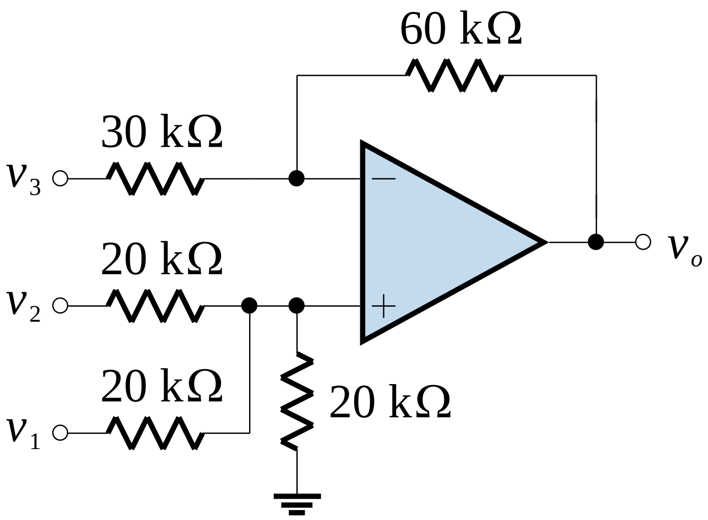
Design an op-amp circuit with a 10 kΩ input resistance which converts a symmetrical square wave at 1 kHz having 2 V peak-to-peak amplitude and zero average value into a triangle wave of 2 V peak-to-peak amplitude.

2.

Design a two op-amp circuit with inputs v_1 and v_2 and input resistances of 100 kΩ whose output is $v_o = v_1 - 10v_2$.

3.

Determine the output voltage of the following circuit:

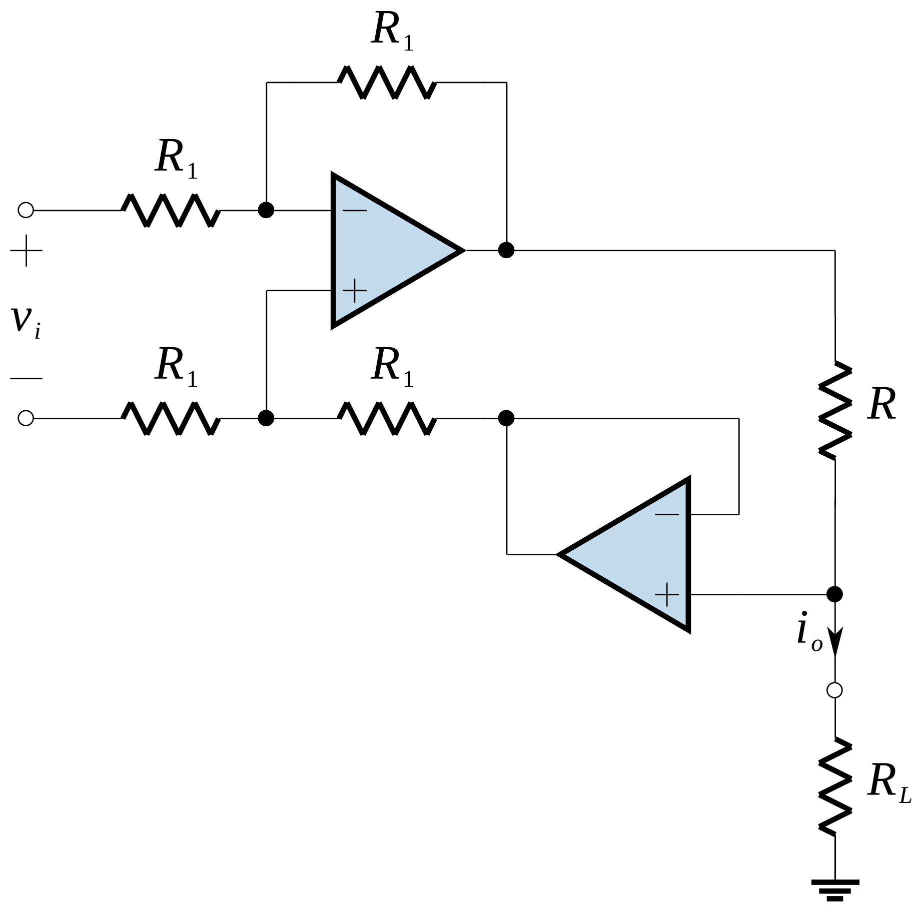


4.

Design a negative impedance converter having an input resistance of -1 kΩ. This circuit is connected to the output terminal of a source whose open-circuit voltage is 1 mV and whose output resistance is 900 Ω. What voltage is then measured at the output of the source?

5.

Express i_o as a function of v_i for the circuit below:



6 Reactive Components

Contents

Introduction

The capacitor is a circuit element whose voltage-current relationship involves the rate of change of voltage. Physically, a capacitor consists of two conducting surfaces on which a charge may be stored, separated by a thin insulating layer which has a very large resistance. Energy is stored in the *electric field* that exists between the capacitor's two conducting surfaces. In addition, the insulating layer may be made of a high permittivity material (such as ceramic) which will dramatically increase the capacitance (compared to air).

The inductor is a circuit element whose voltage-current relationship involves the rate of change of current. Physically, an inductor may be constructed by winding a length of wire into a coil. Energy is stored in the *magnetic field* that exists around the wire in the coil, and the geometry of a coil increases this field compared to a straight wire. In addition, the centre of the coil may be made of a ferromagnetic material (such as iron) which will dramatically increase the inductance (compared to air).

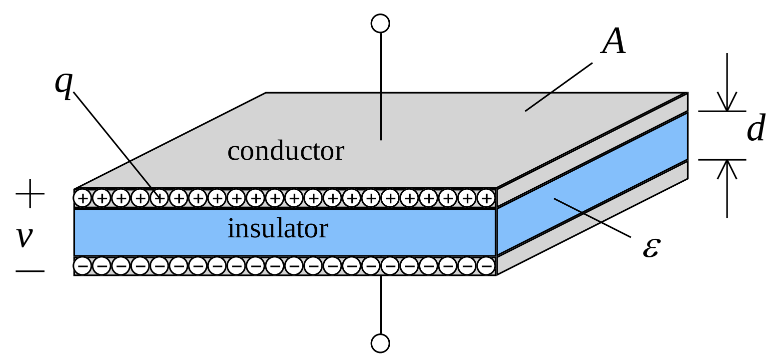
Both the capacitor and the inductor are capable of storing and delivering finite amounts of energy, but they cannot deliver non-zero average power over an infinite time interval. They are therefore *passive* circuit elements, like the resistor.

The capacitor and inductor are *linear* circuit elements. Therefore all the circuit methods previously studied, such as nodal analysis, superposition, Thévenin's theorem, etc., can be applied to circuits containing capacitors and inductors.

Lastly, in dealing with the capacitor and inductor in a circuit, we will note that the equations describing their behaviour bear a similar resemblance – they are the *duals* of each other. It will be shown that the concept of duality is a recurring theme in circuit analysis, and can be readily applied to many simple circuits, saving both time and effort.

6.1 The Capacitor

The simplest capacitor is formed by two conductive plates separated by a dielectric layer:



A parallel plate capacitor

Figure 6.1

One of the plates carries a positive charge, *q*, whilst the other carries an equal but opposite charge, -*q*. Therefore, the capacitor *stores* charge. There is a potential difference, *v*, between the plates. Ideally, the amount of charge *q* deposited on the plates is proportional to the voltage *v* impressed across them. We define a constant[1] called the *capacitance*, *C*, of the structure by the linear relationship:

$$q = Cv \tag{6.1}$$

The definition of capacitance

The unit of capacitance is the *farad*, with symbol F.

There is an *electric field* between the capacitor plates which is established by the charges present on the plates. Energy is stored in the capacitor by virtue of this electric field.

[1] The constant only *models* the behaviour of the structure under certain operating conditions. The capacitance of a structure in the real world will vary with temperature, voltage, pressure, frequency, chemical aging, etc,

The ideal capacitance relationship is a *straight line through the origin*:

The capacitor is a linear circuit element

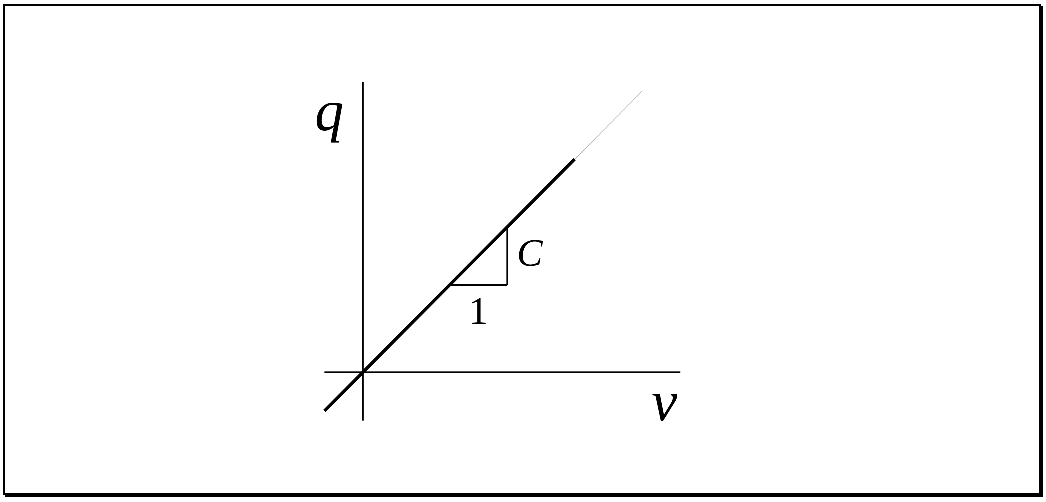


Figure 6.2

Even though capacitance is defined as $C = q/v$, it should be noted that C is a purely geometric property, and depends only on the conductor arrangements and the materials used in the construction. For example, it can be shown for a thin parallel plate capacitor that the capacitance is approximately:

The capacitance of a parallel plate capacitor

$$C \approx \frac{\varepsilon A}{d} \tag{6.2}$$

where A is the area of either of the two parallel plates, and d is the distance between them. The *permittivity*, ε, is a constant of the insulating material between the plates. The permittivity is usually expressed in terms of *relative permittivity*, ε_r:

Relative permittivity defined

$$\varepsilon = \varepsilon_r \varepsilon_0 \tag{6.3}$$

where $\varepsilon_0 = 8.854\ \text{pFm}^{-1}$ is the permittivity of free space (and, for all practical purposes, air).

6.1.1 Capacitor *v-i* Relationships

We now seek a *v-i* relationship for the capacitor. From the definition of current:

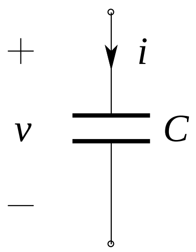
$$i = \frac{dq}{dt} \tag{6.4}$$

we substitute $q = Cv$ and obtain:

$$i = C\frac{dv}{dt} \tag{6.5}$$

The capacitor's *v-i* branch derivative relationship

The circuit symbol for the capacitor is based on the construction of the physical device, and is shown below together with the passive sign convention for the voltage and current:



The circuit symbol for the capacitor

Figure 6.3

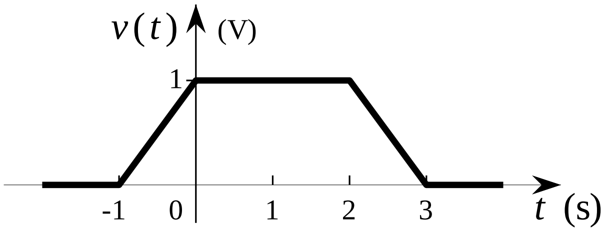
The capacitor voltage may be expressed in terms of the current by integrating Eq. (6.5):

$$v(t) = \frac{1}{C}\int_{t_0}^{t} i\,d\tau + v(t_0) \tag{6.6}$$

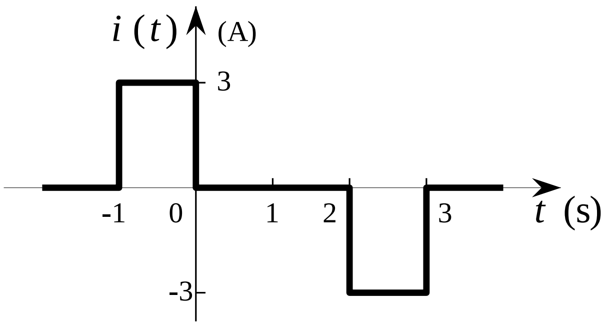
The capacitor's *v-i* branch integral relationship

EXAMPLE 6.1 Capacitor Voltage and Current Waveforms

The voltage waveform across a 3 F capacitor is shown below:



Since the voltage is zero and constant for $t < -1$, the current is zero in this interval. The voltage then begins to increase at the linear rate $dv/dt = 1\,\text{Vs}^{-1}$, and thus a constant current of $i = C\,dv/dt = 3\,\text{A}$ is produced. During the following 2 second interval, the voltage is constant and the current is therefore zero. The final decrease of the voltage causes a negative 3 A and no response thereafter. The current waveform is sketched below:



6.1.2 Energy Stored in a Capacitor

The power delivered to a capacitor is:

$$p = vi = Cv\frac{dv}{dt} \tag{6.7}$$

and the energy stored in its electric field is therefore:

$$\begin{aligned} w_C(t) &= \int_{t_0}^{t} p\,dt + w_C(t_0) \\ &= C\int_{t_0}^{t} v\frac{dv}{dt}dt + w_C(t_0) \\ &= C\int_{v(t_0)}^{v(t)} v\,dv + w_C(t_0) \\ &= \frac{1}{2}C\left[v^2(t) - v^2(t_0)\right] + w_C(t_0) \end{aligned} \tag{6.8}$$

If the capacitor voltage is zero at t_0, then the electric field, and hence the stored capacitor energy, $w_C(t_0)$, is also zero at that instant. We then have:

$$w_C(t) = \frac{1}{2}Cv^2(t) \tag{6.9}$$

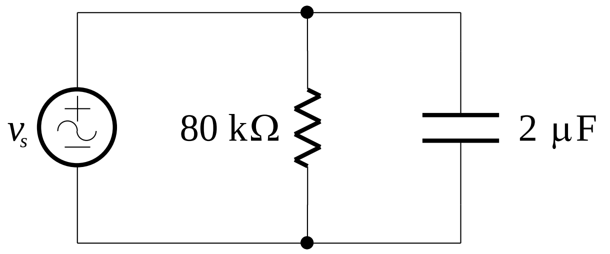
The stored energy in a capacitor

We can see that the energy stored in a capacitor depends only on the capacitance and the voltage. Therefore, a finite amount of energy can be stored in a capacitor even if the current through the capacitor is zero.

Whenever the voltage is not zero, and regardless of its polarity, energy is stored in the capacitor. It follows, therefore, that power must be delivered to the capacitor for a part of the time and recovered from the capacitor later.

EXAMPLE 6.2 Energy Stored in a Capacitor

A sinusoidal voltage source is connected in parallel with an $80\ \mathrm{k\Omega}$ resistor and a 2 μF capacitor, as shown below:



The $80\ \mathrm{k\Omega}$ resistor in the circuit represents the dielectric losses present in a typical $2\ \mu\mathrm{F}$ capacitor. Let $v_s = 325\sin(100\pi t)\,\mathrm{V}$, which corresponds to a typical low voltage 230 V RMS household supply.

The current through the resistor is:

$$i_R = \frac{v}{R} = 4.063\sin(100\pi t)\ \ \mathrm{mA}$$

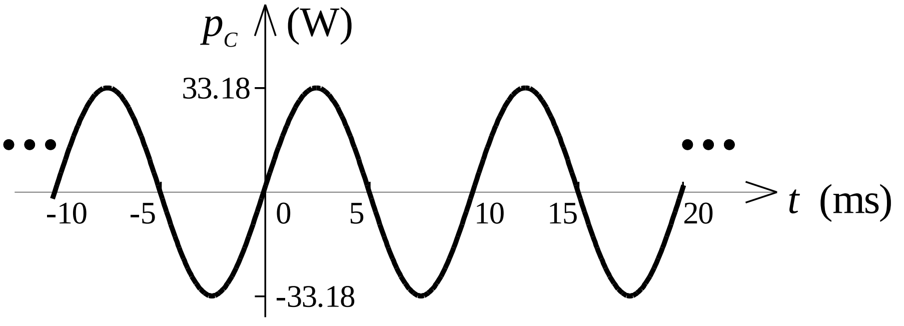
and the current through the capacitor is:

$$i_C = C\frac{dv}{dt} = 2\times10^{-6}\frac{d}{dt}\left[325\sin(100\pi t)\right] = 204.2\cos(100\pi t)\ \ \mathrm{mA}$$

The power delivered to the capacitor is:

$$p_C = vi_C = 325\sin(100\pi t)\times 0.2042\cos(100\pi t) = 33.18\sin(200\pi t)\ \ \mathrm{W}$$

A graph of the power delivered to the capacitor versus time is shown below:

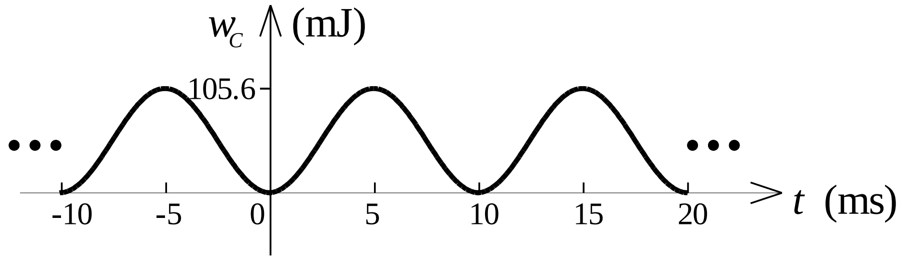


The *instantaneous* power delivered to the capacitor is sinusoidal, and varies at *twice* the frequency of the voltage source. Negative power *delivered to* the capacitor means that power is *sourced from* the capacitor. Note that the *average* power dissipated in the capacitor is zero.

The energy stored in the capacitor is:

$$w_C = \frac{1}{2}Cv^2 = \frac{1}{2}\times 2\times 10^{-6}\left[325\sin(100\pi t)\right]^2 = 105.6\sin^2(100\pi t) \quad \text{mJ}$$

A graph of the stored energy versus time is shown below:



The energy stored is sinusoidal and varies at twice the frequency of the voltage source, but it also has a finite *average* component of 52.81 mJ.

We see that the energy increases from zero at $t = 0$ to a maximum of 105.6 mJ at $t = 5\,\text{ms}$ and then decreases to zero in another 5 ms. During this 10 ms interval, the energy dissipated in the resistor as heat is:

$$w_R = \int_0^{0.01} p_R\,dt = \int_0^{0.01} 1.320\sin^2(100\pi t)\,dt = \int_0^{0.01} 0.6602\left[1-\cos(200\pi t)\right]dt = 6.602 \quad \text{mJ}$$

Thus, an energy equal to 6.25% of the maximum stored energy is lost as heat in the process of storing and removing the energy in the physical capacitor. Later we will formalise this concept by defining a quality factor Q that is proportional to the ratio of the maximum energy stored to the energy lost per period.

6.1.3 Summary of Important Capacitor Characteristics

Several important characteristics of the ideal capacitor are summarised below.

The capacitor is an open-circuit to DC

1. If a constant voltage is held across a capacitor, then $dv/dt = 0$ and subsequently no current enters (or leaves) it. A capacitor is thus **an open-circuit to DC**. This fact is represented by the capacitor symbol.

A capacitor voltage cannot change instantaneously

2. A capacitor **voltage cannot change instantaneously**, for this implies $dv/dt = \infty$, and the capacitor would require infinite current. Thus, capacitor voltage is smooth and continuous. This fact will be used frequently when undertaking *transient analysis* of circuits.

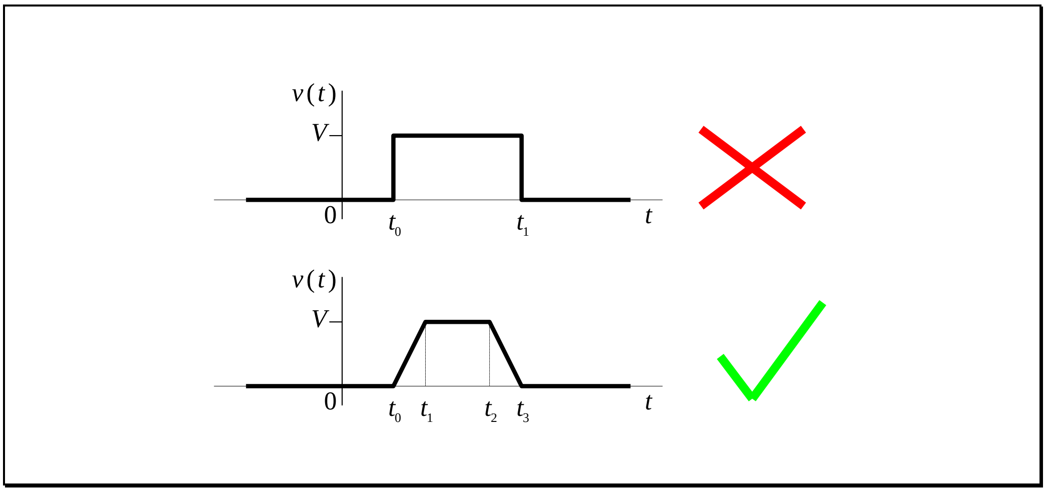


Figure 6.4

3. For an ideal capacitor, **current can change instantaneously**.

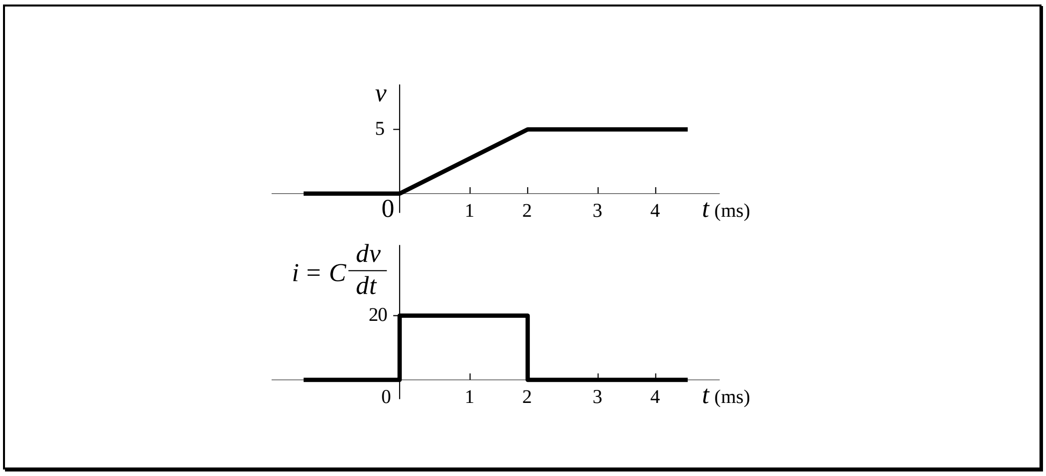


Figure 6.5

4. An ideal capacitor never dissipates energy, but stores it and releases it using its electric field.

6.2 The Inductor

On 21 April 1820, the Danish scientist Ørsted noticed a compass needle deflected from magnetic north when an electric current from a battery was switched on and off, confirming a direct relationship between electricity and magnetism. Three months later he began more intensive investigations and soon thereafter published his findings, showing that an electric current produces a circular magnetic field.

An electric current produces a magnetic field

An inductor is a two-terminal device, usually constructed as a coil of wire, that is designed to store energy in a magnetic field. The coil effectively increases the magnetic field by the number of "turns" of wire, and as we shall see, also increases the number of "circuits" which "link" the magnetic field.

An ideal inductor produces a magnetic field which is wholly confined within it. The closest approximation to an ideal inductor that we can physically produce is a toroid, which has an almost uniform magnetic field confined within it:

A toroidal inductor

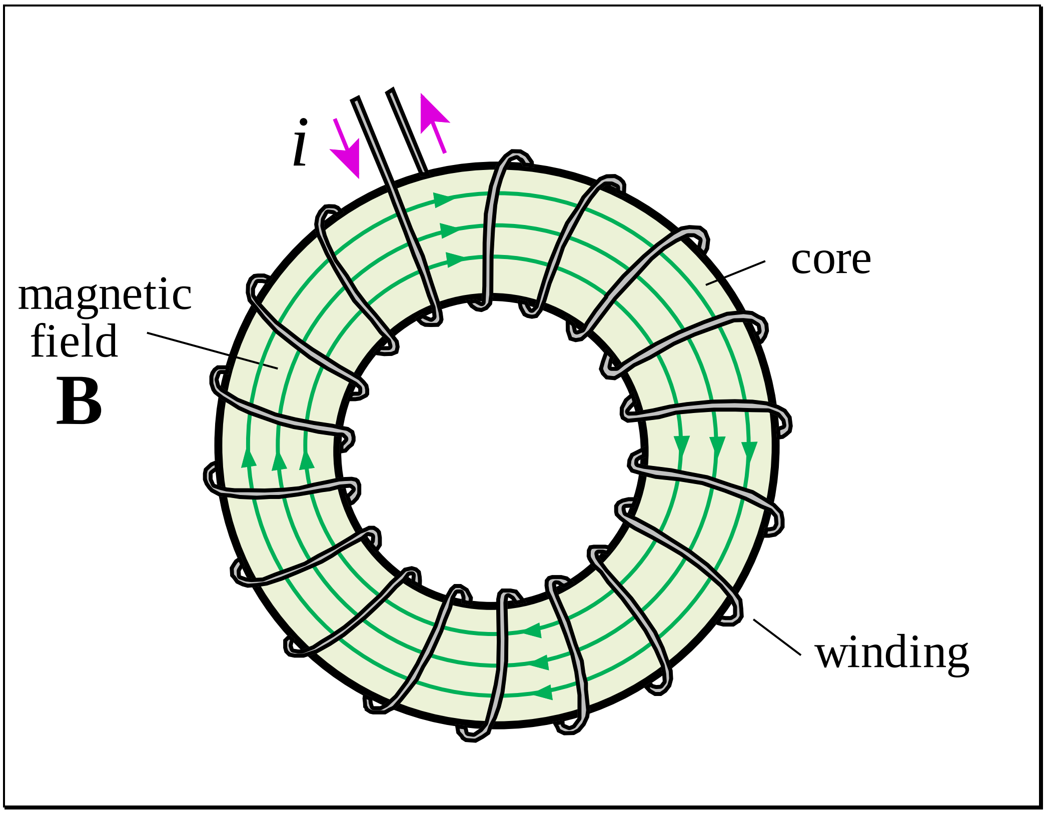


Figure 6.6

Magnetic flux defined

To explain magnetism, 19th century scientists invoked an analogy with fluids and postulated the existence of a magnetic fluid, known as magnetic "flux", Φ, which streamed throughout space and manifested itself as magnetism. Magnetic flux always streams out of north poles and into south poles, and forms a closed loop. We still use this concept of flux today, as the theory has been spectacularly successful.

If a tube of magnetic flux (which is a closed loop) streams through a closed loop of wire (a circuit), then it is said to "link" with the circuit. The amount of "flux linkage", λ, is given by:

Magnetic flux linkage defined

$$\lambda = N\Phi \tag{6.10}$$

where N is the number of loops of wire, or turns, in the circuit, and Φ is the average amount of magnetic flux streaming through each loop.

Magnetic flux linkage shown graphically

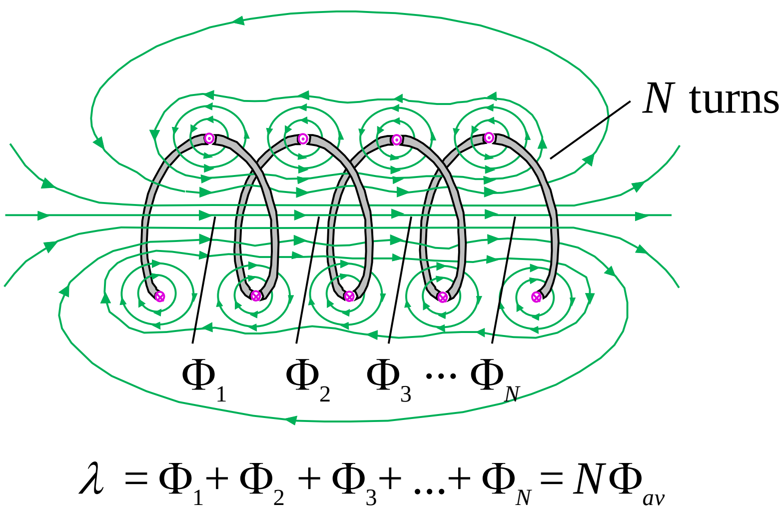


Figure 6.7

An ideal inductor is a structure where the flux linkage (with itself) is directly proportional to the current through it. We define a constant, called the self *inductance*, L, of the structure by the linear relationship:

The definition of inductance

$$\lambda = Li \tag{6.11}$$

The unit of inductance is the *henry*, with symbol H.

The ideal inductance relationship is a *straight line through the origin*:

The inductor is a linear circuit element

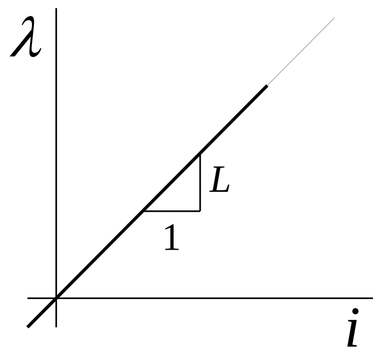


Figure 6.8

Even though inductance is defined as $L = \lambda / i$, it should be noted that L is a purely geometric property, and depends only on the conductor arrangements and the materials used in the construction.

For example, it can be shown for a closely wound toroid that the inductance is approximately:

The inductance of a toroidal inductor

$$L \approx N^2 \frac{\mu A}{l} \tag{6.12}$$

where A is the cross-sectional area of the toroid material, and l is the mean path length around the toroid. The *permeability*, μ, is a constant of the material used in making the toroid. The permeability is usually expressed in terms of *relative permeability*, μ_r:

Relative permeability defined

$$\mu = \mu_r \mu_0 \tag{6.13}$$

where $\mu_0 = 400\pi \text{ nHm}^{-1}$ is the permeability of free space (and, for all practical purposes, air).

6.2.1 Inductor *v-i* Relationships

We now seek a *v-i* relationship for the inductor. In 1840, the great British experimentalist Michael Faraday[2] discovered that a changing magnetic field could induce a voltage in a neighbouring circuit, or indeed the circuit that was producing the magnetic field. Faraday's Law states that the induced voltage is equal to the rate-of-change of magnetic flux linkage:

Faraday's Law

$$v = -\frac{d\lambda}{dt} \tag{6.14}$$

The minus sign comes from the fact that the polarity of the induced voltage is such as to oppose the change in flux. For an inductor, we can figure out that the polarity of the induced voltage must be positive at the terminal where the current enters the inductor. If we know this, then we can mark the polarity on a circuit diagram and deal with the magnitude of the induced voltage by dropping the minus sign (the determination of the voltage polarity is called Lenz's Law).

If we allow the polarity to be set by Lenz's Law, and substitute $\lambda = Li$ into the previous equation, then we get:

The inductor's *v-i* branch derivative relationship

$$v = L\frac{di}{dt} \tag{6.15}$$

[2] The American inventor Joseph Henry discovered this phenomenon independently, but Faraday was the first to publish.

The circuit symbol for the inductor is based on the construction of the physical device, and is shown below together with the passive sign convention for the voltage and current:

The circuit symbol for the inductor

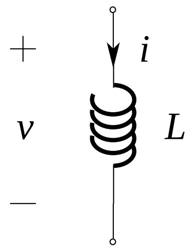


Figure 6.9

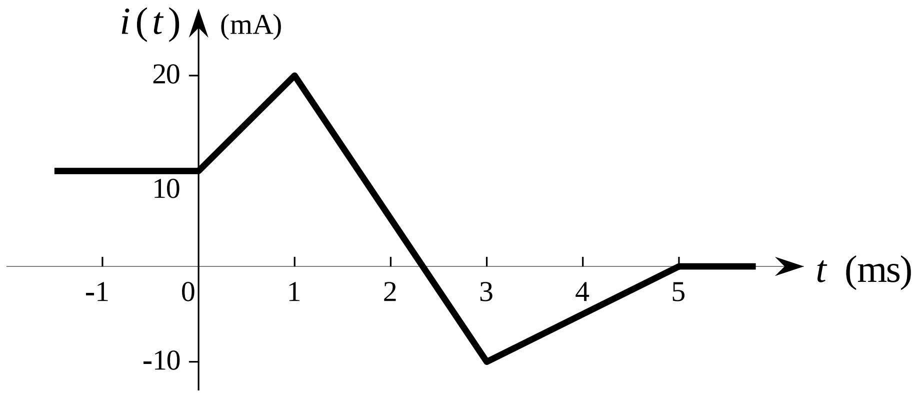
The inductor current may be expressed in terms of the voltage by integrating Eq. (6.15):

$$i(t) = \frac{1}{L}\int_{t_0}^{t} v\,d\tau + i(t_0) \tag{6.16}$$

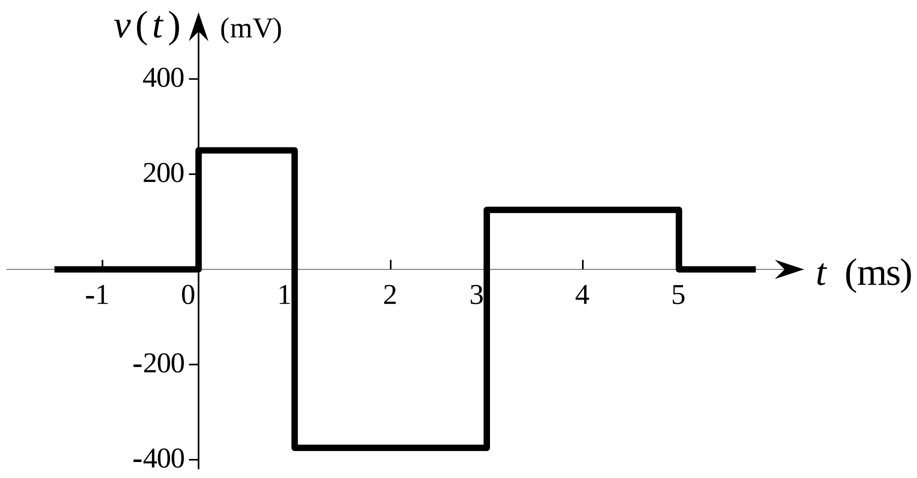
The inductor's v-i branch integral relationship

EXAMPLE 6.3 Inductor Current and Voltage Waveforms

The variation of current through a 25 mH inductor as a function of time is shown below:

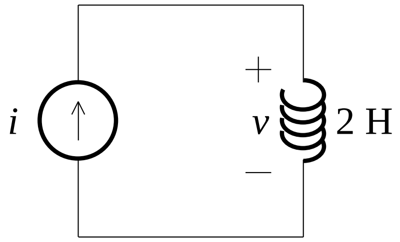


Since the current is constant for $t < 0$, the voltage is zero in this interval. The current then begins to increase at the linear rate $di/dt = 10\ \text{As}^{-1}$, and thus a constant voltage of $v = L\,di/dt = 250\,\text{mV}$ is produced. During the following 2 millisecond interval, the current decreases at the linear rate $di/dt = -15\ \text{As}^{-1}$, and so the voltage is $v = L\,di/dt = -375\,\text{mV}$. The final increase of the current causes a positive 125 mV and no response thereafter. The voltage waveform is sketched below:

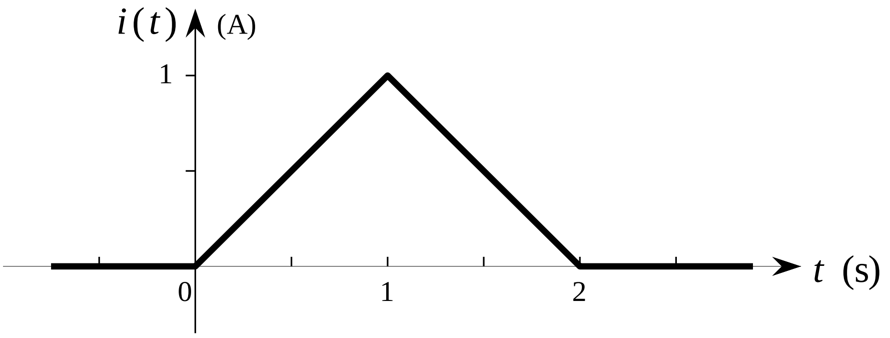


EXAMPLE 6.4 Inductor Current, Voltage and Power Waveforms

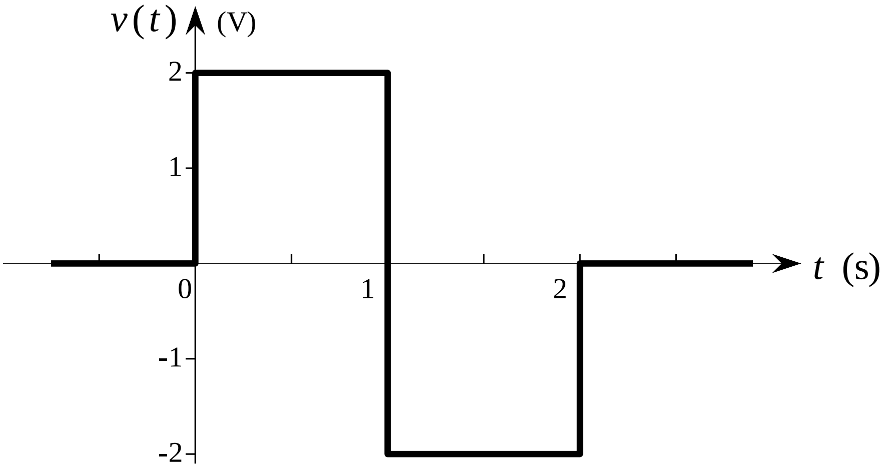
Consider the simple circuit below:



Let the current $i(t)$ which is produced by the source be described by the function of time shown below:

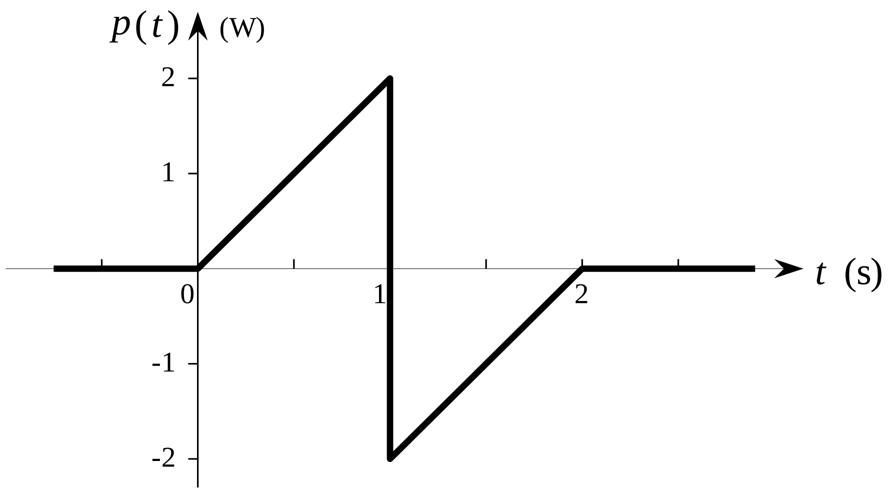


Since $v = L\,di/dt = 2\,di/dt$ then the voltage across the inductor is as shown below:



6.18

We know that a resistor always absorbs power and the energy absorbed is dissipated as heat – but how about an inductor? For the inductor in this example, the instantaneous power $p(t)=v(t)i(t)$ absorbed by the inductor is given by the graph below:



We see that the power absorbed by the inductor is zero for $-\infty<t\leq 0$ and $2\leq t\leq\infty$. For $0<t<1$, since $p(t)$ is a positive quantity, the inductor is absorbing power (which is produced by the source). However, for $1<t<2$, since $p(t)$ is a negative quantity, the inductor is actually supplying power (to the source).

To get the energy absorbed by the inductor, we simply integrate the power absorbed over time. For this example, the energy absorbed increases from 0 to $\frac{1}{2}(1)(2)=1\,\text{J}$ as time goes from $t=0$ to $t=1\,\text{s}$. However, from $t=1$ to $t=2\,\text{s}$, the inductor supplies energy such that at time $t=2\,\text{s}$ and thereafter, the net energy absorbed by the inductor is zero. Since all of the energy absorbed by the inductor is not dissipated but is eventually returned, we say that the inductor *stores* energy. The energy is stored in the magnetic field that surrounds the inductor.

6.2.2 Energy Stored in an Inductor

The power delivered to an inductor is:

$$p = vi = L\frac{di}{dt}i \tag{6.17}$$

and the energy stored in its magnetic field is therefore:

$$\begin{aligned} w_L(t) &= \int_{t_0}^{t} p\,dt + w_L(t_0) \\ &= L\int_{t_0}^{t} i\frac{di}{dt}dt + w_L(t_0) \\ &= L\int_{i(t_0)}^{i(t)} i\,di + w_L(t_0) \\ &= \frac{1}{2}L\left[i^2(t) - i^2(t_0)\right] + w_L(t_0) \end{aligned} \tag{6.18}$$

If the inductor current is zero at t_0, then the magnetic field, and hence the stored inductor energy, $w_L(t_0)$, is also zero at that instant. We then have:

$$w_L(t) = \frac{1}{2}Li^2(t) \tag{6.19}$$

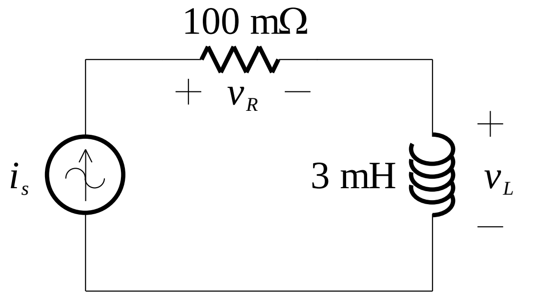
The stored energy in an inductor

We can see that the energy stored in an inductor depends only on the inductance and the current. Therefore, a finite amount of energy can be stored in an inductor even if the voltage across the inductor is zero.

Whenever the current is not zero, and regardless of its direction, energy is stored in the inductor. It follows, therefore, that power must be delivered to the inductor for a part of the time and recovered from the inductor later.

EXAMPLE 6.5 Energy Stored in an Inductor

A sinusoidal current source is connected in series with a $100\ \text{m}\Omega$ resistor and a 3 mH inductor, as shown below:



The $100\ \text{m}\Omega$ resistor in the circuit represents the resistance of the wire which must be associated with the physical coil. Let $i_s = 12\sin(100\pi t)$ A .

The voltage across the resistor is:

$$v_R = Ri = 1.2\sin(100\pi t)\ \text{V}$$

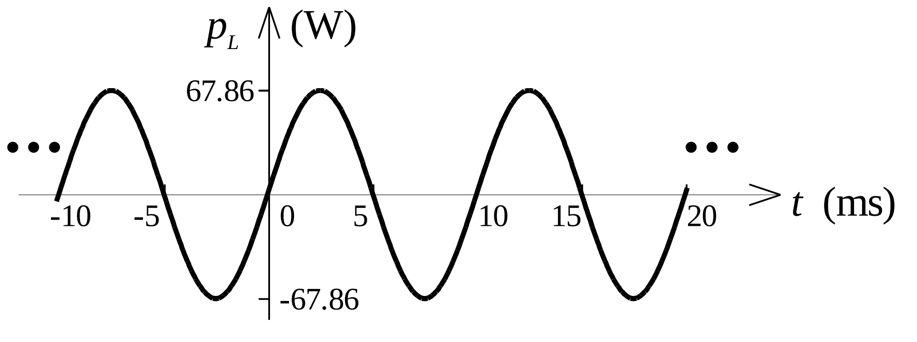
and the voltage across the inductor is:

$$v_L = L\frac{di}{dt} = 3\times 10^{-3}\frac{d}{dt}[12\sin(100\pi t)] = 11.31\cos(100\pi t)\ \text{V}$$

The power delivered to the inductor is:

$$p_L = v_L i = 11.31\cos(100\pi t)\times 12\sin(100\pi t) = 67.86\sin(200\pi t)\ \text{W}$$

A graph of the power delivered to the inductor versus time is shown below:

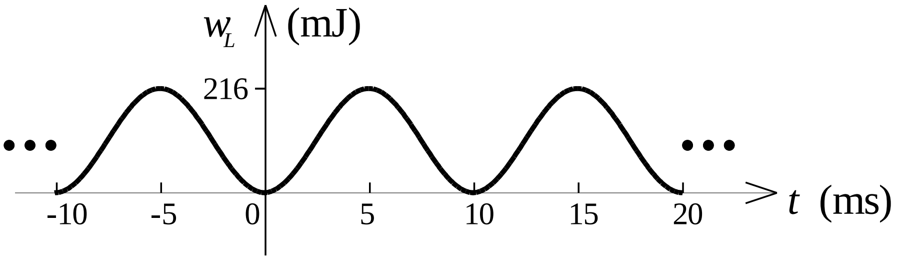


The *instantaneous* power delivered to the inductor is sinusoidal, and varies at *twice* the frequency of the current source. Negative power *delivered to* the inductor means that power is *sourced from* the inductor. Note that the *average* power dissipated in the inductor is zero.

The energy stored in the inductor is:

$$w_L = \frac{1}{2} L i^2 = \frac{1}{2} \times 3 \times 10^{-3} [12 \sin(100\pi t)]^2 = 216 \sin^2(100\pi t) \text{ mJ}$$

A graph of the stored energy versus time is shown below:



The energy stored is sinusoidal and varies at twice the frequency of the current source, but it also has a finite *average* component of 108 mJ.

We see that the energy increases from zero at $t = 0$ to a maximum of 216 mJ at $t = 5$ ms and then decreases to zero in another 5 ms. During this 10 ms interval, the energy dissipated in the resistor as heat is:

$$w_R = \int_0^{0.01} p_R dt = \int_0^{0.01} 14.4 \sin^2(100\pi t) dt = \int_0^{0.01} 7.2[1 - \cos(200\pi t)] dt = 72 \text{ mJ}$$

Thus, an energy equal to 33.33% of the maximum stored energy is lost as heat in the process of storing and removing the energy in the physical inductor.

6.2.3 Summary of Important Inductor Characteristics

Several important characteristics of the ideal inductor are summarised below.

The inductor is a short-circuit to DC

1. If a constant current is passed through an inductor, then $di/dt = 0$ and subsequently no voltage appears across it. An inductor is thus **a short-circuit to DC**.

An inductor current cannot change instantaneously

2. An inductor **current cannot change instantaneously**, for this implies $di/dt = \infty$, and the inductor would require infinite voltage. Thus, inductor current is smooth and continuous. This fact will be used frequently when undertaking *transient analysis* of circuits.

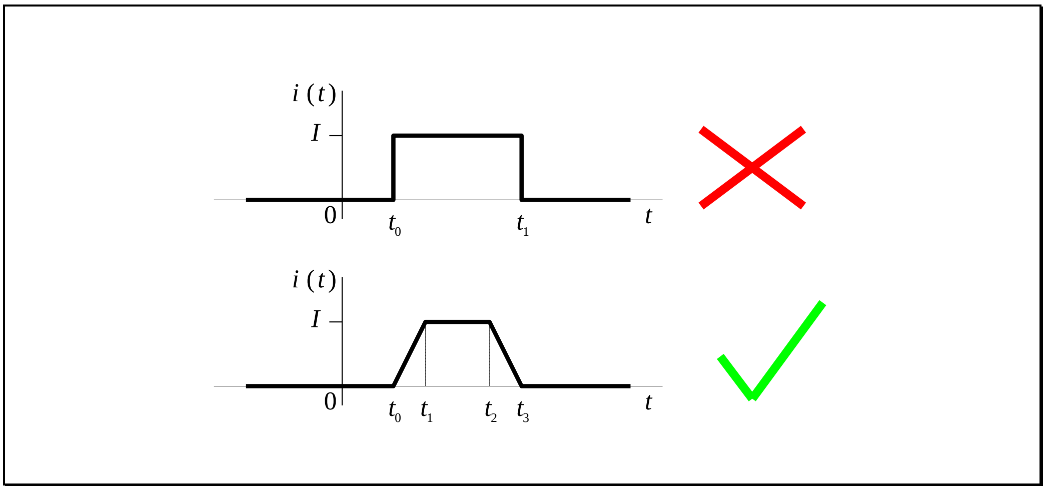


Figure 6.10

3. For an ideal inductor, **voltage can change instantaneously**.

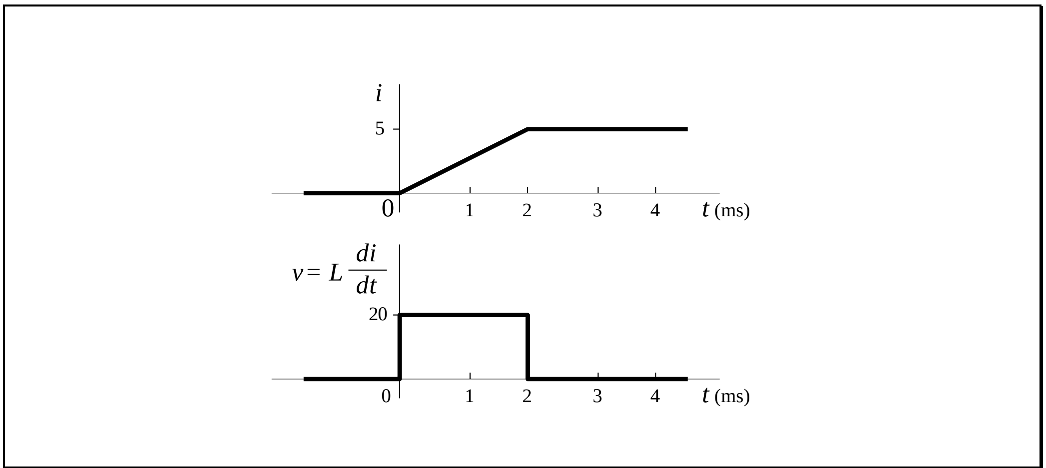


Figure 6.11

4. An ideal inductor never dissipates energy, but stores it and releases it using its magnetic field.

6.3 Practical Capacitors and Inductors

Practical capacitors and inductors are manufactured for different values, voltage and current ratings, accuracy, volumetric efficiency, temperature stability, etc. As such, their construction will play a role in their electrical behaviour. To represent these physical components in a circuit we need to model their non-ideal and parasitic effects.

Practical components have non-ideal and parasitic effects

6.3.1 Capacitors

There are many different types of capacitor construction. Some are shown below, labelled by the type of dielectric:

Some types of capacitors

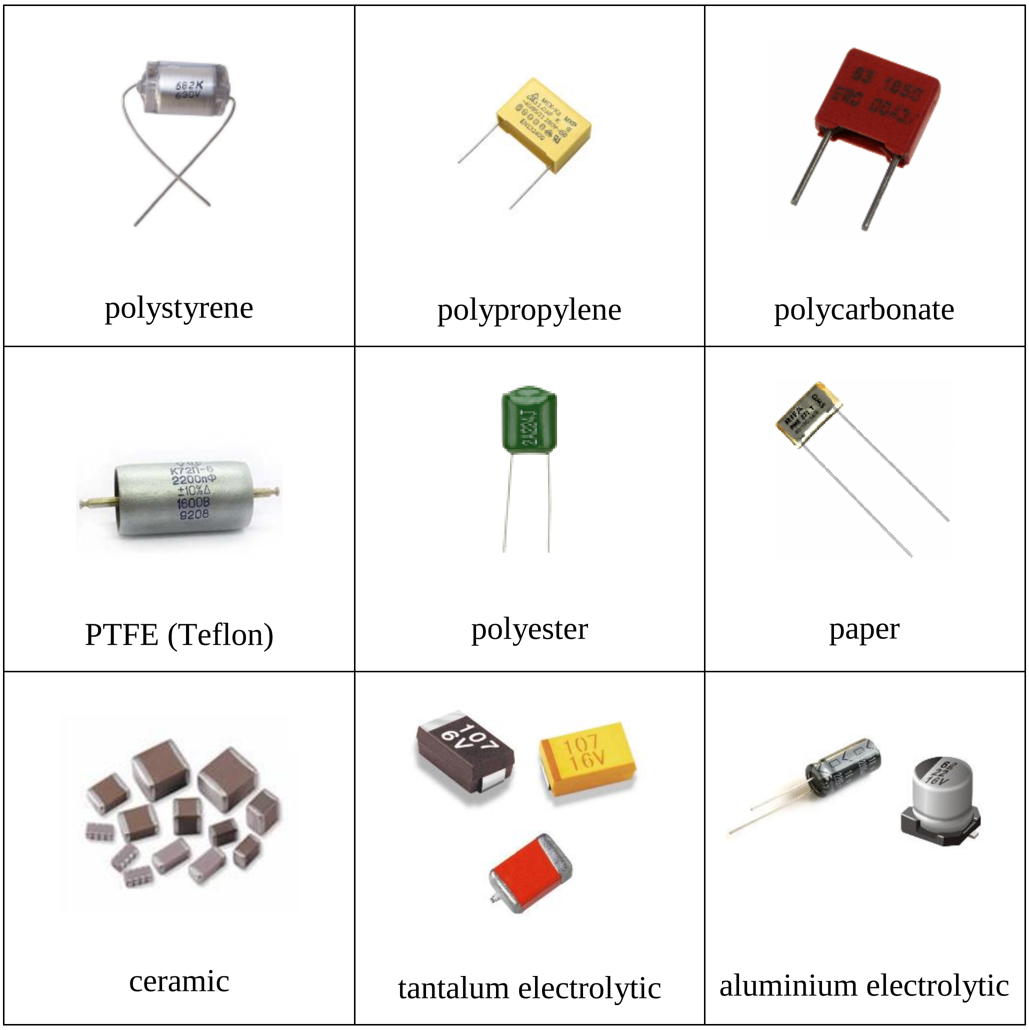


Figure 6.12 – Some types of capacitors

A reasonably accurate circuit model of a real capacitor is shown below:

Model of a real capacitor

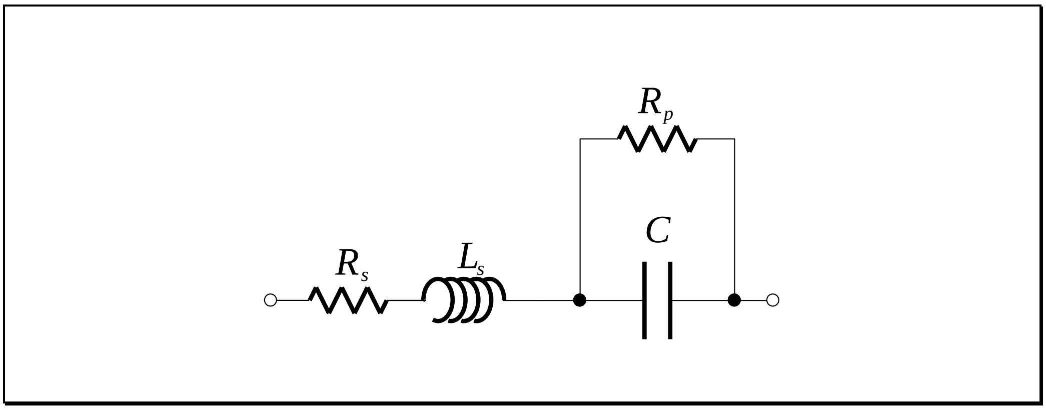


Figure 6.13

In the model, the series resistance R_s takes into account the finite resistance of the plates used to make the capacitor. The series inductance L_s is used to model the fact that a current is required to charge and discharge the plates, and this current must have a magnetic field. Finally, no practical material is a perfect insulator, and the resistance R_p represents conduction through the dielectric.

Parasitic elements defined

We call R_s, L_s and R_p *parasitic elements*. Capacitors are designed to minimize the effects of the parasitic elements consistent with other requirements such as physical size and voltage rating. However, parasitics are always present, and when designing circuits care must be taken to select components for which the parasitic effects do not compromise the proper operation of the circuit.

6.3.2 Electrolytic Capacitors

Electrolytic capacitors have a large capacitance per unit volume, but are polarised

One side (the anode) of an electrolytic capacitor is formed from a foil of aluminium which is anodised to produce an oxide layer which is the dielectric. The oxide-coated foil is immersed in an electrolytic solution which forms the cathode. A large capacitance per unit volume is achieved. However, exposure to a reverse voltage for a long time leads to rapid heating of the capacitor and to breakdown. Thus, electrolytic capacitors are polarised and it must be ensured that the correct voltage polarity is applied to avoid failure.

6.3.3 Inductors

There are many different types of inductor construction. Some are shown below, labelled by the type of core:

Some types of inductors

		
ferrite core	laminated steel core	air core

Figure 6.14 – Some types of inductors

The core of an inductor is generally a magnetic material which exhibits energy losses

The construction of the *core* of the inductor is especially important in determining its properties. Laminated steel cores are common for low frequencies, and ferrites (non-conducting oxides of iron) are used for high frequencies. Since the core is subjected to a changing magnetic field, the induced voltages in the core create what are known as *eddy currents*. Steel laminations and ferrites reduce the energy losses caused by these eddy currents. There are also energy loss mechanisms associated with reversing the "magnetic domains" in a core, and so the losses due to the magnetic characteristic and the induced eddy currents are combined into a term called the *core loss*.

Air-cored inductors do not exhibit core losses

Air-cored inductors are linear and do not exhibit core losses (since there is no conductive core). They can be made by winding a coil on a non-magnetic former, such as plastic, or may be self-supporting if made large enough. Air-cored inductors have lower inductance than ferromagnetic-core inductors, but are often used at high frequencies because they are free of core losses.

A reasonably accurate circuit model of a real inductor is shown below:

Model of a real inductor

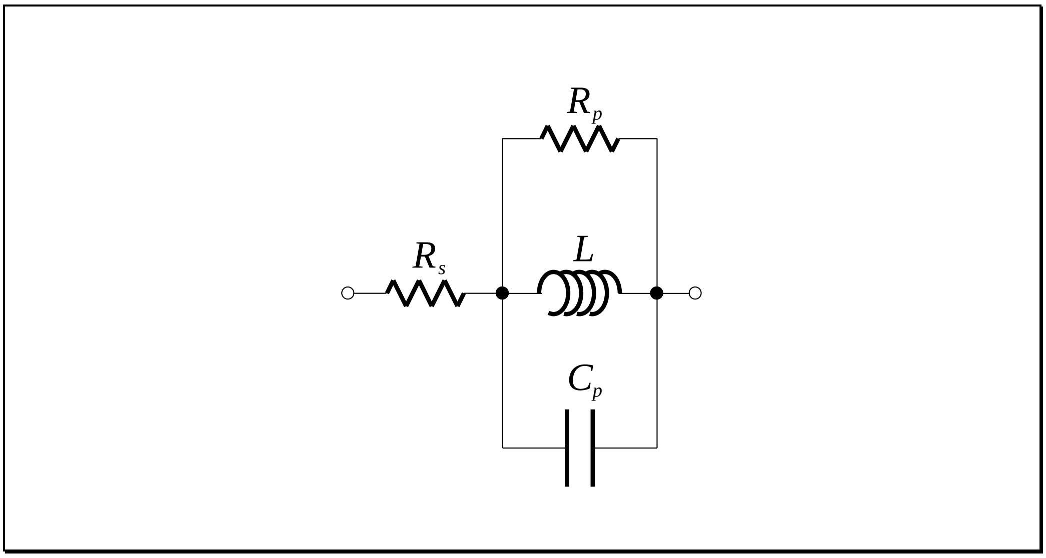


Figure 6.15

The series resistance R_s takes into account the finite resistance of the wire used to create the coil. The parallel capacitance C_p is associated with the electric field in the insulation surrounding the wire, and is called *interwinding capacitance*. The parallel resistance R_p represents the core losses.

The following inductor model, showing just the predominant non-ideal effect of finite winding resistance, is often used at low frequencies:

Low frequency model of a real inductor showing the winding resistance

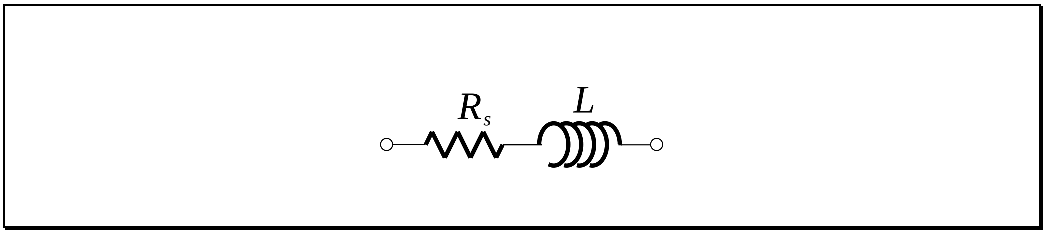


Figure 6.16

6.4 Series and Parallel Connections of Inductors and Capacitors

Resistors connected in series and parallel can be combined into an equivalent resistor. Inductors and capacitors can be combined in the same manner.

6.4.1 Inductors

A series connection of inductors is shown below:

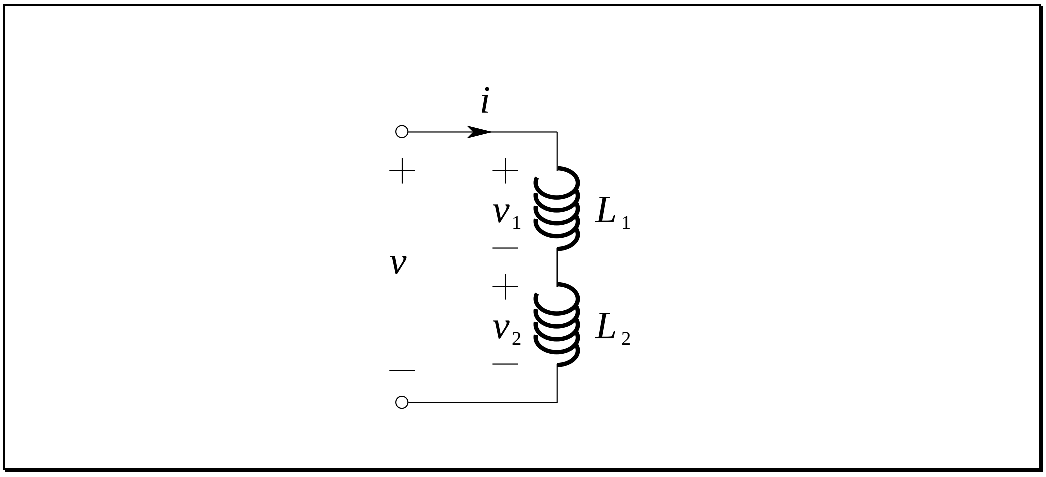


Figure 6.17

We find from KVL that:

$$\begin{aligned} v &= v_1 + v_2 \\ &= L_1 \frac{di}{dt} + L_2 \frac{di}{dt} \\ &= \left(L_1 + L_2\right)\frac{di}{dt} \\ &= L\frac{di}{dt} \end{aligned} \tag{6.20}$$

where:

$$L = L_1 + L_2 \qquad \text{(series)} \tag{6.21}$$

Combining inductors in series

6.28

We depict this below:

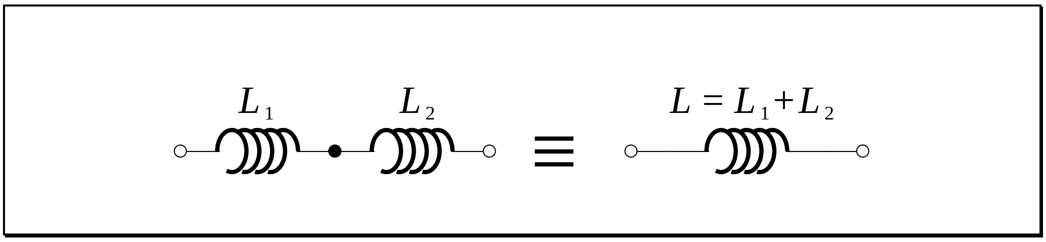


Figure 6.18

For the case of two inductors in parallel, as shown below:

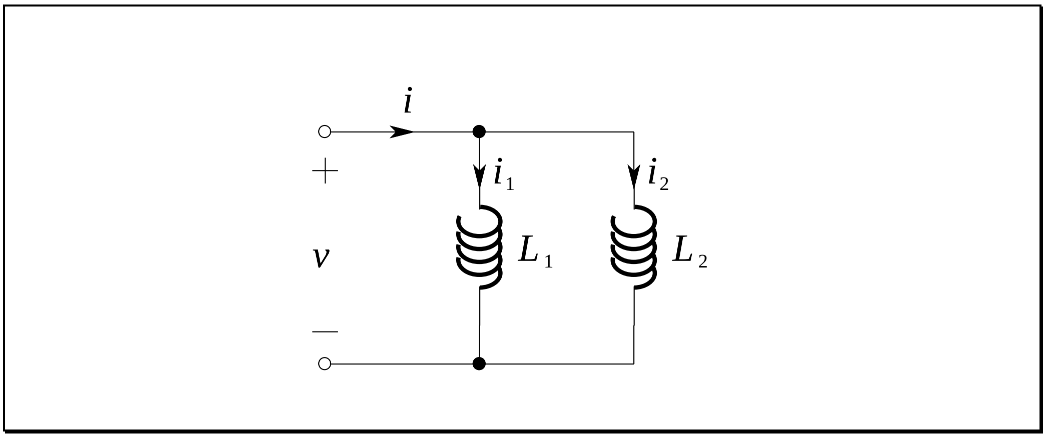


Figure 6.19

we have, by KCL:

$$\begin{aligned} i &= i_1 + i_2 \\ &= \frac{1}{L_1}\int_{t_0}^{t} v\,dt + i_1(t_0) + \frac{1}{L_2}\int_{t_0}^{t} v\,dt + i_2(t_0) \\ &= \left(\frac{1}{L_1} + \frac{1}{L_2}\right)\int_{t_0}^{t} v\,dt + i_1(t_0) + i_2(t_0) \\ &= \frac{1}{L}\int_{t_0}^{t} v\,dt + i(t_0) \end{aligned} \tag{6.22}$$

where:

$$\frac{1}{L} = \frac{1}{L_1} + \frac{1}{L_2} \qquad \text{(parallel)} \tag{6.23}$$

Combining inductors in parallel

and:

$$i(t_0) = i_1(t_0) + i_2(t_0) \tag{6.24}$$

We depict this below:

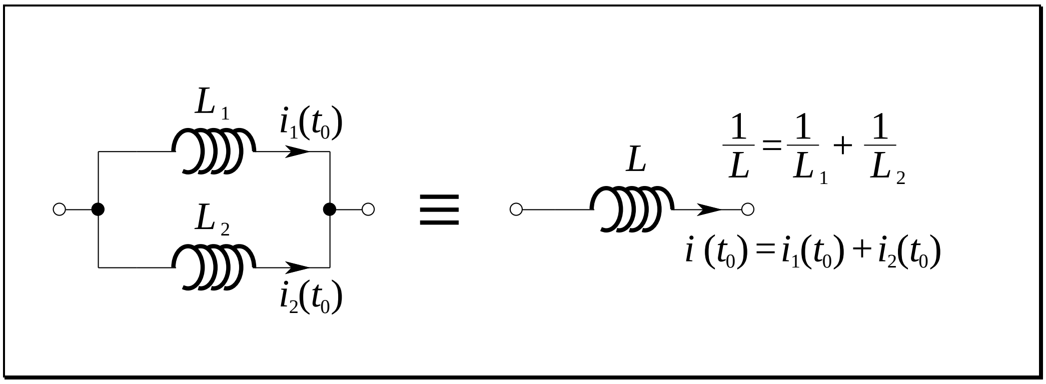


Figure 6.20

6.4.2 Capacitors

Consider two capacitors in parallel as shown below:

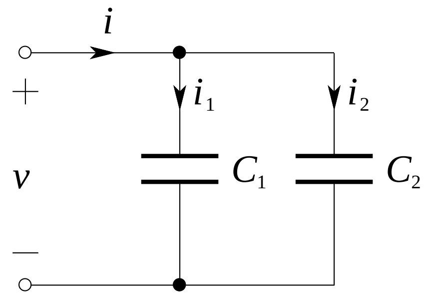


Figure 6.21

By KCL, we have:

$$\begin{aligned} i &= i_1 + i_2 \\ &= C_1 \frac{dv}{dt} + C_2 \frac{dv}{dt} \\ &= (C_1 + C_2) \frac{dv}{dt} \\ &= C \frac{dv}{dt} \end{aligned} \tag{6.25}$$

so that we obtain:

Combining capacitors in parallel

$$C = C_1 + C_2 \qquad \text{(parallel)} \tag{6.26}$$

We depict this below:

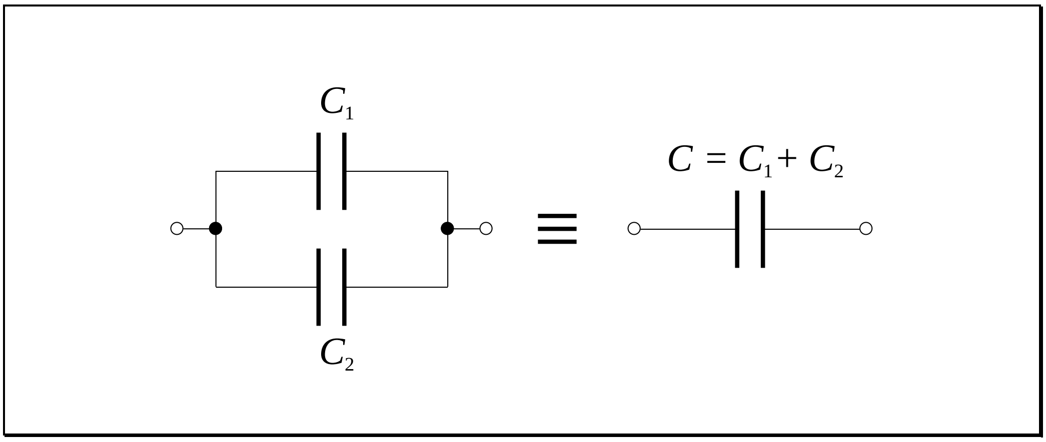


Figure 6.22

For capacitors in series, analogously to inductors in parallel, we get the result:

$$\frac{1}{C} = \frac{1}{C_1} + \frac{1}{C_2} \qquad \text{(series)} \tag{6.27}$$

Combining capacitors in series

which is shown below:

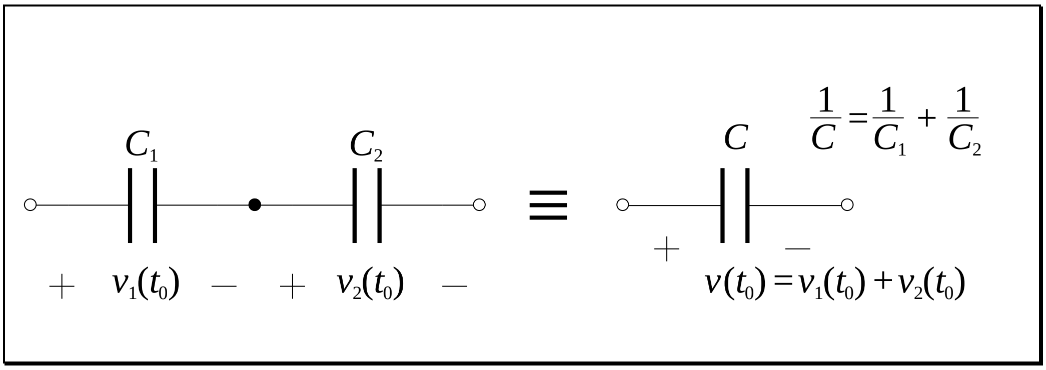


Figure 6.23

In summary, inductors in series and parallel are treated like resistors, whereas capacitors in series and parallel are treated like conductances.

6.5 Circuit Analysis with Inductors and Capacitors

Linear circuit analysis techniques can be applied to circuits with inductors and capacitors

For the ideal inductor and ideal capacitor, all of the relationships between voltage and current are linear relationships (apart from the integral relationships that have an initial condition term). Consequently, circuit analysis techniques which rely on the linearity property (such as nodal analysis, mesh analysis, superposition, Thévenin's theorem and Norton's theorem) can be applied to circuits containing inductors and capacitors.

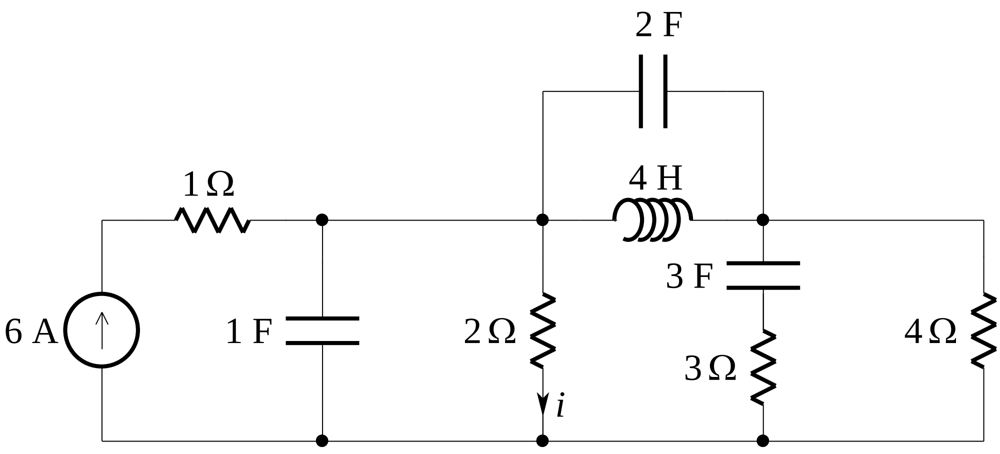
6.5.1 DC Circuits

A DC circuit is treated as a purely resistive circuit in the "steady-state"

An ideal inductor behaves as a short-circuit to DC, and the capacitor behaves as an open-circuit to DC. We can use these facts to determine voltages and currents in DC circuits that contain both inductors and capacitors. Other sources, such as sinusoidal sources, will be treated later.

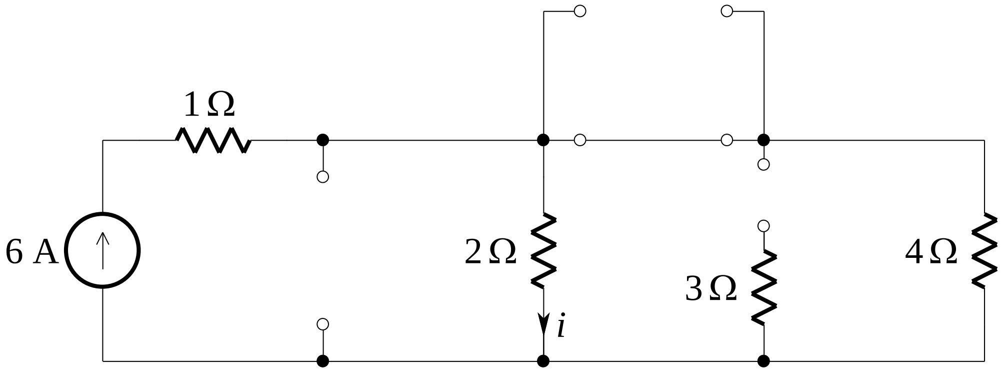
EXAMPLE 6.6 DC Analysis of a Circuit with Storage Elements

Determine the current i in the circuit below:



The circuit has one independent current source whose value is constant. For a resistive circuit we would naturally anticipate that all voltages and currents are constant. However, this is not a resistive circuit. Yet, our intuition suggests that the constant-valued current source produces constant-valued responses. This fact will be confirmed more rigorously later. In the meantime, we shall use the result that a circuit containing only constant-valued sources is a DC circuit.

Since for DC all inductors behave like short-circuits and all capacitors behave like open-circuits, we can replace the original circuit with an equivalent resistive circuit:



By current division we find that:

$$i = \frac{4}{4+2}(6) = 4\text{ A}$$

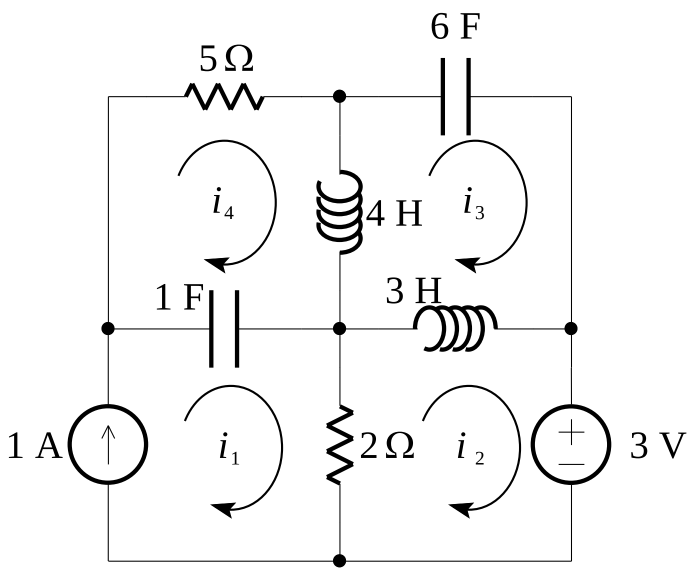
6.5.2 Nodal and Mesh Analysis

Nodal and mesh analysis can be applied to circuits with inductors and capacitors

Just as we analysed resistive circuits with the use of node and mesh equations, we can write a set of equations for circuits that contain inductors and capacitors in addition to resistors and sources. The procedure is similar to that described for the resistive case – the difference being that for inductors and capacitors the appropriate relationship between voltage and current is used in place of Ohm's Law.

EXAMPLE 6.7 Mesh Analysis of a Circuit with Storage Elements

Consider the circuit shown below:



By mesh analysis, for mesh i_1:

$$i_1 = 1$$

For mesh i_2:

$$2(i_2 - i_1) + 3\frac{d}{dt}(i_2 - i_3) = -3$$

For mesh i_3:

$$3\frac{d}{dt}\left(i_3 - i_2\right) + 4\frac{d}{dt}\left(i_3 - i_4\right) + \frac{1}{6}\int_{-\infty}^{t} i_3 dt = 0$$

For mesh i_4:

$$\frac{1}{1}\int_{-\infty}^{t}\left(i_4 - i_1\right)dt + 5i_4 + 4\frac{d}{dt}\left(i_4 - i_3\right) = 0$$

The equation obtained for mesh i_2 is called a *differential equation* since it contains variables and their derivatives. The equations obtained for meshes i_3 and i_4 are called *integrodifferential equations* since they contain integrals as well as derivatives.

Writing the equations for a circuit, as in the preceding example, is not difficult. Finding the solution of equations like these, however, is another matter – it is no simple task. Thus, with the exception of some very simple circuits, we shall have to resort to additional concepts and techniques to be introduced later.

6.6 Duality

Duality is a concept which arises frequently in circuit analysis. To illustrate, consider the two circuits shown below:

Dual circuits

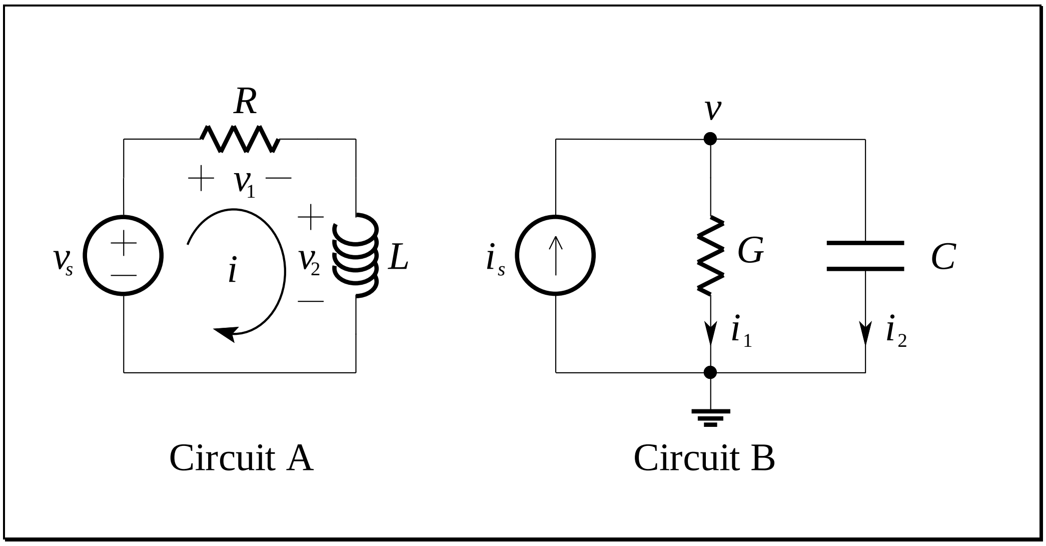


Figure 6.24

Using mesh analysis for circuit A and nodal analysis for circuit B, we get the following results:

Circuit A	**Circuit B**
$v_s = v_1 + v_2$	$i_s = i_1 + i_2$
$v_s = Ri + L\dfrac{di}{dt}$	$i_s = Gv + C\dfrac{dv}{dt}$

In other words, both circuits are described by the same equations:

$$\begin{aligned} x_s &= x_1 + x_2 \\ x_s &= a_1 y + a_2 \frac{dy}{dt} \end{aligned} \tag{6.28}$$

except that a variable that is a current in one circuit is a voltage in the other and vice versa. For these two circuits this result is not a coincidence, but rather is due to a concept known as *duality*, which has its roots in the subject of graph theory.

In actuality, circuits A and B are *dual circuits* since the following roles are interchanged:

Dual circuit relationships

$$\begin{aligned} v &\leftrightarrow i \\ R &\leftrightarrow G \\ L &\leftrightarrow C \\ \text{series} &\leftrightarrow \text{parallel} \end{aligned} \tag{6.29}$$

The usefulness of duality lies in the fact that once a circuit is analysed, its dual is in essence analysed also. Note that if circuit B is the dual of circuit A, then taking the dual of circuit B in essence results in circuit A. Not every circuit, however, has a dual. With the aid of graph theory it can be shown that a circuit has a dual if and only if it is a planar network.

The technique for obtaining the dual of a planar network is:

The technique for obtaining the dual of a planar circuit

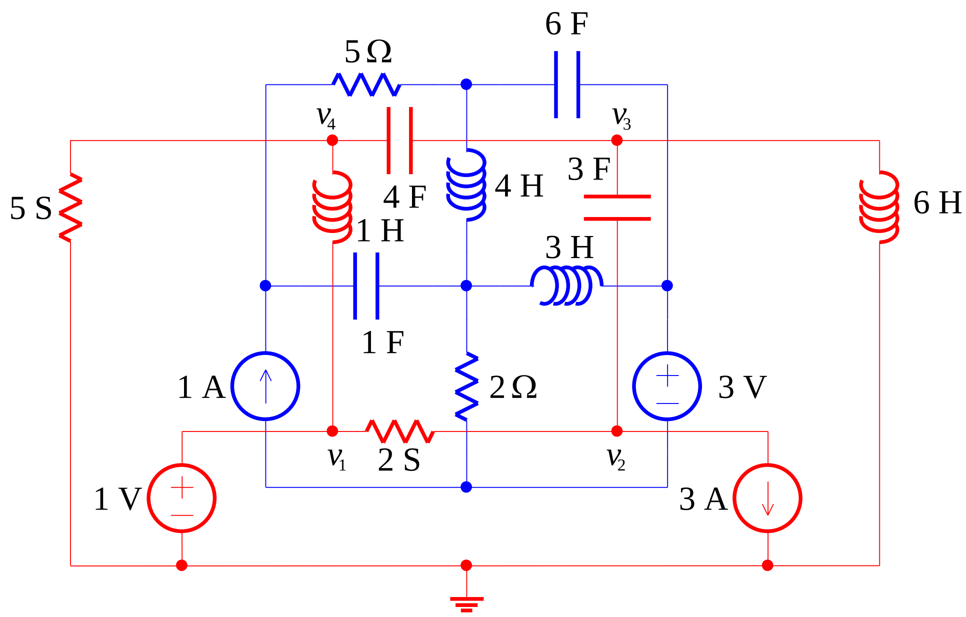
1. Inside of each mesh, and in the infinite region surrounding the circuit, place a node.

2. Suppose two of these nodes, say nodes *a* and *b*, are in adjacent meshes. Then there is at least one element in the boundary common to these two meshes. Place the dual of each common element between nodes *a* and *b*.

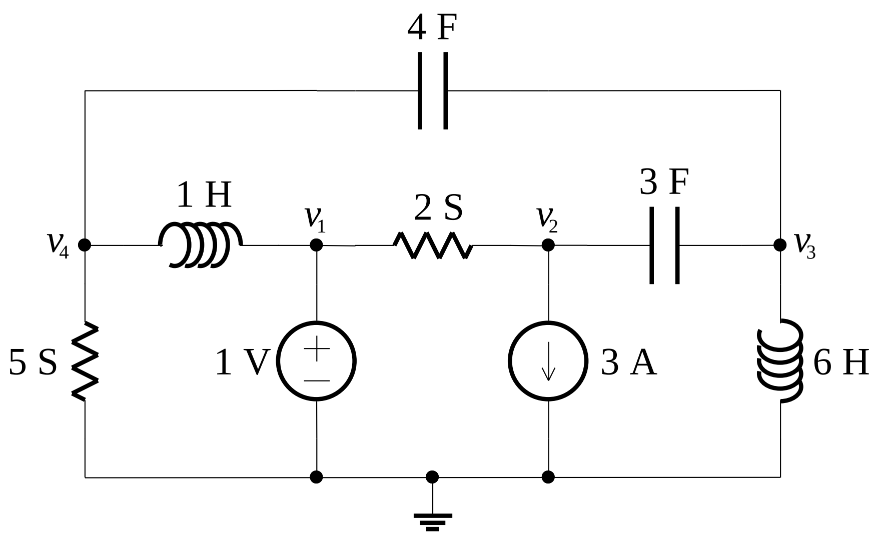
When using this procedure to obtain the dual circuit, in order to get the mesh equations of the original circuit to correspond to the node equations of the dual circuit, place clockwise mesh currents $i_1, i_2, \ldots, i_n$ in the finite regions. The corresponding nodes in the dual are labelled with the voltages $v_1, \mathrm{v}_2, \ldots, v_n$ respectively. The reference node of the dual circuit corresponds to the infinite region of the original circuit.

EXAMPLE 6.8 Analysis of a Dual Circuit

For the circuit given in the previous example (shown in blue), the dual is obtained as (shown in red):



The dual circuit is redrawn for simplicity:



By nodal analysis, at node v_1:

$$v_1 = 1$$

At node v_2:

$$2(v_2 - v_1) + 3\frac{d}{dt}(v_2 - v_3) = -3$$

At node v_3:

$$3\frac{d}{dt}(v_3 - v_2) + 4\frac{d}{dt}(v_3 - v_4) + \frac{1}{6}\int_{-\infty}^{t} v_3 dt = 0$$

At node v_4:

$$\frac{1}{1}\int_{-\infty}^{t}(v_4 - v_1)dt + 5v_4 + 4\frac{d}{dt}(v_4 - v_3) = 0$$

Note that these are the duals of the mesh equations that we obtained earlier for the original circuit.

6.7 Summary

- The *v*-*i* relationship for a capacitor is:

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

- A capacitor behaves as an open-circuit to direct current.
- The voltage across a capacitor cannot change instantaneously.
- The energy stored in a capacitor is:

$$w_C = \frac{1}{2}Cv^2$$

- The *v*-*i* relationship for an inductor is:

$$v = L\frac{di}{dt}$$

- An inductor behaves as a short-circuit to direct current.
- The current through an inductor cannot change instantaneously.
- The energy stored in an inductor is:

$$w_L = \frac{1}{2}Li^2$$

- Inductors in series and parallel are combined in the same way as are resistances. Capacitors in series and parallel are combined in the same way as are conductances.
- Writing node and mesh equations for circuits containing inductors and capacitors is done in the same manner as for resistive circuits. Obtaining solutions of equations in this form will be avoided, except for simple circuits.
- A planar circuit and its dual are in essence described by the same equations.

6.8 References

Bobrow, L.: *Elementary Linear Circuit Analysis*, Holt-Saunders, 1981.

Hayt, W. & Kemmerly, J.: *Engineering Circuit Analysis*, 3rd *Ed.*, McGraw-Hill, 1984.

Exercises

1.

There is a current $i = 5\sin 10t$ A through an inductance $L = 2\,\mathrm{H}$. What is the first instant of time after $t = 0$ when the power entering the inductor is exactly:

(a) 100 W

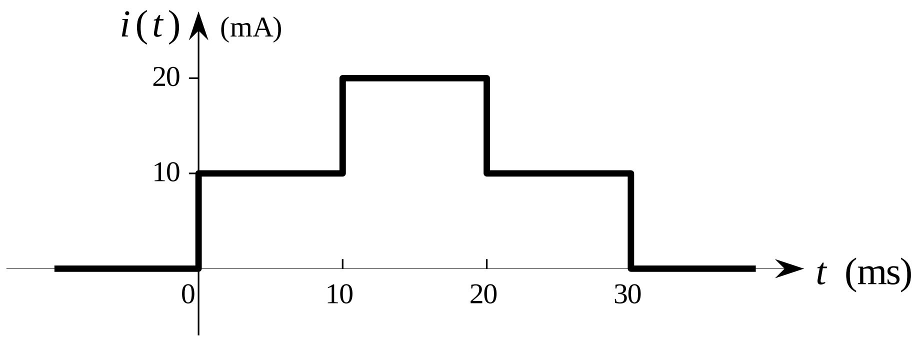
(b) -100 W

2.

The energy stored in a certain 10 mH inductor is zero at $t = 1\,\mathrm{ms}$ and increases linearly by 20 mJ each second thereafter. Find the inductor current and voltage for $t > 1\,\mathrm{ms}$ if neither is ever negative.

3.

A $25\,\mu\mathrm{F}$ capacitor having no voltage across it at $t = 0$ is subjected to the single pulse of current shown below.



Determine the voltage across, the power entering, and the energy stored in C at $t =$:

(a) 17 ms

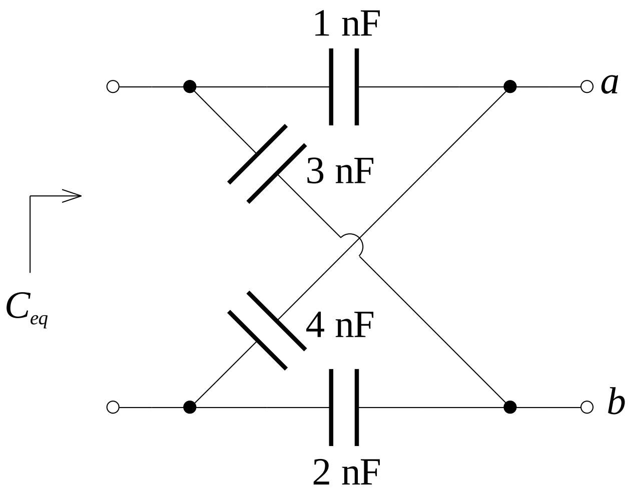
(b) 40 ms

4.

Find C_{eq} for the lattice network shown below if terminals *a* and *b* are:

(a) open-circuited as shown

(b) short-circuited



5.

The series combination of a $4\,\mu\text{F}$ and a $3\,\mu\text{F}$ capacitor is in series with the parallel combination of a $2\,\mu\text{F}$, a $1\,\mu\text{F}$ and a $C\,\mu\text{F}$ capacitor.

(a) What is the maximum possible value for the equivalent capacitance of the five capacitors?

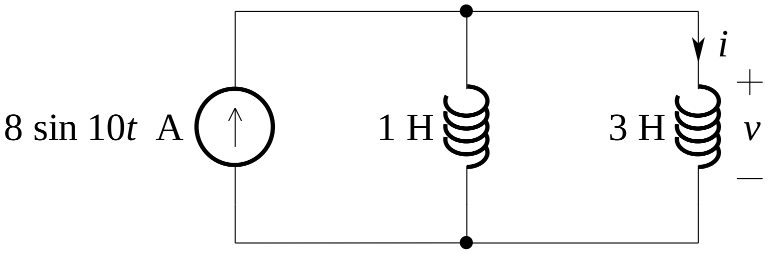
(b) Repeat for the minimum value.

(c) Find C if $C_{eq} = 1.5\,\mu\text{F}$.

6.44

6.

At $t=0$, $i=5\text{ A}$ in the circuit shown below:



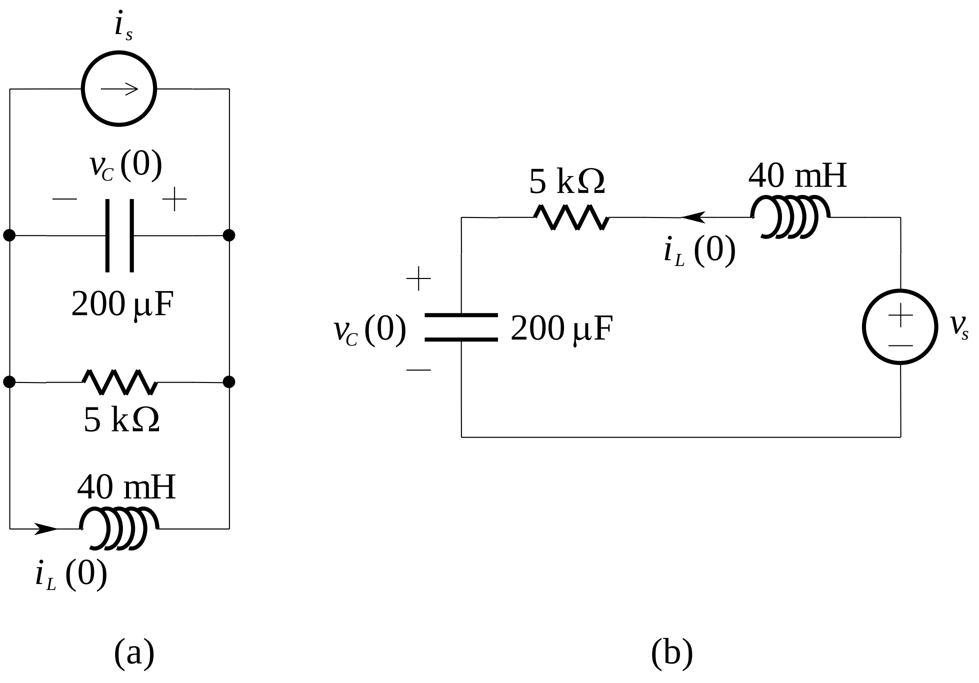
(a) Find $v(t)$ for all t.

(b) Find $i(t)$ for $t \geq 0$.

7.

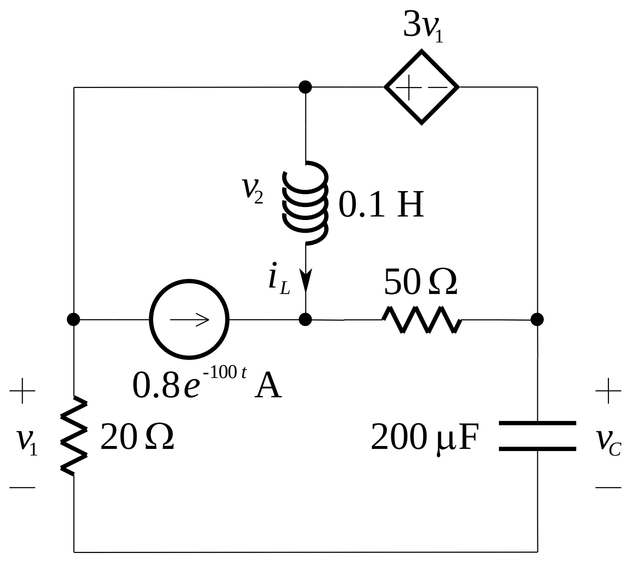
(a) Write the single nodal equation for the circuit (a) below:

(b) Write the single mesh equation for the circuit (b) below:



8.

Using a reference node at the bottom of the circuit, assign node-to-reference voltages in the circuit below and write nodal equations. Let $i_L(0) = 0.5\text{ A}$ and $v_C(0) = 12\text{ V}$.



7 Diodes and Basic Diode Circuits

Contents

7.2

Introduction

Nonlinear circuits play a major role in modern electronics. Examples include signal generators, communication transmitters and receivers, DC power supplies, and digital circuits.

To begin a study of nonlinear circuits, we need to examine the most fundamental two-terminal nonlinear device: the diode. The semiconductor junction diode, or *p-n* junction diode, also forms the basis for other semiconductor devices, such as the transistor.

The terminal characteristics of the diode will be presented, rather than the underlying solid-state physics, so that we can focus on providing techniques for the analysis of diode circuits. There are three types of diode circuit analysis technique – graphical, numerical and use of a linear model. *Graphical analysis* of diode circuits is done using graphs of the diode's terminal characteristic and the connected circuit. *Numerical analysis* can be performed with the nonlinear equations of the diode with a technique known as iteration. Lastly, diodes can be replaced with linear circuit models (of varying complexity), under assumed diode operating conditions, so that we revert to linear circuit analysis. Each analysis technique has its advantages and disadvantages, so it is important to choose the most appropriate technique for a given circuit.

Basic applications of the diode will be introduced, with circuits such as the rectifier, and the limiter. Lastly, with the use of so-called breakdown diodes, we can design circuits that act as voltage regulators – i.e. circuits that provide a steady output voltage when subjected to a wide range of input voltages and output load currents.

7.1 The Silicon Junction Diode

A typical silicon *p-n* junction diode has the following *i-v* characteristic:

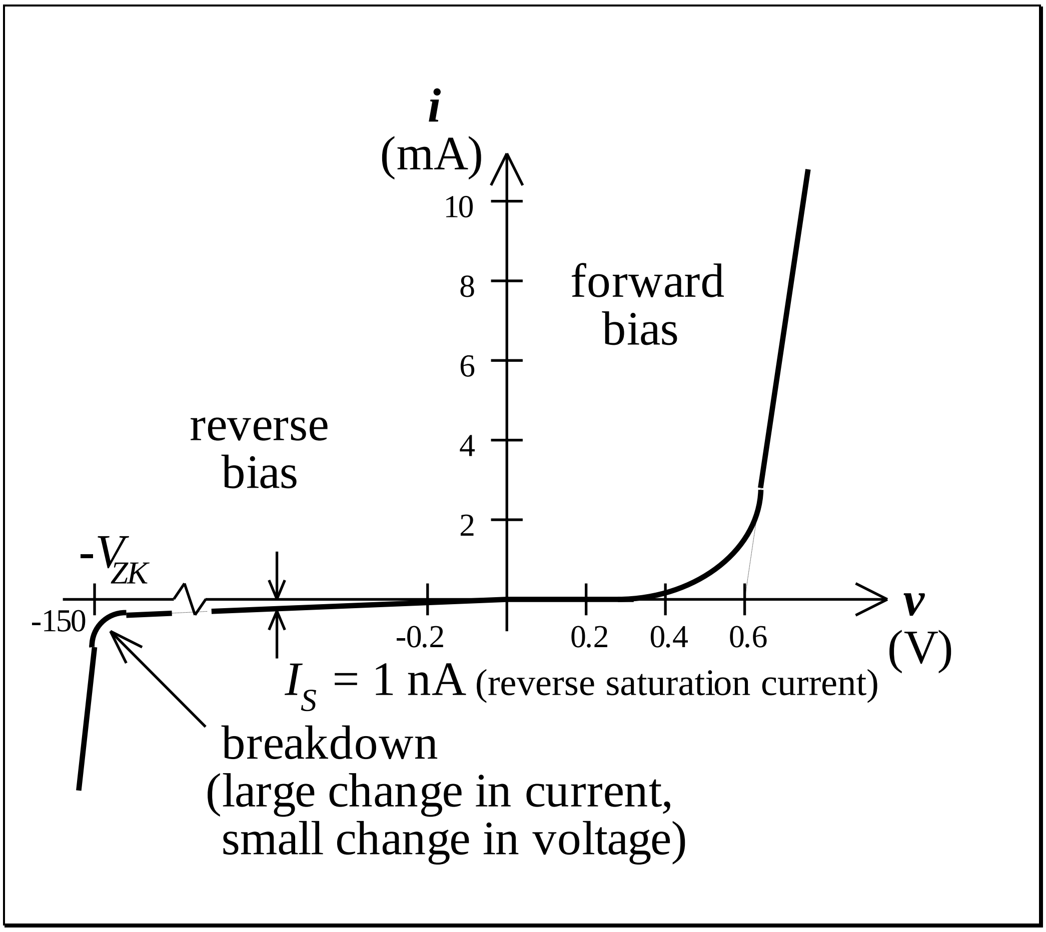


Figure 7.1

The diode is clearly a *nonlinear* element – its characteristic is not a straight line through the origin! The *i-v* characteristic can be divided up into three distinct regions:

1. The forward-bias region, determined by $v > 0$
2. The reverse-bias region, determined by $v < 0$
3. The breakdown region, determined by $v < -V_{ZK}$

7.1.1 The Forward-Bias Region

The forward-bias region of operation is entered when the terminal voltage v is positive. In the forward region the i-v relationship is closely approximated by the Shockley equation, which can be derived from semiconductor physics:

The Shockley equation

$$i = I_S\left(e^{v/nV_T} - 1\right) \tag{7.1}$$

When forward biased, the diode conducts

There is not much increase in current until the "internal barrier voltage" is overcome (approximately 0.6 V in silicon). Then large conduction results.

Saturation current defined

The current I_S is called the *saturation current* and is a constant for a given diode at a given temperature.

Emission coefficient defined

The constant n is called the *emission coefficient*, and has a value between 1 and 2, depending on the material and the physical structure of the diode.

The voltage V_T is a constant called the *thermal voltage*, given by:

Thermal voltage defined

$$V_T = \frac{kT}{q} \tag{7.2}$$

where:

$$\begin{aligned} k &= \text{Boltzmann's constant} \\ &= 1.381 \times 10^{-23}\ \text{JK}^{-1} \end{aligned} \tag{7.3}$$

$$T = \text{temperature in degrees Kelvin} \tag{7.4}$$

$$\begin{aligned} q &= \text{magnitude of electron charge} \\ &= 1.602 \times 10^{-19}\ \text{C} \end{aligned} \tag{7.5}$$

The thermal voltage is approximately equal to 26 mV at 300 K (a temperature that is close to "room temperature" which is commonly used in device simulation software).

For appreciable current in the forward direction ($i >> I_S$), the Shockley equation can be approximated by:

$$i \approx I_S e^{v/nV_T} \tag{7.6}$$

This equation is usually "good enough" for rough hand calculations when we know that the current is appreciable.

From the characteristic we note that the current is negligibly small for v smaller than about 0.5 V (for silicon). This value is usually referred to as the *cut-in voltage*. This apparent threshold in the characteristic is simply a consequence of the exponential relationship.

Another consequence of the exponential relationship is the rapid increase of current for small changes in voltage. Thus for a "fully conducting" diode the voltage drop lies in a narrow range, approximately 0.6 to 0.8 V for silicon. We will see later that this gives rise to a simple model for the diode where it is assumed that a conducting diode has approximately a 0.7 V drop across it (again, for silicon).

The Shockley equation can be rearranged to give the voltage in terms of the current:

$$v = nV_T \ln\left(\frac{i}{I_S} + 1\right) \tag{7.7}$$

This logarithmic form is used in the numerical analysis of diode circuits.

7.1.2 The Reverse-Bias Region

When reverse biased, the diode does not conduct

The reverse-bias region of operation is entered when the diode voltage v is made negative. The Shockley equation predicts that if v is negative and a few times large than V_T in magnitude, the exponential term becomes negligibly small compared to unity and the diode current becomes:

$$i = -I_S \tag{7.8}$$

That is, the current in the reverse direction is constant and equal to I_S. This is the reason behind the term *saturation current*. However, real diodes exhibit reverse currents that, although quite small, are much larger than I_S.

7.1.3 The Breakdown Region

Breakdown occurs eventually for a large enough reverse bias

The breakdown region is entered when the magnitude of the reverse voltage exceeds a threshold value specific to the particular diode and called the *breakdown voltage*. This is the voltage at the "knee" of the *i*-*v* curve and is denoted by V_{ZK}, where the subscript Z stands for Zener (to be explained shortly) and K denotes knee.

Breakdown is not a destructive process unless the device cannot dissipate the heat produced in the breakdown process. Breakdown is actually exploited in certain types of diodes (e.g. the Zener diode) because of the near vertical characteristic in this region.

7.1.4 Diode Symbol

The circuit symbol for the diode is shown below, with the direction of current and polarity of voltage that corresponds to the characteristic:

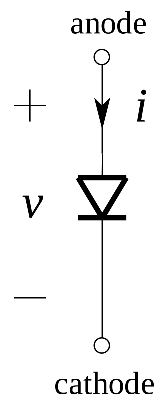


Figure 7.2

7.2 Breakdown Diodes

Some diodes are designed to operate in the breakdown region. It is usually a sharper transition than the forward bias characteristic, and the breakdown voltage is higher than the forward conduction voltage. There are two main types of breakdown.

Some diodes are designed to operate in the breakdown region

7.2.1 Zener Breakdown

The electric field in the depletion layer of a *p*-*n* junction becomes so large that it rips covalent bonds apart, generating holes and electrons. The electrons will be accelerated into the *n*-type material and the holes into the *p*-type material. This constitutes a reverse current. Once the breakdown starts, large numbers of carriers can be produced with negligible increase in the junction voltage.

Zener breakdown is caused by a large internal electric field

7.2.2 Avalanche Breakdown

If the minority carriers are swept across the depletion region of a *p*-*n* junction too fast, they can break the covalent bonds of atoms that they hit. New electron-hole pairs are generated, which may acquire sufficient energy to repeat the process. An avalanche starts.

Avalanche breakdown is caused by electrons with a large kinetic energy

7.3 Other Types of Diode

The silicon junction diode is not the only type of diode. A variety of diode constructions exist, with many of them essential to the modern world, such as the LED.

7.3.1 The Photodiode

A photodiode is controlled by light

In a photodiode, the *p*-*n* junction is very close to the surface of the crystal. The Ohmic contact with the surface material is so thin, it is transparent to light. Incident light (photons) can generate electron-hole pairs in the depletion layer (a process called photoionisation).

7.3.2 The Light Emitting Diode (LED)

An LED emits photons when forward biased

When a light-emitting diode is forward biased, electrons are able to recombine with holes within the device, releasing energy in the form of light (photons). The color of the light corresponds to the energy of the photons emitted, which is determined by the "energy gap" of the semiconductor. LEDs present many advantages over incandescent and compact fluorescent light sources including lower energy consumption, longer lifetime, improved robustness, smaller size, faster switching, and greater durability and reliability. At the moment LEDs powerful enough for room lighting are relatively expensive and require more precise current and heat management than compact fluorescent lamp sources of comparable output.

LEDs are used in diverse applications. The compact size of LEDs has allowed new text and video displays and sensors to be developed, while their high switching rates are useful in advanced communications technology. Infrared LEDs are also used in the remote control units of many commercial products including televisions, DVD players, and other domestic appliances.

7.3.3 The Schottky Diode

A Schottky diode is the result of a metal-semiconductor junction. The Schottky diode is a much faster device than the general purpose silicon diode. There are three main reasons for this: 1) the junction used is a metal-semiconductor junction, which has less capacitance than a *p-n* junction, 2) often the semiconductor used is gallium arsenide (GaAs) because electron mobility is much higher, and 3) the device size is made extremely small. The result is a device that finds applications in high speed switching.

A Schottky diode is a metal-semiconductor junction

7.3.4 The Varactor Diode

This device is also known as a variable capacitance diode. It has a relatively large capacitance, brought about by a large junction area and narrow depletion region. The applied reverse voltage changes the length of the depletion region, which changes the capacitance. Thus, the device can be used in applications that rely on a voltage controlled capacitance. Applications include electronic tuning circuits used in communication circuits, and electronic filters.

7.4 Analysis Techniques

Since the diode's characteristic is nonlinear, we can't apply linear circuit analysis techniques to circuits containing diodes. We therefore have to resort to other analysis methods: graphical, numerical and linear modelling.

7.4.1 Graphical Analysis

Circuits with a single nonlinear element can always be modelled using the Thévenin equivalent of the linear part:

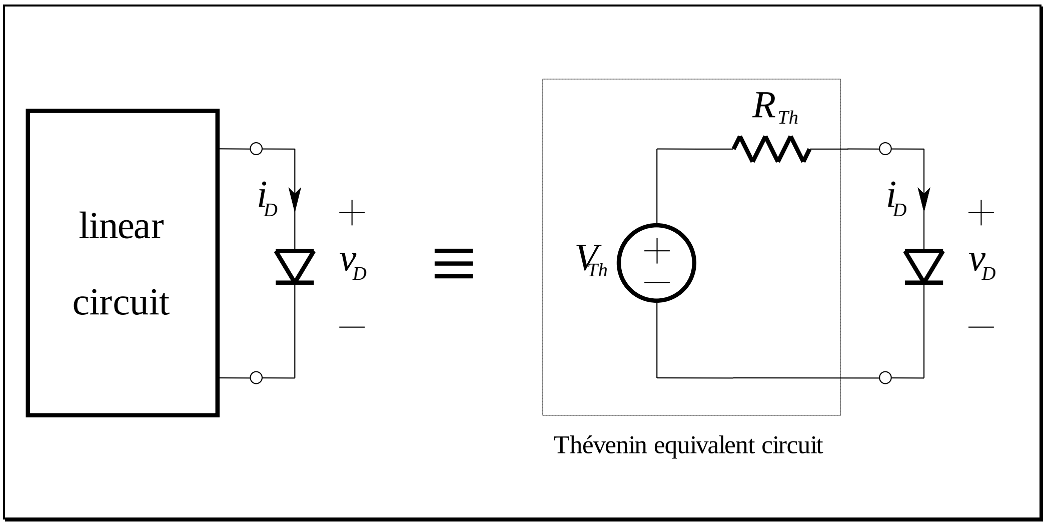


Figure 7.3

KVL around the loop gives:

$$v_D = V_{Th} - R_{Th} i_D \tag{7.9}$$

which, when rearranged to make i_D the subject, gives:

$$i_D = -\frac{1}{R_{Th}}\left(v_D - V_{Th}\right) \tag{7.10}$$

The "load line" is derived using linear circuit theory

When graphed, we call it the *load line*. It was derived from KVL, and so it is always valid. The load line gives a relationship between i_D and v_D that is determined purely by the external circuit. The diode's characteristic gives a relationship between i_D and v_D that is determined purely by the geometry and physics of the diode.

Since both the load line and the characteristic are to be satisfied, the only place this is possible is the point at which they meet. This point is called the quiescent point, or Q point for short:

The "load line" and device characteristic intersect at the Q point

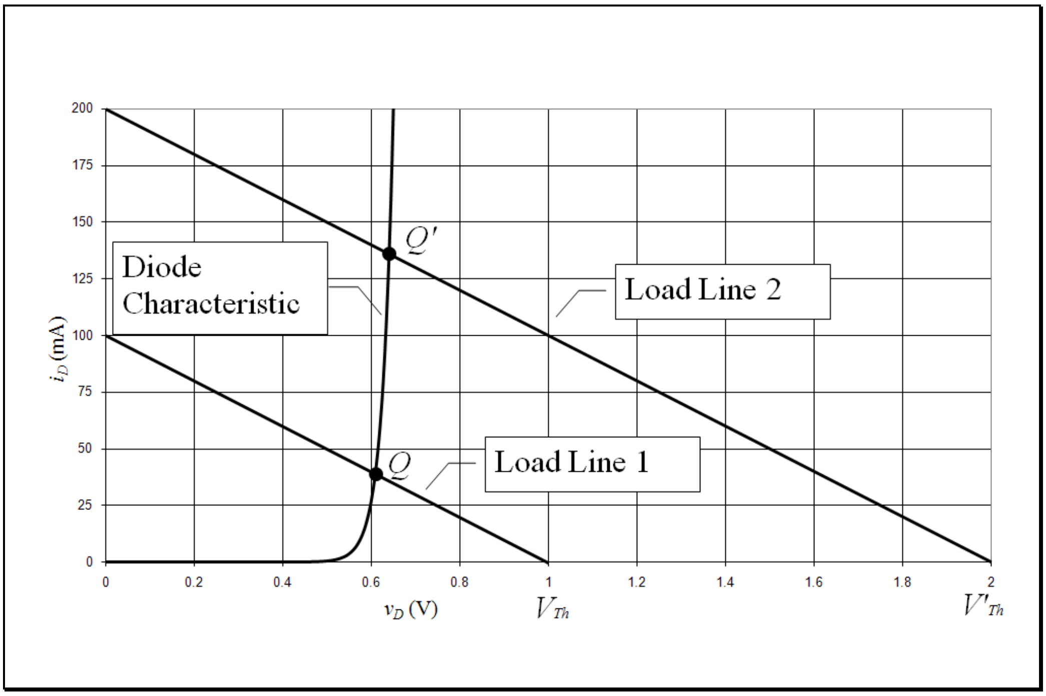


Figure 7.4 – Graphical Analysis Using a Load Line

If the Thévenin voltage changes to V'_{Th}, then the operating point moves to Q' (the DC load line is shifted to the right).

The two end points of the load line are easily determined to enable quick graphing. The two axis intercepts are:

$$v_D = 0, \quad i_D = \frac{V_{Th}}{R_{Th}} \tag{7.11}$$

and:

$$v_D = V_{Th}, \quad i_D = 0 \tag{7.12}$$

Alternatively, we can graph the load line using one known point and the fact that the slope is equal to $-\frac{1}{R_{Th}}$.

7.4.2 Numerical Analysis

Since, in the preceding analysis, we have two equations (the load line and the diode characteristic) and two unknowns, it is tempting to try and solve them simultaneously. If we substitute the voltage from the Shockley equation:

$$v_D = nV_T \ln\left(\frac{i_D}{I_S} + 1\right) \tag{7.13}$$

into the load line equation:

$$i_D = -\frac{1}{R_{Th}}\left(v_D - V_{Th}\right) \tag{7.14}$$

we get:

$$i_D = -\frac{1}{R_{Th}}\left(nV_T \ln\left(\frac{i_D}{I_S} + 1\right) - V_{Th}\right) \tag{7.15}$$

This equation is a *transcendental equation*, and its solution cannot be expressed in term of elementary functions (try it!). With a sufficiently advanced calculator (or mathematical software), we can use a special function called the Lambert W function to solve it, but for engineering purposes, there are usually simpler methods of solution.

We can solve transcendental equations graphically (as shown in the preceding section) but we can also solve them numerically using a technique known as *iteration* – which is suitable for computer simulations.

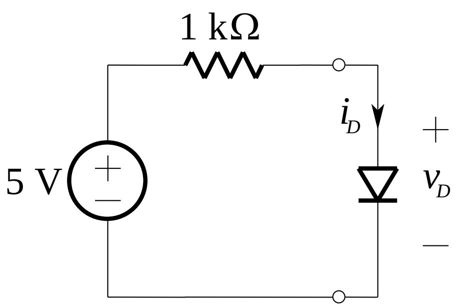
We begin with an initial "guess" for the diode current, labelled $i_{D,0}$, and then compute:

$$\begin{aligned} i_{D,1} &= -\frac{1}{R_{Th}}\left(nV_T \ln\left(\frac{i_{D,0}}{I_S}+1\right)-V_{Th}\right) \\ i_{D,2} &= -\frac{1}{R_{Th}}\left(nV_T \ln\left(\frac{i_{D,1}}{I_S}+1\right)-V_{Th}\right) \\ i_{D,3} &= \cdots \end{aligned} \tag{7.16}$$

and so on until we get "convergence", i.e. $i_{D,k+1} \approx i_{D,k}$. Convergence is not always guaranteed, and depends on the initial "guess". Computer simulation software uses several clever methods to aid in numerical convergence.

EXAMPLE 7.1 Numerical Analysis of a Circuit with a Real Diode

The following circuit contains a diode, with $n=1$ and $I_S = 2.030\ \text{fA}$. We wish to find the diode's operating point, or *Q* point. Assume that $V_T = 26\ \text{mV}$.



Starting with $i_{D,0} = 1\ \text{mA}$ we have:

$$\begin{aligned} i_{D,1} &= -\frac{1}{10^3}\left(0.026\ln\left(\frac{0.001}{2.03\times10^{-15}}+1\right)-5\right) = 4.3\ \text{mA} \\ i_{D,2} &= -\frac{1}{10^3}\left(0.026\ln\left(\frac{0.0043}{2.03\times10^{-15}}+1\right)-5\right) = 4.262\ \text{mA} \end{aligned}$$

Since the second value is very close to the value obtained after the first iteration, no further iterations are necessary, and the solution is $i_D = 4.262\ \text{mA}$ and $v_D = 0.7379\ \text{V}$.

7.5 Diode Models

Why we model the diode

The curve describing the diode's terminal characteristics is non-linear. How can we use this curve to do circuit analysis? We only know how to analyze linear circuits. There is therefore a need for a linear circuit model of the diode.

The concept of modelling

When we model something, we transform it into something else – usually something simpler – which is more amenable to analysis and design using mathematical equations. Modelling mostly involves assumptions and simplifications, and the only requirement of a model is for it to "work" reasonably well. By "work" we mean that it agrees with experimental results to some degree of accuracy.

Models are sometimes only valid under certain operating conditions, as we shall see when modelling the diode.

7.5.1 The Ideal Diode Model

As a first approximation, we can model the diode as an ideal switch:

The diode as an ideal (controlled) switch

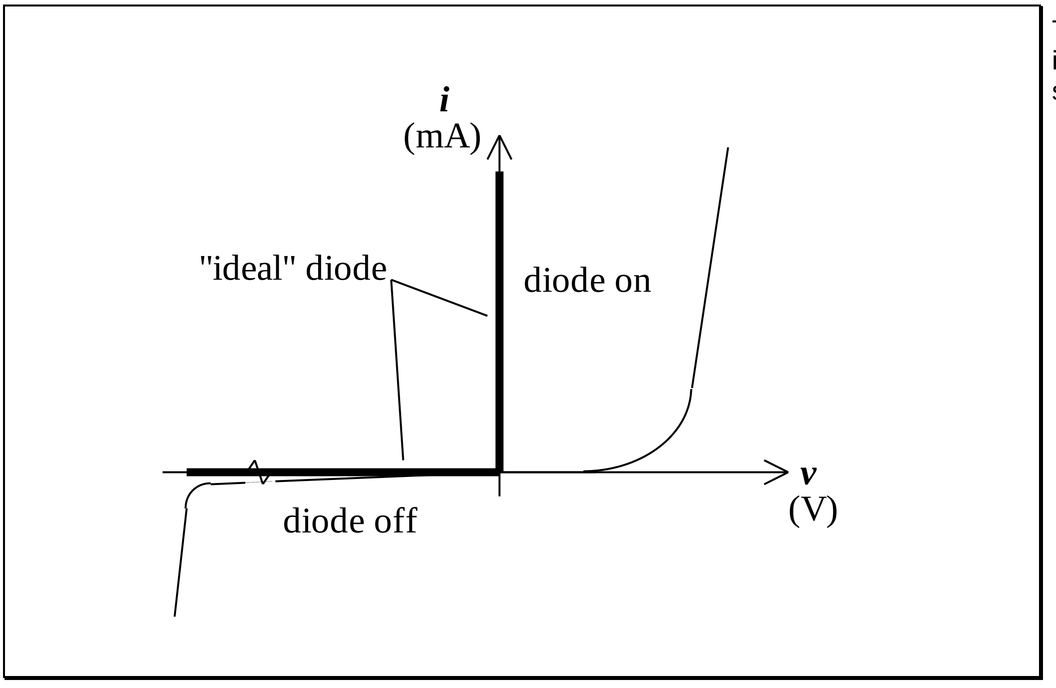


Figure 7.5 – The Ideal Diode Model

The characteristic in this case is approximated by two straight lines – the vertical representing the "on" state of the diode, and the horizontal representing the "off" state. To determine which of these states the diode is in, we have to determine the conditions imposed upon the diode by an external circuit. This model of the diode is used sometimes where a quick "feel" for a diode circuit is needed. The above model can be represented symbolically as:

The ideal diode model

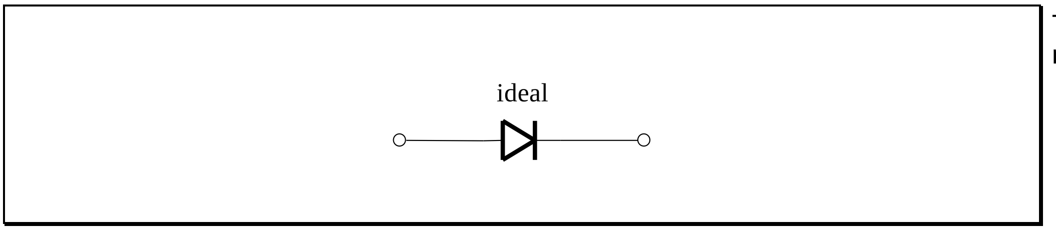
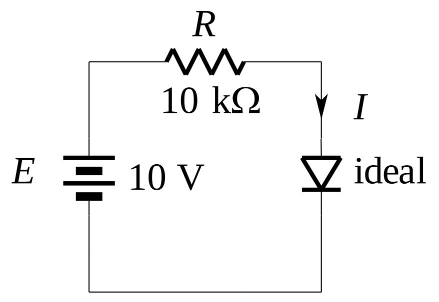


Figure 7.6

EXAMPLE 7.2 Analysis Using the Ideal Diode

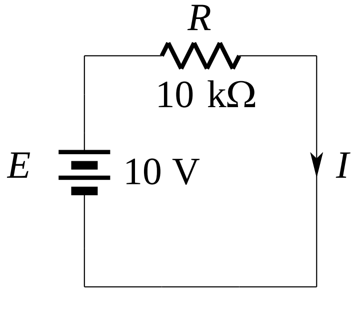
(i) Find the current, *I*, in the circuit shown below, using the ideal diode model.

(ii) If the battery is reversed, what does the current become?



(i) Firstly, we must determine whether the diode is forward biased or reverse biased. In this circuit, the positive side of the battery is connected (via the resistor) to the anode. Therefore, the anode is positive with respect to the cathode, and the diode is *forward biased*. In order to use the ideal diode model, the diode is simply replaced by the ideal diode model (forward bias model), and the simplified circuit is analysed accordingly.

The *equivalent circuit* is shown below, where the diode has now been replaced by a short circuit.



$$I = \frac{E}{R} = \frac{10}{10\text{ k}} = 1\text{ mA}$$

Ohm's Law may be used to determine the current, *I*, as shown:

(ii) If the battery is reversed, the diode becomes *reverse biased*. In this case, the diode is replaced by the ideal diode model for reverse bias. Since the reverse biased ideal diode model is simply an *open circuit*, there is no current, i.e. $I = 0$.

7.5.2 The Constant Voltage Drop Model

A better model is to approximate the forward bias region with a vertical line that passes through some voltage called e_{fd}:

A model that takes into account the forward voltage drop

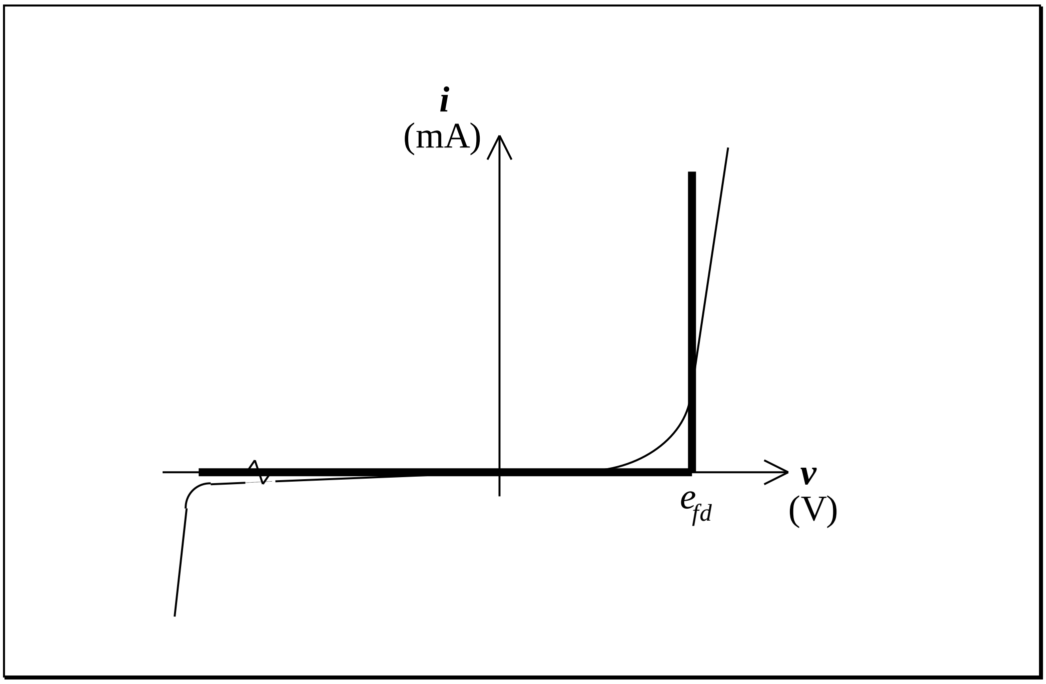


Figure 7.7 – The Constant Voltage Drop Diode Model

This "constant voltage drop" model is better than the ideal model because it more closely approximates the characteristic in the forward bias region. The "voltage drop" is a model for the barrier voltage in the *p*-*n* junction. The model of the diode in this case is:

The constant voltage drop diode model

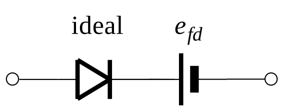


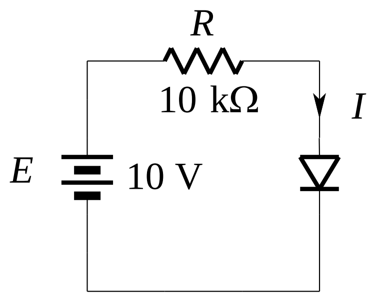
Figure 7.8

This model is the one of the simplest and most widely used. It is based on the observation that a forward-conducting diode has a voltage drop that varies in a relatively narrow range, say 0.6 V to 0.8 V. The model assumes this voltage to be constant, say, 0.7 V. The constant voltage drop model is the one most frequently employed in the initial phases of analysis and design.

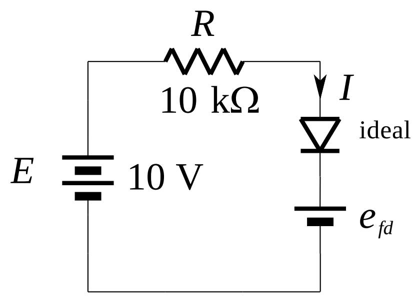
EXAMPLE 7.3 Analysis Using the Constant Voltage Drop Model

(i) Find the current, I, in the circuit shown below, using the constant voltage drop model of the diode (assume $e_{fd} = 0.7\ \text{V}$).

(ii) If the battery is reversed, what does the current become?



(i) Analysis proceeds in exactly the same manner as the previous example, except that the constant voltage drop diode model is used instead. The diode is again forward biased, and so the equivalent circuit is shown below, along with the calculation for I.



$$I = \frac{E - e_{fd}}{R} = \frac{10 - 0.7}{10\text{k}} = 0.93\ \text{mA}$$

(ii) If the battery is reversed, the diode becomes *reverse biased*, resulting in no current, i.e. $I = 0$.

7.5.3 The Piece-Wise Linear Model

An even better approximation to the diode characteristic is called a "piece-wise" linear model. It is made up of pieces, where each piece is a straight line:

A model that approximates the characteristic by using straight lines

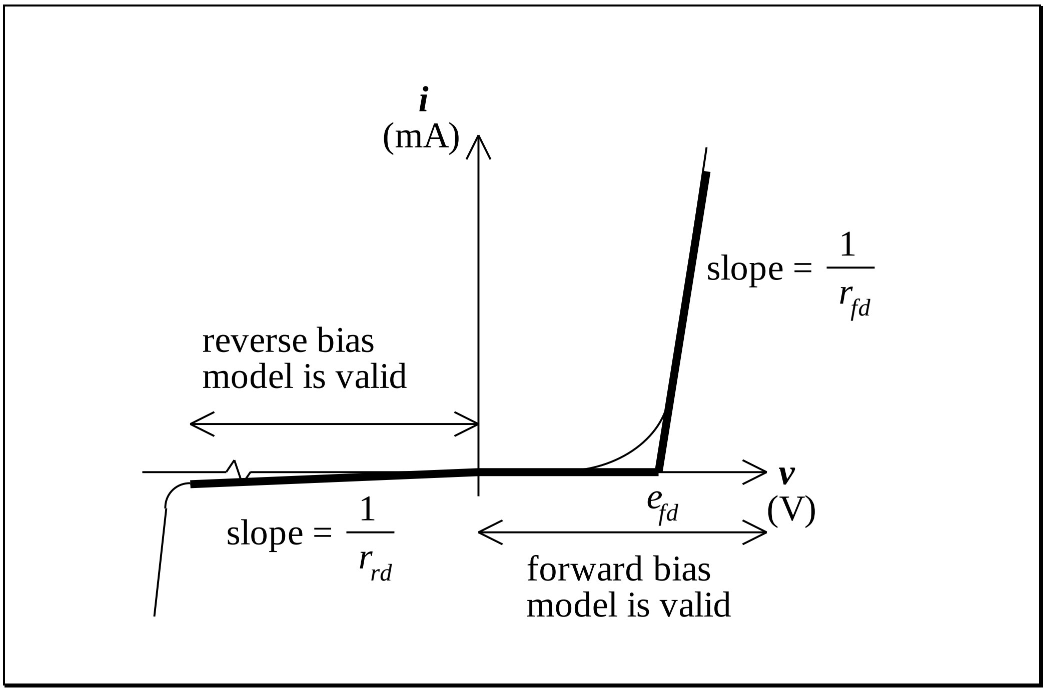


Figure 7.9 – The Piece-Wise Linear Diode Model

For each section, we use a different diode model (one for the forward bias region and one for the reverse bias region):

The piece-wise linear diode model

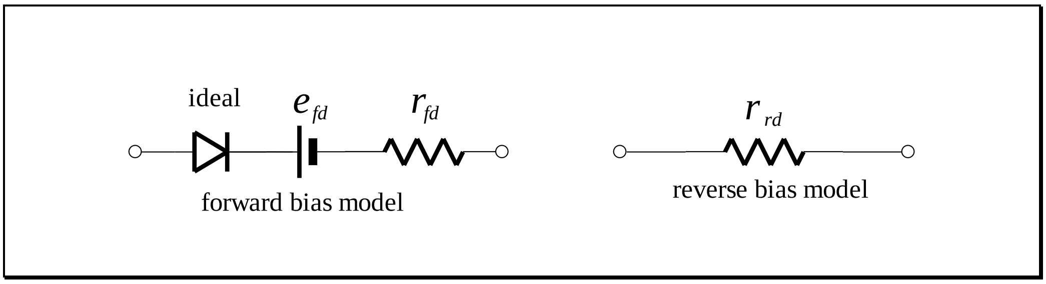


Figure 7.10

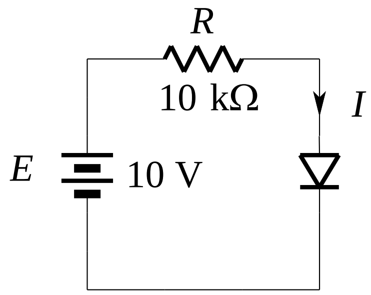
Typical values for the resistances are $r_{fd} = 5\,\Omega$ and $r_{rd} > 10^9\,\Omega$.

Notice how we have done away with the ideal diode part of the model for when the diode is reverse biased. This is because there is a separate equivalent circuit for the forward bias and reverse bias regions, so an ideal diode is not necessary (we apply one equivalent circuit or the other).

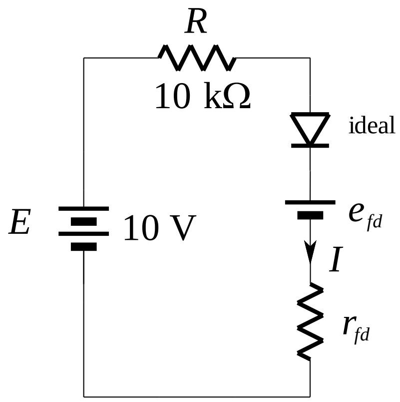
EXAMPLE 7.4 Analysis Using the Piece-Wise Linear Model

(i) Find the current, *I*, in the circuit shown below, using the piece-wise linear model of the diode (assume $e_{fd} = 0.7\text{ V}$, $r_{fd} = 5\,\Omega$ and $r_{rd} = 10^9\,\Omega$).

(ii) If the battery is reversed, what does the current become?



(iii) Analysis proceeds in exactly the same manner as the previous example, except that the piece-wise linear diode model is used instead. The diode is again forward biased, and so the equivalent circuit is shown below, along with the calculation for *I*.



$$I = \frac{E - e_{fd}}{R + r_{fd}} = \frac{10 - 0.7}{10\text{k} + 5} = 0.9295\text{ mA}$$

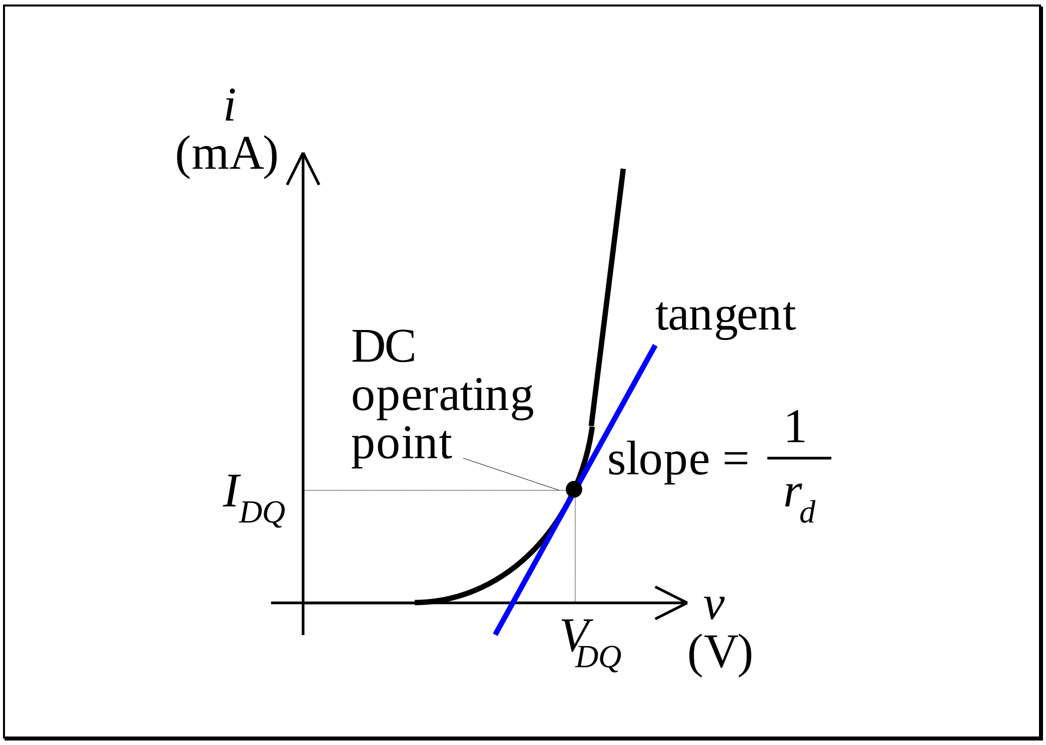
(iv) If the battery is reversed, the diode becomes *reverse biased*, and the diode is replaced by the piece-wise linear model for the reverse region, which is just the resistance r_{rd}. Since $r_{rd} = 10^9\,\Omega$, the reverse current is:

$$I = \frac{E}{R + r_{rd}} = \frac{10}{10^9 + 10^4} \approx 10\text{ nA}$$

which is negligible, i.e. $I \approx 0$.

7.5.4 The Small Signal Model

Suppose we know the diode DC voltage and current exactly. We may want to examine the behaviour of a circuit when we apply a signal (a small AC voltage) to it. In this case we are interested in small excursions of the voltage and current about the “DC operating point” of the diode. The best model in this instance is the following (the forward bias region is used as an example, but the method applies anywhere):



A model that approximates the characteristic by a tangent at a DC operating point

Figure 7.11 – Diode Model as the Tangent to a DC Operating Point

We approximate the curved characteristic by the *tangent* that passes through the operating point. It is only valid for small variations in voltage or current. This is called the *small signal approximation*. A straight line is a good approximation to a curve if we don’t venture too far.

A first look at the small signal approximation

Thus, for a small change Δv_D in the diode voltage, we get a small change in the diode current Δi_D, which can be approximated by the change in current we would get by following the tangent:

$$\Delta i_D \approx \frac{di_D}{dv_D} \Delta v_D \tag{7.17}$$

From the Shockley equation in the forward-biased region, we have:

$$\frac{di_D}{dv_D} = \frac{I_S}{nV_T} e^{v/nV_T} \tag{7.18}$$

We define the *dynamic resistance* of the diode as:

Dynamic resistance

$$r_d = \left. \frac{dv_D}{di_D} \right|_{i_D = I_{DQ}} \approx \frac{nV_T}{I_{DQ}} \tag{7.19}$$

Therefore, using Eq. (7.17) and Eq. (7.19), a small AC signal v_d superimposed upon the DC operating point will result in a small AC current i_d given by:

$$i_d \approx \frac{1}{r_d} v_d \tag{7.20}$$

Thus, for small AC signals, the diode can be modelled as a resistance, r_d.

Now consider the case where the Thévenin equivalent circuit contains both a DC source and a small signal AC source:

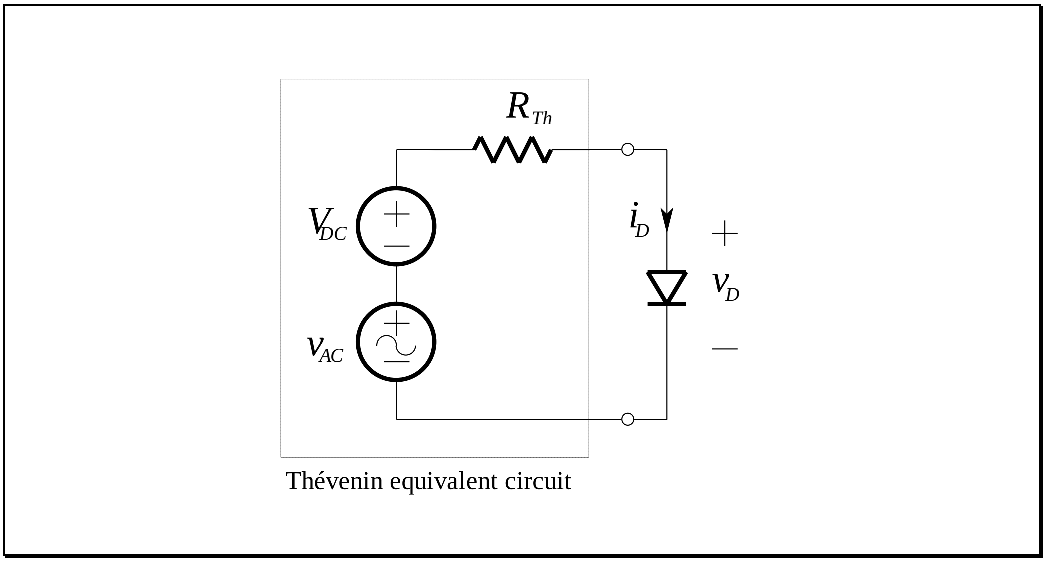


Figure 7.12

In this case we analyse two separate circuits. The first circuit contains the DC source which is used to establish the DC operating point. The diode in this circuit has the standard *i*-*v* characteristic and the DC operating point can be obtained using graphical methods (load line) or numerical methods (iteration).

The second circuit contains the small AC source and uses the dynamic resistance model of the diode. The small AC currents and voltages in this circuit can be superimposed upon the diode's DC current and voltage obtained from the first circuit. Thus we analyse:

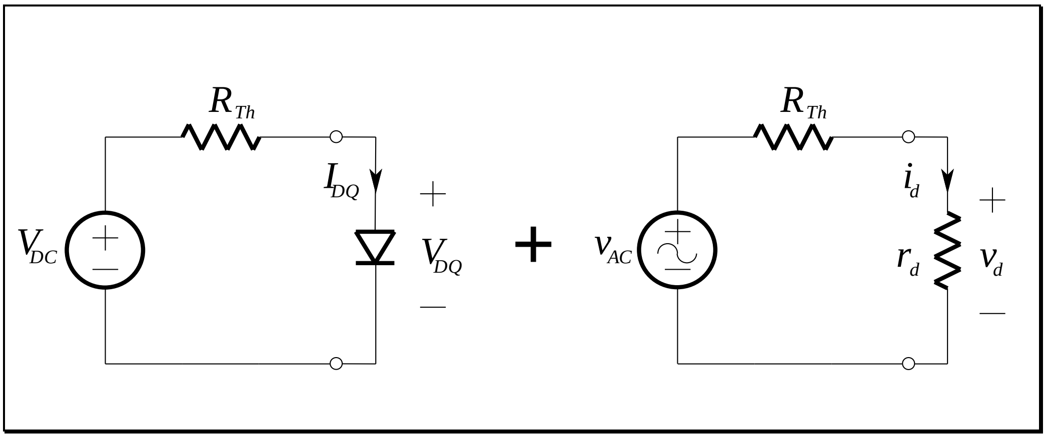


Figure 7.13

Development of the diode small signal model

The analysis can also be illustrated graphically:

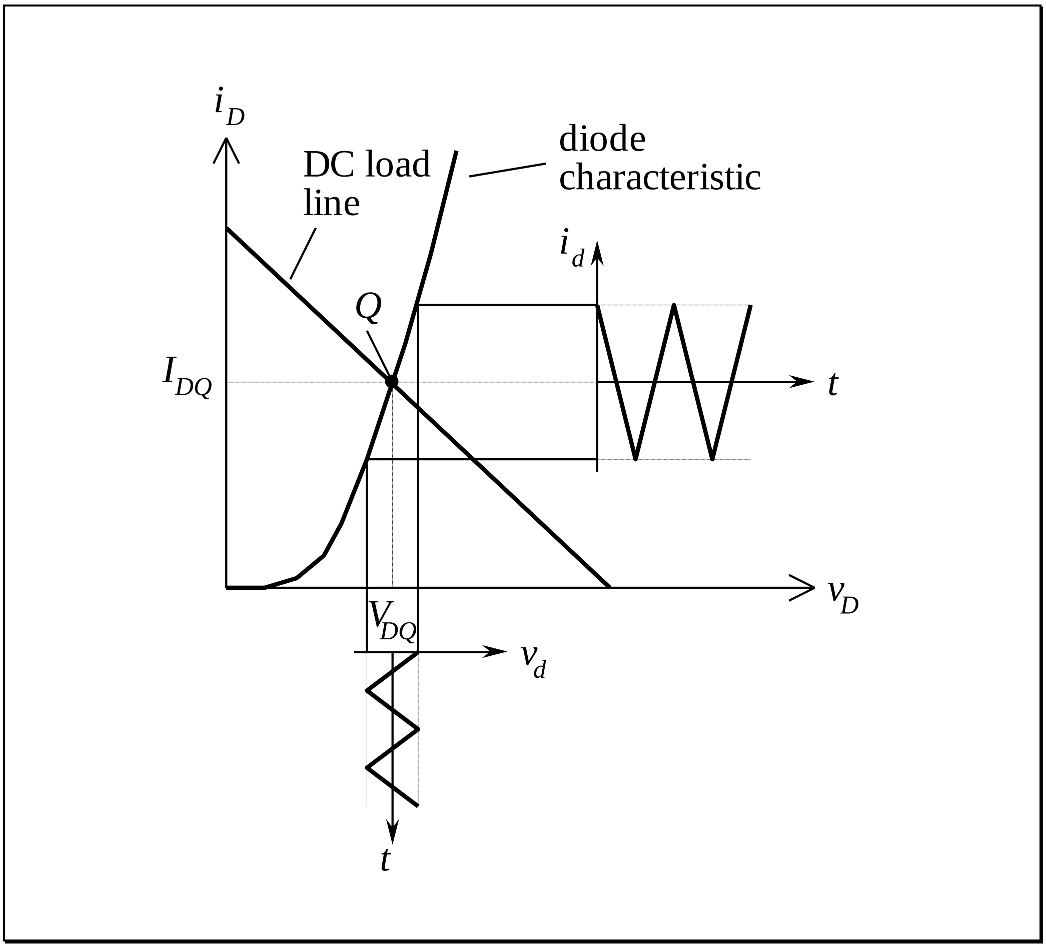


Figure 7.14

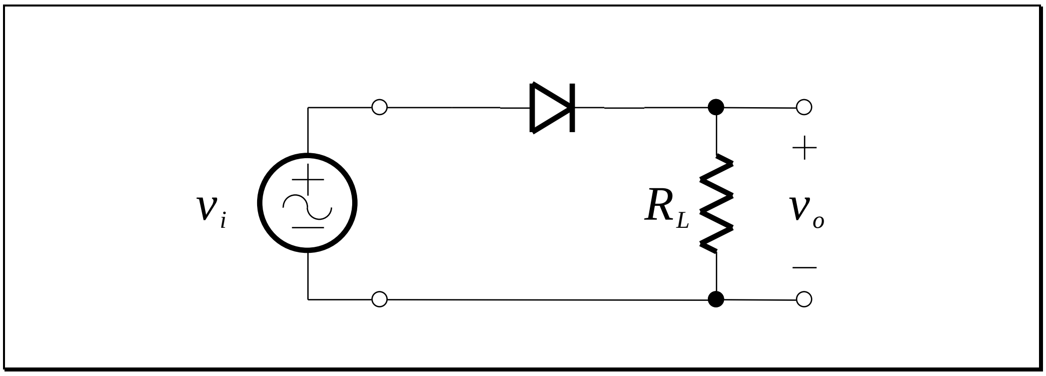
The figure above illustrates the two separate analysis steps. First, the DC operating point, or Q point for short, is found using graphical or numerical techniques. Then the application of an AC voltage on top of the original DC voltage results in a change in current given by the projection of the applied voltage onto the diode *i*-*v* characteristic. If the AC voltage is a “small signal”, then the diode characteristic can be replaced by a straight line. In this context, a “small signal” is defined as one for which the tangent to the curve is a good approximation to the curve, resulting in a linear relationship between the voltage and current.

7.6 Basic Diode Circuits

There are many diode circuits that are used in a wide variety of applications. The most important are summarised below.

7.6.1 Half-Wave Rectifier

A rectifier is a circuit that converts a bipolar (AC) signal into a unidirectional one. The figure below shows a diode rectifier fed by a sine-wave voltage source v_i.



The half-wave rectifier

Figure 7.15

A sinusoidal input waveform, and the resulting output waveform v_o that appears across the "load resistor", R_L, are shown below:

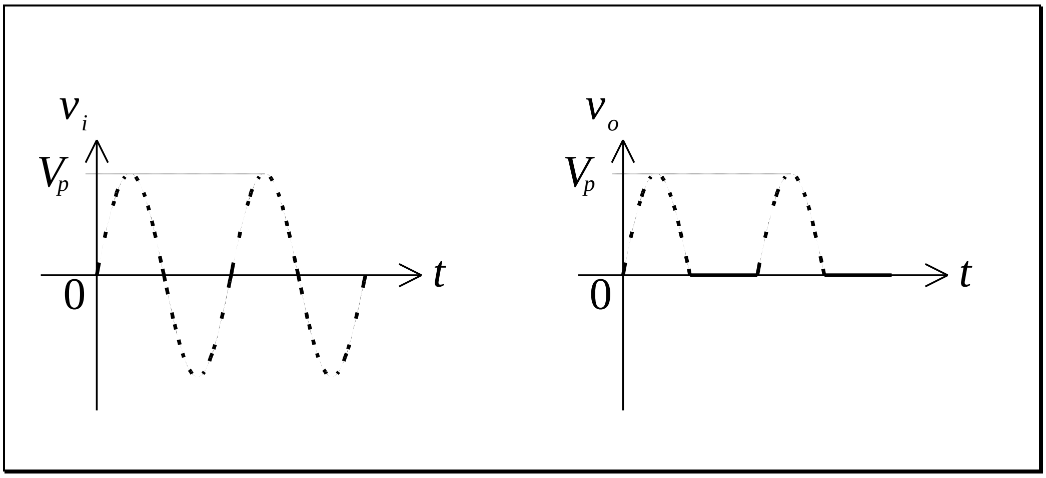


Figure 7.16

Operation of the circuit is straightforward when we assume the diode is ideal: When v_i is positive the diode conducts and acts as a short-circuit. The voltage v_i appears directly at the output – that is, $v_o = v_i$, and the diode forward current is equal to v_i / R_L. On the other hand, when v_i is negative the diode *cuts off* – that is, there is no current. The output voltage v_o will be zero, and the diode becomes reverse-biased by the value of the input voltage v_i. It follows that the output voltage waveform will consist of the positive half cycles of the input sinusoid. Since only half-cycles are utilized, the circuit is called a *half-wave rectifier*.

It should be noted that while the input sinusoid has a zero average value, the output waveform has a finite average value or DC component. Therefore, rectifiers are used to generate DC voltages from AC voltages.

The *transfer characteristic* is often used to describe non-linear circuits – it is simply a plot of the output, v_o, versus the input, v_i. The figure below shows the transfer characteristic of the half-wave rectifier:

Half-wave rectifier transfer characteristic

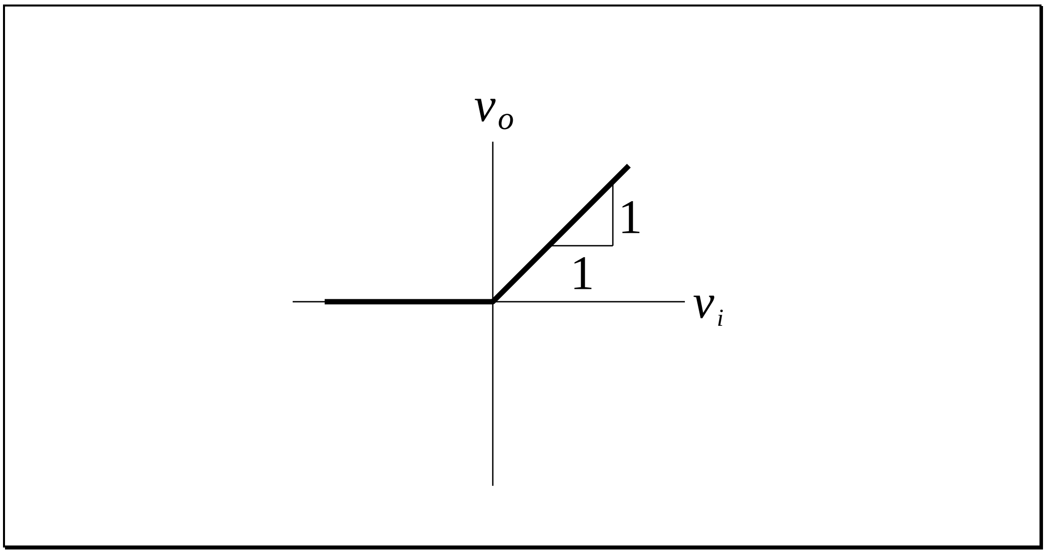


Figure 7.17

As can be seen, the half-wave rectifier produces an output voltage equal to the input voltage when the input voltage is positive and produces zero output voltage when the input voltage is negative.

7.6.2 Full-Wave Rectifier

The full-wave rectifier utilizes both halves of the input signal – it inverts the negative halves of the waveform. One popular implementation is shown below, where the diodes are connected in a bridge configuration:

A full-wave "bridge rectifier"

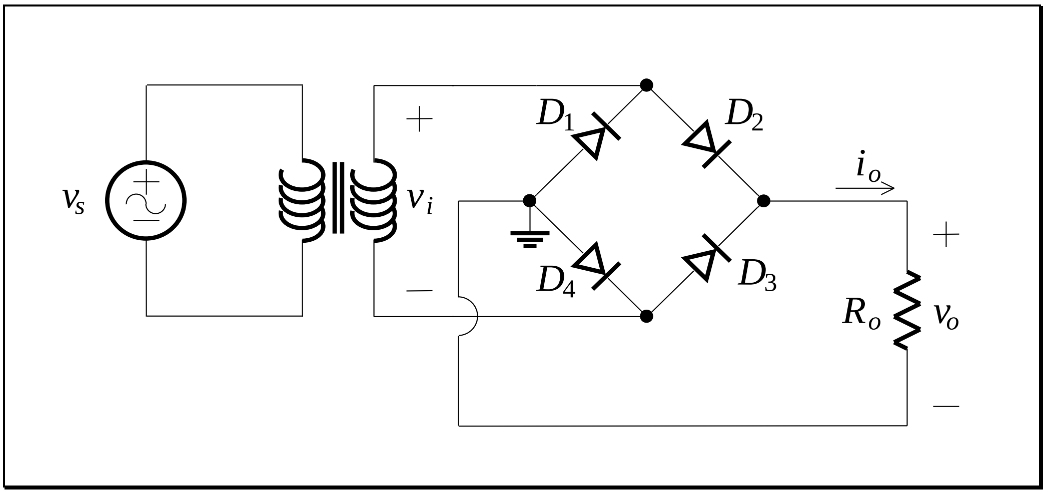


Figure 7.18

We can perform the usual analysis quickly. In the positive half cycle of the input voltage, D_2 and D_4 are on. Meanwhile, D_1 and D_3 will be reverse biased. In the negative half cycle of the input voltage, D_1 and D_3 are on, and D_2 and D_4 are off. The important point to note is that during both half-cycles, the current through the resistor R_o is in the same direction (down), and thus v_o will always be positive. The waveforms are shown below:

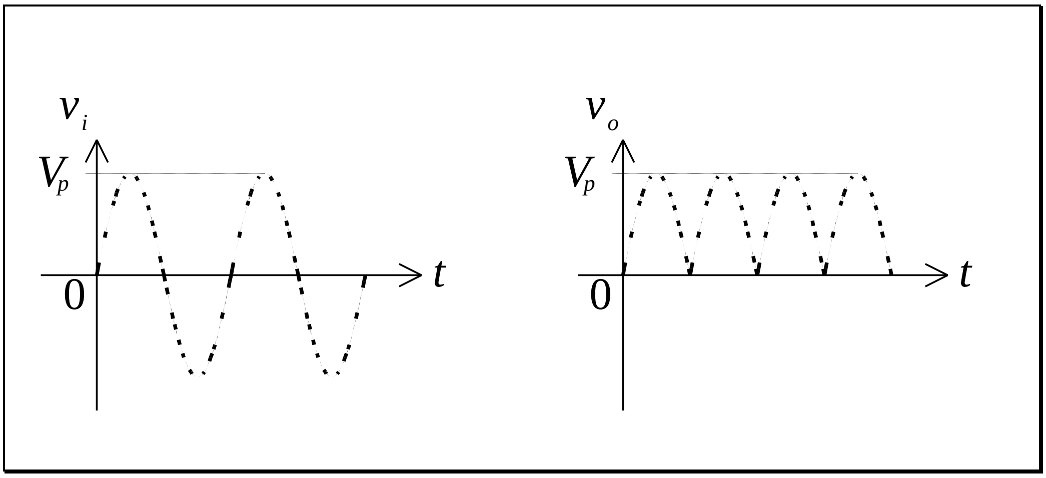


Figure 7.19

7.6.3 Limiter Circuits

Consider the following circuit:

A limiting circuit

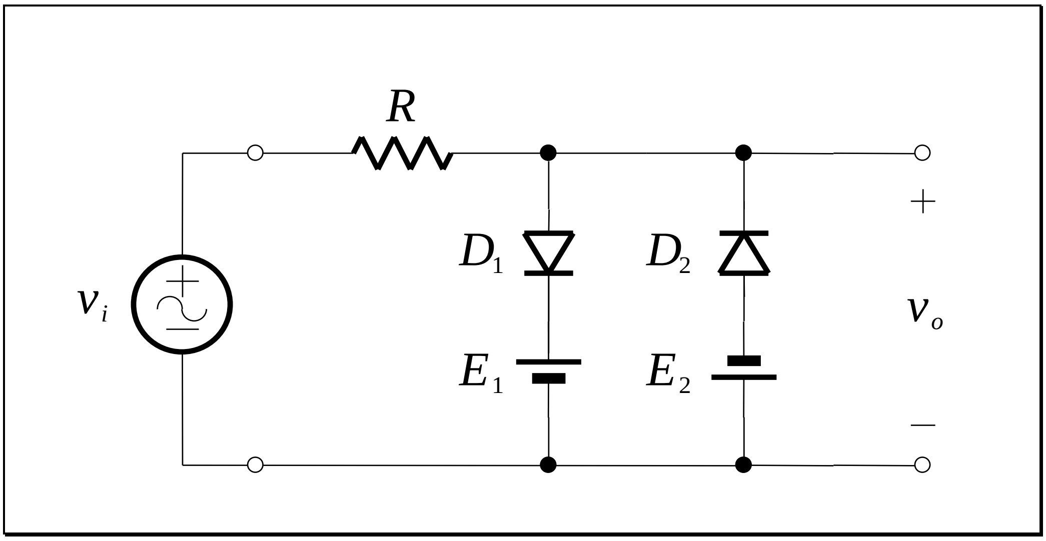


Figure 7.20

The circuit works very simply. Assume both diodes are off. KVL then gives:

$$v_o = v_i \tag{7.21}$$

If the output voltage is greater than E_1, then diode D_1 will be on. This limits or clamps the output voltage to E_1:

$$v_o = E_1 \quad \text{for } v_i > E_1 \tag{7.22}$$

If the output voltage is less than $-E_2$ then diode D_2 will be on, limiting the output voltage to $-E_2$:

$$v_o = -E_2 \quad \text{for } v_i < -E_2 \tag{7.23}$$

A graph of the output is shown below for a sinusoidal input:

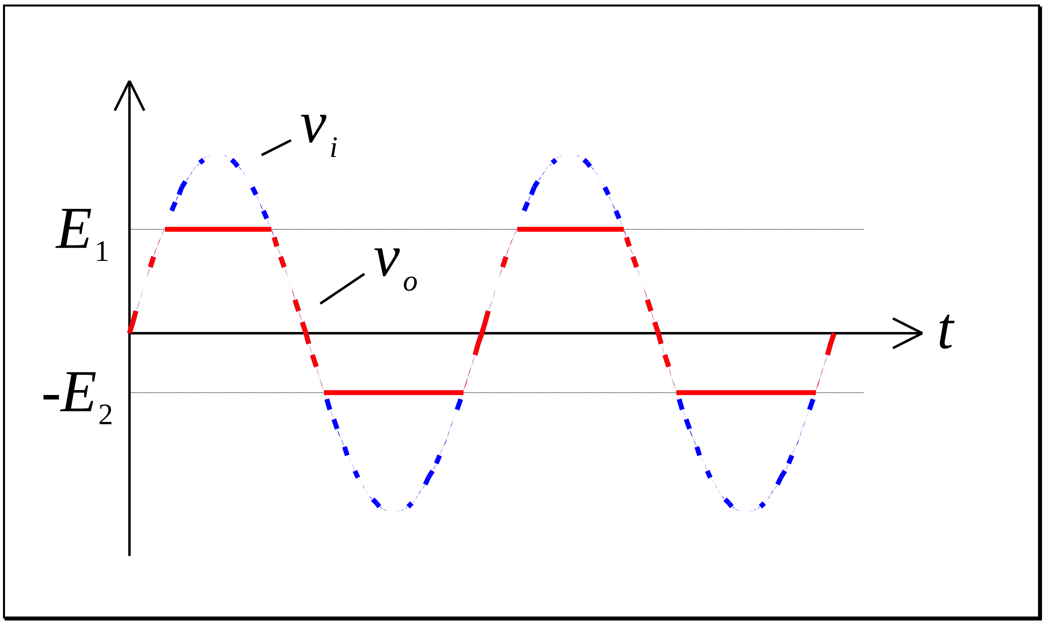


Figure 7.21

The transfer characteristic of this limiter is shown below:

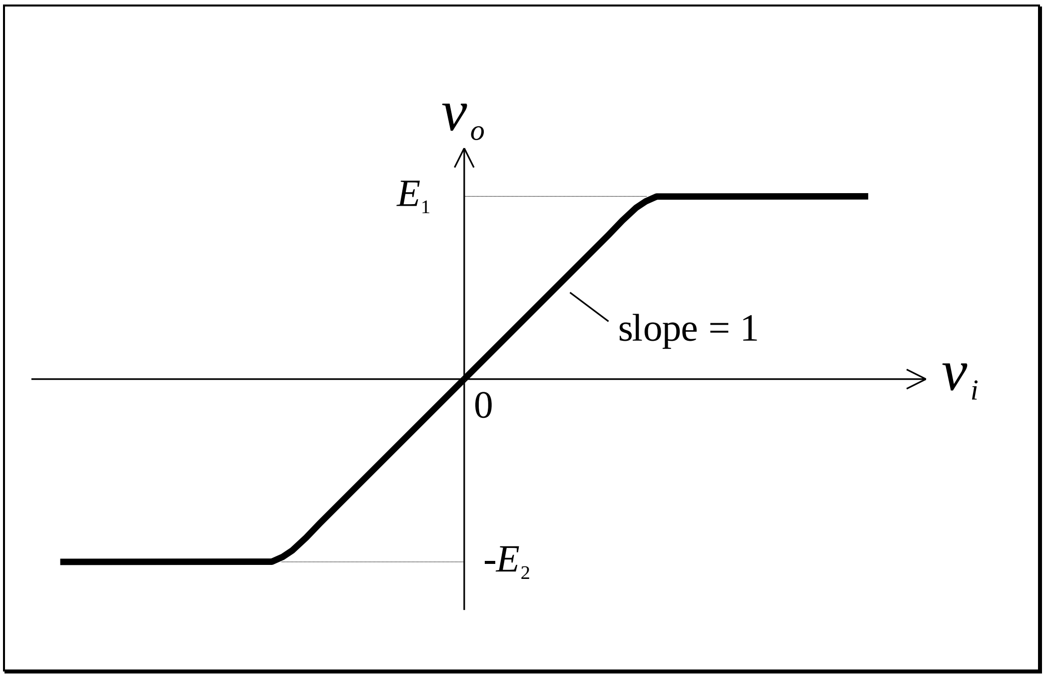


Figure 7.22

7.7 Summary

- The silicon junction diode forms the basis of modern electronics. It is a device that effectively allows a current in only one direction.
- The forward conduction region of practical silicon diodes is accurately characterised by the Shockley equation:

$$i = I_S\left(e^{v/nV_T} - 1\right)$$

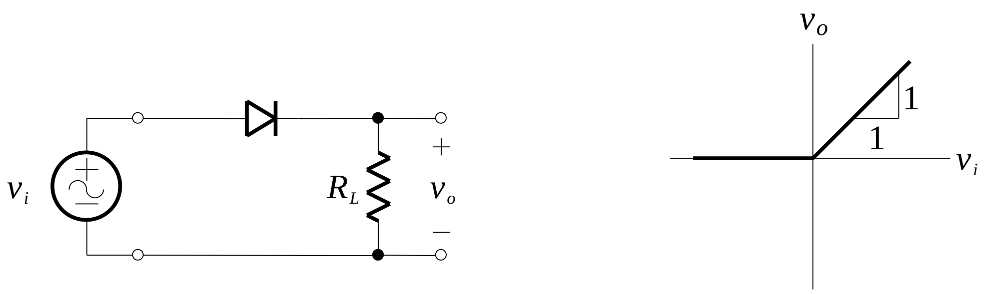
- Beyond a certain value of reverse voltage (that depends on the diode) breakdown occurs, and current increases rapidly with a small corresponding increase in voltage. This property is exploited in diodes known as breakdown diodes.
- A variety of diode constructions exist, with many of them essential to the modern world, such as the LED.
- Since the diode's characteristic is nonlinear, we can't apply linear circuit analysis techniques to circuits containing diodes. We therefore have to resort to other analysis methods: graphical, numerical and linear modelling
- A hierarchy of diode models exists, with the selection of an appropriate model dictated by the application.
- In the forward direction, the ideal diode conducts any current forced by the external circuit while displaying a zero voltage drop. The ideal diode does not conduct in the reverse direction; any applied voltage appears as reverse bias across the diode.
- In many applications, a conducting diode is modelled as having a constant voltage drop, usually approximately 0.7 V.

- A diode biased to operate at a direct current I_D has a small-signal resistance:

$$r_d \approx \frac{nV_T}{I_{DQ}}$$

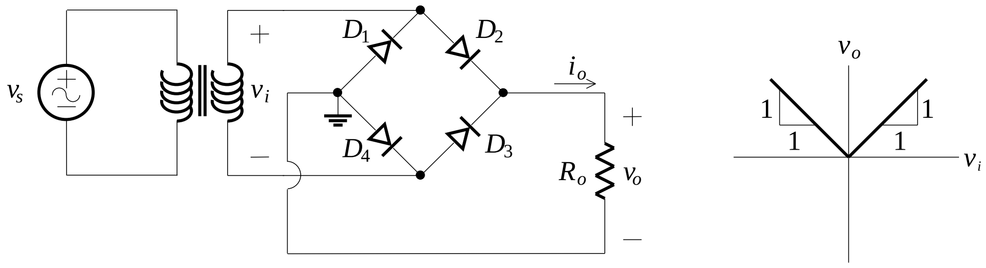
- The unidirectional-current property makes the diode useful in the design of a variety of circuits, such as the half-wave rectifier, the full-wave rectifier, limiting circuits, and many others.

- The half-wave rectifier is:



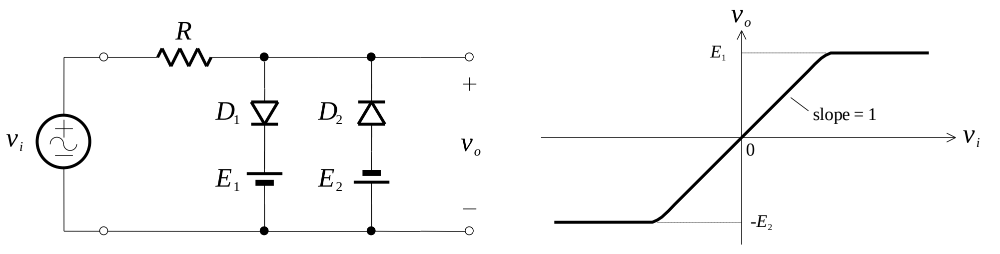
(a) circuit (b) transfer characteristic

- The full-wave rectifier is:



(a) circuit (b) transfer characteristic

- A limiting circuit is:



(a) circuit (b) transfer characteristic

7.8 References

Sedra, A. and Smith, K.: *Microelectronic Circuits*, Saunders College Publishing, New York, 1991.

8 Source-Free RC and RL Circuits

Contents

8.2

Introduction

The analysis of a linear circuit that has storage elements (capacitors and inductors) inevitably gives rise to a linear differential equation. The solution of the differential equation always consists of two parts – one part leads to the so-called *forced response*, the other is the *natural response*. The forced response is due to the application of a source to the circuit. The natural response is due entirely to the circuit's configuration, its initial energy, and the amplitude of the applied source at the instant of application.

Initially, we will study the theory of differential operators, which leads to a rather simple method of solving linear differential equations. For simple *RC* and *RL* circuits, the differential operator seems like overkill – a powerful method applied to a simple situation. However, some of the insights resulting from such a method will put us in good stead when we look at more complicated circuits that have both capacitors and inductors in them.

We shall then study the natural response of some simple source-free *RC* and *RL* circuits. This study will reveal some surprising results, such as the fact that there is only ever one *form* of natural response – an exponential response.

We will become familiar with the exponential response, and we will give some special names to the algebraic terms involved in it, such as *initial condition* and *time constant*.

With practice, we will also see that we can write down the natural response for simple circuits by inspection. This will lead to intuition of circuit behaviour, and we will "get a feel" for the way a circuit behaves by simply looking at it.

8.1 Differential Operators

Let D denote differentiation with respect to t:

$$D \equiv \frac{d}{dt} \tag{8.1}$$

D is just shorthand for d/dt

Then D^2 denotes differentiation twice with respect to t, and so on. That is, for positive integer k:

$$D^k y = \frac{d^k y}{dt^k} \tag{8.2}$$

The expression:

$$A = a_n D^n + a_{n-1} D^{n-1} + \ldots + a_1 D + a_0 \tag{8.3}$$

is called a differential operator of order n. The coefficients a_k in the operator A may be functions of t, but most of the time they will be constant coefficients.

The product AB of two operators A and B is defined as that operator which produces the same result as is obtained by using the operator B followed by the operator A. Thus:

$$ABy = A(By) \tag{8.4}$$

The product of two differential operators always exists and is a differential operator. For operators with *constant coefficients* it is true that $AB = BA$.

Differential operators commute if they have constant coefficients

EXAMPLE 8.1 Product of Differential Operators

Let $A = D + 2$ and $B = 3D - 1$.

Then:

$$By = (3D - 1)y = 3\frac{dy}{dt} - y$$

and:

$$\begin{aligned} A(By) &= (D+2)\left(3\frac{dy}{dt} - y\right) \\ &= 3\frac{d^2y}{dt^2} - \frac{dy}{dt} + 6\frac{dy}{dt} - 2y \\ &= 3\frac{d^2y}{dt^2} + 5\frac{dy}{dt} - 2y \\ &= (3D^2 + 5D - 2)y \end{aligned}$$

Hence $AB = (D+2)(3D-1) = 3D^2 + 5D - 2$.

The sum of two differential operators is obtained by expressing each in the form:

$$a_n D^n + a_{n-1} D^{n-1} + \ldots + a_1 D + a_0 \tag{8.5}$$

and adding corresponding coefficients.

EXAMPLE 8.2 Summation of Differential Operators

Let $A = 3D^3 - D + 2$ and $B = D^2 + 4D + 7$.

Hence $A + B = 3D^3 + D^2 + 3D + 9$.

Differential operators are linear operators that obey the laws of algebra

Differential operators are linear operators. Therefore, differential operators with constant coefficients satisfy all the laws of the algebra of polynomials with respect to the operations of addition and multiplication.

8.2 Properties of Differential Operators

Since for constant s and positive integer k:

$$D^k e^{st} = s^k e^{st} \tag{8.6}$$

it is easy to find the effect that a differential operator has upon e^{st}. Let $f(D)$ be a polynomial in D:

$$f(D) = a_n D^n + a_{n-1} D^{n-1} + \ldots + a_1 D + a_0 \tag{8.7}$$

Then:

$$f(D)e^{st} = a_n s^n e^{st} + a_{n-1} s^{n-1} e^{st} + \ldots + a_1 s e^{st} + a_0 e^{st} \tag{8.8}$$

so:

$$\boxed{f(D)e^{st} = f(s)e^{st}} \tag{8.9}$$

Operating with D and multiplication by s are equivalent for the function e^{st}

This equation does **not** mean that $f(D) = f(s)$. $f(D)$ is an *operator*, and it means "to take a linear sum of derivatives" of a function that it operates on. $f(s)$ is a standard polynomial in s. The equation means that the *effect* of the D operator on e^{st} is the same as multiplication by s. The function e^{st} is the *only* function that enjoys this property – a deeper insight into the special function e^{st} will be developed later.

If s is a root of the equation $f(s) = 0$, then in view of the above result:

$$f(D)e^{st} = 0, \quad \text{if} \quad f(s) = 0 \tag{8.10}$$

EXAMPLE 8.3 Solution of a Homogeneous Differential Equation

Let $f(D) = 2D^2 + 5D - 12$. Then the equation $f(s) = 0$ is:

$$2s^2 + 5s - 12 = 0$$

or:

$$(s+4)(2s-3) = 0$$

of which the roots are $s_1 = -4$ and $s_2 = \frac{3}{2}$.

With the aid of Eq. (8.10), it can be seen that:

$$(2D^2 + 5D - 12)e^{-4t} = 0$$

and that:

$$(2D^2 + 5D - 12)e^{3t/2} = 0$$

In other words, $y_1 = e^{-4t}$ and $y_2 = e^{3t/2}$ are solutions of:

$$(2D^2 + 5D - 12)y = 0.$$

Next consider the effect of the operator $D-s$ on the product of e^{st} and a function *y*. Utilising the "product rule" of differentiation, we have:

$$\begin{aligned}(D-s)(e^{st}y) &= D(e^{st}y) - se^{st}y \\ &= e^{st}Dy\end{aligned} \tag{8.11}$$

and:

$$\begin{aligned}(D-s)^2(e^{st}y) &= (D-s)(e^{st}Dy) \\ &= e^{st}D^2y\end{aligned} \tag{8.12}$$

Repeating the operation, we are led to:

$$(D-s)^n(e^{st}y) = e^{st}D^ny \tag{8.13}$$

Using the linearity of differential operators, we conclude that when $f(D)$ is a polynomial in *D* with constant coefficients, then:

Shifting an exponential factor from the left of a differential operator to the right

$$e^{st}f(D)y = f(D-s)(e^{st}y) \tag{8.14}$$

This relation shows us how to shift an exponential factor from the left of an operator to the right of an operator.

EXAMPLE 8.4 Solution of a Homogeneous Differential Equation

Let us solve the differential equation:

$$(D+3)^4 y = 0$$

First, we multiply the equation by e^{3t} to obtain:

$$e^{3t}(D+3)^4 y = 0$$

Applying the exponential shift as in Eq. (8.14) leads to:

$$D^4\left(e^{3t} y\right) = 0$$

Integrating four times gives us:

$$e^{3t} y = c_0 + c_1 t + c_2 t^2 + c_3 t^3$$

and finally:

$$y = \left(c_0 + c_1 t + c_2 t^2 + c_3 t^3\right)e^{-3t}$$

It can be shown that the four functions e^{-3t}, te^{-3t}, $t^2 e^{-3t}$ and $t^3 e^{-3t}$ are linearly independent – thus the solution given is the general solution of the differential equation.

In Eq. (8.13), if we let $f(D) = D^n$ and $y = t^k$, then:

$$(D-s)^n\left(t^k e^{st}\right) = e^{st} D^n t^k \tag{8.15}$$

But $D^n t^k = 0$ for $k = 0, 1, 2, \ldots, n-1$, and so:

The solution to a special class of differential equation

$$(D-s)^n\left(t^k e^{st}\right) = 0, \quad k = 0, 1, \ldots, (n-1) \tag{8.16}$$

8.3 The Characteristic Equation

The general linear differential equation with constant coefficients of order *n* is an equation that can be written:

$$a_n \frac{d^n y}{dt^n} + a_{n-1} \frac{d^{n-1} y}{dt^{n-1}} + \cdots + a_1 \frac{dy}{dt} + a_0 y = r(t) \tag{8.17}$$

If $r(t)$ is identically zero (i.e. zero for all time, not just a specific time), then we have an equation that is said to be *homogeneous*:

$$a_n \frac{d^n y}{dt^n} + a_{n-1} \frac{d^{n-1} y}{dt^{n-1}} + \cdots + a_1 \frac{dy}{dt} + a_0 y = 0 \tag{8.18}$$

Homogeneous differential equation defined

Any *linear combination* of solutions of a linear homogeneous differential equation is also a solution. If y_i, with $i = 1, 2, \ldots, k$, are solutions of Eq. (8.18) and if c_i, with $i = 1, 2, \ldots, k$, are constants, then:

$$y = c_1 y_1 + c_2 y_2 + \cdots + c_k y_k \tag{8.19}$$

is a solution of Eq. (8.18).

If the solutions y_i are linearly independent (i.e. they cannot be expressed as linear combinations of one another), then there are *n* of them and we say that the *general solution* of the differential equation is:

$$y = c_1 y_1 + c_2 y_2 + \cdots + c_n y_n \tag{8.20}$$

8.10

The linear homogeneous differential equation with constant coefficients may also be written in the form:

A solution to a linear homogeneous differential equation with constant coefficients is $y = e^{st}$

$$f(D)y = 0 \tag{8.21}$$

where $f(D)$ is a linear differential operator. As we saw in the preceding section, if s is any root of the algebraic equation $f(s) = 0$, then:

$$f(D)e^{st} = 0 \tag{8.22}$$

which means simply that $y = e^{st}$ is a solution of Eq. (8.21). The equation:

Characteristic equation defined

$$\boxed{f(s) = 0} \tag{8.23}$$

is called the *characteristic equation* associated with Eq. (8.21).

Let the characteristic equation for Eq. (8.21) be of degree *n*. Let its roots be s_1, s_2, …, s_n. If these roots are all real and distinct, then the *n* solutions $y_1 = e^{s_1 t}$, $y_2 = e^{s_2 t}$, …, $y_n = e^{s_n t}$ are linearly independent and the general solution of Eq. (8.21) can be written at once. It is:

General solution of a linear homogeneous differential equation with constant coefficients

$$y = c_1 e^{s_1 t} + c_2 e^{s_2 t} + \cdots + c_n e^{s_n t} \tag{8.24}$$

in which c_1, c_2, …, c_n are arbitrary constants. Repeated roots of the characteristic equation will be treated later.

EXAMPLE 8.5 General Solution of a Homogeneous Differential Equation

Solve the equation:

$$\frac{d^3y}{dt^3} - 4\frac{d^2y}{dt^2} + \frac{dy}{dt} + 6y = 0$$

First write the characteristic equation:

$$s^3 - 4s^2 + s + 6 = 0$$
$$(s+1)(s-2)(s-3) = 0$$

whose roots are $s = -1, 2, 3$. Then the general solution is seen to be:

$$y = c_1e^{-t} + c_2e^{2t} + c_3e^{3t}$$

EXAMPLE 8.6 General Solution of a Homogeneous Differential Equation

Solve the equation:

$$(3D^3 + 5D^2 - 2D)y = 0$$

The characteristic equation is:

$$3s^3 + 5s^2 - 2s = 0$$
$$s(s+2)(3s-1) = 0$$

whose roots are $s = 0, -2, \frac{1}{3}$. Using the fact that $e^{ot} = 1$, the desired solution may be written:

$$y = c_1 + c_2e^{-2t} + c_3e^{t/3}$$

EXAMPLE 8.7 General Solution of a Homogeneous Differential Equation with Initial Conditions

Solve the equation:

$$\frac{d^2y}{dt^2} - 4y = 0$$

with the conditions that when $t = 0$, $y = 0$ and $dy/dt = 3$.

The characteristic equation is:

$$s^2 - 4 = 0$$

with roots $s = 2, -2$. Hence the general solution of the differential equation is:

$$y = c_1e^{2t} + c_2e^{-2t}$$

It remains to enforce the conditions at $t = 0$.

The condition that $y = 0$ when $t = 0$ requires that:

$$0 = c_1 + c_2$$

Now:

$$\frac{dy}{dt} = 2c_1e^{2t} - 2c_2e^{-2t}$$

The condition that $dy/dt = 3$ when $t = 0$ requires that:

$$3 = 2c_1 - 2c_2$$

From the simultaneous equations for c_1 and c_2 we conclude that $c_1 = \frac{3}{4}$ and $c_2 = -\frac{3}{4}$. Therefore:

$$y = \tfrac{3}{4}\left(e^{2t} - e^{-2t}\right)$$

8.4 The Simple *RC* Circuit

Consider the parallel *RC* circuit shown below.

The simple *RC* circuit

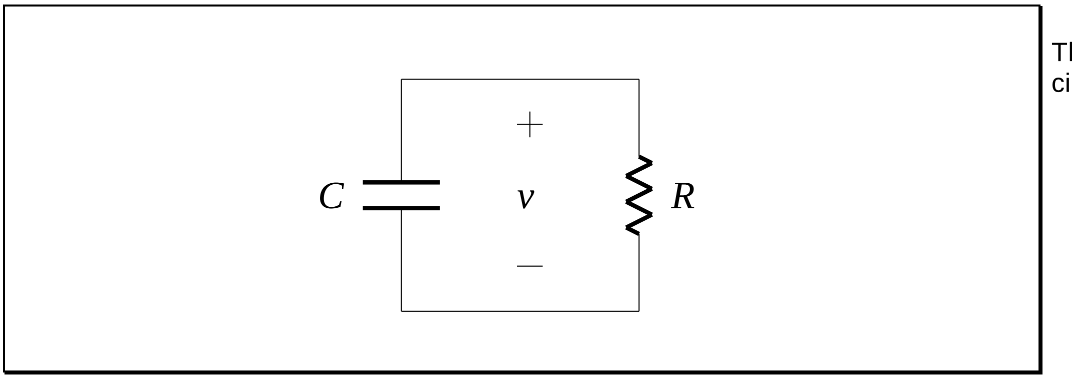


Figure 8.1

We designate the time-varying voltage by $v(t)$, and we shall let the value of $v(t)$ at $t=0$ be prescribed as V_0.

Kirchhoff's Current Law (KCL) applied to the top node gives us:

$$C\frac{dv}{dt}+\frac{v}{R}=0 \tag{8.25}$$

Division by *C* gives us:

The governing differential equation of the simple *RC* circuit

$$\frac{dv}{dt}+\frac{v}{RC}=0 \tag{8.26}$$

and we must determine an expression for $v(t)$ which satisfies this equation and also has the value V_0 at $t=0$.

Written in operator notation, Eq. (8.26) is:

$$\left(D+\frac{1}{RC}\right)v=0 \tag{8.27}$$

for which the characteristic equation is:

$$\left(s+\frac{1}{RC}\right)=0 \tag{8.28}$$

which has a root at $s=-1/RC$. Therefore, the solution to the differential equation is:

$$v=c_1e^{-t/RC} \tag{8.29}$$

It remains to enforce the condition that $v(0)=V_0$. Thus:

$$V_0=c_1 \tag{8.30}$$

and the final form of the solution is:

The solution to the governing differential equation of the simple *RC* circuit

$$v=V_0e^{-t/RC} \tag{8.31}$$

This solution is known as the *natural response* of the circuit. Its mathematical form is Ae^{st}, where $s=-1/RC$ and A is the initial condition. It turns out, for any linear circuit, that this mathematical form is the *only* form for the natural response! That is, the solution to a linear homogeneous differential equation always takes the form Ae^{st}. The actual values of s and A depend only on the initial energy in the circuit (in this case the initial capacitor voltage), and the circuit element configuration and values.

Let us check the power and energy relationships in this circuit. The power being dissipated in the resistor is:

$$p_R = \frac{v^2}{R} = \frac{V_0^2 e^{-2t/RC}}{R} \tag{8.32}$$

Power and energy dissipated in the simple *RC* circuit

and the total energy turned into heat in the resistor is found by integrating the instantaneous power from zero time to infinite time:

$$\begin{aligned} E_R &= \int_0^\infty p_R dt = \frac{V_0^2}{R}\int_0^\infty e^{-2t/RC} dt \\ &= \frac{V_0^2}{R}\left[-\frac{RC}{2}e^{-2t/RC}\right]_0^\infty \\ &= \tfrac{1}{2} C V_0^2 \end{aligned} \tag{8.33}$$

This is the result we expect, because the total energy stored initially in the capacitor is $\tfrac{1}{2} C V_0^2$, and there is no energy stored in the capacitor at infinite time. All the energy is accounted for by dissipation in the resistor.

8.5 Properties of the Exponential Response

We will now consider the nature of the response in the parallel *RC* circuit. We found that the voltage is represented by:

$$v = V_0 e^{-t/RC} \tag{8.34}$$

At zero time, the voltage is the assumed value V_0 and as time increases, the voltage decreases and approaches zero. The shape of this decaying exponential is seen by a plot of v/V_0 versus *t*, as shown below:

The decaying exponential response

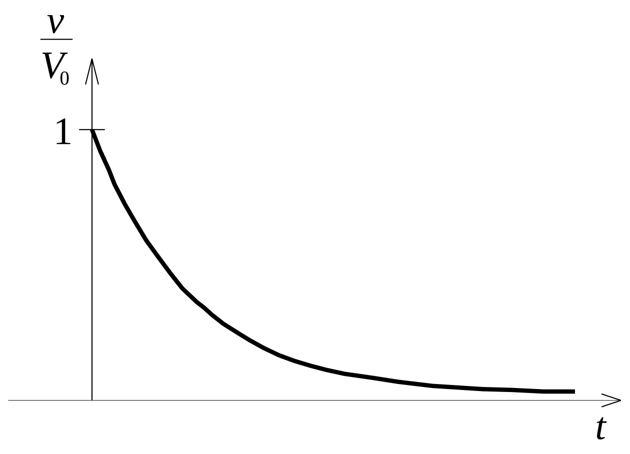


Figure 8.2

Since the function we are plotting is $e^{-t/RC}$, the curve will not change if *RC* does not change. Thus, the same curve must be obtained for every *RC* circuit having the same product of *R* and *C*. Let's see how this product affects the shape of the curve.

If we double the product *RC*, then the exponent will be unchanged if *t* is also doubled. In other words, the original response will occur at a later time, and the new curve is obtained by moving each point on the original curve twice as far to the right. With this larger *RC* product, the voltage takes longer to decay to any given fraction of its original value.

To get a "handle" on the rate at which the curve decays, let's consider the time that would be required for the voltage to drop to zero *if it continued to drop at its initial rate*.

The initial rate of decay is found by evaluating the derivative at zero time:

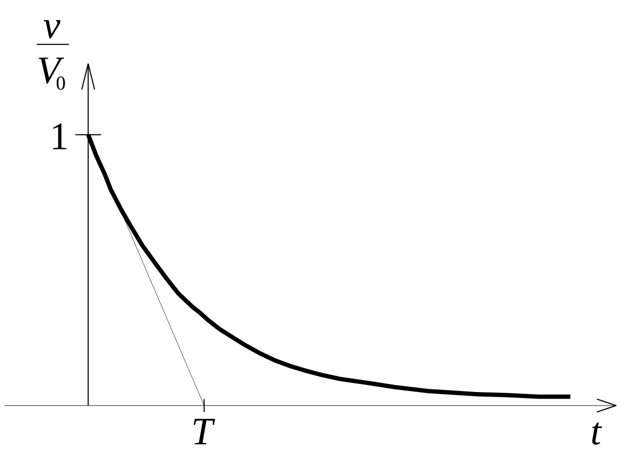
$$\left.\frac{d}{dt}\frac{v}{V_0}\right|_{t=0} = -\frac{1}{RC}e^{-t/RC}\bigg|_{t=0} = -\frac{1}{RC} \tag{8.35}$$

We designate the value of time it takes for v/V_0 to drop from unity to zero, assuming a constant rate of decay, by T (or sometimes the Greek letter τ). Thus $T/RC = 1$, or:

$$T = RC \tag{8.36}$$

Time constant defined for the simple *RC* circuit

The product *RC* has the units of seconds, and therefore the exponent $-t/RC$ is dimensionless (as it must be). The value of time T is called the *time constant*. It is shown below:



Time constant shown graphically on the natural response curve

Figure 8.3

An equally important interpretation of the time constant T is obtained by determining the value of v/V_0 at $t = T$. We have:

$$\frac{v}{V_0} = e^{-1} = 0.3679 \quad \text{or} \quad v = 0.3679 V_0 \tag{8.37}$$

Thus, in one time constant the response has dropped to 36.8 percent of its initial value. The value of T may be determined graphically from this fact from the display on an oscilloscope, as indicated below:

Response curve values at integer multiples of the time constant

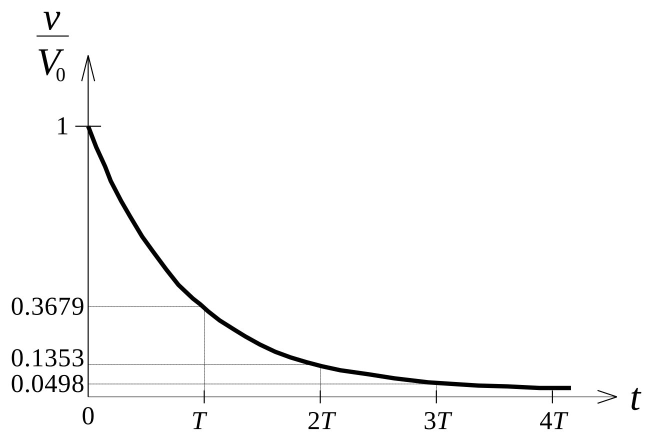


Figure 8.4

At some point three to five time constants after zero time, most of us would agree that the voltage is a negligible fraction of its former self.

Why does a larger value of the time constant RC produce a response curve which decays more slowly? An increase in C allows a greater energy storage for the same initial voltage, and this larger energy requires a longer time to be dissipated in the resistor. For an increase in R, the power flowing into the resistor is less for the same initial voltage; again, a greater time is required to dissipate the stored energy.

In terms of the time constant T, the response of the parallel RC circuit may be written simply as:

$$v = V_0 e^{-t/T} \tag{8.38}$$

8.6 Single Time Constant *RC* Circuits

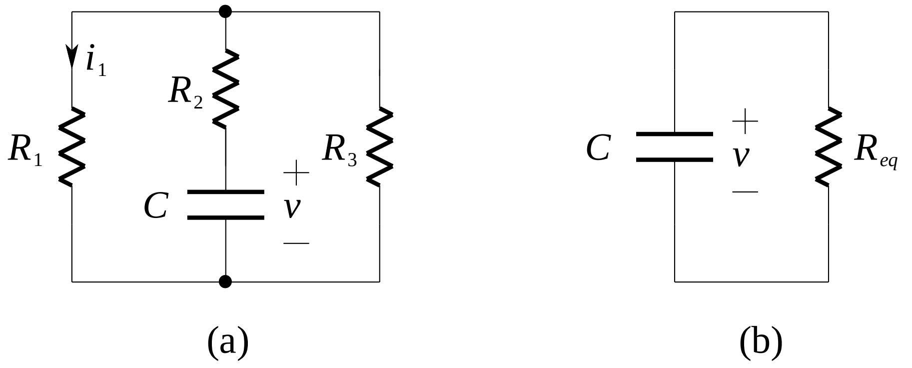
Many of the *RC* circuits for which we would like to find the natural response contain more than a single resistor and capacitor. We first consider those cases in which the given circuit may be reduced to an equivalent circuit consisting of only one resistor and one capacitor. Such a circuit is known as a *single time constant* circuit.

Single time constant *RC* circuit defined

Firstly, consider a circuit with any number of resistors and one capacitor. We fix our attention on the two terminals of the capacitor and determine the equivalent resistance across these terminals. The circuit is thus reduced to the simple parallel case.

EXAMPLE 8.8 Analysis of a Single Time Constant *RC* Circuit

The circuit shown in (a) below may be simplified to that of (b):



enabling us to write:

$$v = V_0 e^{-t/R_{eq}C}$$

where:

$$v(0) = V_0 \quad \text{and} \quad R_{eq} = R_2 + \frac{R_1 R_3}{R_1 + R_3}$$

Every current and voltage in the resistive portion of the circuit must have the form $Ae^{-t/R_{eq}C}$, where A is the initial value of that current or voltage.

Since the current in a resistor may change instantaneously, we shall indicate the instant *after* any change that may have occurred at $t = 0$ by use of the symbol 0^+.

Thus, the current in R_1, for example, may be expressed as:

$$i_1 = i_1(0^+)e^{-t/T}$$

where:

$$T = \left(R_2 + \frac{R_1 R_3}{R_1 + R_3} \right) C$$

and $i_1(0^+)$ remains to be determined from some initial condition. Suppose that $v(0)$ is given. Since v cannot change instantaneously, we may think of the capacitor as being replaced by an independent DC source, $v(0)$, for the instant of time $t = 0$. Thus:

$$i_1(0^+) = \frac{v(0)}{R_2 + R_1 R_3/(R_1 + R_3)} \frac{R_3}{R_1 + R_3}$$

The solution is obtained by collecting all these results:

$$i_1 = \frac{v(0)}{R_2 + R_1 R_3/(R_1 + R_3)} \frac{R_3}{R_1 + R_3} e^{\frac{-t}{(R_2 + R_1 R_3/(R_1 + R_3))C}}$$

Another special case includes those circuits containing one resistor and any number of capacitors. The resistor voltage is easily obtained by establishing the value of the equivalent capacitance and determining the time constant.

Not all *RC* circuits can be converted into single time constant *RC* circuits

Some circuits containing a number of both resistors and capacitors may be replaced by an equivalent circuit containing only one resistor and one capacitor; it is necessary that the original circuit be one which can be broken into two parts, one containing all resistors and the other containing all capacitors. This is not possible in general.

8.7 The Simple *RL* Circuit

Consider the series *RL* circuit shown below.

The simple *RL* circuit

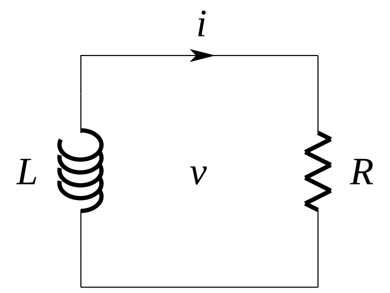


Figure 8.5

We designate the time-varying current by $i(t)$, and we shall let the value of $i(t)$ at $t=0$ be prescribed as I_0.

Kirchhoff's Voltage Law (KVL) applied around the loop gives us:

$$L\frac{di}{dt}+Ri=0 \tag{8.39}$$

Division by *L* gives us:

The governing differential equation of the simple *RL* circuit

$$\frac{di}{dt}+\frac{R}{L}i=0 \tag{8.40}$$

This equation has a familiar form; comparison with:

$$\frac{dv}{dt}+\frac{v}{RC}=0 \tag{8.41}$$

shows that the replacement of v by i and RC by L/R produces an equation identical to Eq. (8.40). It should, for the *RL* circuit we are now analysing is the *dual* of the *RC* circuit we considered first.

This duality forces $i(t)$ for the *RL* circuit and $v(t)$ for the *RC* circuit to have identical expressions if the resistance of one circuit is equal to the conductance of the other and if *L* is numerically equal to *C*. That is, we will obtain the dual circuit (and equation) if we make the substitution:

Duality relations for all dual circuits

$$\begin{aligned} v &\rightarrow i \\ R &\rightarrow G \\ C &\rightarrow L \end{aligned} \tag{8.42}$$

Thus, the response of the *RC* circuit:

$$v(t) = v(0)e^{-t/RC} = V_0 e^{-t/RC} \tag{8.43}$$

enables us to write immediately:

The solution to the governing differential equation of the simple *RL* circuit

$$i(t) = i(0)e^{-tR/L} = I_0 e^{-tR/L} \tag{8.44}$$

for the *RL* circuit.

Let's examine the physical nature of the response of the *RL* circuit as expressed by Eq. (8.44). At $t = 0$ we obtain the correct initial condition, and as t becomes infinite the current approaches zero. This latter result agrees with our thinking that if there were any current remaining through the inductor, then energy would continue to flow into the resistor and be dissipated as heat. Thus a final current of zero is necessary.

The time constant of the *RL* circuit may be found by using the duality relationships on the expression for the time constant of the *RC* circuit, or it may be found by simply noting the time at which the response has dropped to 36.8 percent of its initial value:

$$T = \frac{L}{R} \tag{8.45}$$

Time constant defined for the simple *RL* circuit

Our familiarity with the negative exponential and the significance of the time constant T enables us to sketch the response curve readily:

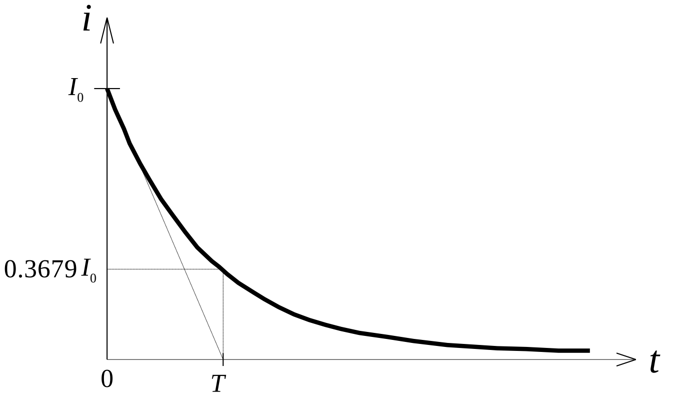


Figure 8.6

An increase in L allows a greater energy storage for the same initial current, and this larger energy requires a longer time to be dissipated in the resistor. If we reduce R, the power flowing into the resistor is less for the same initial current; again, a greater time is required to dissipate the stored energy.

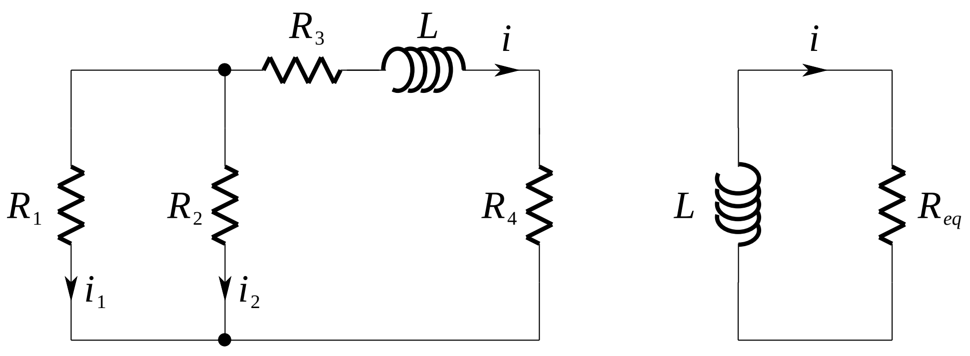
8.8 Single Time Constant *RL* Circuits

Single time constant *RL* circuit defined

It is not difficult to extend the results obtained for the series *RL* circuit to a circuit containing any number of resistors and inductors. We first consider those cases in which the given circuit may be reduced to an equivalent circuit consisting of only one resistor and one inductor; that is, single time constant (STC) circuits.

EXAMPLE 8.9 Analysis of a Single Time Constant *RL* Circuit

The circuit shown in (a) below may be simplified to that of (b):



(a) (b)

The equivalent resistance which the inductor faces is:

$$R_{eq} = R_3 + R_4 + \frac{R_1 R_2}{R_1 + R_2}$$

and the time constant is therefore:

$$T = \frac{L}{R_{eq}}$$

The inductor current is:

$$i = i(0)e^{-t/T}$$

and represents what we might call the basic solution to the problem. It is quite possible that some current or voltage other than i is needed, such as the current i_2 in R_2. We can always apply Kirchhoff's laws and Ohm's law to the resistive portion of the circuit without any difficulty, but current division provides the quickest answer in this circuit:

$$i_2 = -\frac{R_1}{R_1 + R_2} i(0) e^{-t/T}$$

It may also happen that we know the initial value of some current other than the inductor current. Thus, if we are given the initial value of i_1 as $i_1(0^+)$, then it is apparent that the initial value of i_2 is:

$$i_2(0^+) = i_1(0^+)\frac{R_1}{R_2}$$

From these values, we obtain the necessary initial value of $i(0^+)$:

$$i(0^+) = -[i_1(0^+) + i_2(0^+)] = -\frac{R_1 + R_2}{R_2} i_1(0^+)$$

and the expression for i_2 becomes:

$$i_2 = \frac{R_1}{R_2} i_1(0^+) e^{-t/T}$$

We can obtain this last expression more directly. Every current and voltage in the resistive portion of the circuit must have the form $Ae^{-t/T}$, where A is the initial value of that current or voltage. We therefore express i_2 as:

$$i_2 = Ae^{-t/T}$$

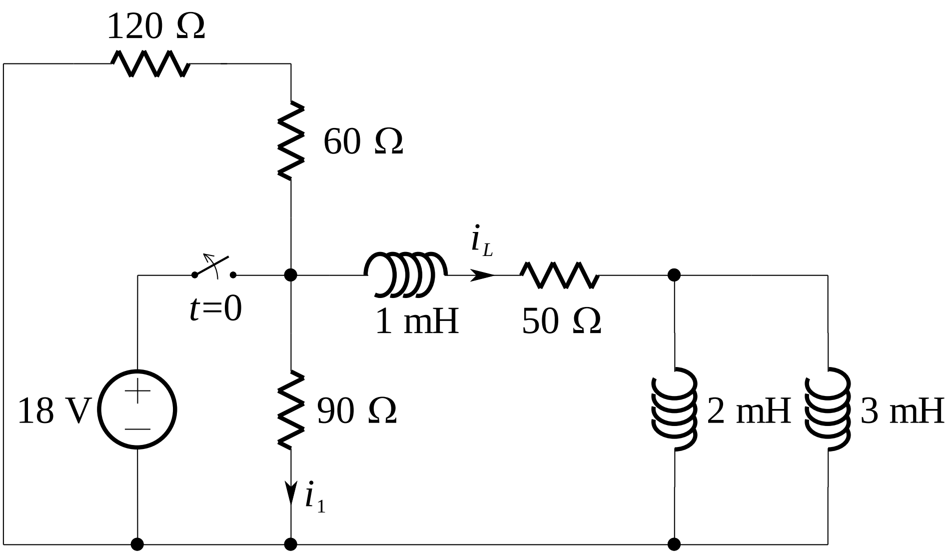
and A must be determined from a knowledge of the initial value of i_2. We found this initial value previously. We therefore have:

$$i_2 = \frac{R_1}{R_2} i_1(0^+) e^{-t/T}$$

A similar sequence of steps provides a solution to a large number of problems. This technique is also applicable to a circuit which contains one resistor and any number of inductors, as well as to those special circuits containing two or more inductors and also two or more resistors that may be simplified by resistance or inductance combination until the simplified circuits have only one inductance or one resistance.

EXAMPLE 8.10 Analysis of a Single Time Constant *RL* Circuit

Consider the circuit shown below:



After $t=0$, when the voltage source is disconnected, we easily calculate an equivalent inductance,

$$L_{eq} = \frac{2\times 3}{2+3} + 1 = 2.2\,\text{mH}$$

an equivalent resistance,

$$R_{eq} = \frac{90(60+120)}{90+180} + 50 = 110\ \Omega$$

and the time constant:

$$T = \frac{L_{eq}}{R_{eq}} = \frac{2.2\times 10^{-3}}{110} = 20\ \mu\text{s}$$

Thus, the form of the natural response is $Ae^{-50000t}$. With the independent source connected $(t<0)$, i_L is $18/50$, or 0.36 A, while i_1 is $18/90$, or 0.2 A. At $t=0^+$, i_L must still be 0.36 A, but i_1 will jump to a new value determined by $i_L(0^+)$. Thus, by the current divider rule:

$$i_1(0^+) = -i_L(0^+)\,{}^{180}\!/_{270} = -0.24 \text{ A}$$

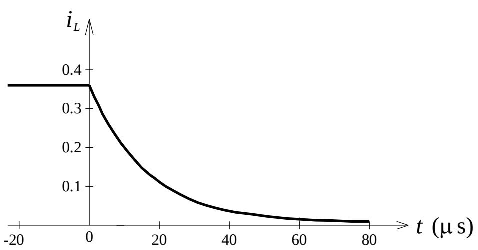
Hence:

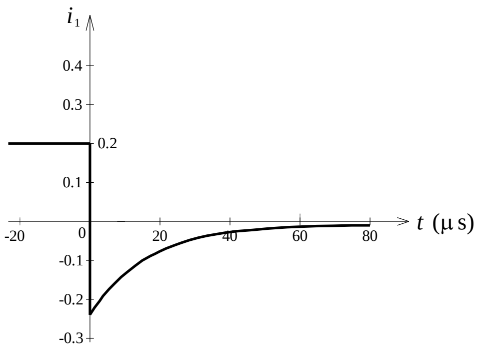
$$\begin{aligned} i_L &= 0.36 && t<0 \\ &= 0.36e^{-50000t} && t \geq 0 \end{aligned}$$

and:

$$\begin{aligned} i_1 &= 0.2 && t<0 \\ &= -0.24e^{-50000t} && t \geq 0 \end{aligned}$$

The responses are sketched below:





A circuit containing several resistances and several inductances does not in general possess a form which allows either the resistances or inductances to be combined into single equivalent elements. There is no single negative exponential term or single time constant associated with the circuit. Rather, in general, there will be several negative exponential terms, the number of terms being equal to the number of inductances that remain after all possible inductor combinations have been made. The natural response of these more complex circuits is obtained using mathematical techniques that will appear later.

Not all *RL* circuits can be converted into single time constant *RL* circuits

8.9 Summary

- The solution to the linear homogeneous differential equation $f(D)y=0$ is $y=c_1e^{s_1t}+c_2e^{s_2t}+\cdots+c_ne^{s_nt}$ where the s_i 's are the roots of the characteristic equation $f(s)=0$ and the c_i 's are arbitrary constants.

- The natural response of any voltage or current in a single time constant circuit always takes the form of $y=Ae^{-t/T}$, where T is the time constant and A is determined from the initial conditions.

- For single time constant *RC* circuits, the time constant is $T=R_{eq}C_{eq}$.

- For single time constant *RL* circuits, the time constant is $T=\frac{L_{eq}}{R_{eq}}$.

- Not all *RC* and *RL* circuits can be reduced to single time constant circuits.

8.10 References

Bedient, P. & Rainville, E.: *Elementary Differential Equations*, 6th *Ed*. Macmillan Publishing Co., 1981.

Hayt, W. & Kemmerly, J.: *Engineering Circuit Analysis*, 3rd *Ed*., Mcgraw-Hill, 1984.

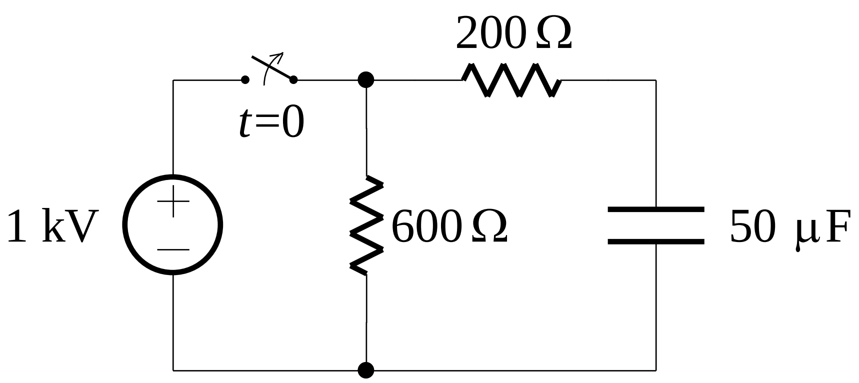
Exercises

1.

The current in a simple source-free series *RC* circuit is given by $i(t) = 20e^{-5000t}$ mA, and the capacitor voltage is 2 V in magnitude at $t = 0$. Find *R* and *C*.

2.

Consider the circuit shown below:



(a) Find the charge in coulombs present on the capacitor at $t = 0^+$.

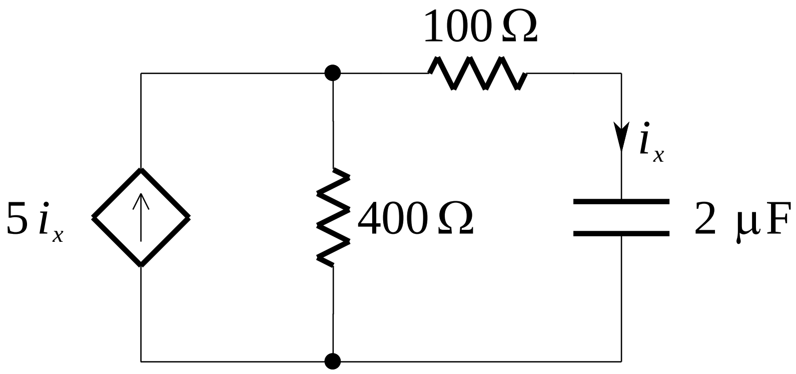
(b) What is the charge at $t = 0.01\text{ s}$?

3.

Let *v* and *i* be the voltage and current variables for a capacitor, assuming the passive sign convention. The capacitor is the only energy-storage element present in a source-free resistive circuit. If $v(0^+) = 80\text{ V}$, $i(0^+) = -0.1\text{ A}$, and $q(0^+)$ for the capacitor is 20 mC, find $v(0.01)$.

4.

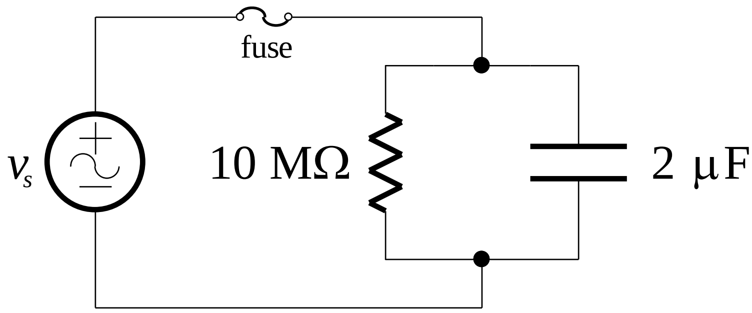
Consider the circuit shown below:



If $i_x(0^+) = 3\text{ mA}$, find $i_x(t)$ for $t > 0$.

5.

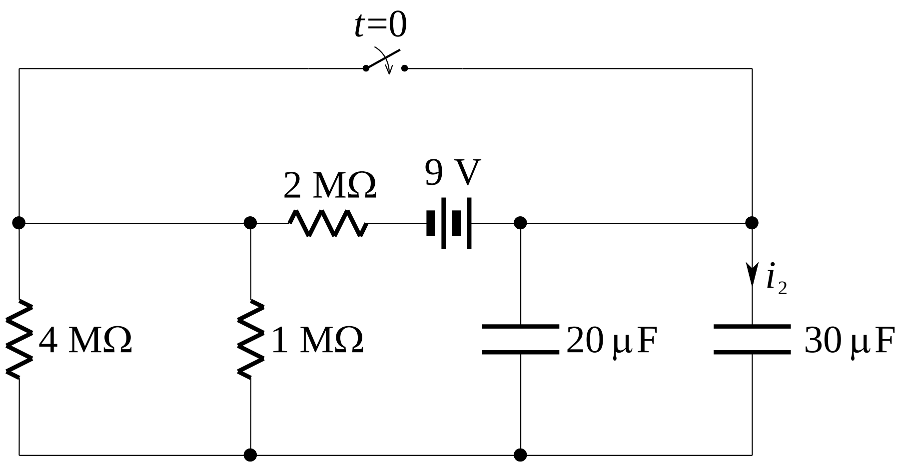
Consider the circuit shown below:



The $10\text{ M}\Omega$ resistor in the circuit represents the leakage resistance present in a typical $2\,\mu\text{F}$ high-voltage capacitor. The fuse blows (opens) at $t = 0$. Let $v_s = 23000\cos(100\pi t)\text{ V}$. Assuming 50 V is relatively nonlethal, at what time is it safe to get your hands across the capacitor?

6.

Consider the circuit shown below:



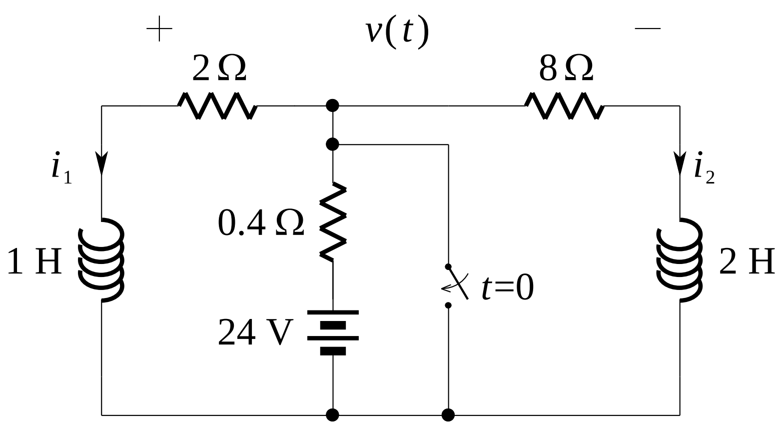
The switch closes at $t = 0$. Find $i_2(t)$ for $t > 0$.

7.

The magnitude of the current in a series *RL* circuit decreases at a rate of $2000\ \text{As}^{-1}$ at $t = 0$ and $100\ \text{As}^{-1}$ at $t = 0.2\ \text{s}$. At what time has the energy stored in the inductor decreased to 1 per cent of its initial value?

8.

Consider the circuit shown below:



Find:

(a) $i_1(0)$ (b) $i_2(0)$ (c) $i_1(t)$, $t > 0$ (d) $i_2(t)$, $t > 0$ (e) $v(t)$, $t > 0$

9.

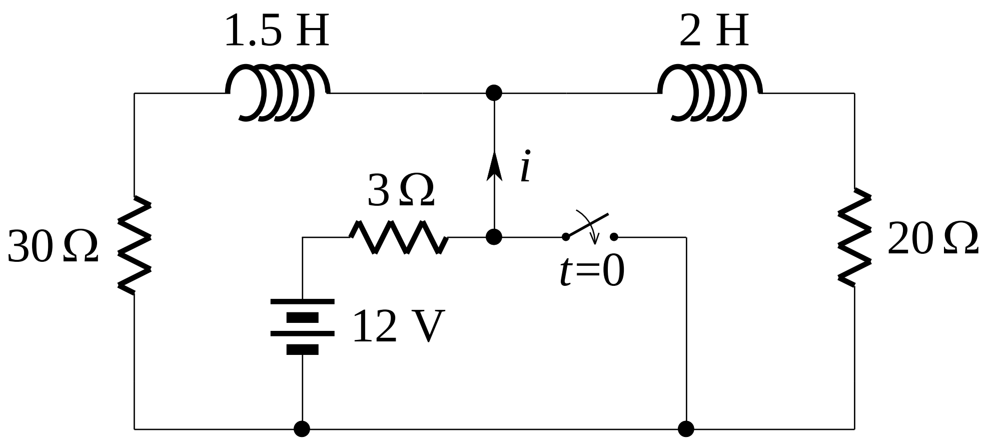
What is the time constant of a series *RL* circuit if:

(a) the current decreases by a factor of 1000 in 0.2 s?

(b) the time required for the current to drop to half of its initial value is 0.1 s less than the time required for it to drop to one-quarter of its initial value?

10.

The switch in the circuit shown below has been open for a long time.



Find i at $t=$:

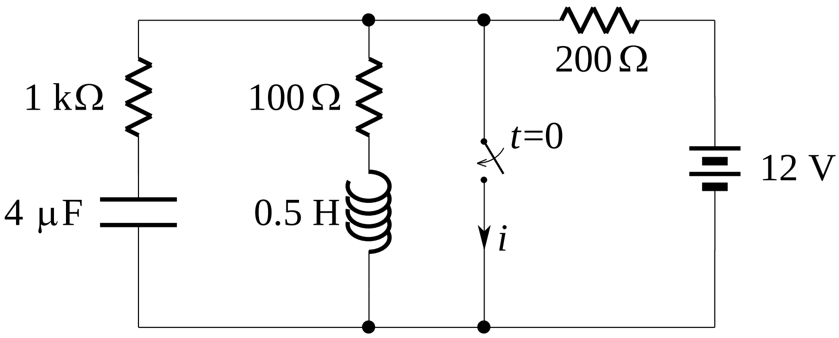
(a) -0.08 s (b) +0.08 s

11.

The voltage across the resistor in a simple source-free series *RL* circuit is $v_R(t)=50e^{-400t}$ V for $t<0$. If the value of resistance changes from $200\ \Omega$ to $40\ \Omega$ at $t=0$ when a second resistor is placed in parallel with it, find $v_R(t)$ for $t>0$.

12.

The switch in the circuit shown below has been open for a long time.



It closes at $t = 0$. Find $i(t)$.

Hint: Superposition can be very useful.

9 Nonlinear Op-Amp Applications

Contents

9.2

Introduction

Nonlinear op-amp circuits play a major role in modern electronics. Examples include comparators, precision rectifiers, peak detectors, limiters and clamps. Each of these circuits can be used as a "building block" in the creation of more advanced signal conditioning circuitry found in a large variety of applications, such as: communication receivers, automatic gain control circuits, oscillators and waveform generators.

In non-linear op-amp circuits the non-linearity can be provided by either (or both) of the following:

- the op-amp's transfer characteristic, i.e. saturation
- a non-linear element, such as a diode or transistor

An op-amp operated *open-loop* can be configured as a comparator – a circuit that can compare one voltage with another, and provide, in effect, a digital output – either "high" or "low".

Precision rectifiers are used in instrumentation and communication systems where the forward voltage drop of a silicon diode (approx. 0.7 V) is larger than the signal being processed. In this case the op-amp's extremely large gain causes the forward diode drop to be effectively reduced to zero, thus allowing precision rectification of signals in the mV range.

Peak detectors are used in a wide variety of instrumentation and communication systems, and again the op-amp allows precision detection.

A limiter circuit can be used to limit a signal to within a certain range – such a circuit can be used to protect following circuitry from overload conditions, and they are often used in signal generation circuitry.

Clamp circuits are used to restore the DC component of an incoming AC waveform. Such circuits find utility in communication systems.

9.1 The Comparator

The op-amp in an open-loop configuration can be used as a basic comparator. When two inputs are applied to the open-loop op-amp, it attempts to amplify the difference, but because the gain is so large, it will *saturate* close to one of the supply voltages (within 1 or 2 V), depending on whether the difference was positive or negative. Thus, the output of a comparator has two voltage levels, either "high" or "low" – it is not linearly proportional to the input voltage.

A basic non-inverting comparator is shown below:

A basic non-inverting comparator

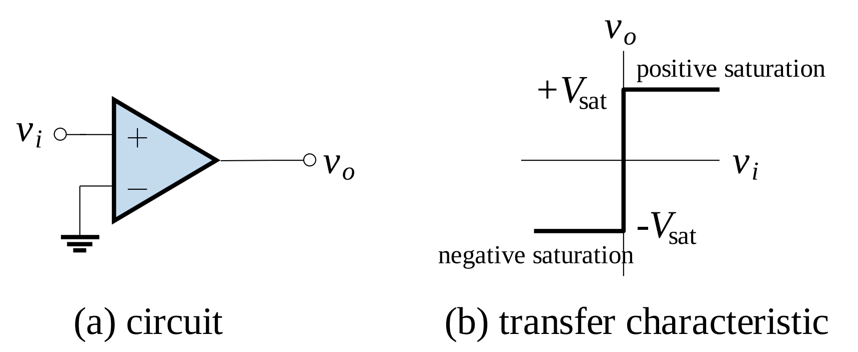


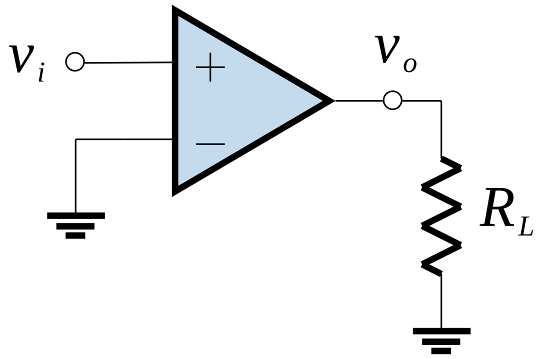
Figure 9.1

Although general purpose op-amps (like the TL071) can be used as comparators, specially designed comparator ICs (like the LM311) can switch faster and have additional features not found on general-purpose op-amps. Some applications of a comparator are:

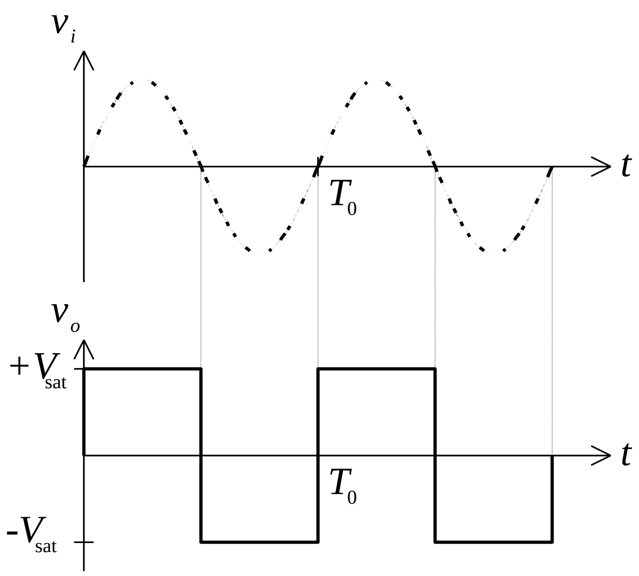
1. Zero crossing detector
2. Level detector
3. Window detector
4. Pulse width modulator
5. Pulse generator

EXAMPLE 9.1 Zero Crossing Detector

The basic noninverting comparator can be used as a zero crossing detector. A typical circuit for such a detector is shown below:



During the positive half-cycle, the input voltage is positive, hence the output voltage is $+V_{sat}$. During the negative half-cycle, the input voltage is negative, hence the output voltage is $-V_{sat}$. Thus the output voltage switches between $+V_{sat}$ and $-V_{sat}$ whenever the input signal crosses the zero level:



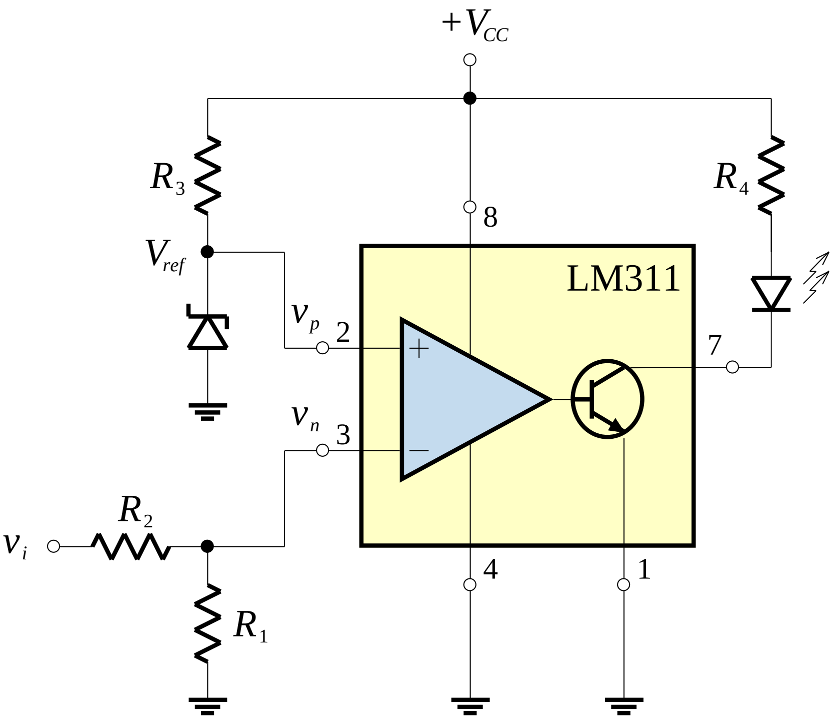
Looking at the waveform shown above, we realize that a zero crossing detector can be used as a sine- to square-wave converter.

This is an impractical circuit, since any noise on the input waveform near the zero crossings will cause multiple level transitions in the output signal.

EXAMPLE 9.2 Level Detector

We would like to monitor a 12 V car battery by having an LED on continuously when the battery voltage is above 10 V. We have a +10 V supply available.

The circuit we would use contains an LM311 comparator:



It consists of a voltage reference to establish a stable switching level, a voltage divider to scale the desired input level to the value of the voltage reference, and a voltage comparator to compare the two.

We choose a Zener diode (with $V_Z = 2.5\,\text{V}$ and $I_Z = 23\,\text{mA}$) to provide a reference voltage. The LED has $V_{LED} = 1.8\,\text{V}$ at $I_{LED} = 3\,\text{mA}$. We calculate the Zener diode's current setting resistor as follows:

$$R_3 = \frac{V_{CC} - V_Z}{I_Z} = \frac{10 - 2.5}{23} = 0.326\ \text{k}\Omega$$
$$\approx 330\ \Omega \quad \text{(standard value)}$$

9.6

Careful examination of the LM311 datasheet reveals that the output transistor will be turned ON when $v_p < v_n$. The voltage at the inverting terminal of the comparator is given by the voltage divider rule:

$$v_n = \frac{R_1}{R_1 + R_2} v_i$$

and the voltage at the noninverting terminal of the comparator is just $v_p = V_Z$.

The threshold voltage of the comparator is given by $v_n = v_p$, or:

$$\frac{R_1}{R_1 + R_2} v_i = V_Z$$

$$v_i = \left(1 + \frac{R_2}{R_1}\right) V_Z$$

Solving for the resistor ratio we have:

$$\frac{R_2}{R_1} = \frac{v_i}{V_Z} - 1$$

The threshold input voltage is given as 10 V, and so:

$$\frac{R_2}{R_1} = \frac{10}{2.5} - 1 = 3$$

We therefore choose $R_1 = 1\,\text{k}\Omega$ and $R_2 = 3\,\text{k}\Omega$.

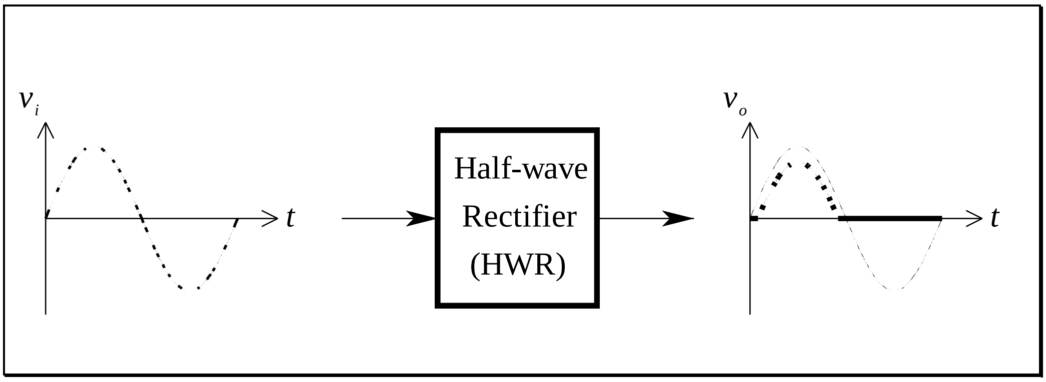
From the datasheet of the LM311, we obtain the saturation voltage of the output transistor as $V_{CE\text{sat}} \approx 0.3\,\text{V}$. The current limiting resistor for the LED is then given by:

$$R_4 = \frac{V_{CC} - V_{LED} - V_{CE\text{sat}}}{I_{LED}} \approx \frac{10 - 1.8 - 0.3}{3} = 2.633\ \text{k}\Omega$$

$$\approx 2.7\ \text{k}\Omega \quad \text{(standard value)}$$

9.2 Precision Rectifiers

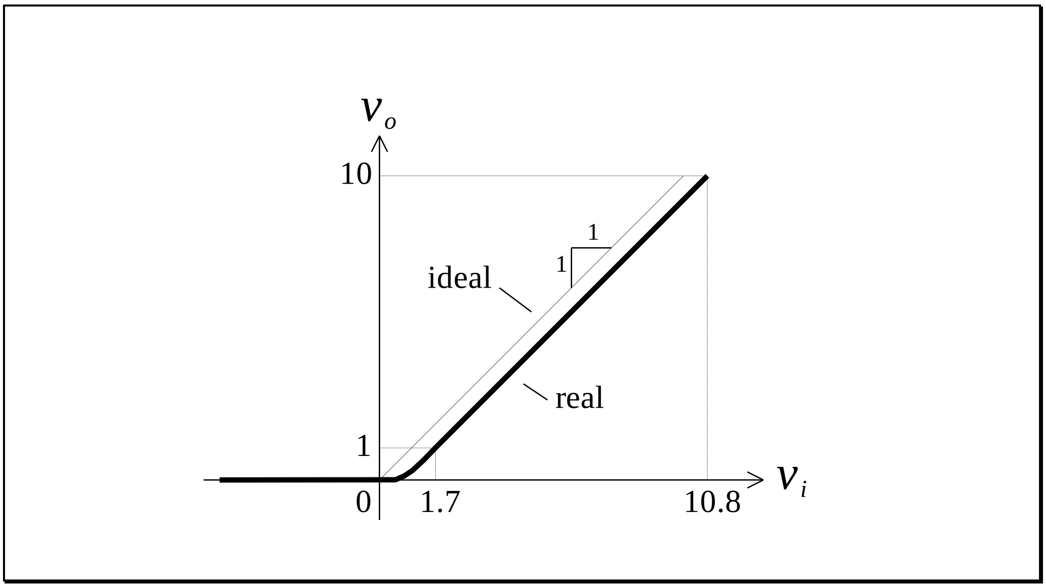
Rectifier circuits can be implemented with silicon junction diodes. Recall that for the diode to conduct appreciably, the voltage across it must be ≈ 0.7 V. Therefore, a major limitation of these circuits is that they cannot rectify voltages below about 0.7 V. In addition, since the input voltage has to rise to about 0.7 V before any appreciable change can be seen at the output, the output is distorted:



The output of a half-wave rectifier is distorted due to real diode characteristics

Figure 9.2

The transfer characteristic of a half-wave rectifier with a real silicon junction diode is shown below:



The transfer characteristic of a half-wave rectifier with real diodes

Figure 9.3

Notice the effect of the finite voltage drop of the diode. To achieve precision rectification we need a circuit that keeps v_o equal to v_i for $v_i > 0$. Such circuits can be made using op-amps and are known as *precision rectifiers*.

9.2.1 The Superdiode

The figure below shows a precision half-wave rectifier circuit consisting of a diode placed in the negative-feedback path of an ideal op-amp, with R_L being the rectifier load resistance:

A precision half-wave rectifier using a superdiode

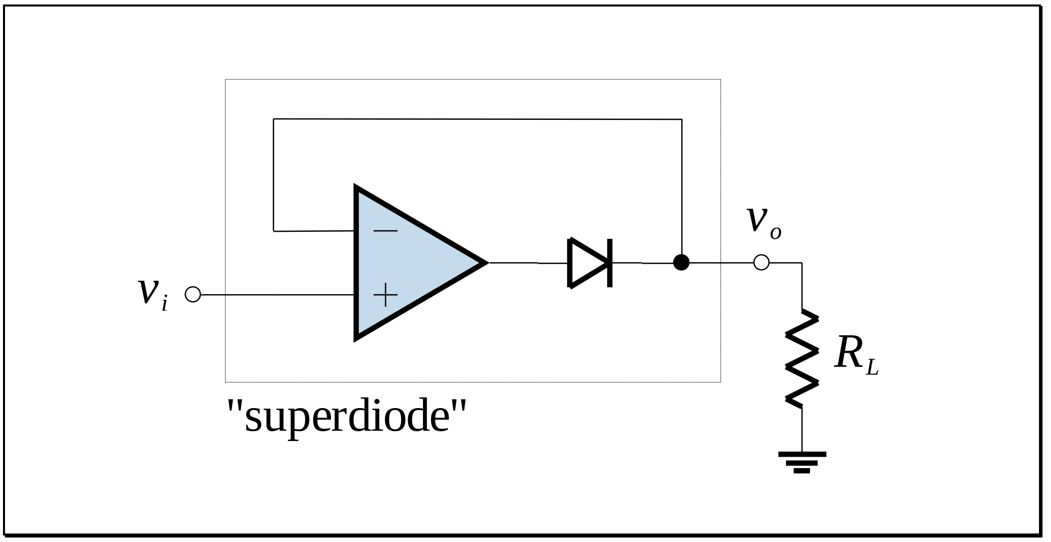


Figure 9.4

If v_i is positive, then assuming there is a virtual short-circuit across the op-amp input terminals (due to a negative feedback path around the op-amp), then the input voltage appears at the output, and v_o equals v_i. In this case the op-amp supplies the load current through the diode, and the output of the op-amp is $v_o + v_d$.

With a real op-amp with a finite open-loop gain, A_{OL}, negative feedback will work to ensure that the output of the op-amp is:

$$A_{OL}(v_i - v_o) = v_o + v_d \tag{9.1}$$

Therefore the output voltage is:

$$v_o = \frac{1}{A_{OL}+1}(A_{OL}v_i - v_d) \tag{9.2}$$

Since $A_{OL} \approx 100\,000$ for a typical op-amp, the transfer characteristic is nearly perfect, with a slope of $\frac{A_{OL}}{A_{OL}+1} \approx 1$, and with $v_o > 0$ when $v_i > v_d / A_{OL} \approx 0$. In other words, the rectified output appears when v_i exceeds a negligibly small voltage equal to the diode drop divided by the op-amp open-loop gain. Thus, for $v_i > 0$, $v_o \approx v_i$ (to a high degree of precision).

Now consider the case when v_i goes negative, and assume that the diode is reverse-biased ("off"). Since there is no current, the output voltage is $v_o = 0$, the voltage appearing across the op-amp input terminals is negative, and the op-amp saturates close to its negative supply rail. A quick check of the assumption that the diode is reverse-biased under these conditions reveals that our analysis is correct. Thus, for $v_i < 0$, $v_o = 0$.

The transfer characteristic of the circuit will be almost identical to the ideal characteristic of a half-wave rectifier. The non-ideal diode characteristic has been almost completely masked by placing it in the negative-feedback path of an op-amp. This is another dramatic application of negative feedback. The combination of diode and op-amp, shown in the dotted box in Figure 9.4, is appropriately referred to as a *superdiode*.

The circuit does have serious disadvantages which make it impractical. When v_i goes negative and $v_o = 0$, the entire magnitude of v_i appears between the two input terminals of the op-amp. If this magnitude is greater than a few volts, the op-amp may be damaged unless it is equipped with what is called "overvoltage protection" (a feature that most modern IC op-amps have). Another disadvantage is that when v_i is negative the op-amp will be saturated close to its negative supply rail. Although not harmful to the op-amp, saturation should usually be avoided, since getting the op-amp out of saturation and back into its linear region of operation takes some time. This time delay is determined by the op-amp's slew rate, and even a very fast op-amp will limit the circuit to a low frequency range of operation.

9.2.2 Precision Inverting Half-Wave Rectifier

An alternative precision half-wave rectifier circuit that does not suffer from the disadvantages of the previous circuit is shown below:

A precision half-wave rectifier

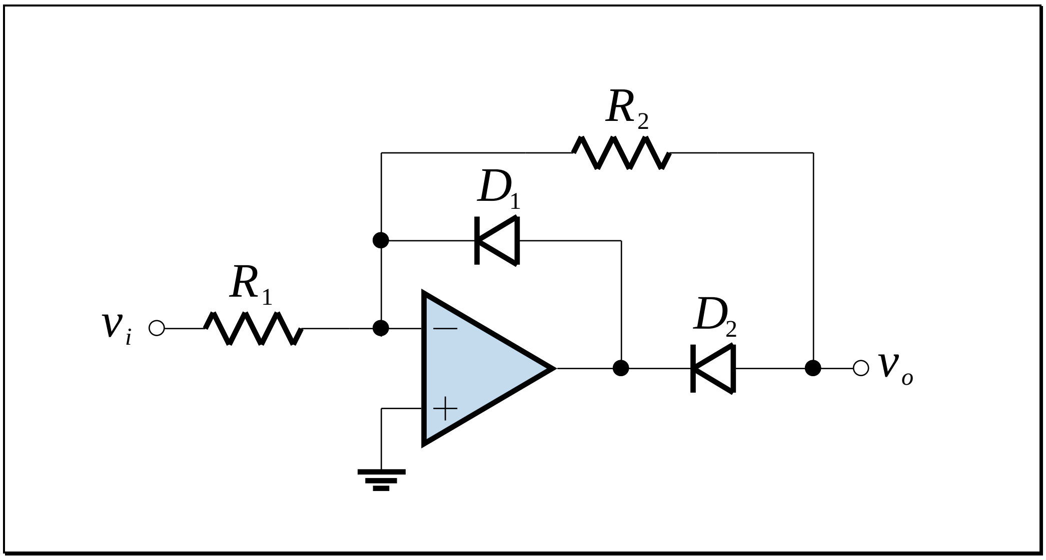


Figure 9.5

To analyze this circuit, we firstly remember that the diode is a nonlinear element so that linear circuit analysis does not apply. We will assume that the op-amp has a negative feedback path around it so that there is a virtual short circuit at its input. Thus, the inverting terminal will be kept at 0 V. We will carry out the analysis with this assumption, and then check that the assumption is true.

Due to the presence of diodes, we will consider two cases: one where the input voltage is positive, the other when the input voltage is negative.

In the positive half cycle ($v_i > 0$), direct analysis on the circuit diagram gives:

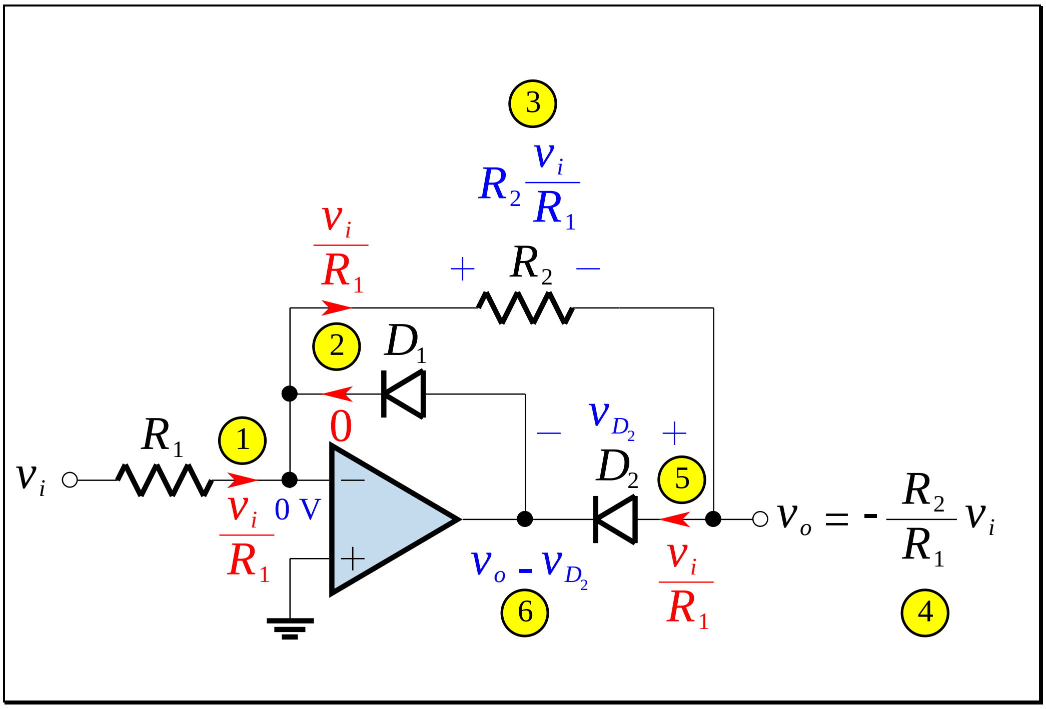


Figure 9.6

The analysis steps are:

1. We assume an ideal op-amp, and also assume that there is a negative feedback path around the op-amp – it is producing a finite output voltage (i.e., the overall amplifier is "working"). Thus, the ideal op-amp must have a virtual short-circuit (VSC) at its input terminals. Since there is no difference in the voltages across the VSC, the voltage at the inverting terminal is $v_- = 0$. The current through resistor R_1, in the direction shown, is given by Ohm's Law, $i_1 = v_i / R_1$.

2. Due to the infinite input resistance of the ideal op-amp, the current entering the inverting terminal is 0 A. We also assume that diode D_1 is "off" (an assumption that we will check later). KCL at the inverting terminal now gives $i_2 = i_1 = v_i / R_1$.

3. The voltage drop across the resistor R_2 is given by Ohm's Law, $v_{R_2} = R_2 i_2 = R_2 \frac{v_i}{R_1}$, with the polarity shown.

4. KVL, from the common, across the VSC, across R_2 and to the output terminal gives $v_o = 0 - R_2 \frac{v_i}{R_1} = -\frac{R_2}{R_1} v_i$.

5. We assume that diode D_2 is "on" (an assumption that we will check later). KCL at the output then gives $i_{D_2} = i_2 = v_i / R_1$. The diode "on" voltage drop is then given by its characteristic, and is labelled v_{D_2}.

6. The output of the op-amp is, by KVL, $v_o - v_{D_2}$, which is a negative voltage (remember that $v_o = -\frac{R_2}{R_1} v_i$ and $v_i > 0$). Thus, diode D_1 is indeed reverse-biased, and our original assumption about it being "off" is correct. Also, the diode D_2 is indeed forward-biased, thus establishing a negative-feedback path around the op-amp and forcing a virtual common to appear at the inverting input terminal. That is, negative feedback works to ensure that the op-amp output voltage is maintained at $v_o - v_{D_2}$.

Thus, in the positive half cycle ($v_i > 0$), the output of the circuit is:

$$v_o = -\frac{R_2}{R_1} v_i \tag{9.3}$$

That is, the output is negative for positive inputs.

In the negative half cycle ($v_i < 0$), direct analysis on the circuit diagram gives:

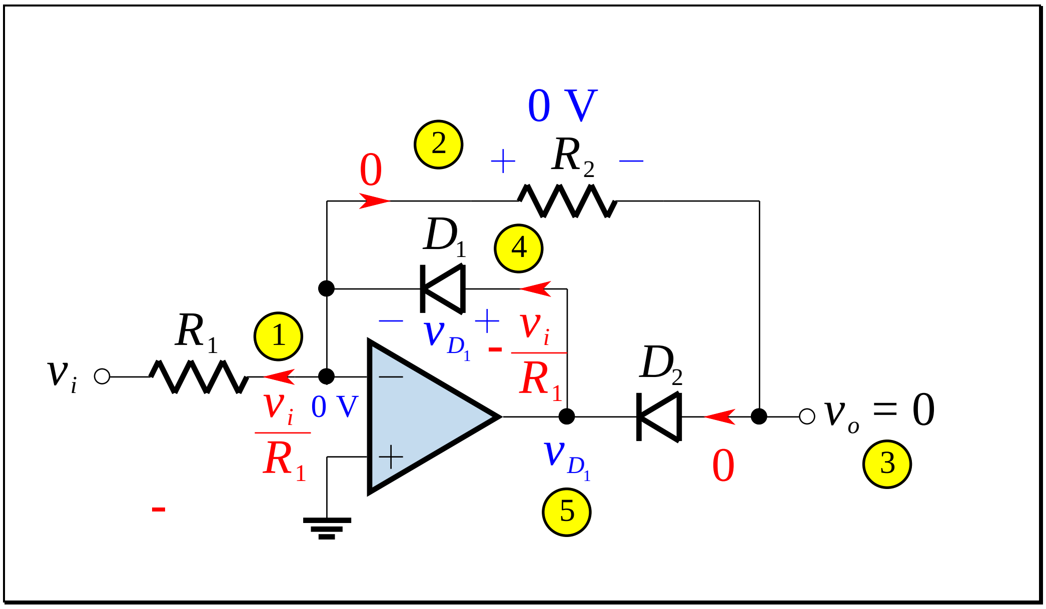


Figure 9.7

The analysis steps are:

1. With the usual assumption that the ideal op-amp has a negative feedback path, then a virtual short-circuit (VSC) exists at its input terminals. The voltage at the inverting terminal is $v_- = 0$. The current through resistor R_1, in the direction shown, is given by Ohm's Law, $i_1 = -v_i / R_1$. Note that this current is *positive* since $v_i < 0$.

2. We assume that diode D_2 is "off" (an assumption that we will check later). Then there is no current in resistor R_2, and hence $v_{R_2} = 0$.

3. KVL, from the common, across the VSC, across R_2 and to the output terminal gives $v_o = 0$.

4. We assume that diode D_1 is "on" (an assumption that we will check later). Due to the infinite input resistance of the ideal op-amp, the current entering the inverting terminal is 0 A. KCL at the inverting terminal now gives $i_{D_1} = i_1 = -v_i / R_1$. The diode "on" voltage drop is then given by its characteristic, and is labelled v_{D_1}.

5. The output of the op-amp is, by KVL, $0 + v_{D_1}$, which is a positive voltage. Thus, diode D_1 is indeed forward-biased, and our original assumption about it being "on" is correct. Diode D_1 thus establishes a negative-feedback path around the op-amp and forces a virtual common to appear at the inverting input terminal. That is, negative feedback works to ensure that the op-amp output voltage is maintained at v_{D_1}. Also, the diode D_2 is indeed reverse-biased.

Thus, in the negative half cycle ($v_i < 0$), the output of the circuit is:

$$v_o = 0 \tag{9.4}$$

The transfer characteristic of the circuit is shown below for the case $R_1 = R_2$:

The transfer characteristic of a precision inverting half-wave rectifier

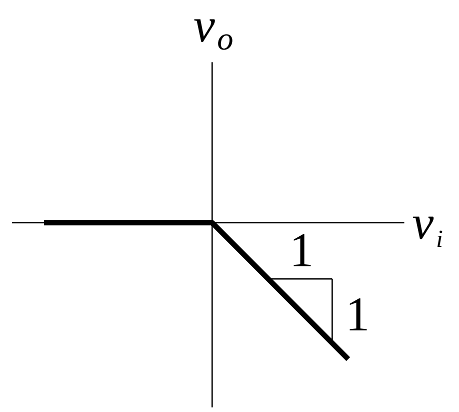


Figure 9.8

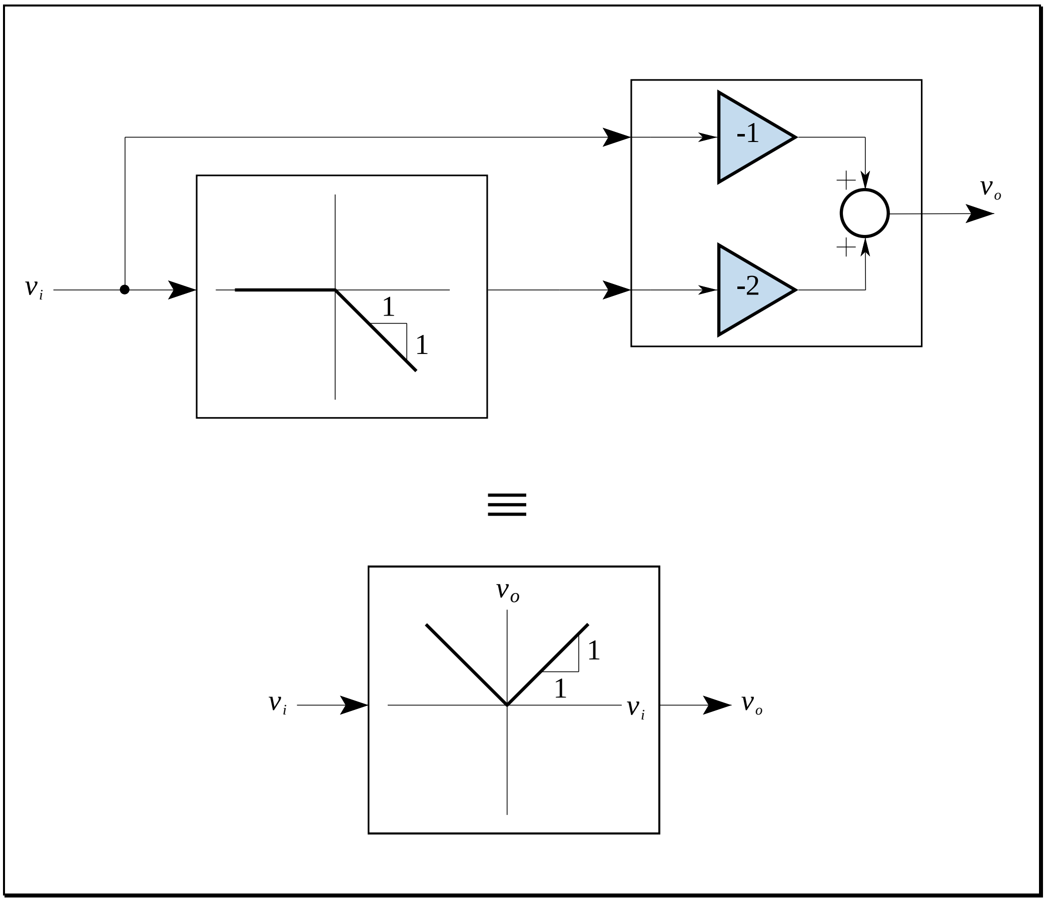
The circuit is seen to be a precision *inverting* half-wave rectifier. Note that unlike the previous circuit, here the slope of the characteristic can be set to any desired value, including unity, by selecting appropriate values for R_1 and R_2.

The major advantage of this circuit is that the feedback loop around the op-amp remains closed at all times. Hence the op-amp remains in its linear operating region, avoiding the possibility of saturation and the associated time delay required to "get out" of saturation.

9.2.3 Precision Full-Wave Rectifier

There are many applications in instrumentation where the information provided by both halves of an AC signal is either useful or necessary, but the signal must be converted into a unipolar waveform. This need can be met by using a precision full-wave rectifier.

There are many possible arrangements for implementing a precision full-wave rectifier. One possible arrangement is depicted symbolically below:



Block diagram of a precision full-wave rectifier

Figure 9.9

As shown, the block diagram consists of two boxes: a precision half-wave rectifier, and a weighted inverting summer. *Convince yourself that this block diagram does in fact realize a precision full-wave rectifier.*

9.16

An implementation of a precision full-wave rectifier

An implementation of the block diagram is shown below:

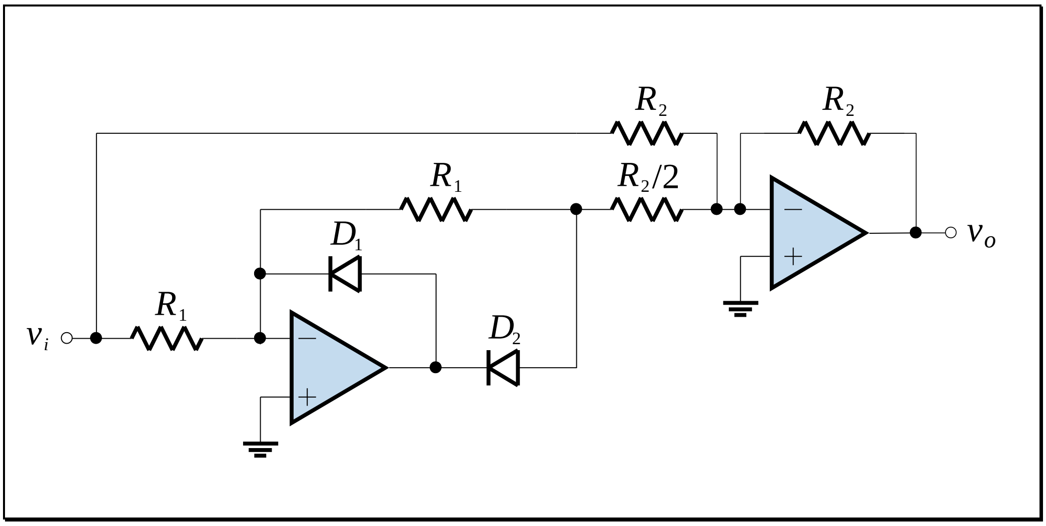


Figure 9.10

The output, which is "buffered" since it comes from the output of an op-amp, is the absolute value of the input voltage:

$$v_o = |v_i| \tag{9.5}$$

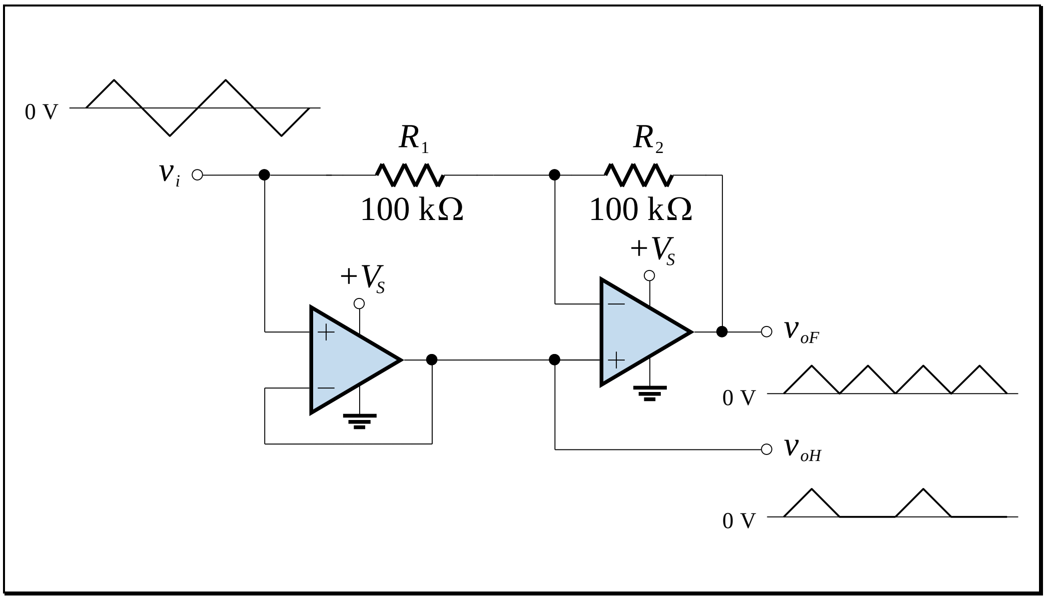
The input resistance of the circuit is:

$$R_{in} = R_1 \parallel R_2 \tag{9.6}$$

If desired, a voltage follower can be placed at the input to buffer the incoming signal.

9.2.4 Single-Supply Half-Wave and Full-Wave Rectifier

There are a number of ways to construct half- and full-wave rectifiers using combinations of op-amps and diodes, but the circuit shown below requires only a dual op-amp, two resistors, and operates on a single supply:



A single-supply half-wave and full-wave rectifier

Figure 9.11

The circuit will work on any single supply op-amp whose inputs can withstand being pulled below 0 V. In addition, the op-amps need to have an output that is capable of swinging “rail-to-rail”, which means that the output can go within a few millivolts of the supply rails under light loading.

When the input signal is above 0 V, the unity-gain follower presents the input signal to the noninverting input of the second op-amp. The feedback around the second op-amp creates a virtual short-circuit across its input terminals, and subsequently the inputs are equal. Thus, there is no voltage across resistor R_1, and no current in R_1 and R_2. The output v_{oF} therefore “tracks” the input.

Conversely, when the input is negative, the output of the first op-amp is forced to zero (it saturates at the limit of its supply). The noninverting input of the second op-amp see the 0 V output, and during this phase operates as a unity-gain inverter, rectifying the negative portion of the input, v_i.

The net output at v_{oF} is therefore a full-wave rectified version of v_i. In addition, a half-wave rectified version is obtained at the output v_{oH} if desired.

9.3 Peak Detector

A peak detector is a circuit that produces an output voltage equal to the positive or negative peak value of the input voltage waveform. The simplest form of positive peak detector is shown below:

A positive peak detector…

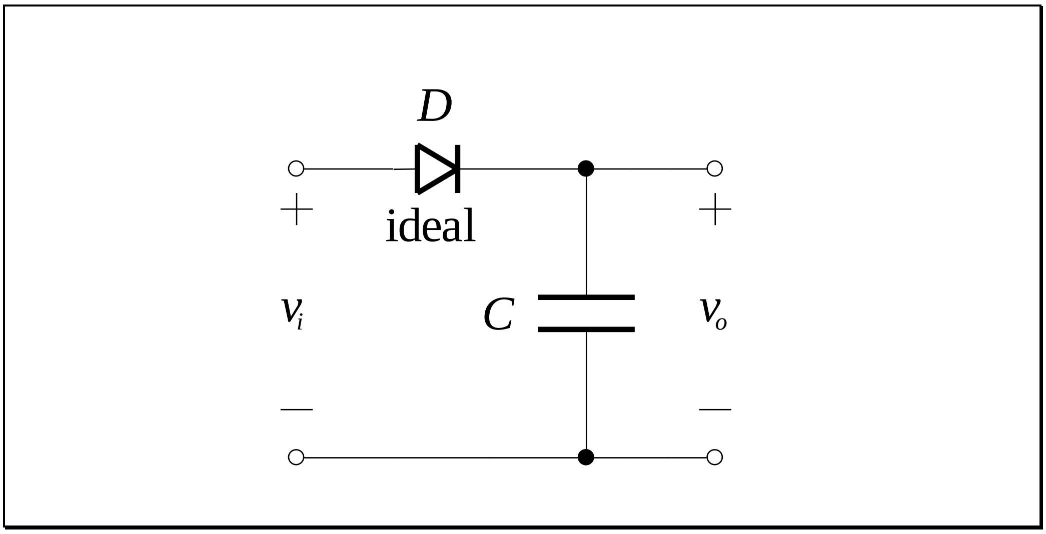


Figure 9.12

The ideal diode allows the capacitor to charge, but not to discharge. Therefore, the capacitor will retain the positive peak value of the input waveform:

… and its input and output waveforms

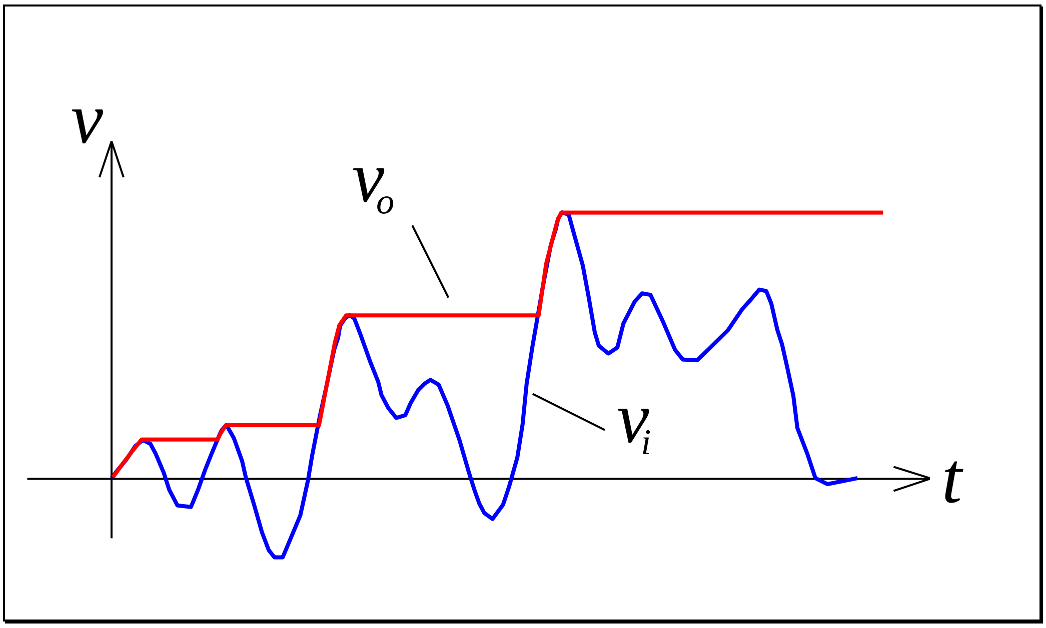
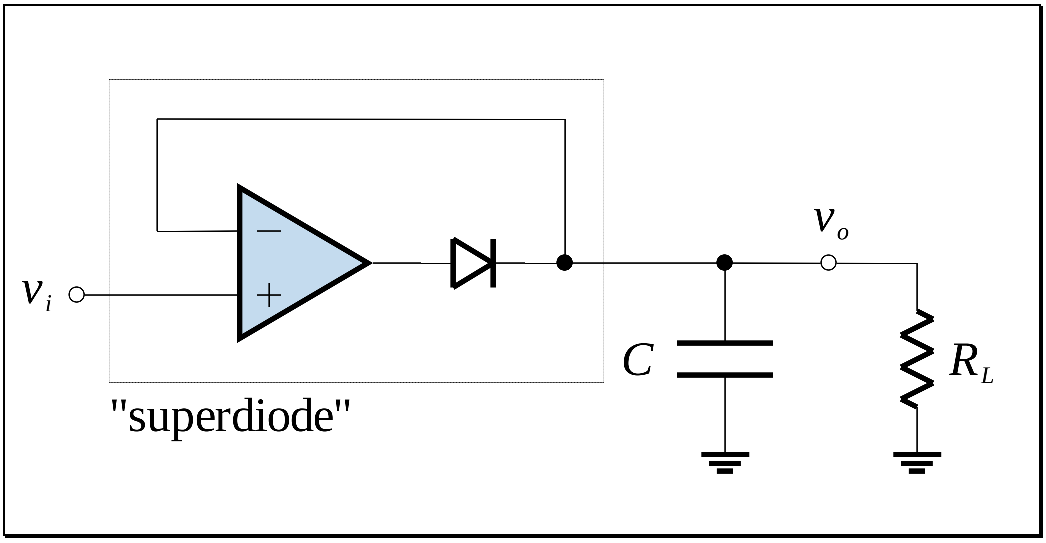


Figure 9.13

A precision peak detector is implemented using a superdiode and a capacitor:



A precision positive peak detector with a decaying output voltage

Figure 9.14

Here we have represented the input resistance of the following circuit as a load resistance, R_L.

For v_i positive and greater than the output voltage the op-amp will drive the diode on, thus closing the negative-feedback path and causing the op-amp to act as a follower. The output voltage will therefore follow that of the input, with the op-amp supplying the capacitor charging current. This process continues until the input reaches its peak value.

Now consider what happens if the input signal falls below the peak value stored on the capacitor. In this case $v^+ < v^-$ at the op-amp's input terminals, and the op-amp enters negative saturation, reverse-basing the diode. The superdiode is effectively in the "off" state and the capacitor will discharge through the load resistance R_L:

$$v_o = \hat{v}_i e^{-\frac{t}{R_L C}} \tag{9.7}$$

The rate of decay of the output voltage is therefore dictated by the capacitor value and the attached load. This decay in output voltage is sometimes desirable – inclusion of a load resistance is essential if the circuit is required to detect reductions in the magnitude of the positive peak.

9.20

In other applications, we may wish to retain the peak voltage for a long period of time. In these cases, we should use a circuit like that shown below:

A precision positive peak detector which retains its output voltage for a long time

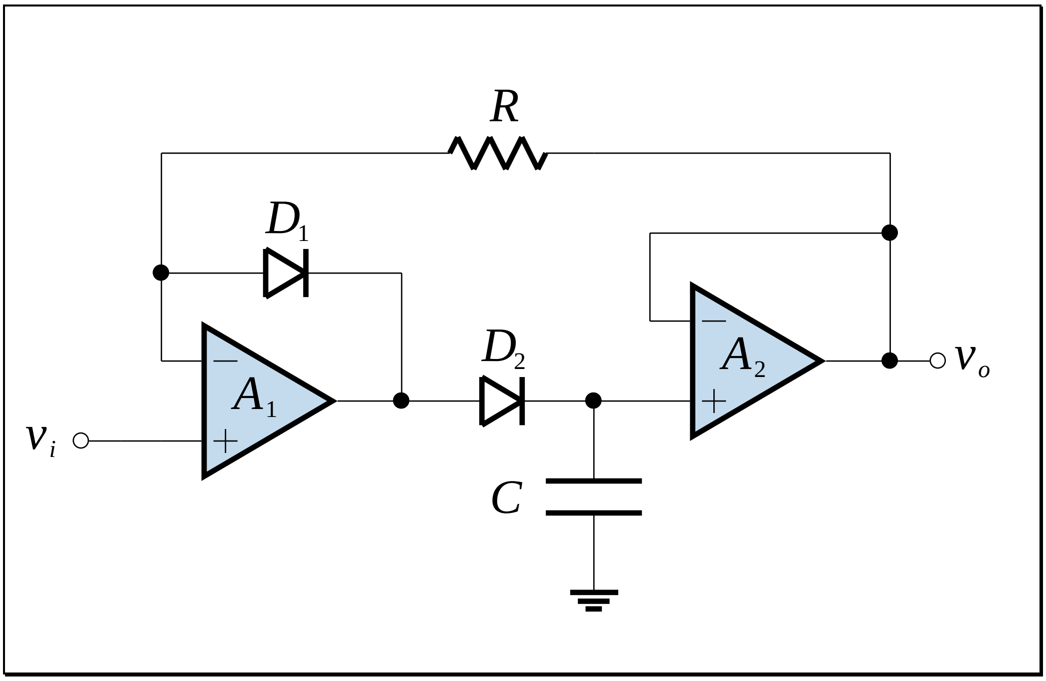


Figure 9.15

The op-amp A_1 offers a high input impedance to the source. The op-amp A_2 acts as a buffer between the capacitor and any attached load, thus preventing it from discharging. The output v_o is equal to the voltage on the capacitor, which equals the positive peak of the input voltage up to that time.

When $v_i > v_o$ we will assume that diode D_1 is "off" and diode D_2 is "on":

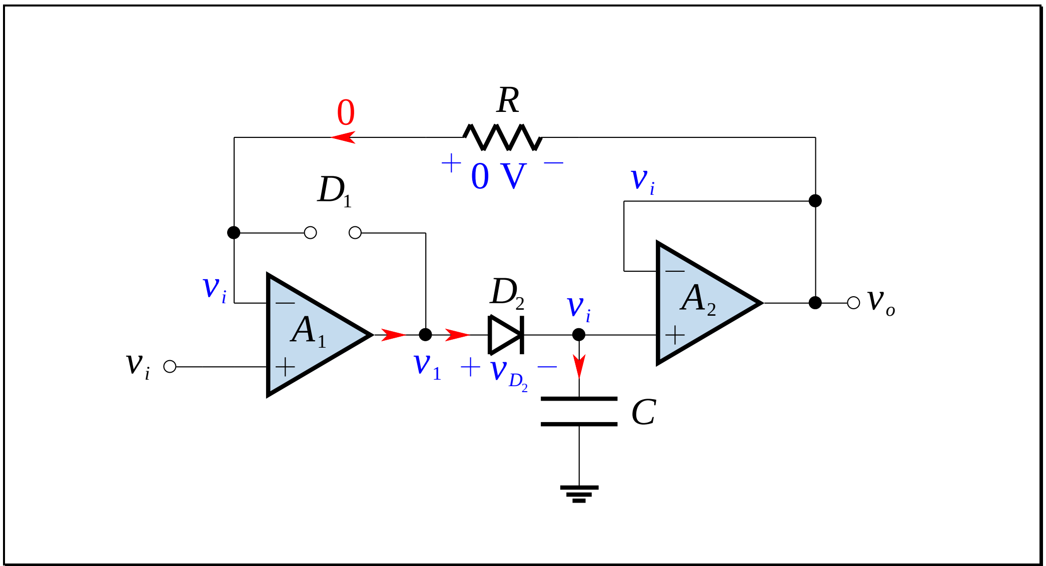


Figure 9.16

The negative feedback path around A_1 consists of diode D_2, the follower A_2, and the resistor *R*. Thus, there is a virtual short-circuit across the input terminals of A_1, and the input voltage v_i appears at the inverting terminal.

Now since diode D_1 is "off", there is no path for current through *R*. Thus the voltage across *R* is 0 V. Thus the follower A_2 must output a voltage such that:

$$v_o = v_i \tag{9.8}$$

Thanks to the virtual short-circuit across the follower A_2, this voltage must appear across the capacitor, and it therefore charges up. The op-amp A_1 provides the charging current, and its output voltage must be:

$$v_1 = v_i + v_{D_2} \tag{9.9}$$

We can now see that the assumption that diode D_1 is "off" is consistent with the voltage across it:

$$v_{D_1} = v_i - v_1 = -v_{D_2} < 0 \tag{9.10}$$

So long as $v_i > v_o$, the circuit will work in this manner, and the output v_o tracks v_i. This mode of operation is called the *track* mode.

It can be observed that placing the diode D_2 and the follower A_2 within the feedback path of A_1 eliminates the possible error due to the diode drop across D_2. We should choose an op-amp for A_2 that has low input bias currents so as to minimize the capacitance discharge. JFET input op-amps are ideal in this case. The op-amp we choose for A_1 must have a high output current capability to charge *C* during fast-occurring input voltage peaks. As will be seen shortly, neither op-amp enters saturation which means the circuit can be operated at relatively high frequencies.

9.22

When $v_i < v_o$ we will assume that diode D_1 is "on" and diode D_2 is "off":

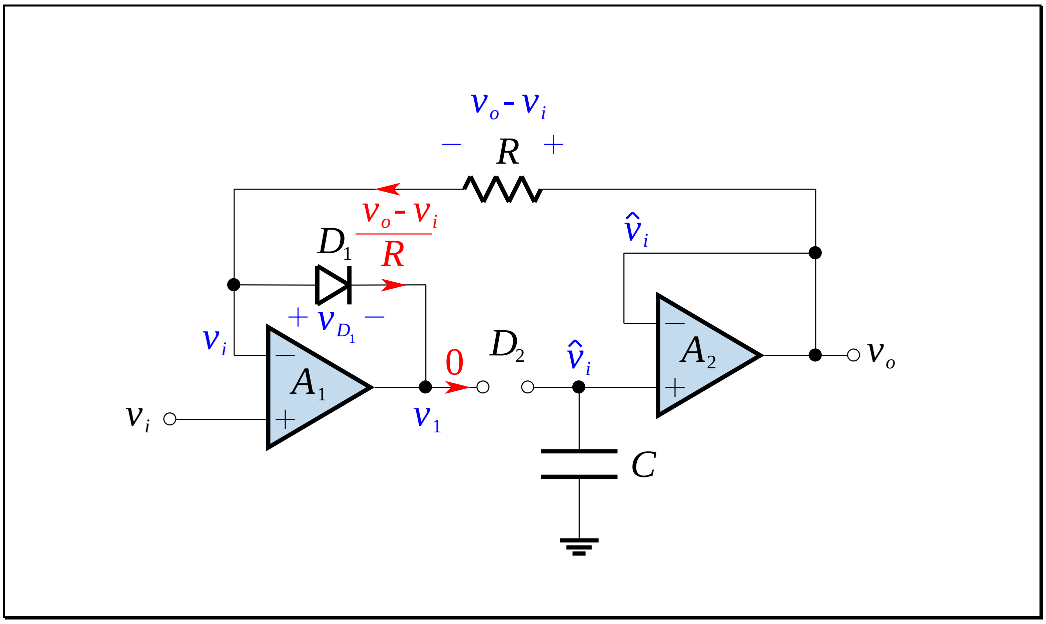


Figure 9.17

Since diode D_2 is "off", there is no path for a current to discharge the capacitor C. Thus, it retains a voltage $\hat{v}_i$, the peak of the input voltage up to that time. The output of follower A_2 is then:

$$v_o = \hat{v}_i \tag{9.11}$$

Since diode D_1 is "on", there is a negative feedback path around A_1 and so there is once again a virtual short-circuit across its input terminals such that the input voltage v_i appears at the inverting terminal. The output of A_1 is then:

$$v_1 = v_i - v_{D_1} \tag{9.12}$$

We can now see that the assumption that diode D_2 is "off" is consistent with the voltage across it:

$$v_{D_2} = v_1 - \hat{v}_i = v_i - v_{D_1} - \hat{v}_i < 0 \tag{9.13}$$

So long as $v_i < v_o$, the circuit will work in this manner, and the output v_o retains the peak value of v_i. This mode of operation is called the *hold* mode.

9.4 Limiter

Circuits which are used to clip off the unwanted portions of the input voltage above or below certain levels, so as to produce limited outputs, are called *limiters* or *clippers*.

A *double limiter* circuit works on both positive and negative peaks of an input waveform. An implementation using an op-amp is shown below:

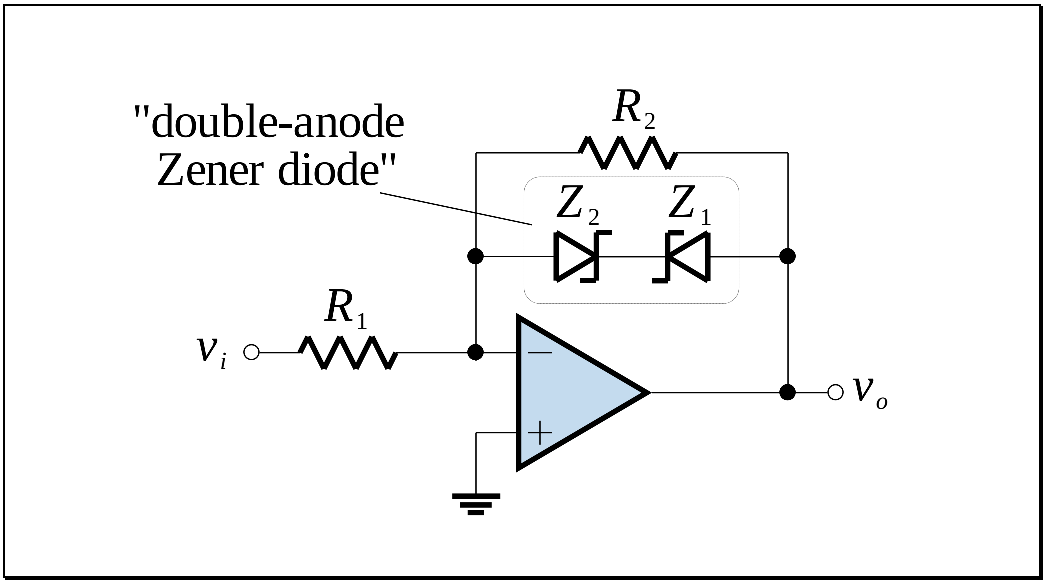


Figure 9.18

A double limiter circuit utilising a double-anode Zener diode…

Its transfer characteristic is:

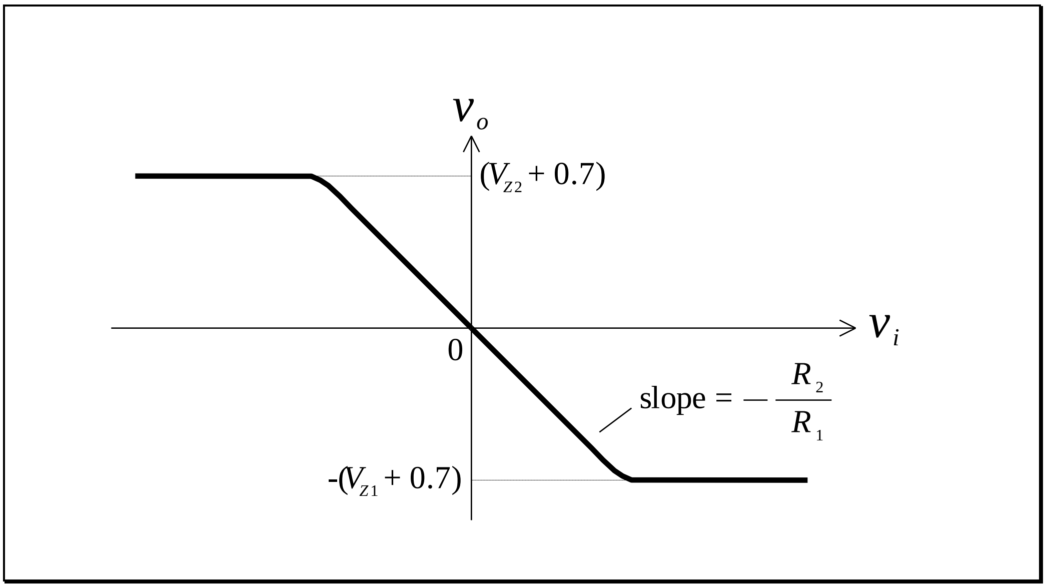


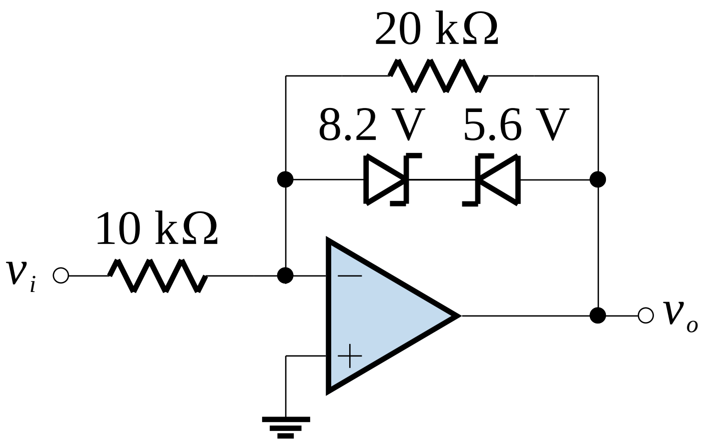
Figure 9.19

…and its transfer characteristic

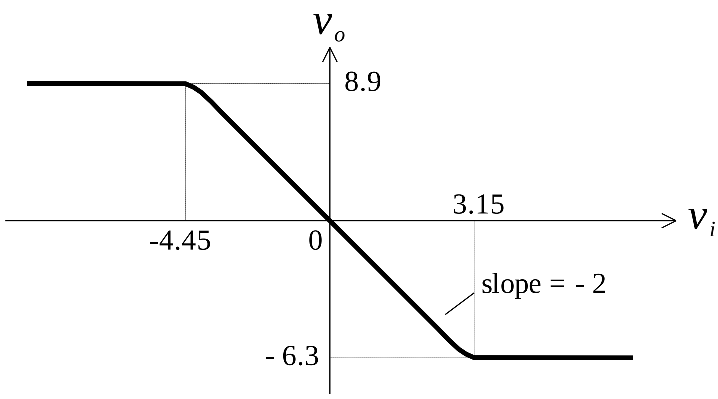
Convince yourself that the transfer characteristic will be that shown.

EXAMPLE 9.3 Asymmetric Double Limiter

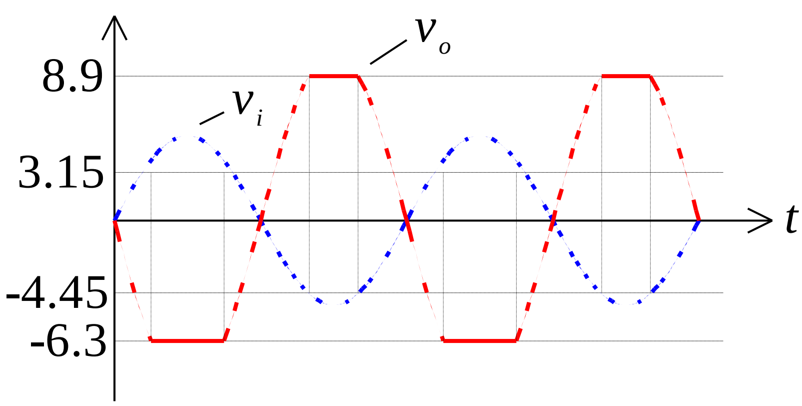
An asymmetric double limiter is shown below:



We assume that when a Zener diode conducts in the forward direction, the voltage drop is approximately 0.7 V. The circuit's transfer characteristic is:

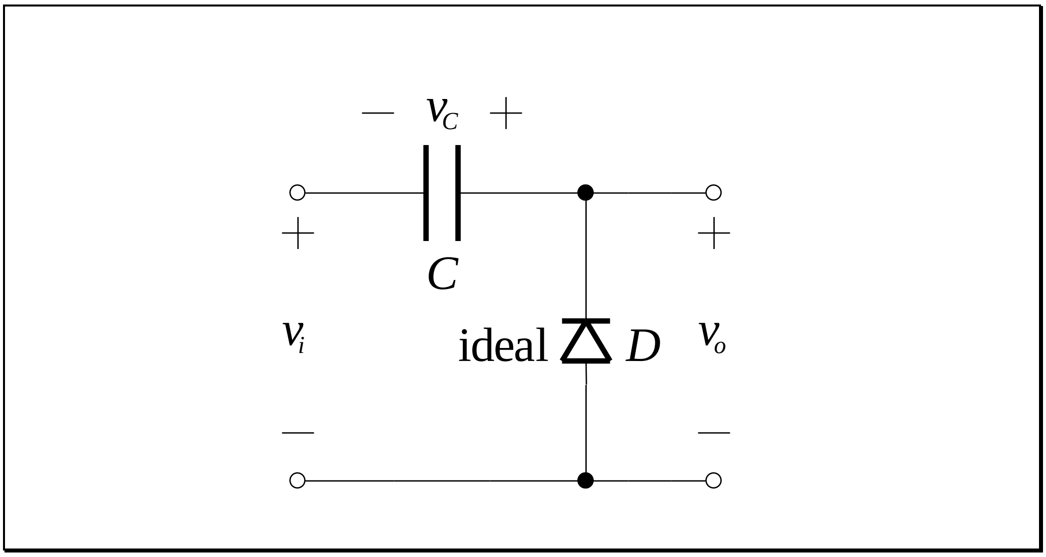


The output waveform resulting from a 5 V peak sinusoidal input is shown below:



9.5 Clamp

A clamp circuit is used to add a DC component to an AC input waveform so that the positive (or negative) peaks are forced to take a specified value – usually zero. In other words, the peaks of the waveform are "clamped" to a specified voltage value. The simplest form of *positive clamp* is shown below:



An ideal positive clamp

Figure 9.20

Because of the polarity in which the diode is connected, it will allow the capacitor to charge to a voltage v_C equal to the magnitude of the most negative peak of the input signal. Subsequently, the diode turns off and the capacitor retains its voltage indefinitely.

Now since the output voltage v_o is given by:

$$v_o = v_i + v_C \qquad (9.14)$$

it follows that the output waveform will be identical to that of the input except shifted upwards by v_C.

Another way of visualizing the operation of the circuit is to note that because the diode is connected across the output with the polarity shown, it prevents the output voltage from going below 0 V (by turning on and charging up the capacitor).

An example of input and output signals is shown below:

Input and output waveforms of an ideal positive clamp, showing DC restoration

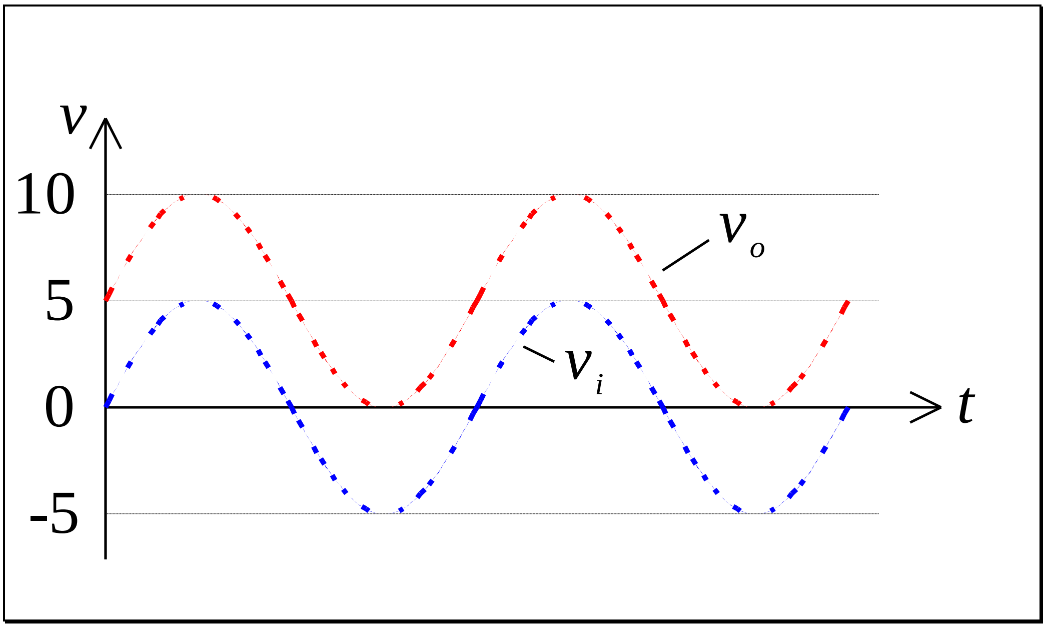


Figure 9.21

As can be seen from the figure above, another appropriate name for the circuit is a *DC restorer*. These circuits find application in communication systems.

It should be obvious that reversing the diode polarity will provide an output waveform whose highest peak is clamped to 0 V – a *negative clamp*.

If we replace the diode in the clamping circuit of Figure 9.20 by a “superdiode”, a precision positive clamping circuit is obtained:

A precision positive clamp

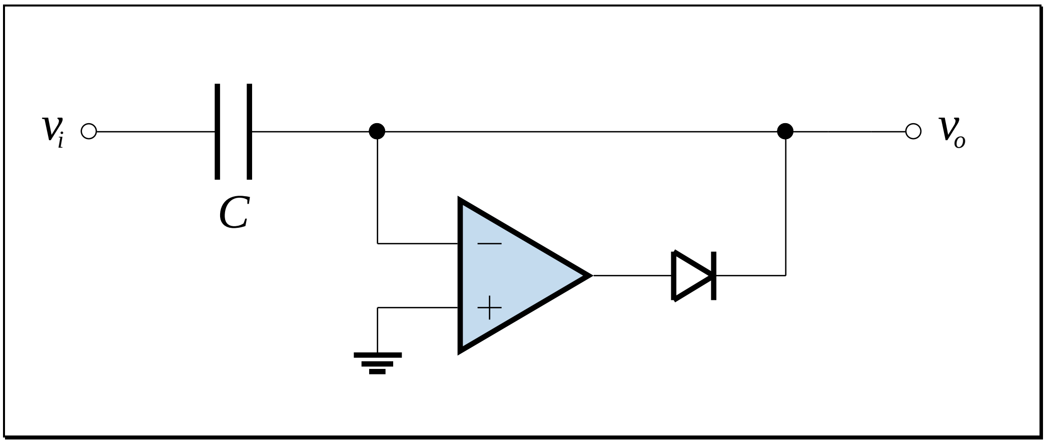
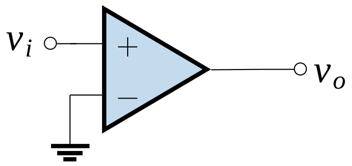


Figure 9.22

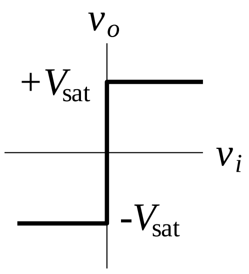
Operation of this circuit should be self-explanatory.

9.6 Summary

- The op-amp in an open-loop configuration can be used as a basic comparator:

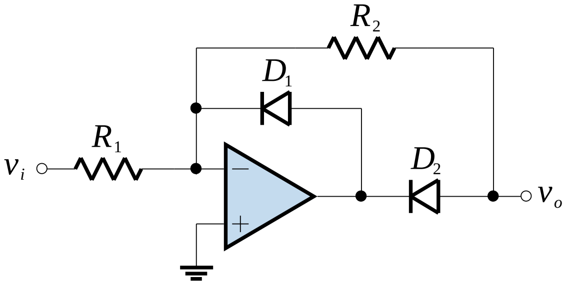


(a) circuit

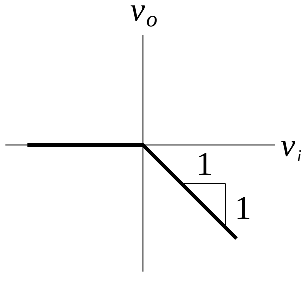


(b) transfer characteristic

- The inverting half-wave precision rectifier is used to rectify signals in the mV range:

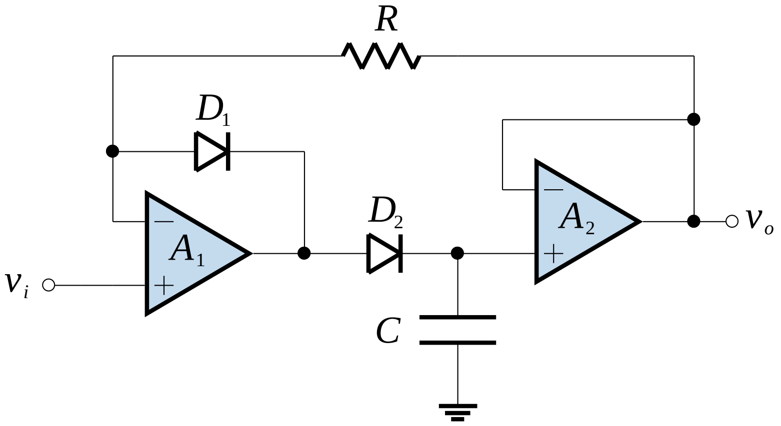


(a) circuit

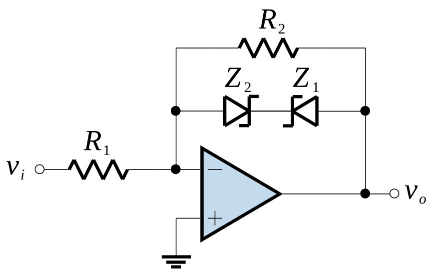


(b) transfer characteristic

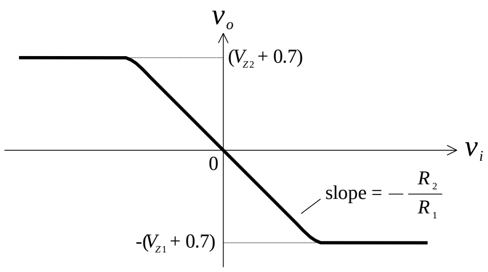
- A precision positive peak detector can retain its output voltage for a long time:



- A limiter is used to clip off the unwanted portions of the input voltage above or below certain levels, so as to produce a limited output:

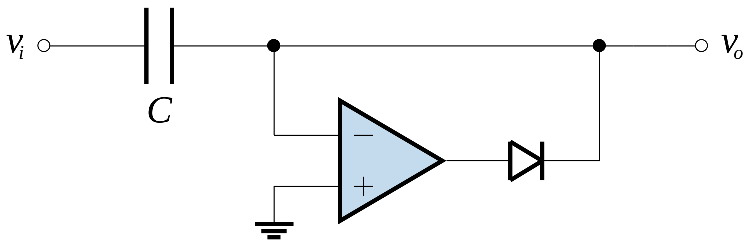


(a) circuit



(b) transfer characteristic

- A precision positive clamp is used to clamp a signal's minimum value to 0 V, thus restoring a DC component:



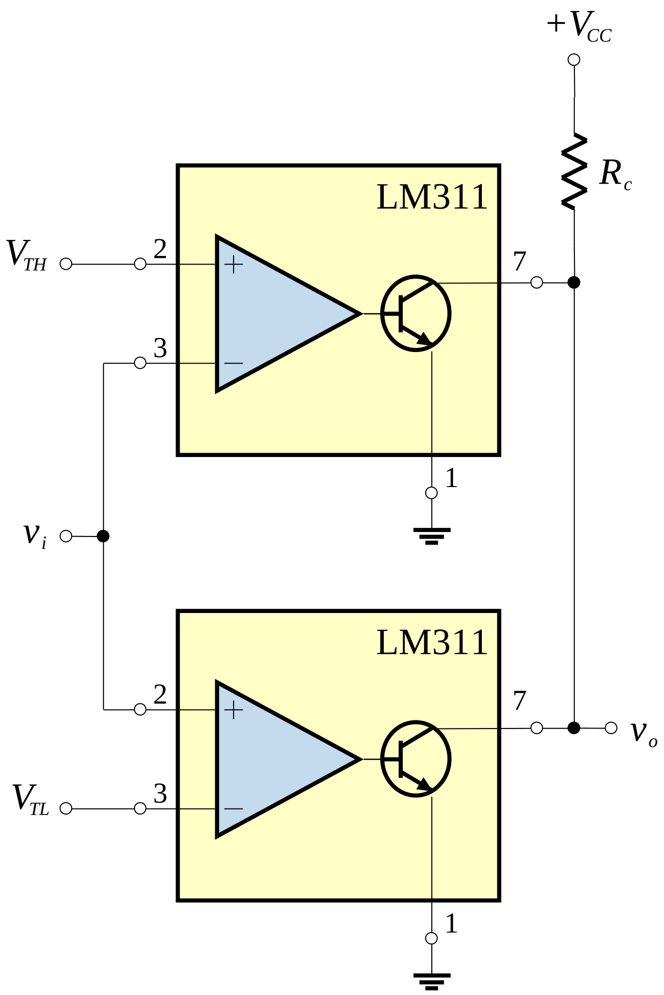
9.7 References

Sedra, A. and Smith, K.: *Microelectronic Circuits*, Saunders College Publishing, New York, 1991.

Exercises

1.

The *window detector* circuit detects when an unknown voltage falls within a specified voltage band or window. It consists of two comparators and two reference voltages V_{TL} and V_{TH} defining the lower and upper limits of the window, as shown below:



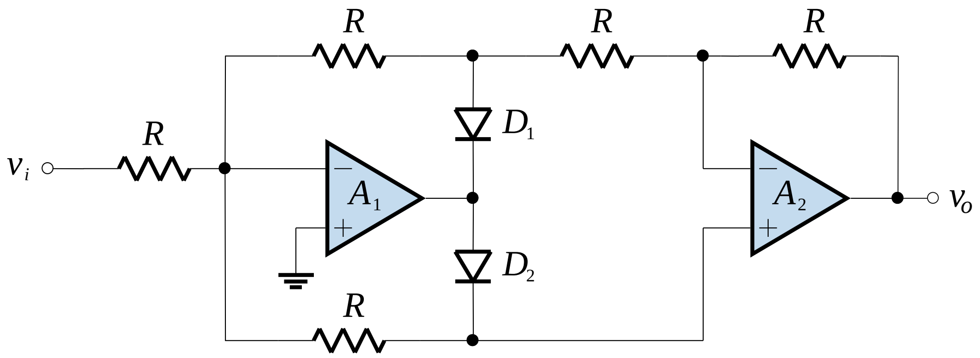
Note the use of a "wired-OR" output. If either comparator turns on then the output will be "pulled low" (i.e. $v_o \approx 0\text{ V}$). With both comparators off there is no current and the output is "pulled high" (i.e. $v_o = V_{CC}$). The resistor R_c in this case is called a "pull-up resistor".

9.30

Design a circuit to monitor an input voltage and turn an LED on when this voltage goes either above 5.5 V or below 4.5 V. Assume that $V_{LED} = 1.8\text{ V}$ and $I_{LED} = 2.\text{mA}$. A supply voltage of 12 V is available.

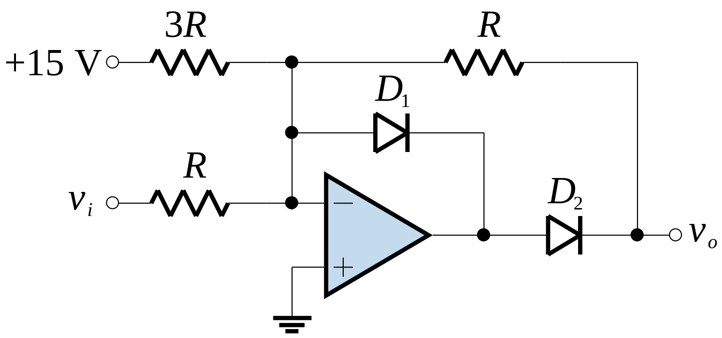
2.

Draw the transfer characteristic of the following op-amp circuit.



3.

Draw the transfer characteristic of the following op-amp circuit.

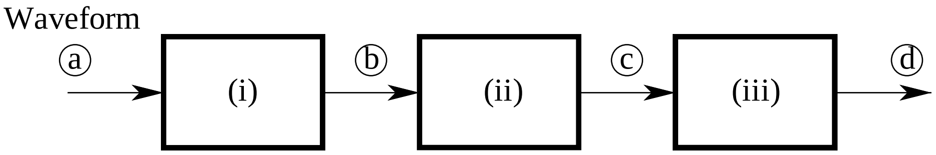


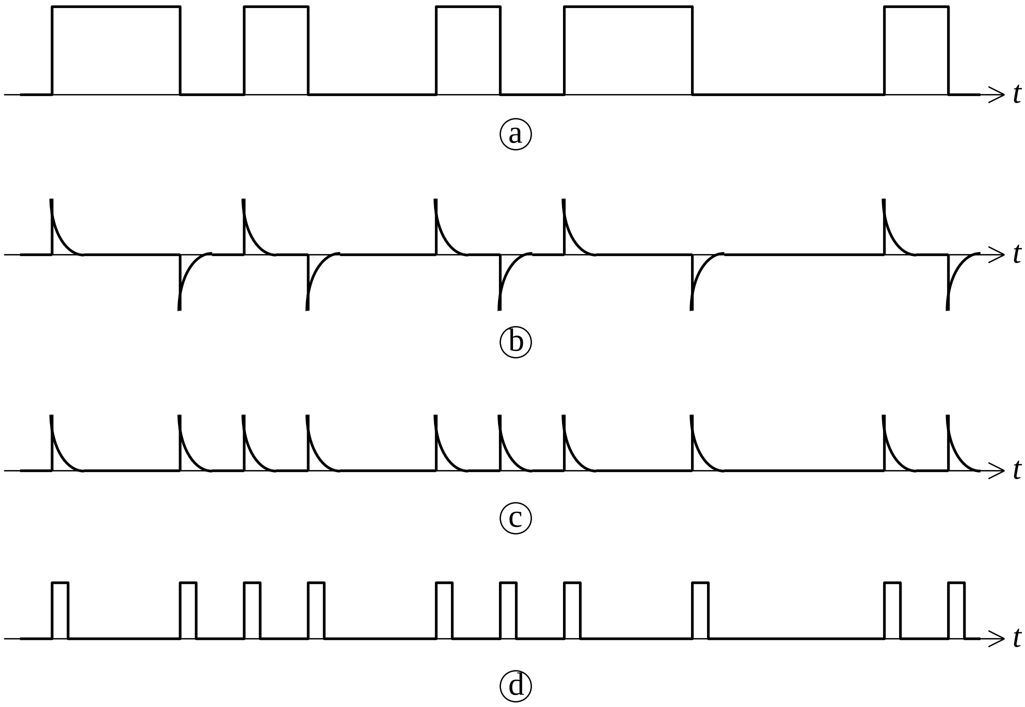
4.

It is desired to clamp a 1 kHz, 5 V peak, sinusoid so that its maximum value is 0 V. Draw the schematic of a circuit that will achieve such an operation with precision.

5.

The following block diagram shows part of a “clock recovery” circuit for a communication system. Determine the function of each block, and therefore give the blocks a name:





10 The First-Order Step Response

Contents

10.2

Introduction

The determination of the *natural* response of a source-free circuit relies solely on the configuration of the circuit elements and on any initial energy storage present in the system (capacitor voltages and inductor currents).

If we apply a DC source to a circuit, it will eventually settle to a state where all voltages and currents are DC. Such a DC source is termed a *forcing function* – it forces the *form* of the response in the circuit voltages and currents.

If we apply a sinusoidal source to a circuit, it will eventually settle to a state where all voltages and currents are sinusoidal. Thus the concept of a forcing function is applicable to other styles of sources. We shall see later that the DC source and the sinusoidal source are special cases of a general forcing function.

The application of a forcing function to a circuit will result in a "response" consisting of two parts. The first part is termed the *forced response*, or *steady-state* response. The second part is the *natural response*. As we know, the natural response dies out after a period of time, leaving only the forced response. We can think of the natural response of a circuit as a necessary *transient* response to move the circuit from one "state" to another.

We will consider circuits that are initially in a known state – any sources have either been off for a very long time, or on for a very long time. Any forcing functions are switched on at $t = 0$.

The forced response can be obtained by considering the response of the circuit after a very long time. The *form* of the natural response will be the same as that obtained for the source-free circuit.

The *complete response* will be obtained by adding the forced response to the natural response.

10.1 The Unit-Step Forcing Function

In an analysis of a circuit we often find the need to turn source voltages or currents on at a specified time. To do this, we define the *unit-step function* to be zero for all values of time less than zero, and unity for all values of time greater than zero:

The unit-step function defined

$$u(t) = \begin{cases} 0, & t < 0 \\ 1, & t > 0 \end{cases} \quad (9.1)$$

At $t = 0$, $u(t)$ changes abruptly from 0 to 1. Its value at $t = 0$ is undefined, but its value is known for all points arbitrarily close to $t = 0$. We often indicate this by writing $u(0^-) = 0$ and $u(0^+) = 1$.

Graphically, the unit-step function looks like:

and graphed

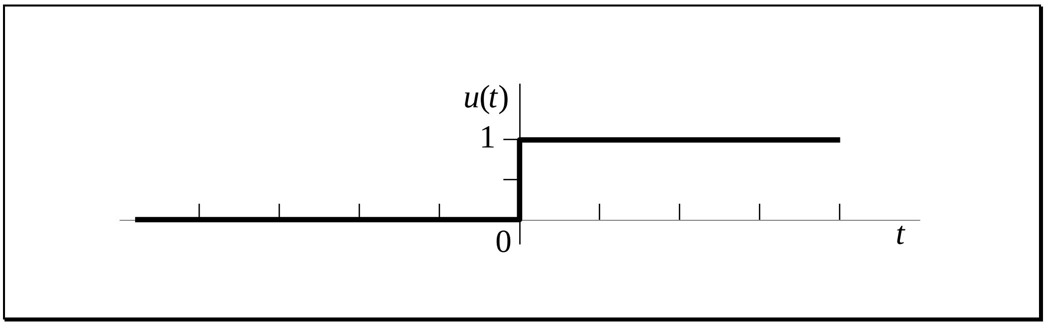


Figure 9.1

Note that a vertical line of unit length is shown at $t = 0$. This “riser” is of course impossible (how can a function take on many values simultaneously?) but is usually shown in drawings.

We cannot always arrange for the switching of a source to occur at $t = 0$. Since the unit-step function provides us with a discontinuity when the value of its argument is zero, then we need to consider a *delayed* version of the unit-step.

The argument of a function determines its position

We will now make a very important observation: it is the *argument* of the function which determines the *position* of the function along the *t*-axis. We therefore have the delayed unit-step function:

$$u(t-t_0)=\begin{cases}0, & t<t_0\\ 1, & t>t_0\end{cases} \tag{9.2}$$

We obtain the conditions on the values of the function by the simple substitution $t \to t - t_0$ in Eq. (9.1). Graphically, we have:

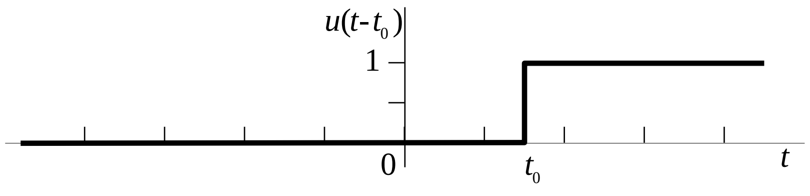


Figure 9.2

We see that the argument $(t-t_0)$ simply shifts the origin of the original function to t_0. A positive value of t_0 shifts the function to the right – corresponding to a *delay* in time. A negative value shifts the function to the left – an *advance* in time.

To represent a constant voltage *V* being turned on in a circuit, we represent it as a source which is zero before $t=t_0$ and a constant *V* after $t=t_0$:

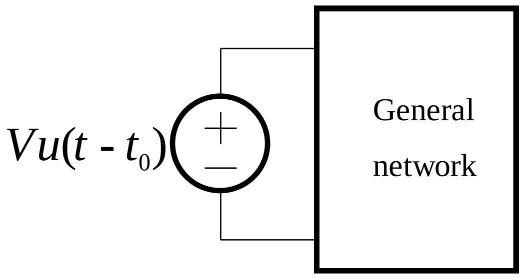


Figure 9.3

The utility of the unit-step function is that it can be used as a "switch" to turn another function (not necessarily a constant) on or off at some point. For example, the product given by $u(t-1)\cos 2\pi t$ is shown below:

The unit-step function as a "switch"

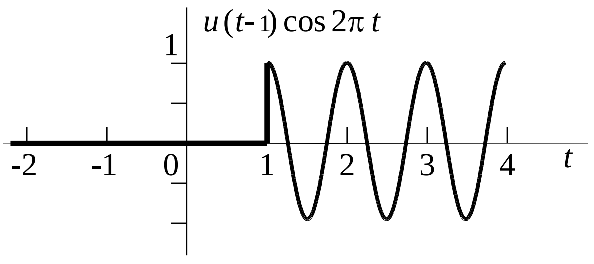


Figure 9.4

Another useful forcing function may be obtained by manipulating the unit-step forcing function. Let us define a rectangular voltage pulse by the following conditions:

The rectangular pulse

$$v(t)=\begin{cases}0, & t<t_0\\ V, & t_0<t<t_1\\ 0, & t_1<t\end{cases} \tag{9.3}$$

The pulse is shown below:

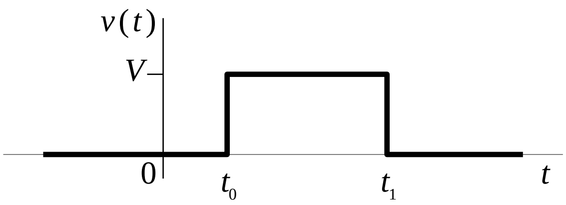


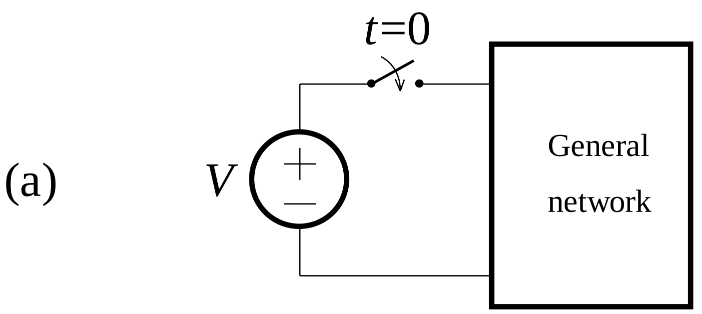
Figure 9.5

This rectangular pulse is the difference of two step functions:

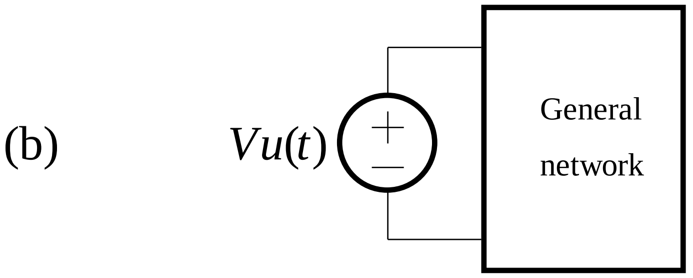
$$v(t)=V\left[u(t-t_0)-u(t-t_1)\right] \tag{9.4}$$

EXAMPLE 10.1 The Unit-Step as a Switch

The unit-step finds application in mathematically representing the application of a source via a switch. Consider the circuit shown below:

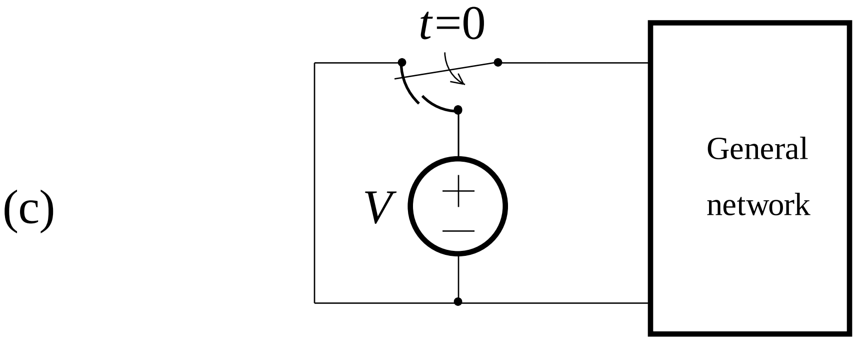


It is tempting to replace the voltage source and switch by a step function:



However, this is incorrect, because the circuit with the step function actually represents:

The circuit equivalent of a step-voltage function



That is, the step function applies $v=0$ for $t<0$, and then $v=V$ for $t>0$. The first circuit applies an *open circuit* for $t<0$, and then $v=V$ for $t>0$.

However, circuit (b) can often be used if we establish that all initial currents and voltages in the original network and in circuit (b) for $t<0$ are equivalent. This is always the case for circuits that start out with zero initial conditions (no stored energy) at $t=0$.

10.2 The Driven *RC* Circuit

Consider a circuit consisting of a battery V in series with a switch, a resistance R, and a capacitance C. The switch is closed at $t=0$, as shown in the circuit diagram (a) below:

A simple *RC* circuit driven by a step-voltage

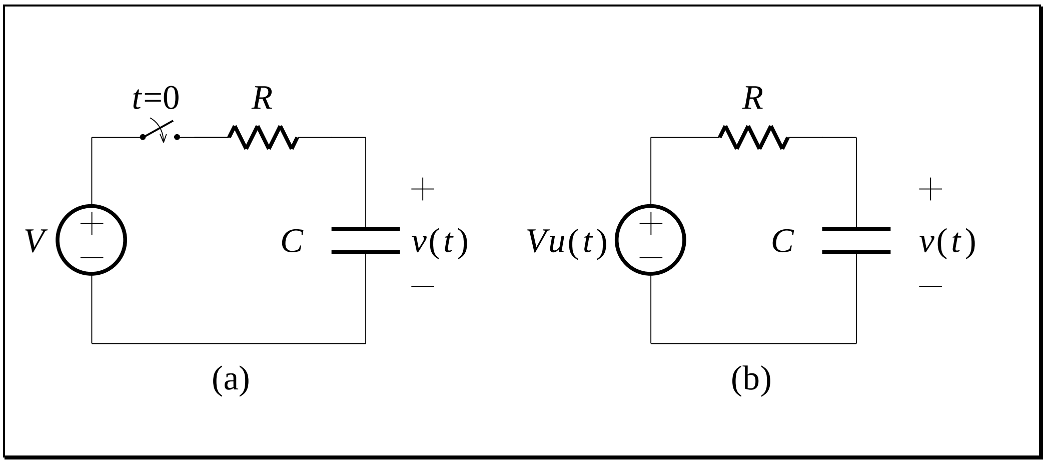


Figure 9.6

We will assume that there is no stored energy in the capacitor before $t=0$, and we are therefore able to replace the battery and switch by a voltage-step forcing function $Vu(t)$, which also produces no response prior to $t=0$. Hence, we will analyse the equivalent circuit shown in circuit diagram (b).

We shall find $v(t)$ by writing down the appropriate differential equation that describes the circuit, and then solve it using Euler's integrating factor. Applying KCL to the top capacitor node, we have:

$$\frac{v-Vu(t)}{R}+C\frac{dv}{dt}=0 \tag{9.5}$$

which can be rewritten as:

The governing differential equation for the *RC* circuit driven by a step-voltage

$$\frac{dv}{dt}+\frac{v}{RC}=\frac{Vu(t)}{RC} \tag{9.6}$$

Since the unit-step function is discontinuous at $t=0$, we first consider the solution for $t<0$ and then for $t>0$. It is obvious that the application of zero voltage since $t=-\infty$ has not produced any response, and therefore:

$$v(t)=0, \quad t<0 \tag{9.7}$$

For positive time $u(t)$ is unity, and we must solve the equation:

$$\frac{dv}{dt}+\frac{v}{RC}=\frac{V}{RC}, \quad t>0 \tag{9.8}$$

To solve, first multiply both sides by an *integrating factor* equal to $e^{t/RC}$. This gives:

$$e^{t/RC}\frac{dv}{dt}+e^{t/RC}\frac{v}{RC}=e^{t/RC}\frac{V}{RC} \tag{9.9}$$

Thus, recognising that the left hand-side is the derivative of $ve^{t/RC}$, we have:

$$\frac{d}{dt}\left[ve^{t/RC}\right]=e^{t/RC}\frac{V}{RC} \tag{9.10}$$

Integrating both sides with respect to time gives:

$$ve^{t/RC}=\int\frac{Ve^{t/RC}}{RC}dt+A \tag{9.11}$$

where A is a constant of integration. Dividing both sides by the integrating factor gives:

$$v=e^{-t/RC}\int\frac{Ve^{t/RC}}{RC}dt+Ae^{-t/RC} \tag{9.12}$$

Noting that *V* is a constant, we perform the integration and obtain:

$$v(t) = V + Ae^{-t/RC} \tag{9.13}$$

Prior to $t = 0$, $v(t) = 0$, and thus $v(0^-) = 0$. Since the voltage across a capacitor cannot change by a finite amount in zero time without being associated with an infinite current, we thus have $v(0^+) = 0$. We thus invoke the initial condition that $v(0^+) = 0$ and get:

$$v(t) = V - Ve^{-t/RC}, \quad t > 0 \tag{9.14}$$

Thus, an expression for the response valid for all t would be:

$$v(t) = V\left(1 - e^{-t/RC}\right)u(t) \tag{9.15}$$

The complete response of an *RC* circuit to a step-voltage

This is the desired solution, but it has not been obtained in the simplest manner. In order to establish a more direct procedure, we will interpret the two terms appearing in Eq. (9.15).

The complete response is composed of two parts…

The exponential term has the functional form of the *natural response* of the *RC* circuit – it is a negative exponential that approaches zero as time increases, and it is characterized by the time constant *RC*. The *functional form* of this part of the response is identical with that which is obtained in the source-free circuit. However, the amplitude of this exponential term depends on *V*, the forcing function.

the natural response and…

Eq. (9.15) also contains a constant term, *V*. Why is it present? The natural response approaches zero as the energy stored in the capacitor gradually reaches its limit. Eventually the capacitor will be fully charged and it will appear as an open circuit – the current will be zero, and the battery voltage *V* will appear directly across the capacitor terminals. This voltage is a part of the response which is directly attributable to the forcing function, and we call it the *forced response*. It is the response which is present a long time after the switch is closed.

the forced response

We see that the complete response is composed of two parts:

The complete response is the sum of the forced response and the natural response

$$\begin{matrix}\text{complete}\\\text{response}\end{matrix} = \begin{matrix}\text{forced}\\\text{response}\end{matrix} + \begin{matrix}\text{natural}\\\text{response}\end{matrix} \tag{9.16}$$

The forced response is determined by forcing function

The forced response has the characteristics of the forcing function; it is found by pretending that all switches have been thrown a long time ago. For circuits with only switches and DC sources, the forced response is just the solution of a simple DC circuit problem (all capacitors are open-circuits, all inductors are short-circuits).

The natural response is determined by the circuit

The natural response is a characteristic of the circuit, and not of the sources. Its form may be found by considering the source-free circuit, and it has an amplitude which depends on the initial amplitude of the source and the initial energy storage.

The natural response provides the link between the initial state and final state of a circuit

The reason for the two responses, forced and natural, may also be seen from physical arguments. We know that our circuit will eventually assume the forced response. However, at the instant the switches are thrown, the initial capacitor voltages (or the currents through the inductors in other circuits) will have values which depend only on the energy stored in these elements. These voltages or currents cannot be expected to be the same as the voltages and currents demanded by the forced response. Hence, there must be a transient period during which the voltages and currents change from their given initial values to their required final values. The portion of the response which provides the transition from initial to final values is the natural response (often called the *transient* response).

10.3 The Forced and the Natural Response

There is an excellent mathematical reason for considering the complete response to be composed of two parts, the forced response and the natural response. The reason is based on the fact that the solution of a linear differential equation may be expressed as the solution of two parts, the *particular solution* (forced response) and the *complementary solution* (natural response).

The solution to a general differential equation consists of two parts…

Recall that a general linear differential equation with constant coefficients of order *n* is an equation that can be written:

$$a_n \frac{d^n y}{dt^n} + a_{n-1} \frac{d^{n-1} y}{dt^{n-1}} + \cdots + a_1 \frac{dy}{dt} + a_0 y = r(t) \tag{9.17}$$

This linear *nonhomogeneous* differential equation with constant coefficients may also be written in the form:

$$f(D)y = r(t) \tag{9.18}$$

We may identify $r(t)$ as a term that is due to the forcing function.

We have already seen that the solution to the homogeneous differential equation with constant coefficients is the solution of the equation:

$$f(D)y_c = 0 \tag{9.19}$$

the complementary solution (natural response) and…

Such a solution, y_c, is called the *complementary* solution. We know that the complementary solution has a form given by:

$$y_c = c_1 e^{s_1 t} + c_2 e^{s_2 t} + \cdots + c_n e^{s_n t} \tag{9.20}$$

where the c_i's are arbitrary constants and the s_i's are the roots of the characteristic equation.

The *particular* solution, y_p, is any solution we happen to come up with that satisfies the original differential equation. That is, the particular solution satisfies:

the particular solution (forced response)

$$f(D)y_p = r(t) \tag{9.21}$$

To see that the general solution of Eq. (9.18) is composed of two parts, let the complete solution be written as the sum of the particular solution and the complementary solution:

The general solution is the sum of the particular solution and the complementary solution

$$\boxed{y = y_p + y_c} \tag{9.22}$$

Substitution into Eq. (9.18) results in:

$$\begin{aligned} f(D)(y_p + y_c) &= f(D)y_p + f(D)y_c \\ &= r(t) + 0 \\ &= r(t) \end{aligned} \tag{9.23}$$

That is, we can safely add y_c to any particular solution, since it contributes *nothing* to the right-hand side.

We now have to find ways to obtain the particular solution of a differential equation. We will look at three methods – one where we utilise the concept of an inverse differential operator, one where the solution can be obtained by inspection, and one where the differential equation is first-order.

10.3.1 Finding a Particular Solution using the Inverse Differential Operator

The concept of the differential operator can be used to find particular solutions for nonhomogeneous differential equations.

In seeking a particular solution of:

$$f(D)y = r(t) \tag{9.24}$$

it is natural to write:

$$y_p = \frac{1}{f(D)} r(t) \tag{9.25}$$

The inverse differential operator is introduced notationally…

and then try to find an operator $1/f(D)$ so that the function y_p will have meaning and satisfy Eq. (9.24). All we require is that:

$$f(D) \cdot \frac{1}{f(D)} r(t) = r(t) \tag{9.26}$$

with the only requirement that it gives a solution to the original differential equation

We proved before that:

$$f(D)e^{st} = f(s)e^{st} \tag{9.27}$$

This suggests that we define:

$$\boxed{\frac{1}{f(D)} e^{st} = \frac{1}{f(s)} e^{st}, \quad f(s) \neq 0} \tag{9.28}$$

The effect of the inverse differential operator on an exponential function

Now from Eq. (9.27) it follows that:

$$f(D)\frac{e^{st}}{f(s)} = \frac{f(s)e^{st}}{f(s)} = e^{st} \tag{9.29}$$

and thus, with the requirement of Eq. (9.26), Eq. (9.28) is verified.

EXAMPLE 10.2 Finding a Particular Solution to a Differential Equation

Solve the equation:

$$\left(D^2 + D\right)y = e^{2t}$$

Here the roots of the characteristic equation are $s = 0, -1$. Further, $f(D) = D^2 + D$ and $f(2) \neq 0$. Hence, using Eq. (9.28):

$$y_p = \frac{1}{D^2 + D} e^{2t} = \frac{e^{2t}}{2^2 + 2} = \frac{1}{6} e^{2t}$$

So the solution is:

$$y = c_1 + c_2 e^{-t} + \tfrac{1}{6} e^{2t}$$

EXAMPLE 10.3 Finding a Particular Solution to a Differential Equation

Find a particular solution of:

$$\left(D^2 - 9\right)y = 5 + 3e^{t}$$

Here the roots of the characteristic equation are $s = 3, -3$ and $f(D) = D^2 - 9$. We also have $f(0) \neq 0$ and $f(1) \neq 0$. Hence, using superposition, we have:

$$y_1 = \frac{1}{D^2 - 9} 5 = \frac{5}{0^2 - 9} = -\frac{5}{9}$$

and:

$$y_2 = \frac{1}{D^2 - 9} 3e^{t} = \frac{3e^{t}}{1^2 - 9} = -\frac{3}{8} e^{t}$$

Hence:

$$y_p = y_1 + y_2 = -\tfrac{5}{9} - \tfrac{3}{8} e^{t}$$

is a particular solution.

10.3.2 Finding a Particular Solution by Inspection

For the second method of finding a particular solution, we note that it is frequently easy to obtain a particular solution of a nonhomogeneous equation:

$$\left(a_n D^n + a_{n-1} D^{n-1} + \cdots + a_1 D + a_0\right) y = r(t) \tag{9.30}$$

by inspection.

For example, if $r(t)$ is a constant R_0 and $a_0 \neq 0$, then:

$$y_p = \frac{R_0}{a_0} \tag{9.31}$$

The particular solution for a constant forcing function is a constant

is a solution because all derivatives of y_p are zero.

EXAMPLE 10.4 Finding a Particular Solution by Inspection

Solve the equation:

$$\left(D^2 + 3D + 2\right) y = 16$$

We obtain the complementary function:

$$y_c = c_1 e^{-t} + c_2 e^{-2t}$$

By inspection a particular solution of the original equation is:

$$y_p = \tfrac{16}{2} = 8$$

Hence the general solution is:

$$y = c_1 e^{-t} + c_2 e^{-2t} + 8$$

10.3.3 Finding a Particular Solution using an Integrating Factor

For the third method of finding a particular solution, let us restrict ourselves to first-order differential equations. The general equation of the type encountered in analysing the *RC* circuit of the previous section can be written as:

$$\frac{dv}{dt} + Pv = Q \tag{9.32}$$

P and *Q* can, in general, be functions of time. We identify *Q* as a term that is due to the forcing function, and *P* as a quantity due solely to the circuit configuration. Following the steps as before, we first multiply both sides by an *integrating factor* equal to e^{Pt}. This gives:

$$e^{Pt}\frac{dv}{dt} + e^{Pt}Pv = e^{Pt}Q \tag{9.33}$$

Thus, recognising that the left hand-side is the derivative of ve^{Pt}, we have:

$$\frac{d}{dt}\left[ve^{Pt}\right] = Qe^{Pt} \tag{9.34}$$

Integrating both sides with respect to time gives:

$$ve^{Pt} = \int Qe^{Pt}dt + A \tag{9.35}$$

where *A* is a constant of integration. Since the constant is explicitly shown, we should remember that no integration constant needs to be added to the remaining integral when it is evaluated.

Dividing both sides by the integrating factor gives:

The general solution of a first-order differential equation…

$$v = e^{-Pt}\int Qe^{Pt}dt + Ae^{-Pt} \tag{9.36}$$

If $Q(t)$, the forcing function, is known, then it remains only to evaluate the integral to obtain the complete response. However, we shall not evaluate such an integral for each problem. Instead we are interested in using Eq. (9.36) to draw several general conclusions.

We should note first that, for a source-free circuit, Q must be zero, and the solution is the *natural response*:

has a natural response…

$$v_n = Ae^{-Pt} \tag{9.37}$$

In linear time-invariant passive circuits, P is always a positive constant that depends only on the circuit elements and their interconnection in the circuit (if a circuit contains a dependent source or a negative resistance, it is possible for P to be negative). The natural response therefore approaches zero as time increases without limit.

We therefore find that one of the two terms making up the complete response has the form of the natural response. It has an amplitude which will depend on the initial energy of the circuit as well as the initial value of the forcing function.

and a forced response

We next observe that the first term of Eq. (9.36) depends on the functional form of $Q(t)$, the forcing function. Whenever we have a circuit in which the natural response dies out as t becomes infinite, then this first term must describe the response completely after the natural response has disappeared. Thus, this term is the *forced response*.

We have only considered those problems involving the sudden application of DC sources, and $Q(t)$ is therefore a constant for all values of time after the switch has been closed. We can now evaluate the integral in Eq. (9.36), obtaining the forced response:

The forced response for DC excitation

$$v_f = \frac{Q}{P} \tag{9.38}$$

and we can write the complete response:

The complete response for DC excitation

$$v = v_f + v_n = \frac{Q}{P} + Ae^{-Pt} \tag{9.39}$$

For the *RC* series circuit, Q/P is the constant voltage *V* and $1/P$ is the time constant *T*. We can see that the forced response might have been obtained without evaluating the integral, because it must be the complete response at infinite time. The forced response is thus obtained by inspection using DC circuit analysis.

In the following section we shall attempt to find the complete response for several *RC* circuits by obtaining the forced and natural responses and adding them.

10.4 Step-Response of *RC* Circuits

We will use the simple *RC* series circuit to illustrate how to determine the complete response by the addition of the forced and natural response. This circuit, shown below, has been analysed earlier, but by a longer method.

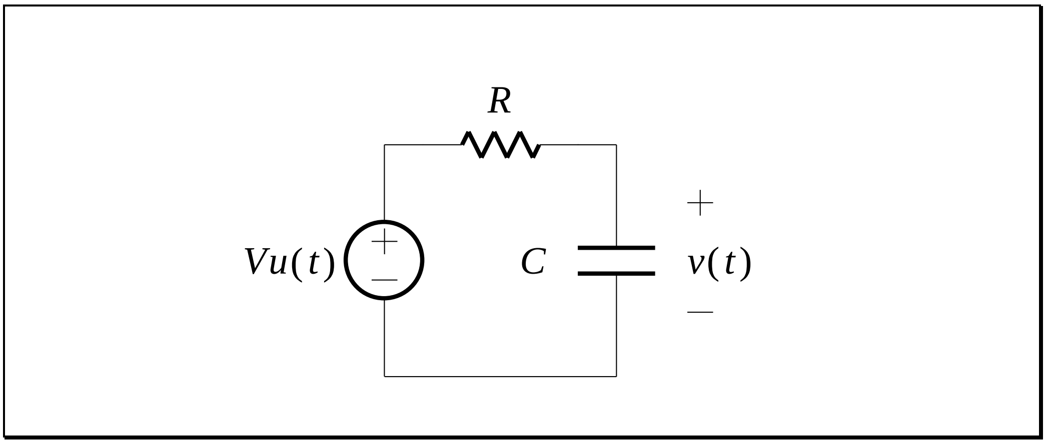


Figure 9.7

The simple *RC* circuit driven by a step-voltage

The desired response is the voltage across the capacitor, $v(t)$, and we first express this voltage as the sum of the forced and natural voltage:

$$v = v_f + v_n \tag{9.40}$$

The complete response expressed as the sum of the forced response and natural response

The functional form of the natural response must be the same as that obtained without any sources. We therefore replace the step-voltage source by a short-circuit and recognize the resulting parallel source-free *RC* circuit. Thus:

$$v_n = Ae^{-t/RC} \tag{9.41}$$

The form of the natural response that results from the source-free circuit

where the amplitude A is yet to be determined.

We next consider the forced response, that part of the response which depends upon the nature of the forcing function itself. In this particular problem the forced response must be constant because the source is a constant V for all positive values of time.

After the natural response has died out the capacitor must be fully charged and the forced response is simply:

The forced response obtained from DC circuit analysis

$$v_f = V \tag{9.42}$$

Note that the forced response is determined completely – there is no unknown amplitude. We next combine the two responses:

$$v = V + Ae^{-t/RC} \tag{9.43}$$

Determining the amplitude of the decaying exponential term

and apply the initial condition to evaluate A. The voltage is zero prior to $t = 0$, and it cannot change value instantaneously since it is the voltage across a capacitor. Thus, the voltage is zero immediately after $t = 0$, and:

$$0 = V + A \tag{9.44}$$

Thus:

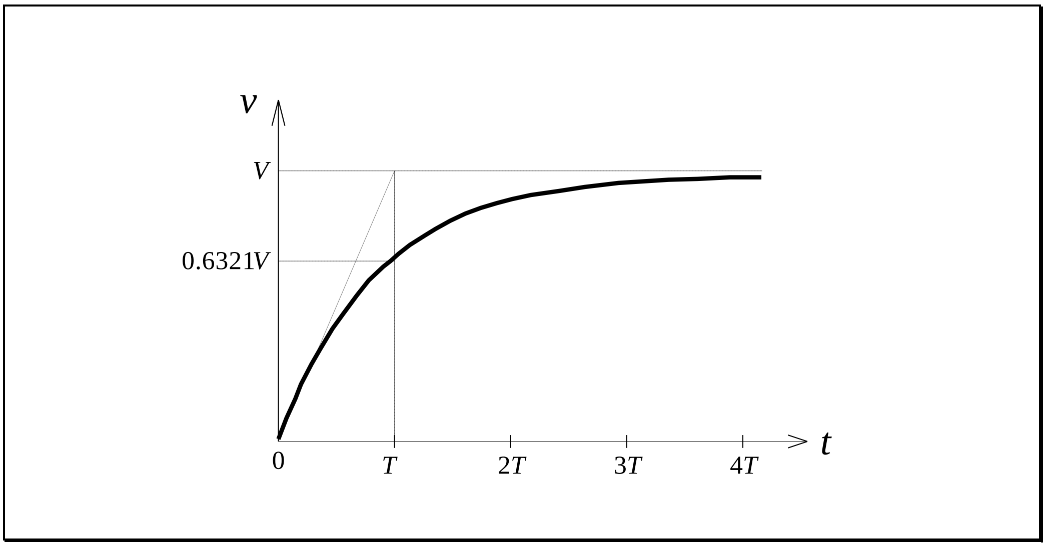
$$A = -V \tag{9.45}$$

Therefore:

The complete response – obtained without solving a differential equation!

$$v = V\left(1 - e^{-t/RC}\right) \tag{9.46}$$

The complete response is plotted below, and we can see the manner in which the voltage builds up from its initial value of zero to its final value of *V*.



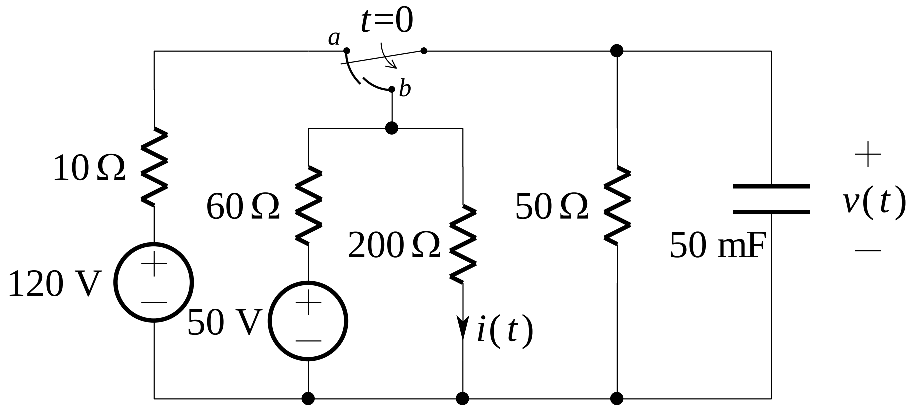
A graph of the step-response of the *RC* circuit

Figure 9.8

The transition is effectively accomplished in a time $4T$. In one time constant, the voltage has attained 63.2% of its final value.

EXAMPLE 10.5 Step Response of an *RC* Circuit

Consider the circuit shown below:



The switch is assumed to have been in position *a* for a long time, or, in other words, the natural response which resulted from the original excitation of the circuit has decayed to a negligible amplitude, leaving only a forced response caused by the 120 V source.

We begin by finding the forced response when the switch is in position *a*. The voltages throughout the circuit are all constant, and there is thus no current through the capacitor (which is treated like an open-circuit). Simple voltage division determines the forced response prior to $t = 0$:

$$v_f = \frac{50}{50+10}120 = 100, \quad t < 0$$

and we thus have the initial condition:

$$v(0) = 100$$

Since the capacitor voltage cannot change instantaneously, this voltage is equally valid at $t = 0^-$ and $t = 0^+$.

The switch is now thrown to *b*, and the complete response is:

$$v = v_f + v_n$$

The form of the natural response is obtained by replacing the 50 V source by a short-circuit and evaluating the equivalent resistance:

$$R_{eq} = \frac{1}{1/50 + 1/200 + 1/60} = 24$$

$$v_n = Ae^{-t/R_{eq}C}$$

or

$$v_n = Ae^{-t/1.2}$$

In order to evaluate the forced response with the switch at *b*, we wait until all the voltages and currents have stopped changing, thus treating the capacitor as an open circuit, and use voltage division once more:

$$v_f = \frac{(50)(200)/(50+200)}{60+(50)(200)/(50+200)} 50 = 20$$

Thus:

$$v = 20 + Ae^{-t/1.2}$$

and from the initial condition already obtained:

$$v(0^+) = 20 + A = 100$$

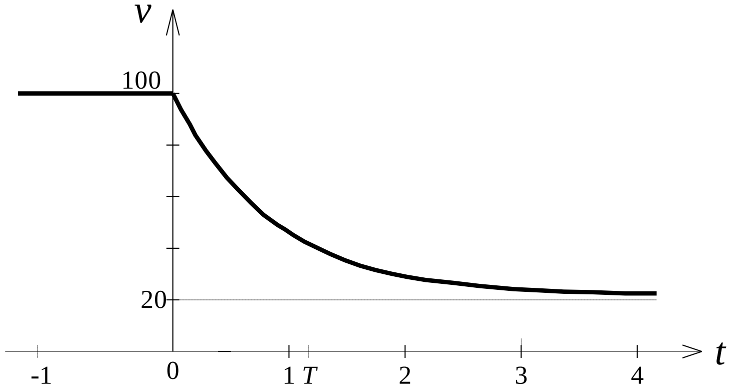
and thus:

$$A = 80$$

Therefore, the complete response is:

$$v = 20 + 80e^{-t/1.2}, \quad t > 0$$

This response is sketched below:



The natural response is seen to form a transition from the initial to the final response.

Finally, let us calculate some response that need not remain constant during the instant of switching, such as $i(t)$ in the circuit diagram. With the contact at *a*, it is evident that $i = 50/260 \approx 0.192\,\mathrm{A}$. When the switch is in position *b*, the forced response for this current now becomes:

$$i_f = \frac{50}{60+(50)(200)/(50+200)}\frac{50}{50+200} = 0.1$$

The form of the natural response is the same as that which we already determined for the capacitor voltage:

$$i_n = Be^{-t/1.2}$$

Combining the forced and natural responses, we obtain:

$$i = 0.1 + Be^{-t/1.2}$$

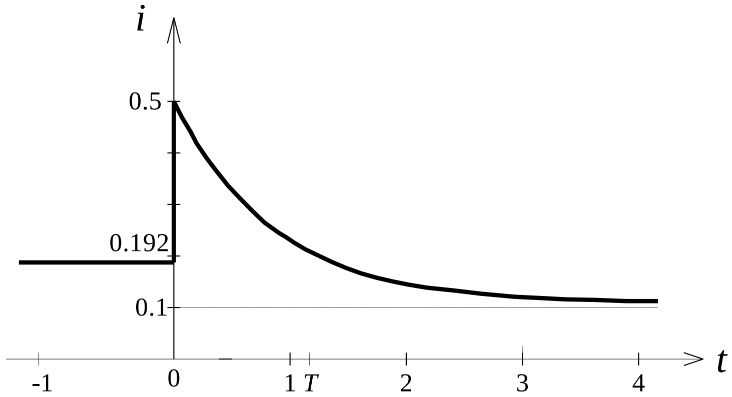
To evaluate B, we need to know $i(0^+)$. This is found by fixing our attention on the energy-storage element, here the capacitor, for the fact that v must remain 100 V during the switching interval is the governing condition establishing other currents and voltages at $t=0^+$. Since $v(0^+)=100\text{ V}$, and since the capacitor is in parallel with the $200\,\Omega$ resistor, we find $i(0^+)=0.5$, $B=0.4$, and thus:

$$\begin{aligned} i(t) &= 0.192 && t<0 \\ i(t) &= 0.1+0.4e^{-t/1.2} && t>0 \end{aligned}$$

or:

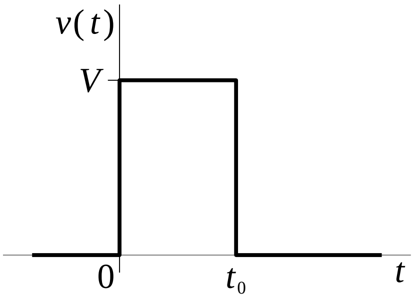
$$i(t)=0.192u(-t)+\left(0.1+0.4e^{-t/1.2}\right)u(t)$$

This response is sketched below:

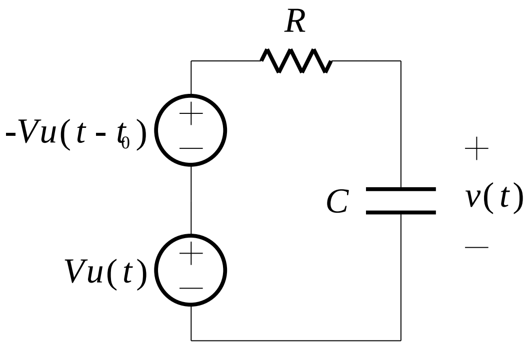


EXAMPLE 10.6 Step Response of an *RC* Circuit using Superposition

Let us apply a rectangular voltage pulse of amplitude V and duration t_0 to the simple *RC* series circuit.



(a)



(b)

We represent the forcing function as the sum of two step-voltage sources $Vu(t)$ and $-Vu(t-t_0)$, and plan to obtain the response by using the superposition principle. We will designate that part of $v(t)$ which is due to the lower source $Vu(t)$ acting alone by the symbol $v_1(t)$ and then let $v_2(t)$ represent that part due to $-Vu(t-t_0)$ acting alone. Then:

$$v(t)=v_1(t)+v_2(t)$$

We now write each of the partial responses v_1 and v_2 as the sum of a forced and a natural response. The response $v_1(t)$ is familiar:

$$v_1(t)=V\left(1-e^{-t/RC}\right),\quad t>0$$

Note that the range of t, $t>0$, in which the solution is valid, is indicated.

We now consider the upper source and its response $v_2(t)$. Only the polarity of the source and the time of its application are different. There is thus no need to determine the form of the natural response and the forced response, the solution for $v_1(t)$ enables us to write:

$$v_2(t) = -V\left(1 - e^{-(t-t_0)/RC}\right), \quad t > t_0$$

where the applicable range of t, $t > t_0$, must again be indicated.

We now add the two solutions carefully, since each is valid over a different interval of time: Thus:

$$\begin{aligned} v(t) &= V\left(1 - e^{-t/RC}\right) & 0 < t < t_0 \\ v(t) &= V\left(1 - e^{-t/RC}\right) - V\left(1 - e^{-(t-t_0)/RC}\right) & t > t_0 \end{aligned}$$

or, simplifying the second equation:

$$\begin{aligned} v(t) &= V\left(1 - e^{-t/RC}\right) & 0 < t < t_0 \\ v(t) &= V\left(e^{t_0/RC} - 1\right)e^{-t/RC} & t > t_0 \end{aligned}$$

The second equation can also be written as:

$$v(t) = V\left(1 - e^{-t_0/RC}\right)e^{-(t-t_0)/RC} \quad t > t_0$$

We can identify the constant at the front of the decaying exponential term as the value of the initial response at time $t = t_0$. If we define:

$$V_0 = V\left(1 - e^{-t_0/RC}\right)$$

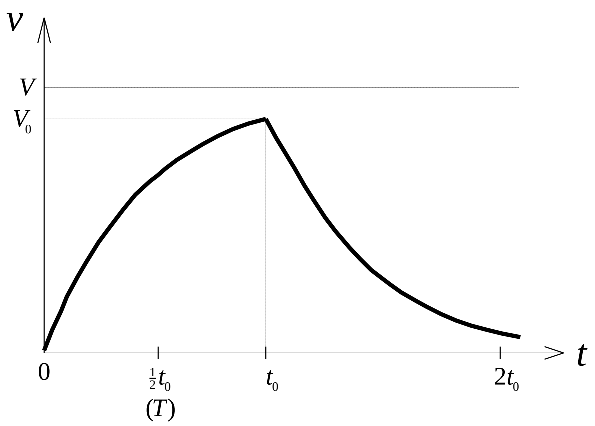
then the second term is just:

$$v(t) = V_0 e^{-(t-t_0)/RC} \quad t > t_0$$

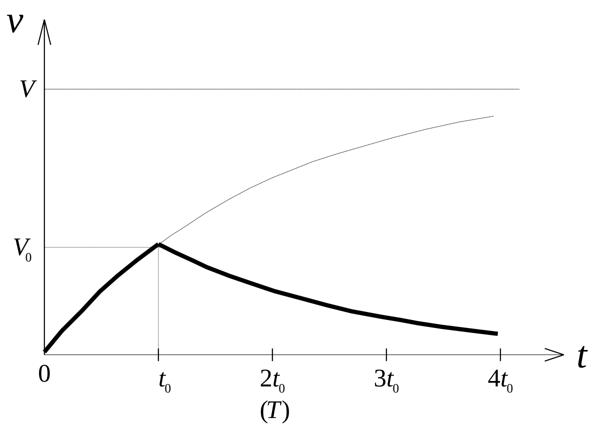
This is a "time shifted" decaying exponential with an initial value given by the response at the switching instant, V_0.

10.28

The solution is completed by stating that $v(t)$ is zero for negative t and sketching the response as a function of time. The type of curve obtained depends upon the relative values of t_0 and the time constant T. Two possible curves are shown below:



(a)



(b)

The top curve (a) is drawn for the case where the time constant is only one-half as large as the length of the applied pulse – the rising portion of the exponential has therefore almost reached V before the decaying exponential begins.

The opposite situation is shown on the bottom (b) – there, the time constant is twice t_0 and the response never has a chance to reach a large amplitude.

10.5 Analysis Procedure for Single Time Constant *RC* Circuits

The procedure we have been using to find the response of an *RC* circuit after DC sources have been switched on or off or in or out of the circuit at some instant of time, say $t = 0$, is summarized in the following. We assume that the circuit is reducible to a single equivalent resistance R_{eq} in parallel with a single equivalent capacitance C_{eq} when all independent sources are set equal to zero, i.e. we have a single time constant (STC) circuit. The response we seek is represented by $f(t)$.

Step-by-step guide to solving STC *RC* circuits with step-sources

1. With all independent sources set to zero, simplify the circuit to determine R_{eq}, C_{eq}, and the time constant $T = R_{eq}C_{eq}$.

2. Viewing C_{eq} as an open circuit, use DC-analysis methods to find $v_C(0^-)$, the capacitor voltage just prior to the discontinuity.

3. Again viewing C_{eq} as an open circuit, use DC-analysis methods to find the forced response. This is the value approached by $f(t)$ as $t \to \infty$; we represent it by $f(\infty)$.

4. Write the total response as the sum of the forced and natural responses: $f(t) = f(\infty) + Ae^{-t/T}$.

5. Find $f(0^+)$ by using the condition that $v_C(0^+) = v_C(0^-)$. If desired, C_{eq} may be replaced by a voltage source $v_C(0^+)$ [a short circuit if $v_C(0^+) = 0$] for this calculation. With the exception of capacitor voltages, other voltages and currents in the circuit may change abruptly.

6. Then $f(0^+) = f(\infty) + A$ and $f(t) = f(\infty) + [f(0^+) - f(\infty)]e^{-t/T}$.

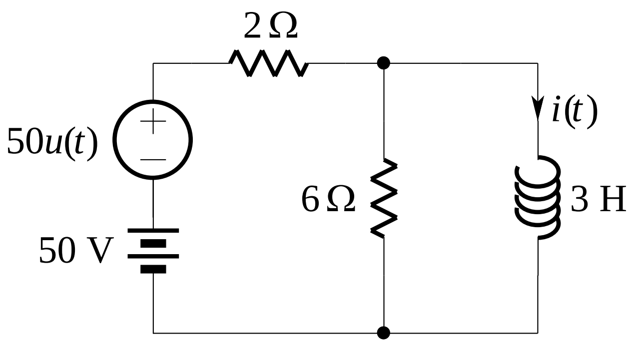
10.6 *RL* Circuits

Solving *RL* circuits uses the same techniques as for *RC* circuits

The complete response of any *RL* circuit may also be obtained as the sum of the forced response and natural response.

EXAMPLE 10.7 Step Response of an *RL* Circuit

Consider the circuit shown below:



The circuit contains a DC voltage source as well as a step-voltage source. Let us determine $i(t)$ for all values of time. We might choose to replace everything to the left of the inductor by the Thévenin equivalent, but instead let us merely recognize the form of that equivalent as a resistor in series with some voltage source. The circuit contains only one energy-storage element, the inductor, and the natural response is therefore a negative exponential:

$$T = \frac{L}{R_{eq}} = \frac{3}{1.5} = 2$$

and:

$$i_n = Ae^{-t/2}$$

The forced response must be that produced by a constant voltage of 100 V. The forced response is constant, and no voltage is present across the inductor (it behaves like a short-circuit) and therefore:

$$i_f = \frac{100}{2} = 50$$

Thus:

$$i = i_f + i_n = 50 + Ae^{-t/2}$$

In order to evaluate *A*, we must establish the initial value of the inductor current. Prior to $t = 0$, this current is 25 A, and it cannot change instantaneously. Thus:

$$25 = 50 + A \quad \text{or} \quad A = -25$$

Hence:

$$i = 50 - 25e^{-t/2}, \quad t > 0$$

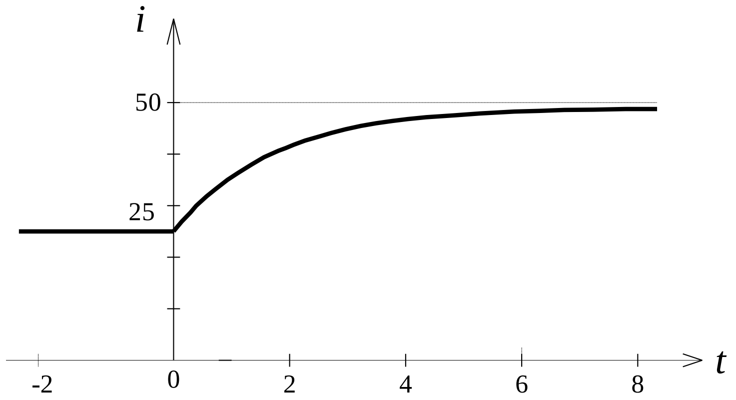
We complete the solution by also stating:

$$i = 25, \quad t < 0$$

or by writing a single expression valid for all *t*:

$$i = 25 + 25\left(1 - e^{-t/2}\right)u(t)\ \text{A}$$

The complete response is sketched below:



Note how the natural response serves to connect the response for $t < 0$ with the constant forced response.

10.7 Analysis Procedure for Single Time Constant *RL* Circuits

This is the dual of the statements given for the analysis procedure for single time constant *RC* circuits.

Step-by-step guide to solving STC *RL* circuits with step-sources

The procedure we have been using to find the response of an *RL* circuit after DC sources have been switched on or off or in or out of the circuit at some instant of time, say $t=0$, is summarized in the following. We assume that the circuit is reducible to a single equivalent resistance R_{eq} in series with a single equivalent inductance L_{eq} when all independent sources are set equal to zero, i.e. we have a single time constant (STC) circuit. The response we seek is represented by $f(t)$.

1. With all independent sources set to zero, simplify the circuit to determine R_{eq}, L_{eq}, and the time constant $T = L_{eq}/R_{eq}$.

2. Viewing L_{eq} as a short circuit, use DC-analysis methods to find $i_L(0^-)$, the inductor current just prior to the discontinuity.

3. Again viewing L_{eq} as a short circuit, use DC-analysis methods to find the forced response. This is the value approached by $f(t)$ as $t \to \infty$; we represent it by $f(\infty)$.

4. Write the total response as the sum of the forced and natural responses: $f(t) = f(\infty) + Ae^{-t/T}$.

5. Find $f(0^+)$ by using the condition that $i_L(0^+) = i_L(0^-)$. If desired, L_{eq} may be replaced by a current source $i_L(0^+)$ [an open circuit if $i_L(0^+) = 0$] for this calculation. With the exception of inductor currents, other voltages and currents in the circuit may change abruptly.

6. Then $f(0^+) = f(\infty) + A$ and $f(t) = f(\infty) + \left[f(0^+) - f(\infty)\right]e^{-t/T}$.

10.8 Summary

- The unit-step function can be used to simulate the opening and closing of switches under certain conditions.
- The complete response of a circuit to a forcing function consists of two parts: the forced response, and the natural response.
- Mathematically, the forced response is the particular solution of a nonhomogeneous linear differential equation with constant coefficients. The natural response is the complementary solution to the corresponding homogeneous equation.
- For circuits driven by constant voltages or currents, the forced response can be obtained by undertaking DC circuit analysis.
- For single time constant circuits, the natural response is a decaying exponential, $y = Ae^{-t/T}$, where T is the time constant.
- The complete response of single time constant circuits to DC excitation always takes the form of $f(t) = f(\infty) + \left[f(0^+) - f(\infty)\right]e^{-t/T}$. We just need to determine the time constant of the circuit T, the forced response $f(\infty)$, and the initial response $f(0^+)$.

10.9 References

Bedient, P. & Rainville, E.: *Elementary Differential Equations*, 6th *Ed*. Macmillan Publishing Co., 1981.

Hayt, W. & Kemmerly, J.: *Engineering Circuit Analysis*, 3rd *Ed*., Mcgraw-Hill, 1984.

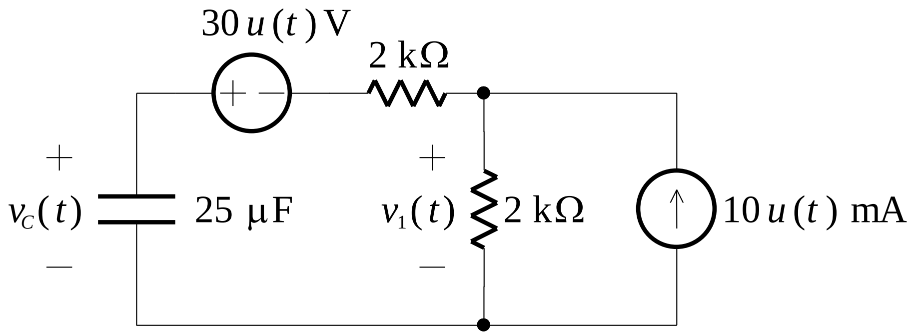
Exercises

1.

A current source of 5 A, a $4\,\Omega$ resistor, and a closed switch are in parallel. The switch opens at $t=0$, closes at $t=0.2\,\text{s}$, opens at $t=0.4\,\text{s}$, and continues in this periodic pattern. Express the voltage across the switch as an infinite summation of step functions.

2.

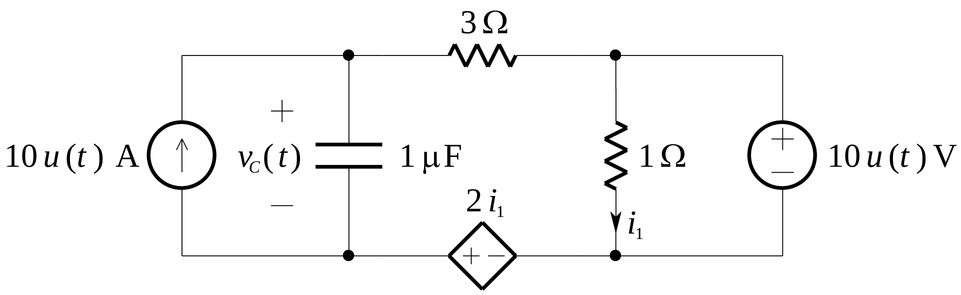
Consider the circuit shown below:



Find $v_C(t)$ and $v_1(t)$.

3.

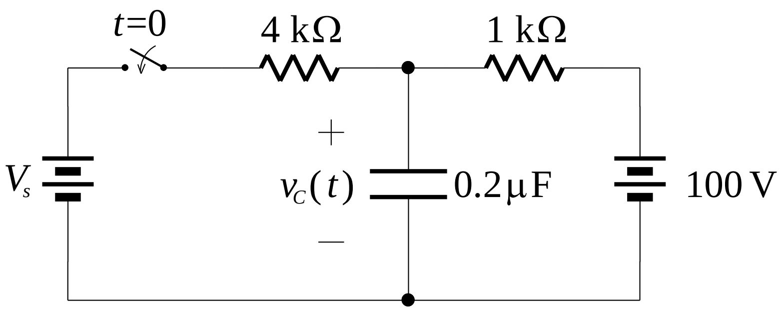
Consider the circuit shown below:



Find $v_C(t)$.

4.

After being open for several minutes, the switch in the circuit below closes at $t = 0$.



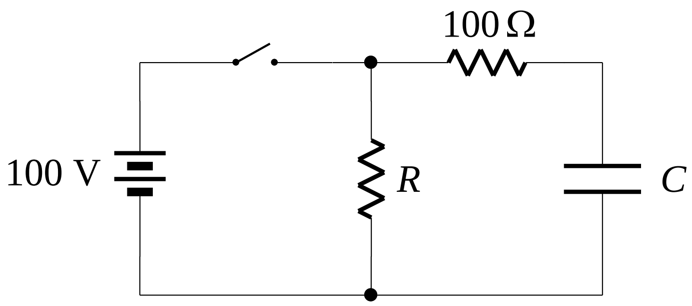
Find $v_C(t)$ for all t if $V_s =$:

(a) -200 V

(b) +100 V

5.

Consider the circuit shown below:

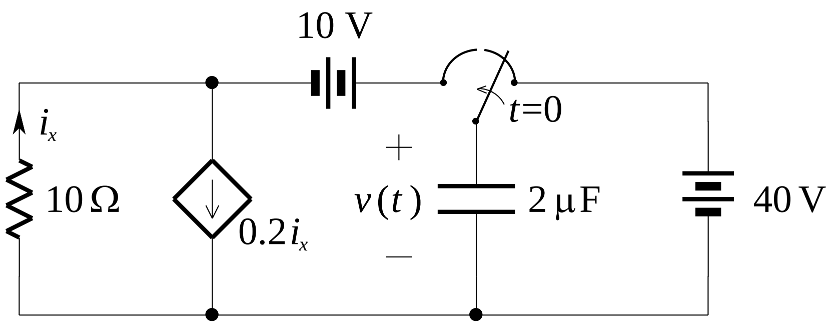


Specify values for R and C in the circuit so that the capacitor voltage will reach 80 V, 10 ms after the switch is closed, but will not drop below 90 V until 0.5 s after the switch is opened, assuming that it has been closed for a very long time.

10.36

6.

Consider the circuit shown below:



Find $v(t)$ for $t > 0$.

7.

A $250\ \Omega$ resistor and a source, $12u(t) - 12u(t - 10^{-3})\,\mathrm{V}$, are in series with a $0.2\ \mathrm{H}$ inductor. Find the inductor current magnitude at $t = 0.5$ and 2 ms.

8.

The general solution to the driven series *RL* circuit is given by Eq. (9.36). Use it to find $i(t)$ for $t > 0$ if $R = 250\ \Omega$, $L = 0.2\ \mathrm{H}$, and the source voltage is:

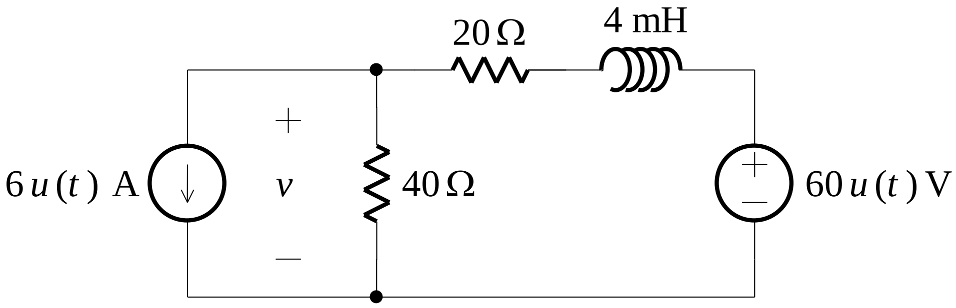
(a) $100u(t)\,\mathrm{V}$

(b) $100e^{-100t}u(t)\,\mathrm{V}$

(c) $100\cos(1250t)u(t)\,\mathrm{V}$

9.

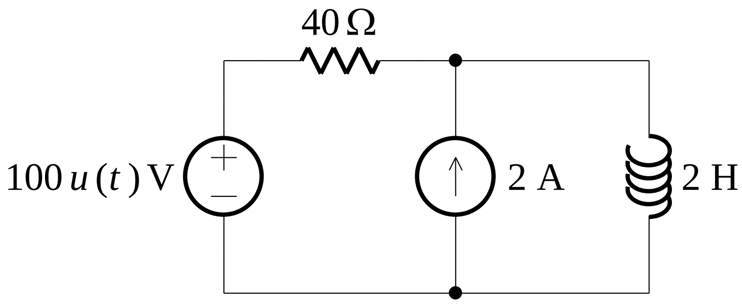
Consider the circuit shown below:



Find v as a function of time.

10.

Consider the circuit shown below:



Find the power being absorbed by the inductor at $t =$:

(a) 0^- (b) 0^+ (c) 0.05 (d) ∞

Leonhard Euler (1707-1783) *(Len´ ard Oy´ ler)*

The work of Euler built upon that of Newton and made mathematics the tool of analysis. Astronomy, the geometry of surfaces, optics, electricity and magnetism, artillery and ballistics, and hydrostatics are only some of Euler's fields. He put Newton's laws, calculus, trigonometry, and algebra into a recognizably modern form.

Euler was born in Switzerland, and before he was an adolescent it was recognized that he had a prodigious memory and an obvious mathematical gift. He received both his bachelor's and his master's degrees at the age of 15, and at the age of 23 he was appointed professor of physics, and at age 26 professor of mathematics, at the Academy of Sciences in Russia.

Among the symbols that Euler initiated are the sigma (Σ) for summation (1755), e to represent the constant 2.71828…(1727), i for the imaginary $\sqrt{-1}$ (1777), and even a, b, and c for the sides of a triangle and A, B, and C for the opposite angles. He transformed the trigonometric ratios into functions and abbreviated them $\sin$, $\cos$ and $\tan$, and treated logarithms and exponents as functions instead of merely aids to calculation. He also standardised the use of π for 3.14159…

His 1736 treatise, *Mechanica*, represented the flourishing state of Newtonian physics under the guidance of mathematical rigor. An introduction to pure mathematics, *Introductio in analysin infinitorum*, appeared in 1748 which treated algebra, the theory of equations, trigonometry and analytical geometry. In this work Euler gave the formula $e^{ix} = \cos x + i \sin x$. It did for calculus what Euclid had done for geometry. Euler also published the first two complete works on calculus: *Institutiones calculi differentialis*, from 1755, and *Institutiones calculi integralis*, from 1768.

Euler's work in mathematics is vast. He was the most prolific writer of mathematics of all time. After his death in 1783 the St Petersburg Academy continued to publish Euler's unpublished work for nearly 50 more years!

Some of his phenomenal output includes: books on the calculus of variations; on the calculation of planetary orbits; on artillery and ballistics; on analysis; on shipbuilding and navigation; on the motion of the moon; lectures on the differential calculus. He made decisive and formative contributions to geometry, calculus and number theory. He integrated Leibniz's differential calculus and Newton's method of fluxions into mathematical analysis. He introduced beta and gamma functions, and integrating factors for differential equations. He studied continuum mechanics, lunar theory, the three body problem, elasticity, acoustics, the wave theory of light, hydraulics, and music. He laid the foundation of analytical mechanics. He proved many of Fermat's assertions including Fermat's Last Theorem for the case $n=3$. He published a full theory of logarithms of complex numbers. Analytic functions of a complex variable were investigated by Euler in a number of different contexts, including the study of orthogonal trajectories and cartography. He discovered the Cauchy-Riemann equations used in complex variable theory.

Euler made a thorough investigation of integrals which can be expressed in terms of elementary functions. He also studied beta and gamma functions. As well as investigating double integrals, Euler considered ordinary and partial differential equations. The calculus of variations is another area in which Euler made fundamental discoveries.

He considered linear equations with constant coefficients, second order differential equations with variable coefficients, power series solutions of differential equations, a method of variation of constants, integrating factors, a method of approximating solutions, and many others. When considering vibrating membranes, Euler was led to the Bessel equation which he solved by introducing Bessel functions.

Euler made substantial contributions to differential geometry, investigating the theory of surfaces and curvature of surfaces. Many unpublished results by Euler in this area were rediscovered by Gauss.

Euler considered the motion of a point mass both in a vacuum and in a resisting medium. He analysed the motion of a point mass under a central force and also considered the motion of a point mass on a surface. In this latter topic he had to solve various problems of differential geometry and geodesics.

He wrote a two volume work on naval science. He decomposed the motion of a solid into a rectilinear motion and a rotational motion. He studied rotational problems which were motivated by the problem of the precession of the equinoxes.

He set up the main formulas for the topic of fluid mechanics, the continuity equation, the Laplace velocity potential equation, and the Euler equations for the motion of an inviscid incompressible fluid.

He did important work in astronomy including: the determination of the orbits of comets and planets by a few observations; methods of calculation of the parallax of the sun; the theory of refraction; consideration of the physical nature of comets.

Euler also published on the theory of music...

Euler did not stop working in old age, despite his eyesight failing. He eventually went blind and employed his sons to help him write down long equations which he was able to keep in memory. Euler died of a stroke after a day spent: giving a mathematics lesson to one of his grandchildren; doing some calculations on the motion of balloons; and discussing the calculation of the orbit of the planet Uranus, recently discovered by William Herschel.

His last words, while playing with one of his grandchildren, were: “I die.”

11 Op-Amp Imperfections

Contents

11.2

Introduction

The initial analysis phase of op-amp circuits assumes the op-amps to be ideal. Although in many applications such an assumption is not a bad one, a circuit designer has to be thoroughly familiar with the characteristics of practical op-amps and the effects of such characteristics on the performance of op-amp circuits.

Although no op-amp is ideal, modern processing techniques yield devices that come close, at least in some parameters. This is by design. In fact, different op-amps are optimized to be close to ideal for some parameters, while other parameters for the same op-amp may be quite ordinary (some parameters can be improved, but only at the expense of others). It is the designer's function to select the op-amp that is closest to ideal in ways that matter to the application, and to know which parameters can be discounted or ignored.

For this reason, it is very important to understand the specifications and to compare the limitations of the different commercially available op-amps, in order to select the right op-amp for a specific application.

DC imperfections in the op-amp's internal circuit give rise to DC voltages appearing at the op-amp output that are independent of the input signal.

The most serious real op-amp deficiency is finite gain and limited bandwidth. This causes "gain error" at low frequencies, and the op-amp to stop "working" altogether at high frequencies.

Large signal operation of op-amps is limited by the power supplies, output current limiting, and the "slew rate" – the maximum rate of change at the output.

11.1 DC Imperfections

The standard input stage of an integrated circuit op-amp is a “DC coupled differential pair” of transistors. Without delving deeply into the topology and analysis of such a configuration, we can imagine that for it to amplify only the “difference” in voltage appearing at its inputs that it must, in some sense, be perfectly “balanced”. This requires the transistors that make up the “differential pair” to be perfectly matched – i.e. each transistor should have exactly the same characteristics. In real devices it is impossible to perfectly match transistors – so there must be some “inherent imbalance” in the differential pair. This imbalance causes an output even when the inputs are connected to the same voltage (so that there is no differential input). The result is a DC offset appearing at the output of the op-amp. To take this into account, we refer the voltage back to the input, and define the *offset voltage* as that voltage which must be applied at the *input* of an op-amp to cause the output to be zero.

In addition, each of the differential pair’s transistors is required to be “biased” at a certain “operating point” on the transistor’s characteristic. Thus, each input of the op-amp draws a DC *bias current*. Components of this current inevitably must pass through the resistors connected to the op-amp’s input terminals, with the result that they generate a DC output voltage (in addition to that generated by the offset voltage).

Practical circuits therefore need to take these DC considerations into account so that non-intentional DC output voltages are minimised.

11.1.1 Offset Voltage

The input offset voltage, V_{OS}, is a DC voltage which must be applied to the op-amp's noninverting *input* terminal to drive the *output* voltage to 0 V, as illustrated below:

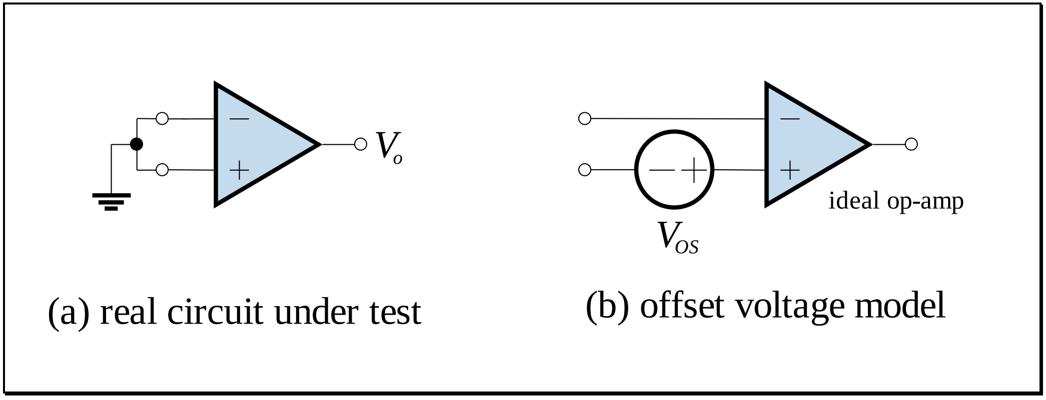


Figure 11.1

To analyse the effect of the offset voltage on the closed-loop performance of the inverting and noninverting amplifiers, we can use superposition (since the offset voltage effectively appears as another independent source). In either configuration, the circuit we obtain just by considering the offset voltage is:

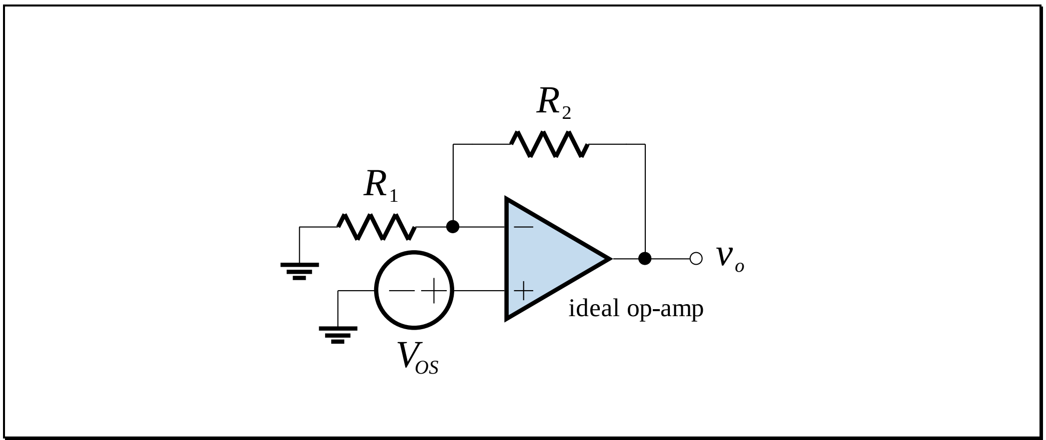


Figure 11.2

It is easily seen that the resulting DC output voltage is given by:

$$V_o = V_{OS}\left(1+\frac{R_2}{R_1}\right) \tag{11.1}$$

11.1.2 Input Bias Currents

For op-amps that have bipolar junction transistors (BJTs) at their inputs, a finite DC current is required for proper biasing of the internal differential amplifier. For op-amps that have junction field effect transistors (JFETs) at their inputs, the DC input currents are junction leakage currents. In either case, there are DC input bias currents which enter the input terminals of the op-amp:

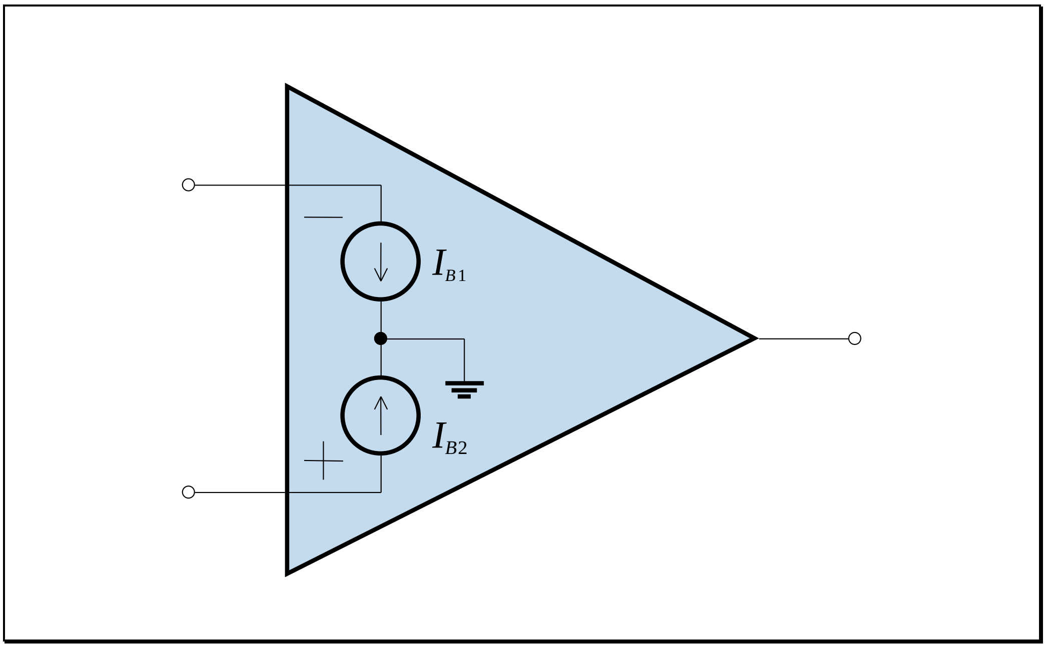


Figure 11.3

The two currents are represented by two current sources, I_{B1} and I_{B2}, connected to the two input terminals. The input bias currents are *independent* of the fact that the op-amp has a finite (though large) input resistance.

We call the average value of I_{B1} and I_{B2} the *input bias current*:

$$I_B = \frac{I_{B1} + I_{B2}}{2} \tag{11.2}$$

while the difference is called the *input offset current*:

$$I_{OS} = \left| I_{B1} - I_{B2} \right| \tag{11.3}$$

11.6

To analyse the effect of the bias currents on the closed-loop performance of the inverting and noninverting amplifiers, we can analyse the following circuit:

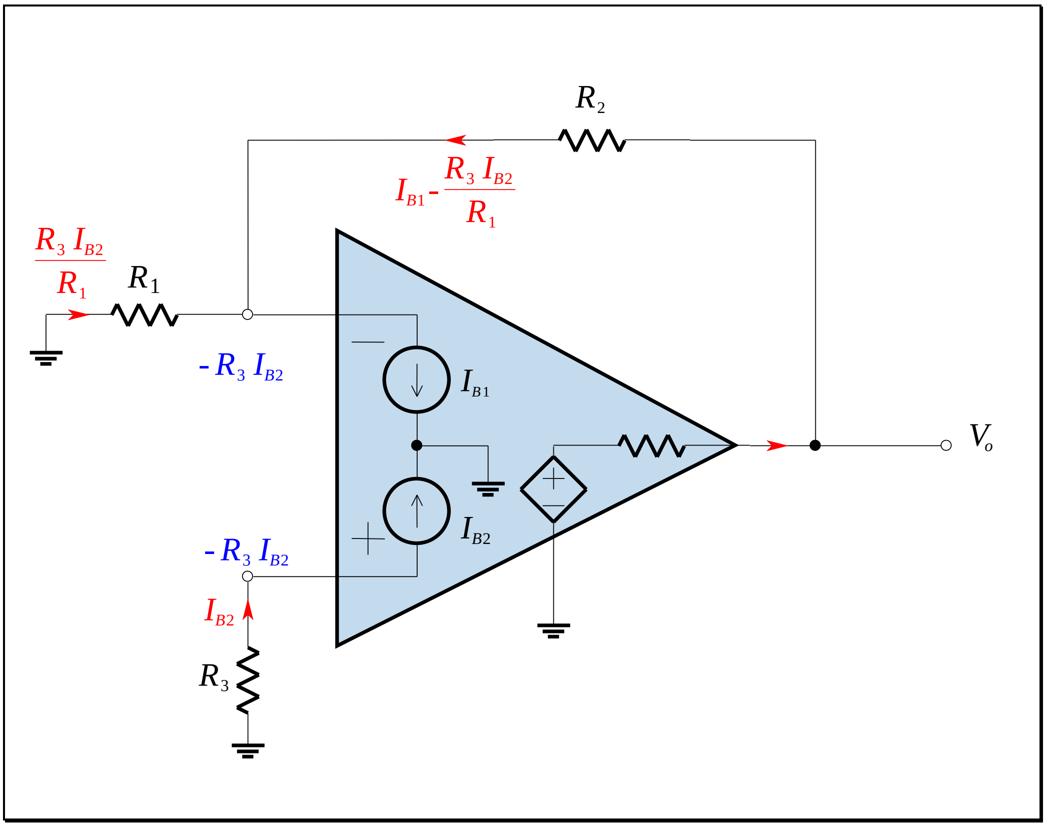


Figure 11.4

With $R_3 = 0$, the output DC voltage is given by:

$$V_o = R_2 I_{B1} \tag{11.4}$$

which can be significant if R_2 is large. Fortunately, a technique exists for reducing the value of the output DC voltage due to the input bias currents. The method consists of introducing a resistance $R_3 \neq 0$ in series with the noninverting terminal. From a signal point of view (i.e. when we analyse the circuit considering the normal input voltages and an ideal op-amp), the inclusion of R_3 has no effect.

With $R_3 \neq 0$, the output DC voltage is given by:

$$V_o = -R_3 I_{B2} + R_2 \left(I_{B1} - I_{B2} R_3 / R_1 \right) \quad (11.5)$$

For the case $I_{B1} = I_{B2} = I_B$, we get:

$$V_o = \left[R_2 - R_3 \left(1 + R_2 / R_1 \right) \right] I_B \quad (11.6)$$

We may reduce V_o to zero by selecting R_3 such that:

$$R_3 = \frac{R_2}{1 + R_2 / R_1} = \frac{R_1 R_2}{R_1 + R_2} = R_1 \parallel R_2 \quad (11.7)$$

Therefore, to reduce the effect of DC bias currents, we should select R_3 to be equal to the parallel equivalent of R_1 and R_2. Having selected this value, substitution into Eq. (11.5) gives:

$$\left| V_o \right| = R_2 \left| I_{B1} - I_{B2} \right| = R_2 I_{OS} \quad (11.8)$$

which is usually about an order of magnitude smaller than the value obtained without R_3.

Thus, to minimise the effect of the bias currents, we place a resistance in series with the noninverting terminal that is equal to the DC resistance seen by the inverting terminal.

11.2 Finite Open-Loop Gain

To illustrate the effect of finite open-loop gain, we will consider what happens to the closed-loop gain of the standard noninverting and inverting amplifier configurations.

11.2.1 Noninverting Amplifier

For the noninverting amplifier, we have already derived the closed-loop gain due to a finite open-loop gain. The result was:

$$A_{CL} = \frac{A_{OL}}{1 + A_{OL}\beta} \tag{11.9}$$

where:

$$\beta = \frac{R_1}{R_1 + R_2} \tag{11.10}$$

Dividing the gain expression's numerator and denominator by $A_{OL}\beta$, we get:

$$A_{CL} = \frac{1/\beta}{1 + (1/\beta)/A_{OL}} \tag{11.11}$$

Substituting β we then get:

$$A_{CL} = \frac{(1 + R_2/R_1)}{1 + (1 + R_2/R_1)/A_{OL}} \tag{11.12}$$

We note, with reassurance, that as $A_{OL} \to \infty$, the closed-loop gain approaches the ideal or nominal closed-loop gain, $1 + R_2/R_1$.

11.2.2 Inverting Amplifier

For the inverting configuration, we will use the following model for the op-amp, which has an infinite input resistance, zero output resistance, but a finite open-loop gain, A_{OL}:

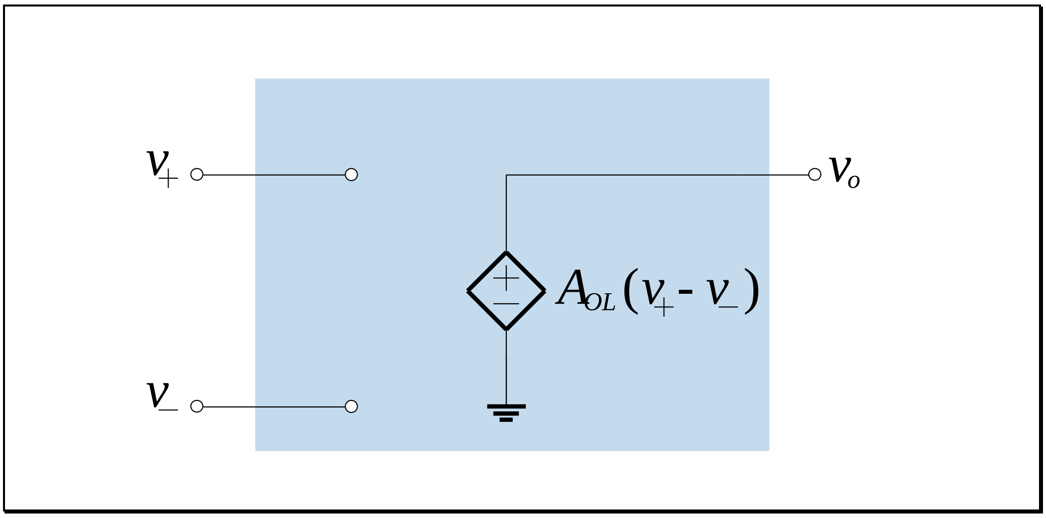


Figure 11.5

If we substitute this model into the inverting amplifier configuration, we get:

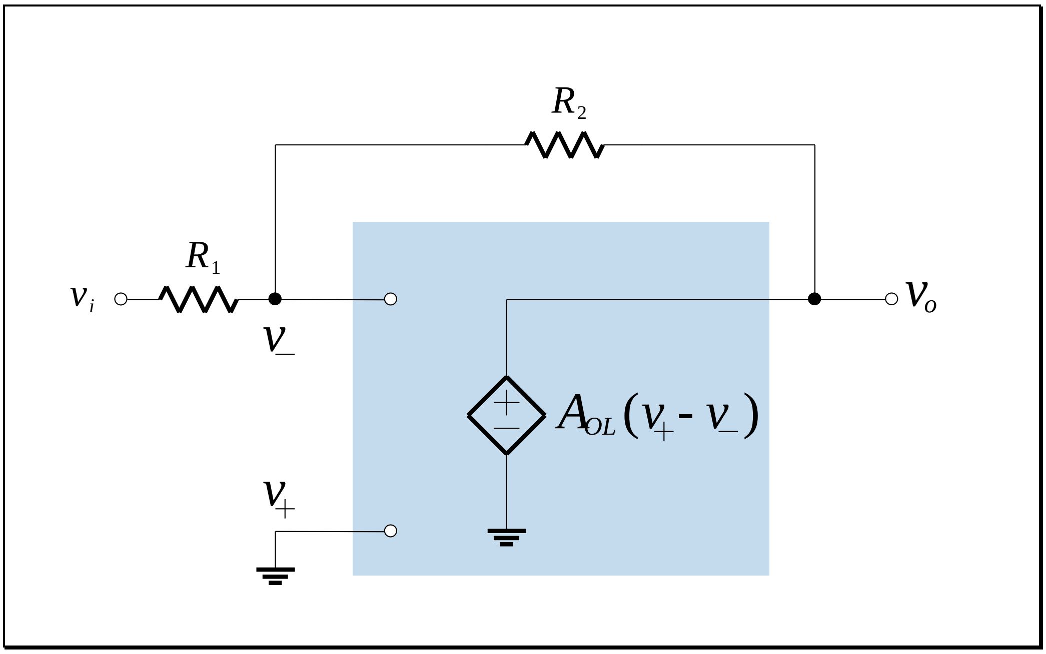


Figure 11.6

11.10

Since $v_+ = 0$, the voltage at the inverting input terminal must be $v_- = -v_o / A_{OL}$. The current i_1 through R_1 is therefore:

$$i_1 = \frac{v_i - (-v_o / A_{OL})}{R_1} = \frac{v_i + v_o / A_{OL}}{R_1} \tag{11.13}$$

The infinite input resistance of the op-amp causes this current to go through R_2. The output voltage is then:

$$\begin{aligned} v_o &= -\frac{v_o}{A_{OL}} - R_2 i_1 \\ &= -\frac{v_o}{A_{OL}} - R_2 \left(\frac{v_i + v_o / A_{OL}}{R_1} \right) \end{aligned} \tag{11.14}$$

Collecting terms, the closed-loop gain is found to be:

$$A_{CL} = \frac{v_o}{v_i} = \frac{-R_2 / R_1}{1 + (1 + R_2 / R_1) / A_{OL}} \tag{11.15}$$

Again, we note with reassurance, that as $A_{OL} \to \infty$, the closed-loop gain approaches the ideal or nominal closed-loop gain, $-R_2 / R_1$.

11.2.3 Percent Gain Error

We observe that the denominators of both the noninverting and inverting closed-loop gain expressions are identical. This is a result of the fact that both configurations have the same feedback loop, which can be seen if the input signal sources are set to zero.

The numerators, however, are different, for the numerator gives the ideal or nominal closed-loop gain, which we will denote *G*. The percent gain error between the actual and ideal gain, for either configuration, is then:

$$\begin{aligned}
\varepsilon &= \frac{A_{CL} - G}{G} \\
&= \frac{\dfrac{G}{1+\left(1+R_2/R_1\right)/A_{OL}} - G}{G} \\
&= \frac{1}{1+\left(1+R_2/R_1\right)/A_{OL}} - 1 \\
&= -\frac{\left(1+R_2/R_1\right)/A_{OL}}{1+\left(1+R_2/R_1\right)/A_{OL}} \\
&= -\frac{\left(1+R_2/R_1\right)}{A_{OL}+\left(1+R_2/R_1\right)} \times 100\,\%
\end{aligned} \tag{11.16}$$

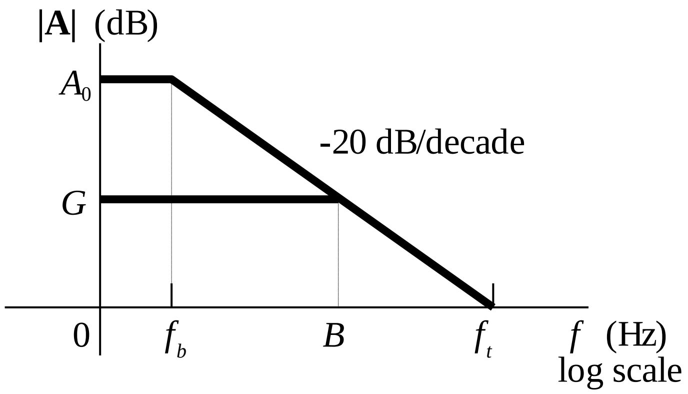
For example, if an op-amp with $A_{OL} = 10^4$ is used to design a noninverting amplifier with a nominal closed-loop gain of $G = 100$, we would expect the closed-loop gain to be about 1% below the nominal value.

Thus, in either configuration, to achieve gain accuracy we must have:

$$A_{OL} >> \left(1+R_2/R_1\right) \tag{11.17}$$

11.3 Finite Bandwidth

The open-loop gain of an op-amp is finite and decreases with frequency. The gain is quite high at DC and low frequencies, but it starts to fall off at a rather low frequency (10's of Hz). Most op-amps have a capacitor included within the IC whose function is to cause the op-amp to have a single-time-constant (STC) lowpass response shown:



This process of modifying the open-loop gain is termed frequency compensation, and its purpose is to ensure that op-amp circuits will be stable (as opposed to oscillating).

For frequencies $f >> f_b$ (about 10 times and higher), the magnitude of the open-loop gain **A** can be approximated as:

$$|\mathbf{A}| \approx \frac{f_t}{f}$$

The frequency f_t where the op-amp has a gain of 1 (or 0 dB) is known as the *unity-gain bandwidth*. Datasheets of internally compensated op-amps normally call f_t the *gain-bandwidth product*, since:

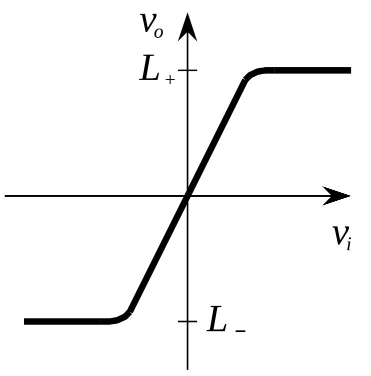
$$f_t = A_0 f_b$$

The noninverting amplifier configuration exhibits a constant gain-bandwidth product equal to f_t of the op-amp. Thus, you can easily determine the "bandwidth", B, of a non-inverting amplifier with a gain, G, since the gain-bandwidth product is a constant:

$$GB = f_t = \text{constant}$$

11.4 Output Voltage Saturation

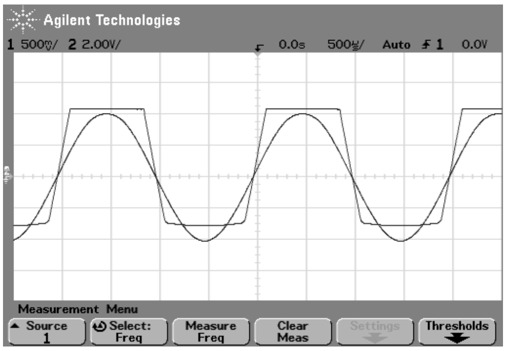
Real amplifiers can only output a voltage signal that is within the capabilities of the internal circuitry and the external DC power supplies. When amplifier outputs approach their output limitation, they are said to *saturate* – they cannot provide the output that is required by a linear characteristic. The resulting transfer characteristic, with the positive and negative saturation levels denoted L_+ and L_- respectively, is shown below:



The transfer characteristic of a real amplifier, showing that it saturates eventually

Figure 11.7

Each of the two saturation levels is usually within a volt or so of the voltage of the corresponding power supply. Obviously, in order to avoid distorting the output signal waveform, the input signal swing must be kept within the linear range of operation. If we don't, then the output waveform becomes distorted and eventually gets *clipped* at the output saturation levels.



The input signal and the output signal of a saturated amplifier showing clipping

Figure 11.8

11.5 Output Current Limits

The output current of an op-amp is normally limited by design to prevent excessive power dissipation within the device which would destroy it. For example, the popular TL071 op-amp has a resistor in series with the output:

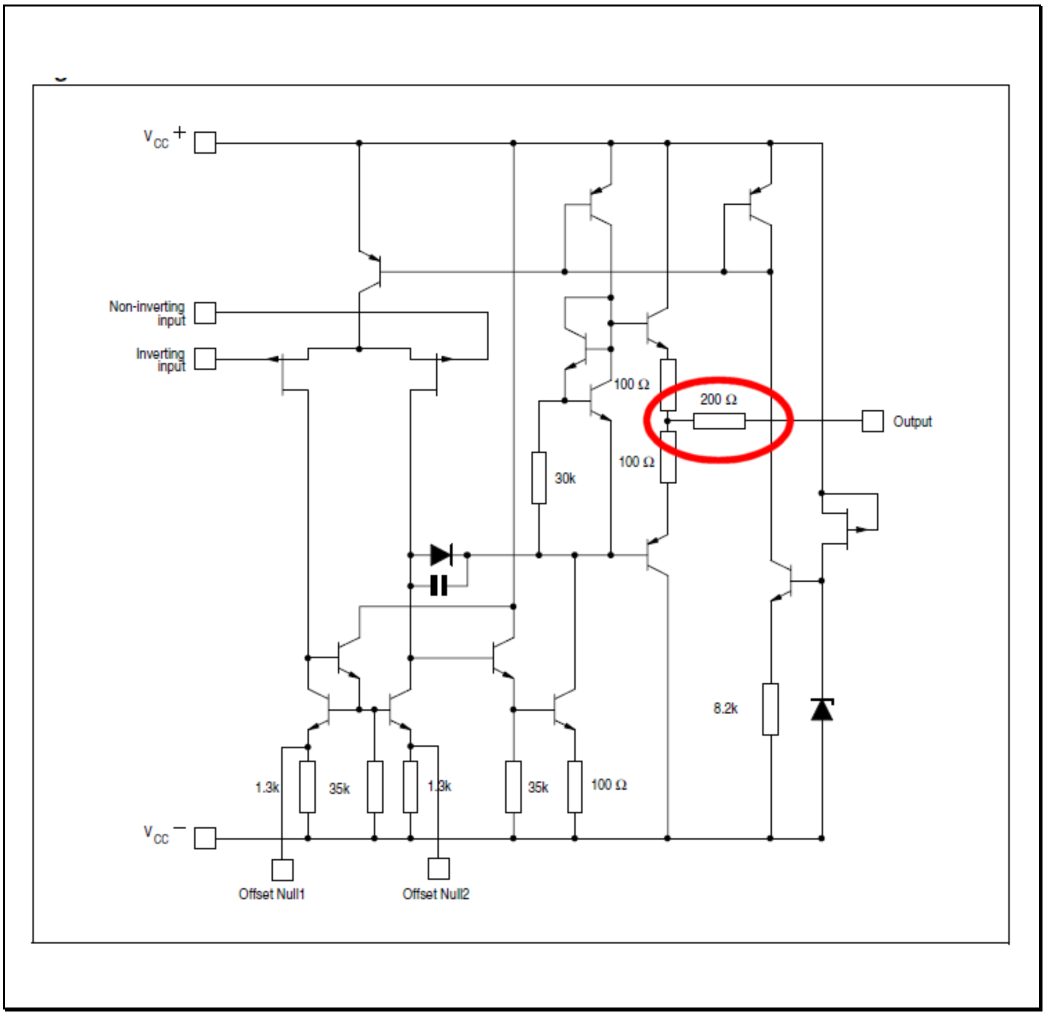


Figure 11.9

The datasheet specifies that, typically, the output current is limited to 40 mA. Thus, in designing closed-loop circuits utilizing the TL071, the designer has to ensure that under no conditions will the op-amp be required to supply an output current, in either direction, exceeding 40 mA. This current has to include both the current in the feedback circuit as well as the current supplied to any load. If the circuit requires a larger current, the op-amp output voltage will saturate at the level corresponding to the maximum allowed output current.

11.6 Slew Rate

The slew-rate limit of an op-amp is caused by a current source within the amplifier that limits the amount of current that can be supplied by the first stage of the amplifier. When the amplifier is pushed to the point where this limit is reached, it can no longer function properly. The slew-rate limit manifests itself as a maximum value of dv_o/dt for the amplifier because there is an internal amplifier capacitance that must be charged by the first-stage output current and a first-stage current limit thus corresponds to a maximum dv/dt for this capacitor. We therefore define:

$$\mathrm{SR} = \left.\frac{dv_o}{dt}\right|_{\max} \tag{11.18}$$

If the input signal applied to an op-amp circuit is such that it demands an output response that is faster than the specified value of SR, the op-amp will not comply. Rather, its output will change at the maximum possible rate, which is equal to its SR. The amplifier is then said to be *slewing*.

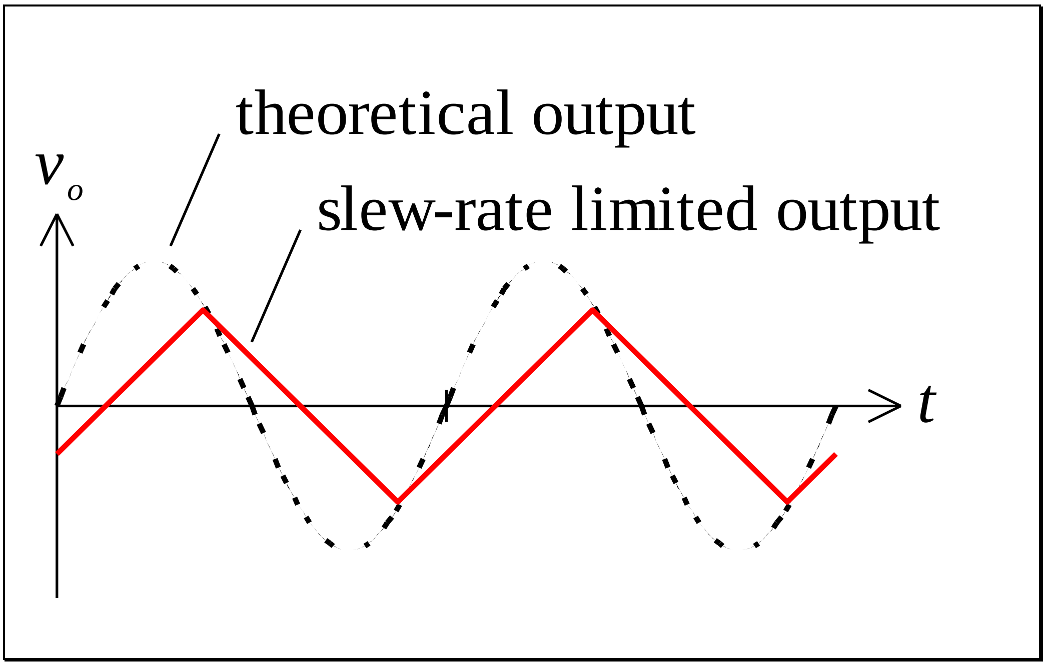


Figure 11.10

11.6.1 Full-Power Bandwidth

Slew rate can be related to the *full-power bandwidth* f_M, which is defined as the frequency at which a sine-wave output whose amplitude is equal to the rated op-amp output voltage starts to show distortion (due to op-amp slewing).

If the op-amp is producing a maximum amplitude output sinusoid, then:

$$v_o = V_m \sin(\omega t) \tag{11.19}$$

Then:

$$\frac{dv_o}{dt} = \omega V_m \cos(\omega t)$$

$$\left.\frac{dv_o}{dt}\right|_{\max} = \omega V_m \tag{11.20}$$

Denoting the slew rate by SR, it follows that:

$$\mathrm{SR} = \omega_M V_m \tag{11.21}$$

or:

$$f_M = \frac{\mathrm{SR}}{2\pi V_m} \tag{11.22}$$

11.7 Summary

- The input offset voltage, V_{OS}, is the magnitude of DC voltage that, when applied between the op-amp input terminals, with appropriate polarity, reduces the DC offset voltage at the output to zero.

- Direct currents exist at the input of the op-amp terminals. The average of these currents, I_B, is termed the bias current. The difference between these currents, I_{OS}, is called the offset current. These currents produce a DC voltage at the output. The effect of bias currents can be minimised by organising for the two op-amp inputs to "see" the same resistance.

- The finite open-loop gain, A_{OL}, of an op-amp causes a gain error. The open-loop gain also drop off with frequency. For internally compensated op-amps at high frequencies (> 1 kHz), the magnitude of the open-loop gain A_{OL} can be approximated as:

$$A_{OL} \approx \frac{f_t}{f}$$

  where f_t is specified on the op-amp datasheet and is called the *gain-bandwidth product*.

- Real amplifiers can only output a voltage signal that is within the capabilities of the internal circuitry and the external DC power supplies.

- The output current of an op-amp is normally limited by design to prevent excessive power dissipation within the device which would destroy it.

- The maximum rate at which the op-amp output voltage can change is called the slew rate:

$$\mathrm{SR} = \left.\frac{dv_o}{dt}\right|_{\mathrm{max}}$$

The slew rate is usually specified in op-amp datasheets in V/μs. Op-amp slewing can result in nonlinear distortion of output signal waveforms.

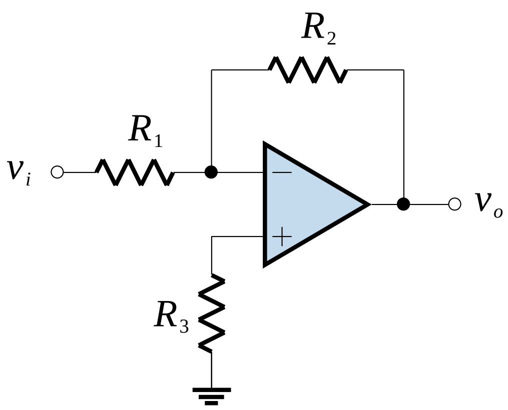
11.8 References

Sedra, A. and Smith, K.: *Microelectronic Circuits*, Saunders College Publishing, New York, 1991.

Exercises

1.

The op-amp circuit below is to be used at DC and very low frequencies.



A closed-loop gain of -200 is required. Specifications indicate that:

(i) the error due to finite open-loop gain cannot exceed 0.1%

(ii) DC output voltage due to input offset voltage ≤ 100 mV

(iii) DC output voltage due to input offset current ≤ 5 mV

Determine:

(a) the minimum open-loop gain required of the op-amp

(b) the maximum input offset voltage required of the op-amp

(c) assuming that the op-amp's input offset current is 10 nA, and a suitable compensating resistor R_3 is used, calculate the maximum value of R_2 that can be permitted.

12 The Phasor Concept

Contents

Introduction

The sinusoid is the most important function in electrical engineering. There are several reasons for this.

It was the great Swiss mathematician Euler (1707-1783) who first identified the fact that:

The special relationship enjoyed by sinusoids and linear circuits

For circuits described by linear differential equations a sinusoidal source yields a sinusoidal response. (11.1)

The response sinusoid has the same frequency as the source, it is however altered in *amplitude* and *phase*.

Only sinusoids have this property with respect to linear systems. For example, applying a square wave input does not produce a square wave output.

The source / response form-invariance of the sinusoid can be attributed to the fact that derivatives and integrals of a sinusoid yield sinusoids. Euler recognised that this was because the sinusoid is really composed of exponential components, and it is the exponential function which enjoys the rather peculiar property that:

$$\frac{d}{dt}\left(e^{t}\right)=e^{t} \tag{11.2}$$

That is, the derivative of the exponential function is itself an exponential function. We have already seen that this relationship plays a vital role in determining the solution to differential equations, where the *form* of each term has to be the same for the differential equation to "work", i.e. so that each term, such as $3dy/dt$ or $-5d^2y/dt^2$, can actually be added together.

The second major reason for the importance of the sinusoid can be attributed to Joseph Fourier (1768-1830), who in 1807 recognised that:

Any periodic function can be represented as the weighted sum of a family of sinusoids. (11.3)

Periodic signals are made up of sinusoids - Fourier Series

With this observation, we can analyse the behaviour of a linear circuit for any periodic forcing function by determining the response to the individual sinusoidal (or exponential) components and adding them up (superposition).

This decomposition of a periodic forcing function into a number of appropriately chosen sinusoidal forcing functions is a very powerful analytical method. It is called *frequency-domain* analysis, and we shall use it extensively.

The third reason is that the sinusoid is easy to generate, transmit and utilise. Most of our electric power is generated as a sinusoid, and the functional form of the sinusoid is needed to make most of our motors turn (it doesn't really matter about the lights or heaters). It is therefore of great practical importance to engineers who specialise in "power and machines".

Sinusoids are utilised in a lot of practical applications

It is also extremely important to engineers who specialise in communications. Our modern-day communications are "carried" through the air (and the vacuum of space) on "sinusoidal carriers". In fact, the basis of analog communication is to "modulate" some aspect of a sinusoid such as its amplitude (AM), its frequency (FM), or its phase (PM).

With digital communications, we encode the binary ones and zeros into some aspect of the sinusoid, e.g. amplitude shift keying (ASK), frequency shift keying (FSK), phase shift keying (PSK), and Gaussian minimum-shift keying (GMSK) which is used in the Groupe Spécial Mobile (GSM) mobile telephone network.

Thus, the sinusoid plays a prominent role in electrical engineering due to both its theoretical and practical importance.

12.4

12.1 Sinusoidal Signals

Consider the sinusoid:

$$v(t) = V_m \cos(\omega t) \tag{11.4}$$

which is shown graphically below:

The cosinusoid graphed

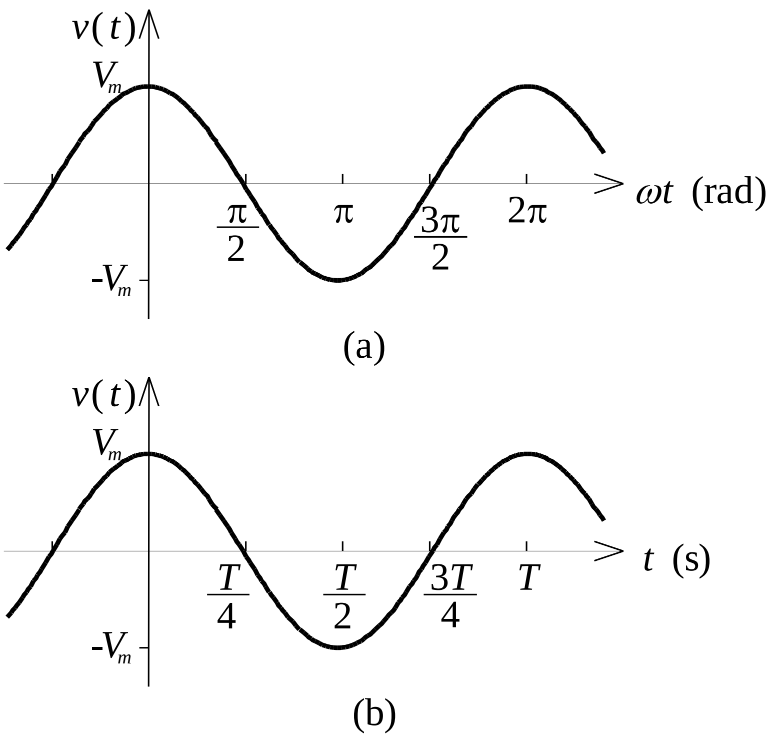


Figure 11.1

The *amplitude* of the cosine wave is V_m, and the *argument* is ωt. The *radian frequency* or *angular frequency* is ω. In Figure 11.1 (a) the function repeats itself every 2π radians, and its *period* is therefore 2π radians. In Figure 11.1 (b), the function is plotted against t and the period is now T where $\omega T = 2\pi$. A cosine wave having a period T seconds per period must execute $1/T$ periods each second – its *frequency* is therefore:

The relationship between frequency and period

$$f = \frac{1}{T} \tag{11.5}$$

Since $\omega T = 2\pi$, we have the relationship between frequency and radian frequency:

$$\omega = 2\pi f \tag{11.6}$$

The relationship between frequency and radian frequency

A more general form of the sinusoid is:

$$v(t) = V_m \cos(\omega t + \theta) \tag{11.7}$$

which now includes a *phase angle* θ in its argument $(\omega t + \theta)$. Eq. (11.7) is plotted below as a function of ωt, and the phase angle appears as the number of radians by which the original cosine wave, shown as a dotted line, is shifted to the *left* or earlier in time.

The phase angle defined

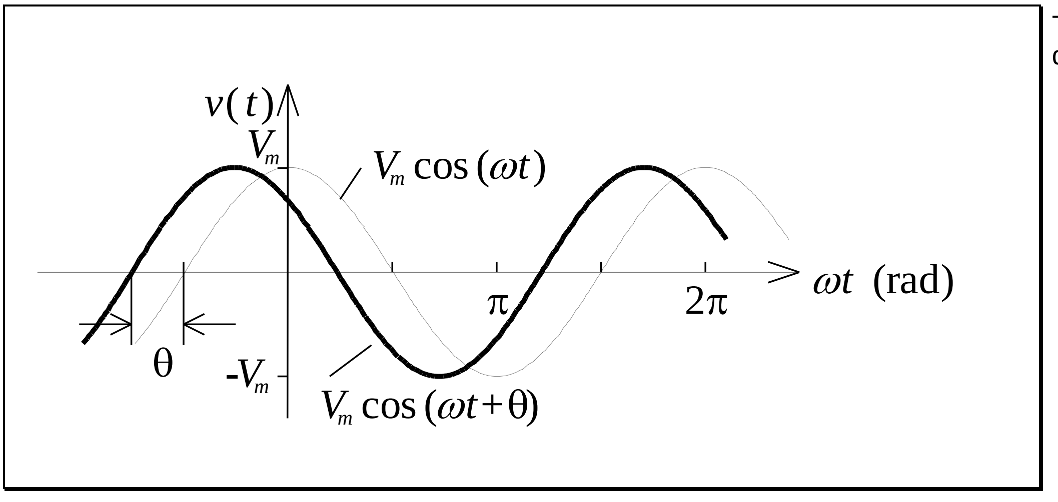


Figure 11.2

Since corresponding points on the cosinusoid $\cos(\omega t + \theta)$ occur θ rad earlier compared to $\cos(\omega t)$, we say that $\cos(\omega t + \theta)$ *leads* $\cos(\omega t)$ by θ rad. Conversely, $\cos(\omega t - \theta)$ *lags* $\cos(\omega t)$ by θ rad. In either case, leading or lagging, we say that the cosinusoids are *out of phase*. If the phase angles are equal, they are said to be *in phase*.

In electrical engineering the phase angle is commonly given in degrees rather than radians. Thus, instead of writing $v = 100\cos(100\pi t - \pi/6)$ we write $v = 100\cos(100\pi t - 30°)$.

12.2 Sinusoidal Steady-State Response

We are now ready to apply a sinusoidal forcing function to a simple circuit and obtain the forced response. We shall first write the differential equation which applies to the given circuit. We know that the complete solution is composed of the natural response and the forced response.

The natural response is determined by the circuit

The form of the natural response is independent of the mathematical form of the forcing function and depends only upon the type of circuit and the element values. We have already determined the natural response of simple *RC* and *RL* circuits.

The forced response for a sinusoid is termed the sinusoidal steady-state response

The forced response has the mathematical form of the forcing function. We therefore expect the forced response to be sinusoidal. The term *steady-state* is used synonymously with forced response, and the circuits we are about to analyse are commonly said to be in the "sinusoidal steady-state". Unfortunately, steady-state implies "not changing with time", but this is not correct – the sinusoidal forced response definitely changes with time. The steady-state simply refers to the condition which is reached after the natural response has died out.

Consider the series *RL* circuit below:

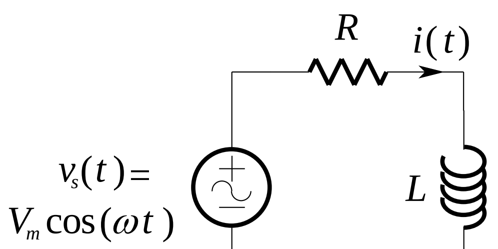


Figure 11.3

The sinusoidal source voltage $v_s = V_m \cos(\omega t)$ has been switched into the circuit at some remote time in the past, and the natural response has died out completely.

We seek the forced response, or steady-state response, or particular solution, and it must satisfy the differential equation:

$$L\frac{di}{dt}+Ri=V_m\cos(\omega t) \tag{11.8}$$

We now invoke the inverse differential operator to find a particular solution to the differential equation and write:

$$\begin{aligned}(LD+R)i&=V_m\cos(\omega t)\\ i&=\frac{1}{LD+R}V_m\cos(\omega t)\end{aligned} \tag{11.9}$$

We only know the effect of the inverse differential operator on exponentials. We therefore use Euler's formula:

$$e^{j\theta}=\cos\theta+j\sin\theta \tag{11.10}$$

Euler's identity

and note also that:

$$e^{-j\theta}=\cos\theta-j\sin\theta \tag{11.11}$$

Adding the above two equations, we get:

$$\cos\theta=\frac{e^{j\theta}+e^{-j\theta}}{2} \tag{11.12}$$

A cosinusoid expressed as a sum of complex exponentials

If we subtract the two equations, we get:

$$\sin\theta=\frac{e^{j\theta}-e^{-j\theta}}{j2} \tag{11.13}$$

A sinusoid expressed as a sum of complex exponentials

We can now express both cos and sin as complex exponentials.

Then Eq. (11.9) can be written:

$$i = \frac{V_m}{LD + R}\left[\frac{e^{j\omega t} + e^{-j\omega t}}{2}\right] \tag{11.14}$$

Now remembering that:

$$\frac{1}{f(D)}e^{st} = \frac{1}{f(s)}e^{st}, \quad f(s) \neq 0 \tag{11.15}$$

we can use superposition and rewrite Eq. (11.14) as:

$$i = \frac{V_m}{R + j\omega L}\frac{e^{j\omega t}}{2} + \frac{V_m}{R - j\omega L}\frac{e^{-j\omega t}}{2} \tag{11.16}$$

Realizing the denominators, we get:

$$\begin{aligned} i = {} & \frac{R - j\omega L}{R^2 + \omega^2 L^2}V_m \frac{e^{j\omega t}}{2} \\ & + \frac{R + j\omega L}{R^2 + \omega^2 L^2}V_m \frac{e^{-j\omega t}}{2} \end{aligned} \tag{11.17}$$

Collecting terms gives:

$$\begin{aligned} i = {} & \frac{RV_m}{R^2 + \omega^2 L^2}\left[\frac{e^{j\omega t} + e^{-j\omega t}}{2}\right] \\ & + \frac{\omega L V_m}{R^2 + \omega^2 L^2}\left[\frac{e^{j\omega t} - e^{-j\omega t}}{j2}\right] \end{aligned} \tag{11.18}$$

Using Euler's identities for $\cos\theta$ and $\sin\theta$, the forced response is obtained:

$$i(t) = \frac{RV_m}{R^2 + \omega^2 L^2}\cos(\omega t) + \frac{\omega L V_m}{R^2 + \omega^2 L^2}\sin(\omega t) \tag{11.19}$$

This expression is quite unwieldy, and a clearer picture of the response can be obtained by expressing the response as a single sinusoid with a phase angle. We will therefore let:

$$i(t) = A\cos(\omega t + \theta) \tag{11.20}$$

After expanding the function $\cos(\omega t + \theta)$, we have:

$$A\cos\theta\cos\omega t - A\sin\theta\sin\omega t = \frac{RV_m}{R^2 + \omega^2 L^2}\cos(\omega t) + \frac{\omega L V_m}{R^2 + \omega^2 L^2}\sin(\omega t) \tag{11.21}$$

Equating like coefficients of $\cos(\omega t)$ and $\sin(\omega t)$, we find:

$$\begin{aligned} A\cos\theta &= \frac{RV_m}{R^2 + \omega^2 L^2} \\ -A\sin\theta &= \frac{\omega L V_m}{R^2 + \omega^2 L^2} \end{aligned} \tag{11.22}$$

To find A and θ, we square both equations and add the results:

$$A^2\cos^2\theta + A^2\sin^2\theta = A^2$$

$$\frac{R^2 V_m^{\;2}}{\left(R^2+\omega^2 L^2\right)^2} + \frac{\omega^2 L^2 V_m^{\;2}}{\left(R^2+\omega^2 L^2\right)^2} = \frac{V_m^{\;2}}{R^2+\omega^2 L^2} \tag{11.23}$$

and also divide one equation by the other:

$$\frac{-A\sin\theta}{A\cos\theta} = -\tan\theta = \frac{\omega L}{R} \tag{11.24}$$

Hence:

$$A = \frac{V_m}{\sqrt{R^2+\omega^2 L^2}} \tag{11.25}$$

and:

$$\theta = -\tan^{-1}\frac{\omega L}{R} \tag{11.26}$$

The alternative form of the forced response therefore becomes:

The forced current response for the series *RL* circuit to a sinusoidal voltage source

$$i(t) = \frac{V_m}{\sqrt{R^2+\omega^2 L^2}}\cos\left(\omega t - \tan^{-1}\frac{\omega L}{R}\right) \tag{11.27}$$

We can see that the amplitude of the response is proportional to the amplitude of the forcing function – thus linearity between input and output holds (e.g. doubling the input leads to a doubling of the output). We can also see that the current decreases for any increase in R, ω or L, but not proportionately.

Also, the current lags the applied voltage by $\tan^{-1}(\omega L/R)$, an angle between $0°$ and $90°$. When $\omega = 0$ or $L = 0$, the current must be in phase with the voltage since the former situation is DC and the inductor appears as a short-circuit, and the latter situation is a resistive circuit. If $R = 0$ then the current lags the voltage by $90°$.

Note also that the frequency of the response is the same as the forcing function.

The applied voltage and the resultant current are shown below:

The forced response graphed, showing only an amplitude and phase change

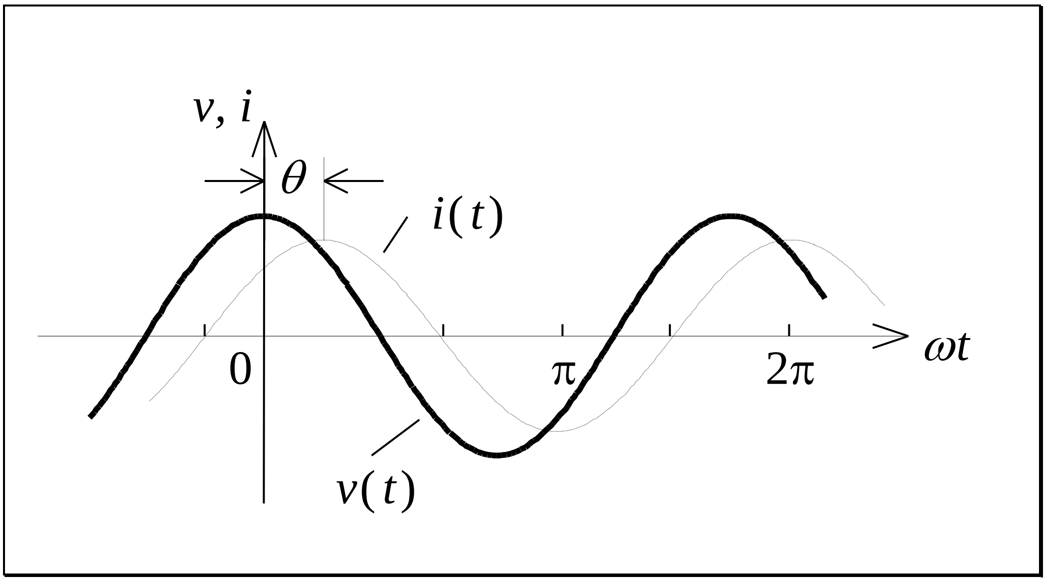


Figure 11.4

The fact that current lags the voltage in this simple *RL* circuit is now visually apparent.

The method by which we found the sinusoidal steady-state response for the simple *RL* circuit is quite intricate. It would be impractical to analyse every circuit by this method. We shall see in the next section that there is a way to simplify the analysis. It involves the formulation of complex algebraic equations instead of differential equations, but the advantage is that we can produce a set of complex algebraic equations for a circuit of any complexity. Sinusoidal steady-state analysis becomes almost as easy as the analysis of resistive circuits.

12.3 The Complex Forcing Function

It seems strange at first, but the use of complex quantities in sinusoidal steady-state analysis leads to methods which are simpler than those involving only real quantities.

Consider a sinusoidal source:

$$V_m \cos(\omega t + \theta) \tag{11.28}$$

which is connected to a general, passive, linear, time-invariant (LTI) circuit as shown below:

Excitation of a passive LTI circuit by a real sinusoid produces a real sinusoidal response

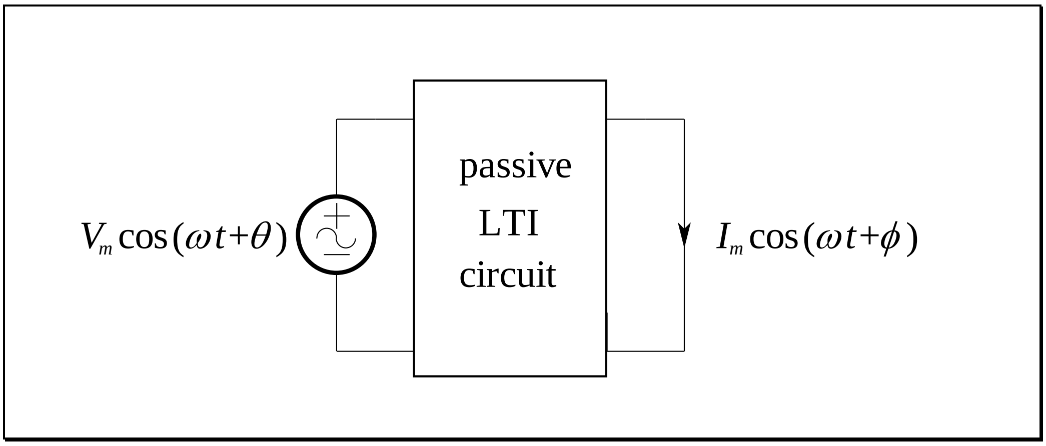


Figure 11.5

A current response in some other branch of the circuit is to be determined, and we know for a sinusoidal forcing function that the forced response is sinusoidal. Let the sinusoidal forced response be represented by:

$$I_m \cos(\omega t + \phi) \tag{11.29}$$

Note that the frequency stays the same – only the amplitude and phase are unknown.

If we delay the forcing function by $90°$, then since the system is time-invariant, the corresponding forced response must be delayed by $90°$ also (because the frequencies are the same). Thus, the forcing function:

$$V_m \cos(\omega t + \theta - 90°) = V_m \sin(\omega t + \theta) \tag{11.30}$$

will produce a response:

$$I_m \cos(\omega t + \phi - 90°) = I_m \sin(\omega t + \phi) \tag{11.31}$$

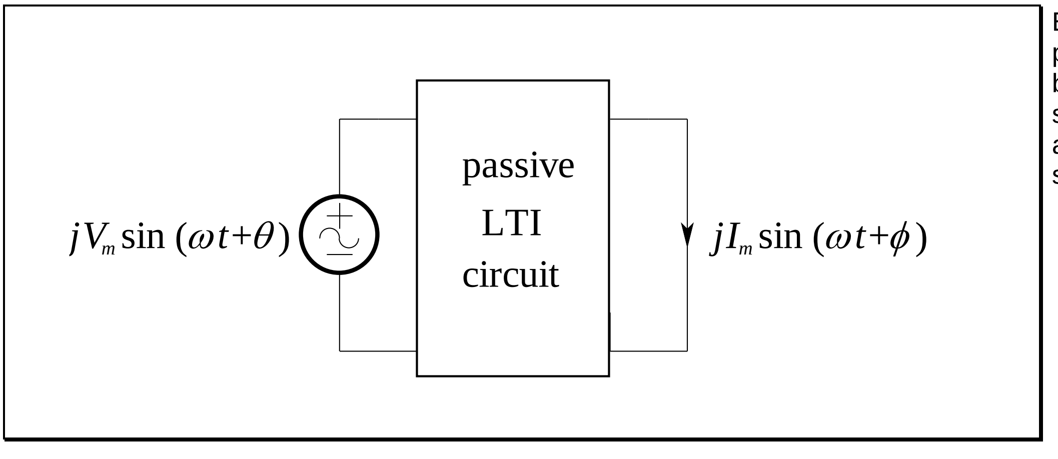
Since the circuit is linear, if we double the source, we double the response. In fact, if we multiply the source by any constant *k*, we achieve a response which is *k* times bigger. We now construct an imaginary source – we multiply the source by $j = \sqrt{-1}$. We thus apply:

$$jV_m \sin(\omega t + \theta) \tag{11.32}$$

and the response is:

$$jI_m \sin(\omega t + \phi) \tag{11.33}$$

The imaginary source and response are shown below:



Excitation of a passive LTI circuit by an imaginary sinusoid produces an imaginary sinusoidal response

Figure 11.6

We have applied a real source and obtained a real response, and we have applied an imaginary source and obtained an imaginary response. We can now use the superposition theorem (the circuit is linear) to find the response to a complex forcing function which is the sum of the real and imaginary forcing functions. Thus, the sum of the forcing functions of Eqs. (11.28) and (11.32) is:

$$V_m \cos(\omega t + \theta) + jV_m \sin(\omega t + \theta) \tag{11.34}$$

and it produces a response which is the sum of Eqs. (11.29) and (11.33):

$$I_m \cos(\omega t + \phi) + jI_m \sin(\omega t + \phi) \tag{11.35}$$

The complex source and response may be represented more simply by applying Euler's identity. Thus, the forcing function:

$$V_m e^{j(\omega t + \theta)} \tag{11.36}$$

produces:

$$I_m e^{j(\omega t + \phi)} \tag{11.37}$$

The complex source and response are illustrated below:

Excitation of a passive LTI circuit by a complex source produces a complex response

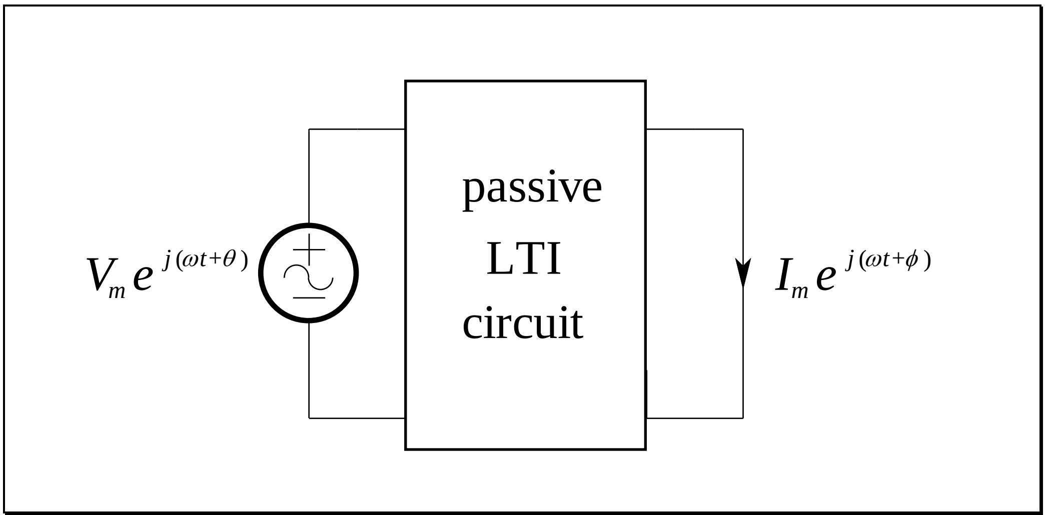


Figure 11.7

We are now ready to see how this helps with sinusoidal analysis. We first note that the real part of the complex response is produced by the real part of the complex forcing function, and the imaginary part of the complex response is produced by the imaginary part of the complex forcing function.

Our strategy for sinusoidal analysis will be to apply a complex forcing function whose real part is the given real forcing function – we should then obtain a complex response whose real part is the desired real response.

We analyse circuits in the sinusoidal steady-state by using a complex forcing function whose real part is the given real forcing function

We will try this strategy on the previous *RL* circuit:

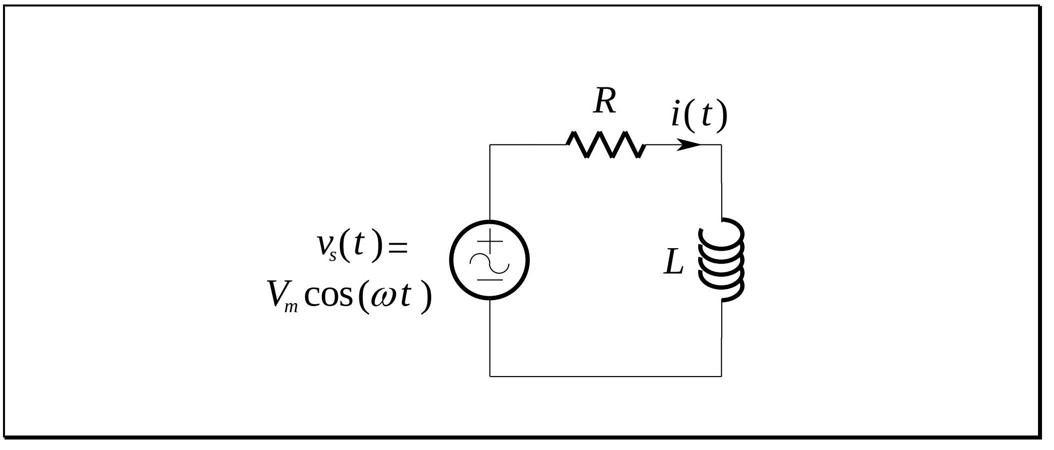


Figure 11.8

The real source $V_m \cos(\omega t)$ is applied, and the real response $i(t)$ is desired.

We first construct the complex forcing function by adding an appropriate imaginary component to the given real forcing function. The necessary complex source is:

$$V_m \cos(\omega t) + jV_m \sin(\omega t) = V_m e^{j\omega t} \tag{11.38}$$

The complex response which results is expressed in terms of an unknown amplitude I_m and an unknown phase angle ϕ:

$$I_m e^{j(\omega t+\phi)} \tag{11.39}$$

Writing the differential equation for this circuit:

$$Ri + L\frac{di}{dt} = v_s \tag{11.40}$$

we insert our complex expressions for v_s and i:

$$RI_m e^{j(\omega t+\phi)} + L\frac{d}{dt}\left(I_m e^{j(\omega t+\phi)}\right) = V_m e^{j\omega t} \tag{11.41}$$

Taking the indicated derivative gives:

Using complex sources and responses reduces the original differential equation to a complex algebraic equation

$$RI_m e^{j(\omega t+\phi)} + j\omega L I_m e^{j(\omega t+\phi)} = V_m e^{j\omega t} \tag{11.42}$$

which is a complex *algebraic* equation. This is a considerable advantage – we have turned a *differential* equation into an *algebraic* equation. The only "penalty" is that the algebraic equation uses complex numbers. It will be seen later that this is not a significant disadvantage.

In order to determine the value of I_m and ϕ, we divide through by the common factor $e^{j\omega t}$:

$$RI_m e^{j\phi} + j\omega L I_m e^{j\phi} = V_m \tag{11.43}$$

Factoring the left side gives:

$$(R + j\omega L) I_m e^{j\phi} = V_m \tag{11.44}$$

Rearranging, we have:

The complex response expressed in rectangular form

$$I_m e^{j\phi} = \frac{V_m}{R + j\omega L} \tag{11.45}$$

This is the complex response, and it was obtained in a few easy steps. If we express the response in exponential or polar form, we have:

$$I_m e^{j\phi} = \frac{V_m}{\sqrt{R^2 + \omega^2 L^2}} e^{j\left[-\tan^{-1}(\omega L/R)\right]} \tag{11.46}$$

The complex response expressed in polar form

Thus, by comparison:

$$I_m = \frac{V_m}{\sqrt{R^2 + \omega^2 L^2}} \tag{11.47}$$

and:

$$\phi = -\tan^{-1}(\omega L/R) \tag{11.48}$$

We said that the complex response was:

$$I_m e^{j(\omega t + \phi)} = I_m e^{j\phi} e^{j\omega t} \tag{11.49}$$

and that the real response was just the real part of the complex response. Therefore, the real response $i(t)$ is obtained by multiplying both sides of Eq. (11.46) by $e^{j\omega t}$ and taking the real part. Thus:

The real sinusoidal steady-state response

$$\begin{aligned} i(t) &= I_m \cos(\omega t + \phi) \\ &= \frac{V_m}{\sqrt{R^2 + \omega^2 L^2}} \cos\left(\omega t - \tan^{-1}\frac{\omega L}{R}\right) \end{aligned} \tag{11.50}$$

This agrees with the response derived before using the *D* operator.

Although the analysis was straightforward, we have not yet taken advantage of the full power of the complex representation. In order to do so, we must introduce the concept of the *phasor*.

12.4 The Phasor

A sinusoidal voltage or current *at a given frequency* is characterized by only two parameters, an amplitude and a phase. The complex representation of the voltage or current is characterized by a magnitude and an angle. For example, the assumed sinusoidal form of the current response in the previous example was:

$$I_m \cos(\omega t + \phi) \tag{11.51}$$

and the corresponding representation of this current in complex form is:

$$I_m e^{j(\omega t + \phi)} \tag{11.52}$$

Once I_m and ϕ are specified, the current is exactly defined. Throughout any linear circuit operating in the sinusoidal steady-state at a single frequency ω, every voltage and current may be characterized completely by a knowledge of its amplitude and phase angle.

A sinusoid of a given frequency is specified by an amplitude and phase

All sinusoidal responses in a linear circuit have a frequency of ω. Therefore, instead of writing $I_m \cos(\omega t + \phi)$, we could just say "amplitude I_m" and "phase ϕ".

A complex response of a given frequency is specified by a magnitude and angle

All complex responses in a linear circuit have the factor $e^{j\omega t}$. Therefore, instead of writing $I_m e^{j(\omega t + \phi)}$, we could just say "magnitude I_m" and "angle ϕ".

Thus, we can simplify the voltage source and current response of the example by representing them concisely as complex numbers:

$$V_m e^{j0^\circ} \quad \text{and} \quad I_m e^{j\phi} \tag{11.53}$$

We usually write the complex representation in polar form. Thus, the source voltage:

$$v(t) = V_m \cos(\omega t) \tag{11.54}$$

is represented in complex form as:

$$\mathbf{V} = V_m \angle 0° \tag{11.55}$$

and the current response:

$$i(t) = I_m \cos(\omega t + \phi) \tag{11.56}$$

A general sinusoid…

as:

$$\mathbf{I} = I_m \angle \phi \tag{11.57}$$

…and its phasor representation

The abbreviated complex representation is called a *phasor*. Phasors are printed in boldface because they are effectively like a *vector*, they have a magnitude and direction (angle). In hand writing, we normally place a tilde underneath:

$$\underset{\sim}{V} = V_m \angle 0° \quad \text{and} \quad \underset{\sim}{I} = I_m \angle \phi \tag{11.58}$$

Capital letters are used to represent phasors because they are constants – they are not functions of time.

In general, we refer to $x(t)$ as a *time-domain representation* and the corresponding phasor $\mathbf{X}$ as a *frequency-domain representation*.

Introducing the "frequency-domain"

We can see that the *magnitude* of the complex representation is the *amplitude* of the sinusoid and the *angle* of the complex representation is the *phase* of the sinusoid.

It is a simple matter to convert a signal from the time-domain to the frequency-domain – it is achieved by inspection:

The time-domain and frequency-domain relationships for a sinusoid

$$x(t) = A\cos(\omega t + \phi) \quad \Leftrightarrow \quad \mathbf{X} = Ae^{j\phi}$$

$$\text{amplitude} \quad \Leftrightarrow \quad \text{magnitude}$$

$$\text{phase} \quad \Leftrightarrow \quad \text{angle} \qquad (11.59)$$

EXAMPLE 12.1 Phasor Representation

If $x(t) = 3\sin(\omega t - 30°)$ then we have to convert to our cos notation: $x(t) = 3\cos(\omega t - 120°)$. Therefore $\mathbf{X} = 3\angle -120°$.

Note carefully that $\mathbf{X} \neq 3\cos(\omega t - 120°)$. All we can say is that $x(t) = 3\cos(\omega t - 120°)$ is *represented* by $\mathbf{X} = 3\angle -120°$.

Phasors make manipulating sinusoids of the same frequency easy

The convenience of complex numbers extends beyond their compact representation of the amplitude and phase. The sum of two phasors corresponds to the sinusoid which is the sum of the two component sinusoids represented by the phasors. That is, if $x_3(t) = x_1(t) + x_2(t)$ where $x_1(t)$, $x_2(t)$ and $x_3(t)$ are sinusoids with the same frequency, then $\mathbf{X}_3 = \mathbf{X}_1 + \mathbf{X}_2$.

EXAMPLE 12.2 Phasor Representation

If $x_3(t) = \cos\omega t - 2\sin\omega t$ then $\mathbf{X}_3 = 1\angle 0° - 2\angle -90° = 1 + j2 = 2.24\angle 63°$ which corresponds to $x_3(t) = 2.24\cos(\omega t + 63°)$.

12.4.1 Formalisation of the Relationship between Phasor and Sinusoid

Using Euler's identity:

$$e^{j\theta} = \cos\theta + j\sin\theta \tag{11.60}$$

we have:

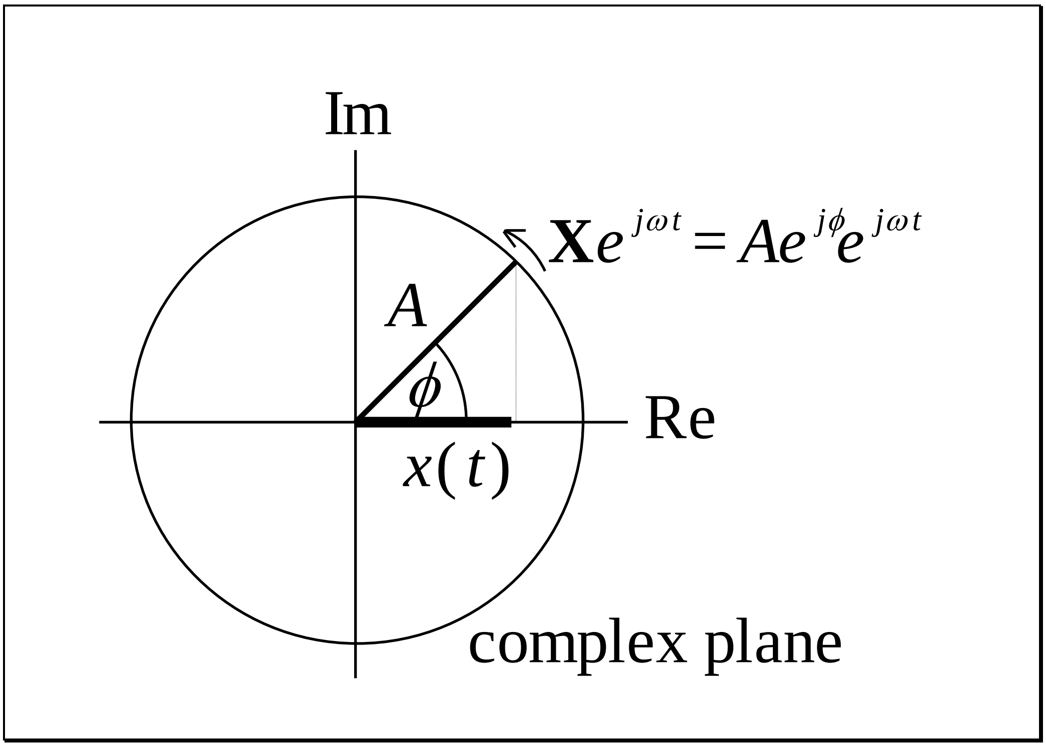
$$Ae^{j\phi}e^{j\omega t} = Ae^{j(\omega t+\phi)} = A\cos(\omega t+\phi) + jA\sin(\omega t+\phi) \tag{11.61}$$

We can see that the sinusoid $A\cos(\omega t+\phi)$ represented by the phasor $\mathbf{X} = Ae^{j\phi}$ is equal to the real part of $\mathbf{X}e^{j\omega t}$. Therefore:

$$x(t) = \mathrm{Re}\left\{\mathbf{X}e^{j\omega t}\right\} \tag{11.62}$$

The phasor / time-domain relationship

This can be visualised as:



Graphical interpretation of rotating phasor / time-domain relationship

Figure 11.9

Run the Phasor simulation program to see this view of phasors in action!

12.4.2 Graphical Illustration of the Relationship between a Phasor and its Corresponding Sinusoid

Consider the representation of a sinusoid by its phasor: $x(t) = \mathrm{Re}\{\mathbf{X}e^{j\omega t}\}$. Graphically, $x(t)$ can be "generated" by taking the projection of the rotating phasor formed by multiplying $\mathbf{X}$ by $e^{j\omega t}$, onto the real axis:

A sinusoid can be generated by taking the real part of a rotating complex number

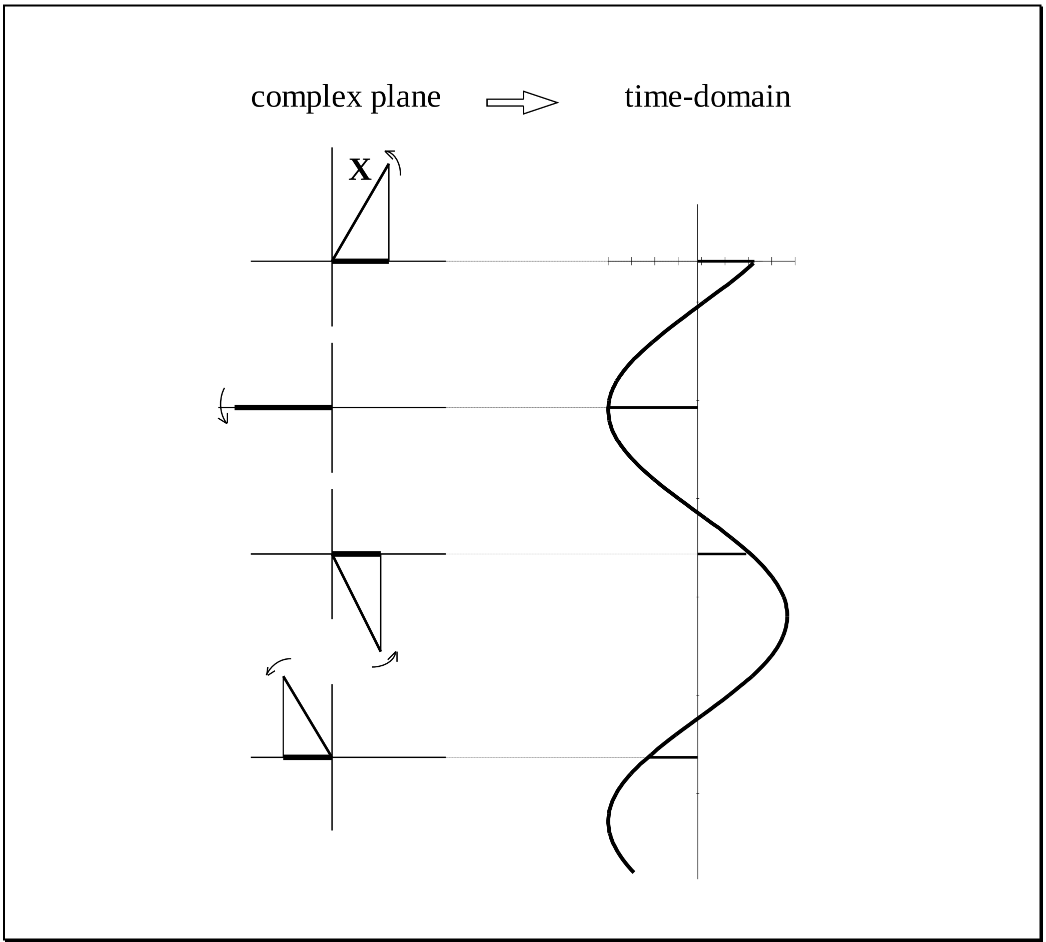


Figure 11.10

Phasor Representations

Phasors can be represented in four different ways:

$\mathbf{X} = X_m\angle\phi$	polar form
$\mathbf{X} = X_m e^{j\phi}$	exponential form
$\mathbf{X} = X_m(\cos\phi + j\sin\phi)$	trigonometric form
$\mathbf{X} = a + jb$	rectangular form

12.5 Phasor Relationships for *R*, *L* and *C*

Now that we can transform into and out of the frequency-domain, we can derive the phasor relationships for each of the three passive circuit elements. This will lead to a great simplification of sinusoidal steady-state analysis.

Phasor relationships for the passive elements

We will begin with the defining time-domain equation for each of the elements, and then let both the voltage and current become complex quantities. After dividing throughout the equation by $e^{j\omega t}$, the desired relationship between the phasor voltage and phasor current will become apparent.

12.5.1 Phasor Relationships for a Resistor

The resistor provides the simplest case. The defining time-domain equation is:

$$v(t) = Ri(t) \tag{11.63}$$

If we apply a complex voltage $V_m e^{j(\omega t+\theta)}$ and assume a complex current $I_m e^{j(\omega t+\phi)}$, we obtain:

$$V_m e^{j(\omega t+\theta)} = RI_m e^{j(\omega t+\phi)} \tag{11.64}$$

By dividing throughout by $e^{j\omega t}$, we find:

$$V_m e^{j\theta} = RI_m e^{j\phi} \tag{11.65}$$

or in polar form:

$$V_m \angle\theta = RI_m \angle\phi \tag{11.66}$$

But $V_m\angle\theta$ and $I_m\angle\phi$ are just the voltage and current phasors $\mathbf{V}$ and $\mathbf{I}$. Thus:

$$\boxed{\mathbf{V} = R\mathbf{I}} \tag{11.67}$$

Phasor **V-I** relationship for a resistor

Equality of the angles θ and ϕ is apparent, and the current and voltage are thus in phase.

The voltage-current relationship in phasor form for a resistor has the same form as the relationship between the time-domain voltage and current as illustrated below:

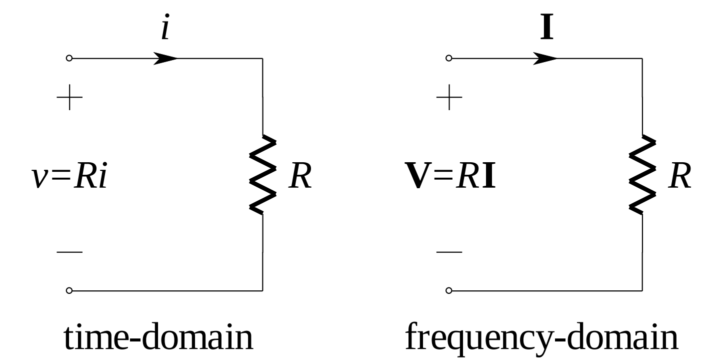


Figure 11.11

EXAMPLE 12.3 Phasor Analysis with a Resistor

Assume a voltage of $8\cos(100t - 50°)$ across a $4\,\Omega$ resistor. Working in the time-domain, the current is:

$$i(t) = \frac{v(t)}{R} = 2\cos(100t - 50°)$$

The phasor form of the same voltage is $8\angle -50°$, and therefore:

$$\mathbf{I} = \frac{\mathbf{V}}{R} = 2\angle -50°$$

If we transform back to the time-domain, we get the same expression for the current.

No work is saved for a resistor by analysing in the frequency-domain – because the resistor has a linear relationship between voltage and current.

12.5.2 Phasor Relationships for an Inductor

The defining time-domain equation is:

$$v(t) = L\frac{di(t)}{dt} \tag{11.68}$$

After applying the complex voltage and current equations, we obtain:

$$V_m e^{j(\omega t+\theta)} = L\frac{d}{dt}\left(I_m e^{j(\omega t+\phi)}\right) \tag{11.69}$$

Taking the indicated derivative:

$$V_m e^{j(\omega t+\theta)} = j\omega L I_m e^{j(\omega t+\phi)} \tag{11.70}$$

By dividing throughout by $e^{j\omega t}$, we find:

$$V_m e^{j\theta} = j\omega L I_m e^{j\phi} \tag{11.71}$$

Thus the desired phasor relationship is:

$$\boxed{\mathbf{V} = j\omega L\mathbf{I}} \tag{11.72}$$

Phasor **V-I** relationship for an inductor

The time-domain equation Eq. (11.68) has become an algebraic equation in the frequency-domain. The angle of $j\omega L$ is exactly $+90°$ and you can see from Eq. (11.71) that $\theta = 90° + \phi$. $\mathbf{I}$ must therefore *lag* $\mathbf{V}$ by $90°$ in an inductor.

The phasor relationship for an inductor is indicated below:

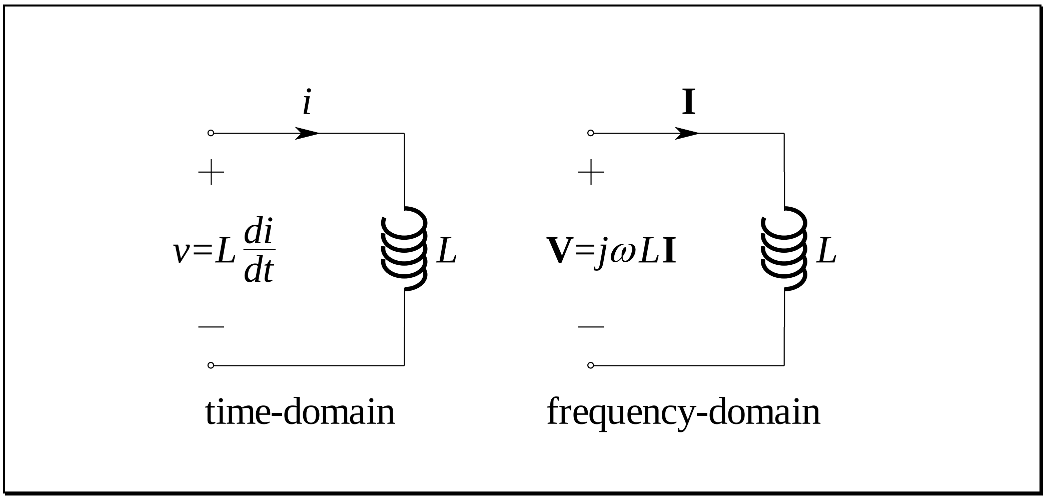


Figure 11.12

EXAMPLE 12.4 Phasor Analysis with an Inductor

Assume a voltage of $8\cos(100t-50°)$ across a $4\,\mathrm{H}$ inductor. Working in the time-domain, the current is:

$$\begin{aligned} i(t) &= \int \frac{v(t)}{L}\,dt \\ &= \int 2\cos(100t-50°)\,dt \\ &= 0.02\sin(100t-50°) \\ &= 0.02\cos(100t-140°) \end{aligned}$$

The phasor form of the same voltage is $8\angle-50°$, and therefore:

$$\mathbf{I}=\frac{\mathbf{V}}{j\omega L}=\frac{8\angle-50°}{100(4)\angle 90°}=0.02\angle-140°$$

If we transform back to the time-domain, we get the same expression for the current.

12.5.3 Phasor Relationships for a Capacitor

The defining time-domain equation is:

$$i(t) = C\frac{dv(t)}{dt} \tag{11.73}$$

After applying the complex voltage and current equations, we obtain:

$$I_m e^{j(\omega t+\phi)} = C\frac{d}{dt}\left(V_m e^{j(\omega t+\theta)}\right) \tag{11.74}$$

Taking the indicated derivative:

$$I_m e^{j(\omega t+\phi)} = j\omega C V_m e^{j(\omega t+\theta)} \tag{11.75}$$

By dividing throughout by $e^{j\omega t}$, we find:

$$I_m e^{j\phi} = j\omega C V_m e^{j\theta} \tag{11.76}$$

Thus the desired phasor relationship is:

$$\mathbf{I} = j\omega C\mathbf{V} \tag{11.77}$$

Phasor **V-I** relationship for a capacitor

Thus $\mathbf{I}$ *leads* $\mathbf{V}$ by $90°$ in a capacitor.

The time-domain and frequency-domain representations are compared below:

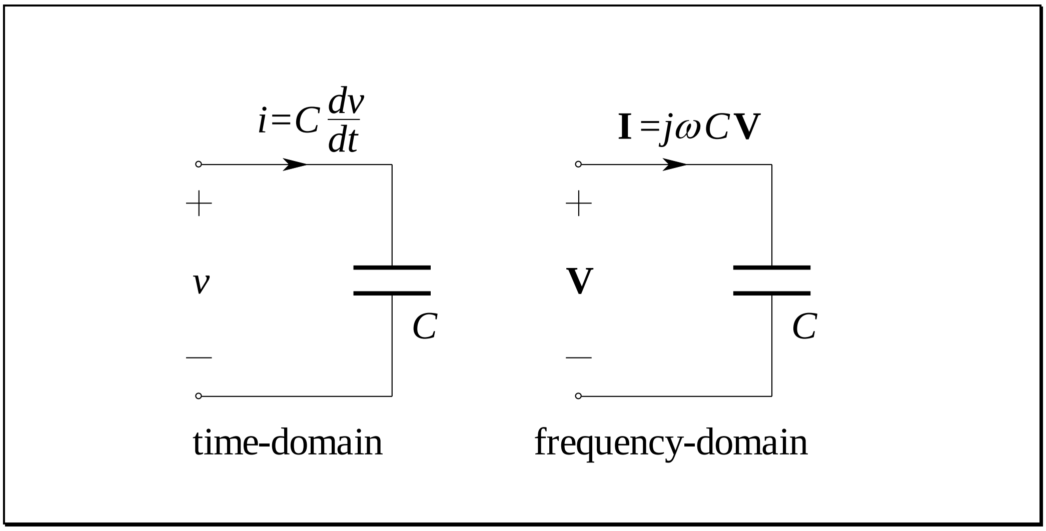


Figure 11.13

EXAMPLE 12.5 Phasor Analysis with a Capacitor

Assume a voltage of $8\cos(100t-50°)$ across a $4\,\mathrm{F}$ capacitor. Working in the time-domain, the current is:

$$\begin{aligned} i(t) &= C\frac{dv(t)}{dt} \\ &= 4\frac{d}{dt}8\cos(100t-50°) \\ &= -3200\sin(100t-50°) \\ &= 3200\cos(100t+40°) \end{aligned}$$

The phasor form of the same voltage is $8\angle-50°$, and therefore:

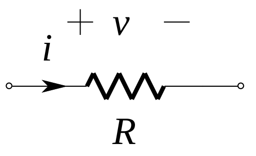
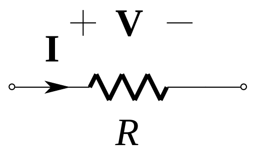
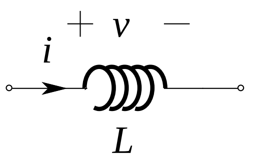
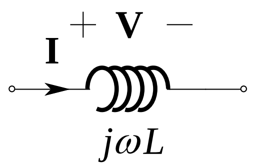
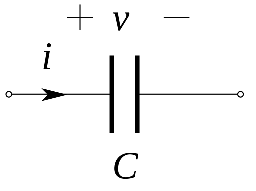
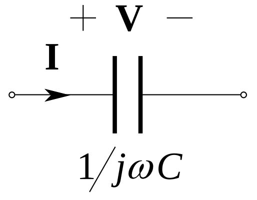
$$\mathbf{I} = j\omega C\mathbf{V} = 100(4)\angle 90° \times 8\angle-50° = 3200\angle 40°$$

If we transform back to the time-domain, we get the same expression for the current.

12.5.4 Summary of Phasor Relationships for *R*, *L* and *C*

We have now obtained the phasor $\mathbf{V}-\mathbf{I}$ relationships for the three passive elements. These results are summarized in the table below:

Summary of phasor **V-I** relationships for the passive elements

Time-domain		*Frequency-domain*	
	$v = Ri$	$\mathbf{V} = R\mathbf{I}$	
	$v = L\dfrac{di}{dt}$	$\mathbf{V} = j\omega L\mathbf{I}$	
	$v = \dfrac{1}{C}\int i\,dt$	$\mathbf{V} = \dfrac{1}{j\omega C}\mathbf{I}$	

All the phasor equations are algebraic. Each is also linear, and the equations relating to inductance and capacitance bear a great similarity to Ohm's Law.

Before we embark on using the phasor relationships in circuit analysis, we need to verify that KVL and KCL work for phasors. KVL in the time-domain is:

$$v_1(t) + v_2(t) + \cdots + v_n(t) = 0 \tag{11.78}$$

If all voltages are sinusoidal, we can now use Euler's identity to replace each real sinusoidal voltage by the complex voltage having the same real part, divide by $e^{j\omega t}$ throughout, and obtain:

$$\mathbf{V}_1 + \mathbf{V}_2 + \cdots + \mathbf{V}_n = 0 \tag{11.79}$$

KVL and KCL are obeyed by phasors

Thus KVL holds. KCL also holds by a similar argument.

12.5.5 Analysis Using Phasor Relationships

We now return to the series *RL* circuit that we considered several times before, shown as (a) in the figure below. We draw the circuit in the frequency-domain, as shown in (b):

A circuit and its frequency-domain equivalent

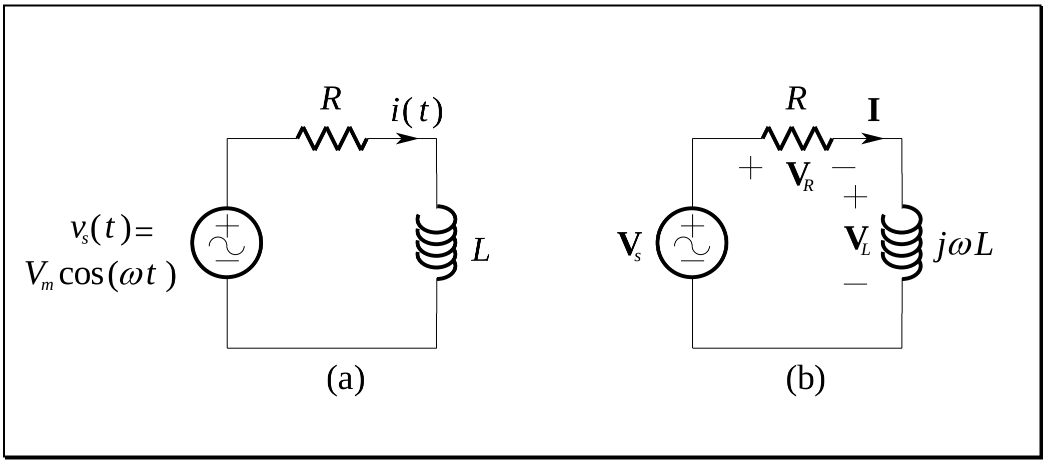


Figure 11.14

From KVL in the frequency-domain:

$$\mathbf{V}_R + \mathbf{V}_L = \mathbf{V}_s \tag{11.80}$$

We now insert the recently obtained $\mathbf{V}-\mathbf{I}$ relationships for the elements:

$$R\mathbf{I} + j\omega L\mathbf{I} = \mathbf{V}_s \tag{11.81}$$

The phasor current is then found:

$$\mathbf{I} = \frac{\mathbf{V}_s}{R + j\omega L} \tag{11.82}$$

The source has a magnitude of V_m and a phase of $0°$ (it is the reference by which all other phase angles are measured). Thus:

The response of the circuit in the frequency-domain

$$\mathbf{I} = \frac{V_m \angle 0°}{R + j\omega L} \tag{11.83}$$

The current may be transformed to the time-domain by first writing it in polar form:

$$\begin{aligned}\mathbf{I} &= \frac{V_m}{\sqrt{R^2+\omega^2 L^2}}\angle -\tan^{-1}(\omega L/R) \\ &= I_m \angle \phi\end{aligned} \tag{11.84}$$

Transforming back to the time-domain we get:

The response of the circuit in the time-domain

$$\begin{aligned}i(t) &= I_m \cos(\omega t+\phi) \\ &= \frac{V_m}{\sqrt{R^2+\omega^2 L^2}}\cos\left(\omega t - \tan^{-1}\frac{\omega L}{R}\right)\end{aligned} \tag{11.85}$$

which is the same result as we obtained before the "hard way".

12.6 Impedance

The voltage-current relationships for the three passive elements in the frequency-domain are:

$$\mathbf{V} = R\mathbf{I} \qquad \mathbf{V} = j\omega L\mathbf{I} \qquad \mathbf{V} = \frac{\mathbf{I}}{j\omega C} \tag{11.86}$$

If these equations are written as phasor-voltage phasor-current ratios, we get:

Phasor **V-I** relationships for the passive elements

$$\frac{\mathbf{V}}{\mathbf{I}} = R \qquad \frac{\mathbf{V}}{\mathbf{I}} = j\omega L \qquad \frac{\mathbf{V}}{\mathbf{I}} = \frac{1}{j\omega C} \tag{11.87}$$

These ratios are simple functions of the element values, and in the case of the inductor and capacitor, frequency. We treat these ratios in the same manner we treat resistances, with the exception that they are complex quantities and all algebraic manipulations must be those appropriate for complex numbers.

We define the ratio of the phasor voltage to the phasor current as *impedance*, symbolized by the letter $\mathbf{Z}$:

Impedance defined

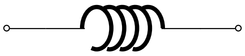
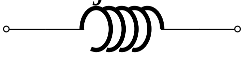
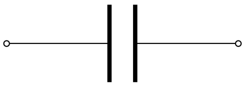
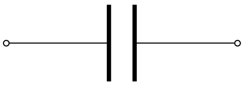
$$\mathbf{Z} = \frac{\mathbf{V}}{\mathbf{I}} \tag{11.88}$$

The impedance is a complex quantity having the dimensions of ohms.

Impedance is *not* a phasor and cannot be transformed to the time-domain by multiplying by $e^{j\omega t}$ and taking the real part.

In the table below, we show how we can represent a resistor, inductor or capacitor in the time-domain with its frequency-domain impedance:

Impedances of the three passive elements

Time-domain	*Frequency-domain*
R 	R 
L 	$j\omega L$ 
C 	$1/j\omega C$ 

Impedances may be combined in series and parallel by the same rules we use for resistances.

In a circuit diagram, a general impedance is represented by a rectangle:

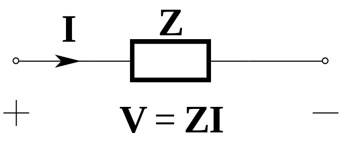
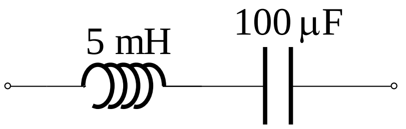


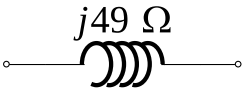
Figure 11.15

EXAMPLE 12.6 Impedance of an Inductor and Capacitor in Series

We have an inductor and capacitor in series:



At $\omega = 10^4 \text{ rads}^{-1}$, the impedance of the inductor is $\mathbf{Z}_L = j\omega L = j50\,\Omega$ and the impedance of the capacitor is $\mathbf{Z}_C = 1/j\omega C = -j1\,\Omega$. Thus the series combination is equivalent to $\mathbf{Z}_{eq} = \mathbf{Z}_L + \mathbf{Z}_C = j50 - j1 = j49\,\Omega$:



The impedance of inductors and capacitors is a function of frequency, and this equivalent impedance is only valid at $\omega = 10^4 \text{ rads}^{-1}$. For example, if $\omega = 5000 \text{ rads}^{-1}$, then the impedance would be $\mathbf{Z}_{eq} = j23\,\Omega$.

Impedance may be expressed in either polar or rectangular form.

In polar form an impedance is represented by:

$$\mathbf{Z} = |\mathbf{Z}|\angle\theta \tag{11.89}$$

No special names or symbols are assigned to the magnitude and angle. For example, an impedance of $100\angle -60°\,\Omega$ is described as having an impedance magnitude of $100\ \Omega$ and an angle of $-60°$.

Impedance is composed of a resistance (real part) and a reactance (imaginary part)

In rectangular form an impedance is represented by:

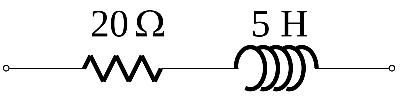
$$\mathbf{Z} = R + jX \tag{11.90}$$

The real part, *R*, is termed the *resistive component*, or *resistance*. The imaginary component, *X*, including sign, but excluding *j*, is termed the *reactive component*, or *reactance*. The impedance $100\angle -60^\circ\ \Omega$ in rectangular form is $50 - j86.6\ \Omega$. Thus, its resistance is $50\ \Omega$ and its reactance is $-86.6\ \Omega$.

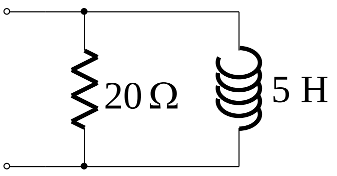
It is important to note that the resistive component of the impedance is not necessarily equal to the resistance of the resistor which is present in the circuit.

EXAMPLE 12.7 Impedance of a Resistor and Inductor in Series

Consider a resistor and an inductor in series:



At $\omega = 4\ \text{rads}^{-1}$, the equivalent impedance is $\mathbf{Z}_{eq} = 20 + j20\ \Omega$. In this case the resistive component of the impedance is equal to the resistance of the resistor because the network is a simple series network. Now consider the same elements placed in parallel:



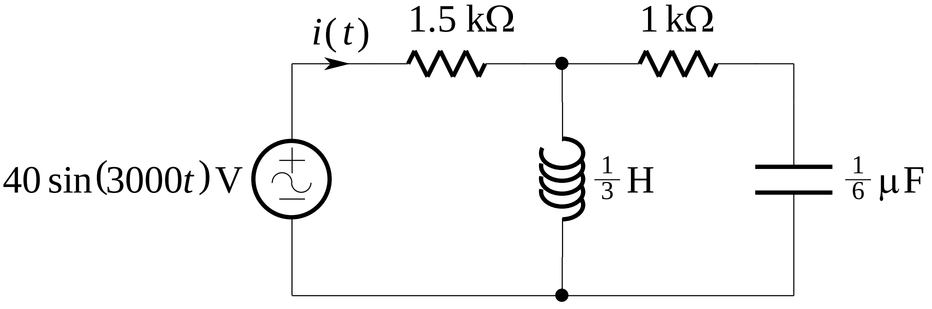
The equivalent impedance is:

$$\mathbf{Z}_{eq} = \frac{20(j20)}{20 + j20} = 10 + \text{j}10\ \Omega$$

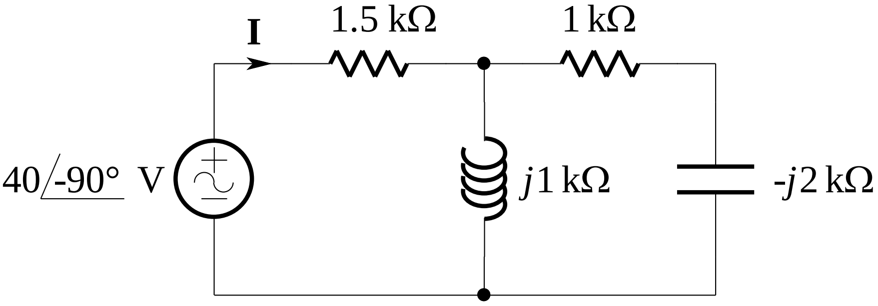
The resistive component of the impedance is now $10\ \Omega$.

EXAMPLE 12.8 Circuit Analysis using Impedance

We will use the impedance concept to analyse the *RLC* circuit shown below:



The circuit is shown in the time-domain, and a time-domain response is required. However, analysis should be carried out in the frequency-domain. We therefore begin by drawing the frequency-domain circuit – the source is transformed to the frequency-domain, becoming $40\angle -90°$, the response is transformed to the frequency-domain, being represented as $\mathbf{I}$, and the impedances of the inductor and capacitor, determined at $\omega = 3000 \text{ rads}^{-1}$, are $j1\,\text{k}\Omega$ and $-j2\,\text{k}\Omega$ respectively. The frequency-domain circuit is shown below:



The equivalent impedance offered to the source is:

$$\begin{aligned}\mathbf{Z}_{eq} &= 1.5 + \frac{(j1)(1-j2)}{j1+1-j2} = 1.5 + \frac{2+j}{1-j} \\ &= 1.5 + \frac{2+j}{1-j}\frac{1+j}{1+j} = 1.5 + \frac{1+j3}{2} \\ &= 2 + j1.5 = 2.5\angle 36.9° \quad \text{k}\Omega\end{aligned}$$

The phasor current is thus:

$$\mathbf{I} = \frac{\mathbf{V}_s}{\mathbf{Z}_{eq}} = \frac{40\angle -90°}{2.5\angle 36.9°} = 16\angle -126.9° \quad \text{mA}$$

Upon transforming the current to the time-domain, the desired response is obtained:

$$i(t) = 16\cos(3000t - 126.9°) \quad \text{mA}$$

12.7 Admittance

The reciprocal of impedance can offer some convenience in the sinusoidal steady-state analysis of circuits. We define *admittance* as the ratio of phasor current to phasor voltage:

Admittance defined

$$\mathbf{Y} = \frac{\mathbf{I}}{\mathbf{V}} \tag{11.91}$$

and thus:

Admittance is the reciprocal of impedance

$$\mathbf{Y} = \frac{1}{\mathbf{Z}} \tag{11.92}$$

The real part of the admittance is the *conductance* G, and the imaginary part of the admittance is the *susceptance* B. Thus:

Admittance is composed of a conductance (real part) and a susceptance (imaginary part)

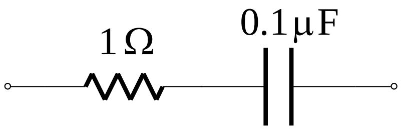
$$\mathbf{Y} = G + jB = \frac{1}{\mathbf{Z}} = \frac{1}{R + jX} \tag{11.93}$$

This formula should be scrutinized carefully. It does *not* mean that $G = 1/R$ (unless $\mathbf{Z} = R$, a pure resistance), nor does it mean $B = 1/X$.

Admittance, conductance and susceptance are all measured in siemens (S).

EXAMPLE 12.9 Admittance of a Resistor and Capacitor in Series

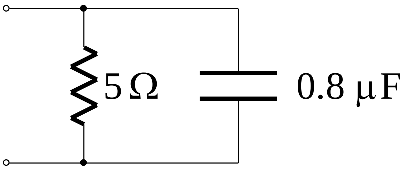
Consider a resistor and a capacitor in series:



At $\omega = 5\ \text{Mrads}^{-1}$, the equivalent impedance is $\mathbf{Z}_{eq} = 1 - j2\,\Omega$. Its admittance is:

$$\begin{aligned}\mathbf{Y} &= \frac{1}{\mathbf{Z}} = \frac{1}{1-j2} = \frac{1}{1-j2}\frac{1+j2}{1+j2} \\ &= 0.2 + j0.4\ \text{S}\end{aligned}$$

It should be apparent that the equivalent admittance of a circuit consisting of a number of parallel branches is the sum of the admittances of the individual branches. Thus, the admittance obtained above is equivalent to:



only at $\omega = 5\ \text{Mrads}^{-1}$. As a check , the equivalent impedance of the parallel network at $\omega = 5\ \text{Mrads}^{-1}$ is:

$$\mathbf{Z}_{eq} = \frac{5(-j2.5)}{5-j2.5} = 1 - j2\ \ \Omega$$

as before.

12.8 Summary

- Sinusoids are important theoretically and practically. A sinusoidal source yields a sinusoidal response.
- The sinusoidal forced response is also known as the sinusoidal steady-state – the condition which is reached after the natural response has died out.
- A complex forcing function produces a complex response – the real part of the forcing function creates the real part of the response.
- The application of a complex forcing function to a linear circuit turns the describing differential equation into a complex algebraic equation.
- The phasor representation of a sinusoid captures the amplitude and phase information in a complex number – amplitude corresponds to magnitude, and phase corresponds to angle.
- The phasor corresponding to $x(t) = A\cos(\omega t + \phi)$ is $\mathbf{X} = Ae^{j\phi} = A\angle\phi$.
- Phasor $\mathbf{V} - \mathbf{I}$ relationships for the three passive elements lead to the concept of frequency-domain impedance. The impedances of the three passive elements are: $\mathbf{Z}_R = R$, $\mathbf{Z}_L = j\omega L$, $\mathbf{Z}_C = 1/j\omega C$. Impedances can be combined and manipulated like resistors except we use complex algebra.
- Impedance consists of a real resistive component and an imaginary reactive component: $\mathbf{Z} = R + jX$.
- Admittance is defined as the inverse of impedance: $\mathbf{Y} = 1/\mathbf{Z} = G + jB$.

12.9 References

Bedient, P. & Rainville, E.: *Elementary Differential Equations*, 6th *Ed*. Macmillan Publishing Co., 1981.

Hayt, W. & Kemmerly, J.: *Engineering Circuit Analysis*, 3rd *Ed*., McGraw-Hill, 1984.

Exercises

1.

A sinusoidal voltage is zero and increasing at $t = -1.6\,\text{ms}$. The next zero crossing occurs at $t = 4.65\,\text{ms}$.

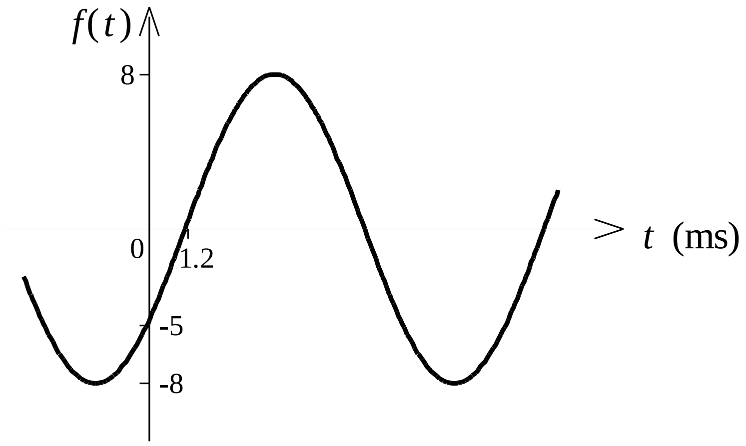
(a) Calculate T, f and ω.

(b) If $v(0) = 20\,\text{V}$, find $v(t)$.

(c) By what angle does $v(t)$ lead the current $i = 5\cos(\omega t - 110°)\,\text{A}$?

2.

For the sinusoidal waveform shown below:



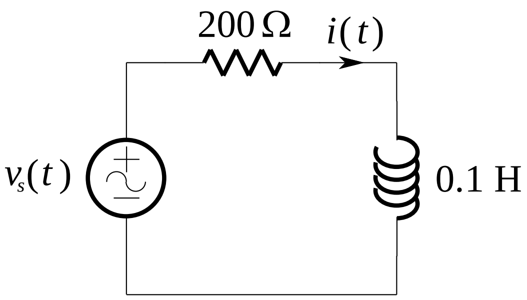
Find:

(a) A and ϕ if $f(t) = A\sin(\omega t + \phi)$ (b) T (c) f (d) ω

(e) A and θ if $f(t) = A\cos(\omega t + \theta)$ (f) $f(0.0014)$

3.

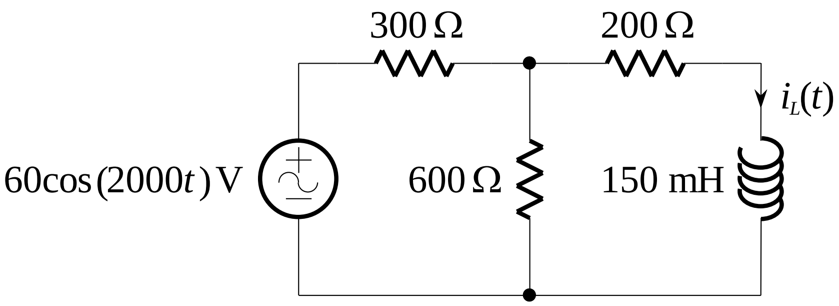
Consider the circuit shown below:



Let $i = 2\sin(500t - 40°)$ A. Determine the source voltage $v_s(t)$.

4.

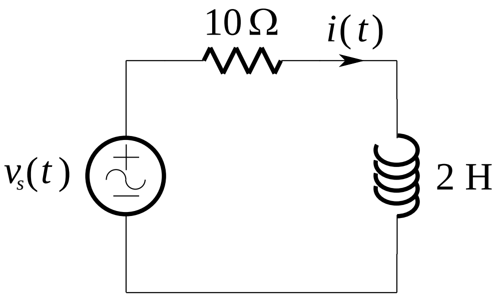
Consider the circuit shown below:



Find $i_L(t)$.

5.

Consider the circuit shown below:



Let $v_s = 20\cos 5t + 30\cos 10t$ V. Find $i(t)$.

6.

Give in polar form:

(a) $20\angle 110° - 8\angle 40°$ (b) $(6.1 - j3.82)/(1.17 + j0.541)$

Give in rectangular form:

(c) $j/(6.3 - j9.71)$ (d) $e^{j2.1\angle 52°}$

7.

A box contains a voltage source v_{s1} and a current source i_{s2}. The voltage between an available pair of terminals is labelled v_{AB}.

If $v_{s1} = 5\cos(1000t + 40°)$ V and $i_{s2} = 0.1\cos(500t - 20°)$ A, then $v_{AB} = 2\cos(1000t - 10°) + 3\cos(500t - 30°)$ V.

However, if $v_{s1} = 5\cos(500t + 40°)$ V while $i_{s2} = 0.1\cos(1000t - 20°)$ A, then $v_{AB} = 3\cos(1000t - 20°) + 2\cos(500t - 10°)$ V.

Find v_{AB} if $v_{s1} = 20\cos(1000t) + 10\cos(500t)$ V and $i_{s2} = 0.3\sin(1000t) - 0.2\sin(500t)$ A.

8.

Assume that only three currents, i_1, i_2 and i_3, enter a certain node.

(a) Find $i_1(t)$ if $\mathbf{I}_2 = 65\angle -110°$ mA and $\mathbf{I}_3 = 45\angle -50°$ mA

(b) Find $\mathbf{I}_2$ if $i_1(t) = 55\cos(400t + 40°)$ mA and $i_3(t) = 35\sin(400t - 70°)$ mA.

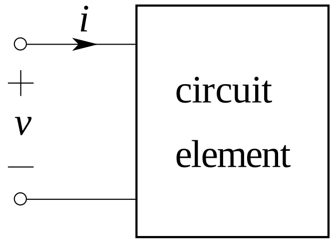
9.

At the input of an *RLC* circuit it is found that $\mathbf{V} = 41 - j18 \text{ V}$ and $\mathbf{I} = 3.5 + j7.5 \text{ A}$. Assuming $\mathbf{V}$ and $\mathbf{I}$ satisfy the passive sign convention, determine the power entering the network at:

(a) $t = 0$ (b) $\omega t = \pi/2$

10.

In the figure below, the voltage v is given as the phasor, $\mathbf{V} = 60 - j25 \text{ V}$.

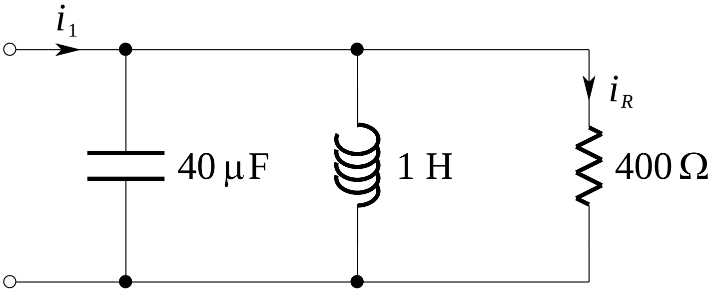


If $\omega = 2000 \text{ rads}^{-1}$, find the power being delivered to the element at $t = 0.1\pi \text{ ms}$ if the element is:

(a) a $100 \ \Omega$ resistor (b) a 50 mH inductor (c) a $5 \ \mu\text{F}$ capacitor

11.

The circuit below is operated at $\omega = 100 \text{ rads}^{-1}$.



(a) If $\mathbf{I}_R = 0.01\angle 20° \text{ A}$, find $\mathbf{I}_1$.

(b) If $\mathbf{I}_1 = 2\angle -30° \text{ A}$, find $\mathbf{I}_R$.

12.

Using a 1 H inductor and a 1 μF capacitor, at what frequency (in hertz) may an impedance be obtained having a magnitude of $2000\ \Omega$ if the two elements are combined in:

(a) series (b) parallel

13.

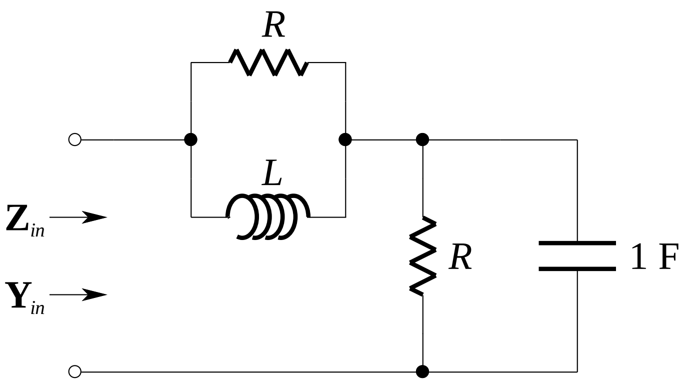
A 10 μF capacitor and a $25\ \Omega$ resistor are in parallel. What size inductor should be placed in series with this parallel combination so that the impedance of the final series network has zero reactance at $8\ \text{krads}^{-1}$?

14.

What size capacitor should be placed in series with the series combination of $800\ \Omega$ and 20 mH to give an admittance whose magnitude is 1 mS at $\omega = 10\ \text{krads}^{-1}$?

15.

If the input admittance and impedance of the network shown below are equal at every frequency, find *R* and *L*.



Joseph Fourier (1768-1830) *(Jo´ sef Foor´ yay)*

Fourier is famous for his study of the flow of heat in metallic plates and rods. The theory that he developed now has applications in industry and in the study of the temperature of the Earth's interior. He is also famous for the discovery that many functions could be expressed as infinite sums of sine and cosine terms, now called a trigonometric series, or Fourier series.

Fourier first showed talent in literature, but by the age of thirteen, mathematics became his real interest. By fourteen, he had completed a study of six volumes of a course on mathematics. Fourier studied for the priesthood but did not end up taking his vows. Instead he became a teacher of mathematics. In 1793 he became involved in politics and joined the local Revolutionary Committee. As he wrote:-

> *As the natural ideas of equality developed it was possible to conceive the sublime hope of establishing among us a free government exempt from kings and priests, and to free from this double yoke the long-usurped soil of Europe. I readily became enamoured of this cause, in my opinion the greatest and most beautiful which any nation has ever undertaken.*

Fourier became entangled in the French Revolution, and in 1794 he was arrested and imprisoned. He feared he would go to the guillotine but political changes allowed him to be freed. In 1795, he attended the Ecole Normal and was taught by, among others, Lagrange and Laplace. He started teaching again, and began further mathematical research. In 1797, after another brief period in prison, he succeeded Lagrange in being appointed to the chair of analysis and mechanics. He was renowned as an outstanding lecturer but did not undertake original research at this time.

In 1798 Fourier joined Napoleon on his invasion of Egypt as scientific adviser. The expedition was a great success (from the French point of view) until August 1798 when Nelson's fleet completely destroyed the French fleet in the Battle of the Nile, so that Napoleon found himself confined to the land he was occupying. Fourier acted as an administrator as French type political institutions and administrations were set up. In particular he helped establish educational facilities in Egypt and carried out archaeological explorations.

While in Cairo, Fourier helped found the Institute d'Égypte and was put in charge of collating the scientific and literary discoveries made during the time in Egypt. Napoleon abandoned his army and returned to Paris in 1799 and soon held absolute power in France. Fourier returned to France in 1801 with the remains of the expeditionary force and resumed his post as Professor of Analysis at the Ecole Polytechnique.

The Institute d'Égypte was responsible for the completely serendipitous discovery of the Rosetta Stone in 1799. The three inscriptions on this stone in two languages and three scripts (hieroglyphic, demotic and Greek) enabled Thomas Young and Jean-François Champollion, a protégé of Fourier, to invent a method of translating hieroglyphic writings of ancient Egypt in 1822.

Napoleon appointed Fourier to be Prefect at Grenoble where his duties were many and varied – they included draining swamps and building highways. It was during his time in Grenoble that Fourier did his important mathematical work on the theory of heat. His work on the topic began around 1804 and by 1807 he had completed his important memoir *On the Propagation of Heat in Solid Bodies*. It caused controversy – both Lagrange and Laplace objected to Fourier's expansion of functions as trigonometric series.

> *…it was in attempting to verify a third theorem that I employed the procedure which consists of multiplying by* $\cos x \, dx$ *the two sides of the equation*
>
> $$\phi(x) = a_0 + a_1 \cos x + a_2 \cos 2x + \ldots$$
>
> *and integrating between* $x = 0$ *and* $x = \pi$. *I am sorry not to have known the name of the mathematician who first made use of this method because I would have cited him. Regarding the researches of d'Alembert and Euler could one not add that if they knew this expansion they made but a very imperfect use of it. They were both persuaded that an arbitrary…function could never be resolved in a series of this kind, and it does not seem that any one had developed a constant in cosines of multiple arcs* [i.e. found a_1, $a_2, \ldots$, with $1 = a_1 \cos x + a_2 \cos 2x + \ldots$ for $-\pi/2 < x < \pi/2$] *the first problem which I had to solve in the theory of heat.*

This extract is from a letter found among Fourier's papers, and unfortunately lacks the name of the addressee, but was probably intended for Lagrange.

Other people before Fourier had used expansions of the form $f(x) \sim \sum_{r=-\infty}^{\infty} a_r \exp(irt)$ but Fourier's work extended this idea in two totally new ways. One was the "Fourier integral" (the formula for the Fourier series coefficients) and the other marked the birth of Sturm-Liouville theory (Sturm and Liouville were nineteenth century mathematicians who found solutions to many classes of partial differential equations arising in physics that were analogous to Fourier series).

Napoleon was defeated in 1815 and Fourier returned to Paris. Fourier was elected to the Académie des Sciences in 1817 and became Secretary in 1822. Shortly after, the Academy published his prize winning essay *Théorie analytique de la chaleur* (*Analytical Theory of Heat*). In this he obtains for the first time the equation of heat conduction, which is a partial differential equation in three dimensions. As an application he considered the temperature of the ground at a certain depth due to the sun's heating. The solution consists of a yearly component and a daily component. Both effects die off exponentially with depth but the high frequency daily effect dies off much more rapidly than the low frequency yearly effect. There is also a phase lag for the daily and yearly effects so that at certain depths the temperature will be completely out of step with the surface temperature.

All these predictions are confirmed by measurements which show that annual variations in temperature are imperceptible at quite small depths (this accounts for the permafrost, i.e. permanently frozen subsoil, at high latitudes) and that daily variations are imperceptible at depths measured in tenths of metres. A reasonable value of soil thermal conductivity leads to a prediction that annual temperature changes will lag by six months at about 2–3 metres depth. Again this is confirmed by observation and, as Fourier remarked, gives a good depth for the construction of cellars.

As Fourier grew older, he developed at least one peculiar notion. Whether influenced by his stay in the heat of Egypt or by his own studies of the flow of heat in metals, he became obsessed with the idea that extreme heat was the natural condition for the human body. He was always bundled in woollen clothing, and kept his rooms at high temperatures. He died in his sixty-third year, "thoroughly cooked".

References

Körner, T.W.: *Fourier Analysis*, Cambridge University Press, 1988.

13 Circuit Simulation

Contents

13.2

Introduction

The most popular circuit simulation "package" is known as PSpice®. SPICE is an acronym for Simulation Program with Integrated Circuit Emphasis – a program developed at the University of California, Berkeley in the 1970's that was used to simulate analog electronic circuits. It quickly became an industry standard – versions became available in the public domain and large companies developed their own versions – and continue to do so. The "P" prefix came about in 1984 when the first version for a PC became available. Most professional computer-aided electronic design tools now incorporate PSpice® compatible simulators.

There are many software packages which use PSpice® (or PSpice® compatible) simulators, and they are updated frequently. It is not the purpose of this document to outline how each version of the various software packages are used. There are many other resources that are available that do this – online tutorials, textbook appendices, student and demo versions of software, etc.

This document will give an overview of the procedure of circuit simulation, and then highlight particular topics that are important for us in generating time-domain and frequency-domain analyses of simple analog circuits.

This document will be based on OrCAD® – an integrated software package used for electronic design automation (EDA). It "captures" schematic designs, simulates them, and allows a printed circuit board to be designed for manufacture. Demo versions are freely available.

Australians tend to use Altium Designer – EDA software originally developed in Hobart at the University of Tasmania. Altium is now a large international company listed on the ASX, and has a large share of the EDA market.

PSpice and OrCAD are registered trademarks of Cadence Design Systems, Inc.

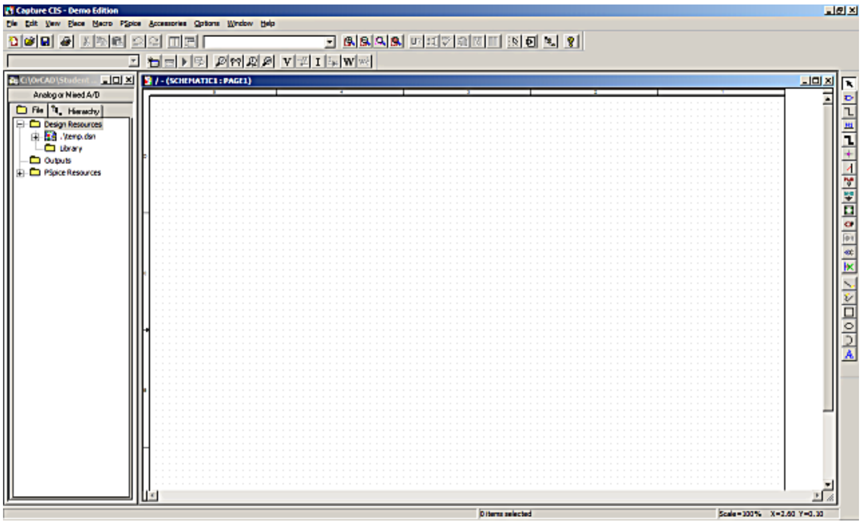
13.1 Project Flow

The basic steps to running a circuit simulation are:

- Start a new project
- Draw the schematic
- Set up and run simulation profiles

13.1.1 Starting a New Project

1. Open OrCAD Capture CIS - Demo Edition (freely available).
2. Go to **File ▶ New ▶ Project…**.
3. Enter a name for the project.
4. Choose “Analog or Mixed A/D”.
5. Set the location. You should create a new directory for your project since PSpice will generate a lot of project files in this folder.
6. Click OK.
7. Choose “Create a blank project” and click OK.
8. You should see a window where you can draw the schematic (i.e. your circuit diagram).



13.1.2 Drawing the Schematic

To add parts for your circuit (i.e. resistors, etc.)

1. Go to **Place ▶ Parts**.

2. Click on the library you want to use, or select multiple libraries by holding Ctrl or dragging the mouse. In the part window you should see at least the ANALOG, BIPOLAR, EVAL, SOURCE and SPECIAL libraries.

3. Find the part you want to add and press OK.

4. Click where you want to place the part on your schematic. Press R to rotate the part by 90 degrees.

5. Use wires to connect the part to complete your circuit.

13.1.3 Simulation

Simulations run via "simulation profiles". You need to set up a simulation profile for each type of simulation you want to run.

1. Go to **PSpice ▶ New Simulation Profile**.

2. Enter a name for the simulation profile, e.g. "Step Response" or "Frequency Response".

3. Press Create.

4. In the "Analysis Type" drop-down list, choose the simulation type, e.g. "Time Domain (Transient)" for a step response simulation, "AC Sweep / Noise" for a frequency response simulation.

5. Set up the parameters associated with the type of simulation.

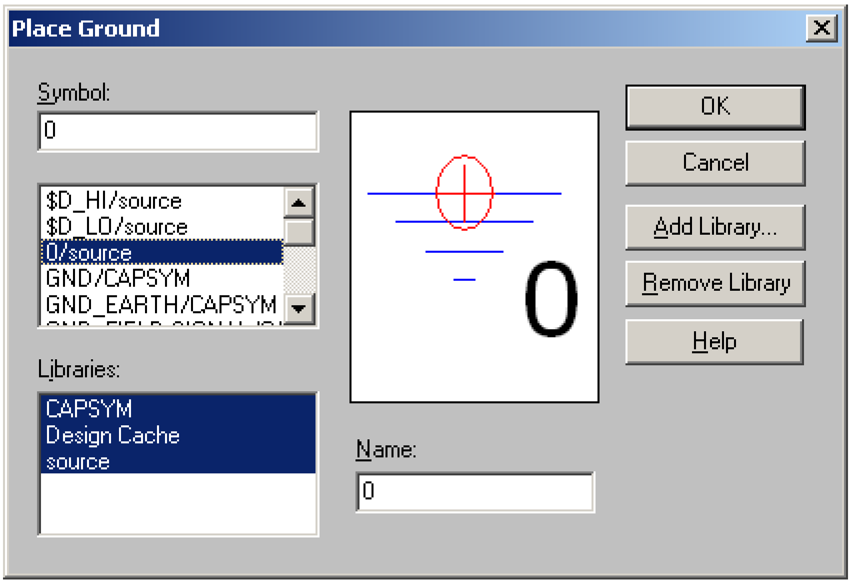
6. To run the simulation, press F11 or choose **PSpice ▶ Run**.

13.2 Schematic Capture

Here are some tips for laying out the schematic for simulation purposes.

13.2.1 Ground

There are many types of grounds (common points in the circuit) and there may be more than one common point, e.g. an “analog ground” and a “digital ground” in a mixed circuit. PSpice uses nodal analysis for circuit simulation and therefore needs a reference node with “zero voltage”. This is provided by a special ground symbol called **0/SOURCE**. You need to have it in your circuit! It looks like a ground symbol with a zero. If you don't have it, PSpice may complain of "floating nodes" even if you have a ground.



To place the ground on the circuit Go to **Place ▶ Ground** and choose **0/source** (if you don't see “source” in the Libraries section, you will need to add the source library). Alternatively, click on the ground button in the schematic toolbar and choose **0/source**. Or just push “G”.

13.2.2 SI Unit Prefixes

PSpice uses letters to represent the most common SI unit prefixes:

SI Unit Prefix	PSpice letter
giga	G
mega	Meg
kilo	k
milli	m
micro	u
nano	n
pico	p
femto	f

The two differences are “mega” and “micro”.

The “mega” prefix is written “Meg” (case does not matter). “M” is NOT “mega”, it is “milli”.

Example: For 6.5 MHz, enter "6.5 Meg", for 3 mV, enter "3 m".

The “micro” prefix is not the Greek letter mu (μ) as in the SI system of units, because it is not supported on a standard keyboard. We use the letter “u” instead.

Example: For $680\ \mu\text{H}$, enter "680u".

13.2.3 All Parts Must Have Unique Names

You can't have two parts named "R1" in your circuit. If you are copying and pasting parts or circuits, OrCAD usually increments the part number automatically. However, you may accidentally name two parts the same, which will cause a “netlist” error.

13.2.4 Labeling Nodes

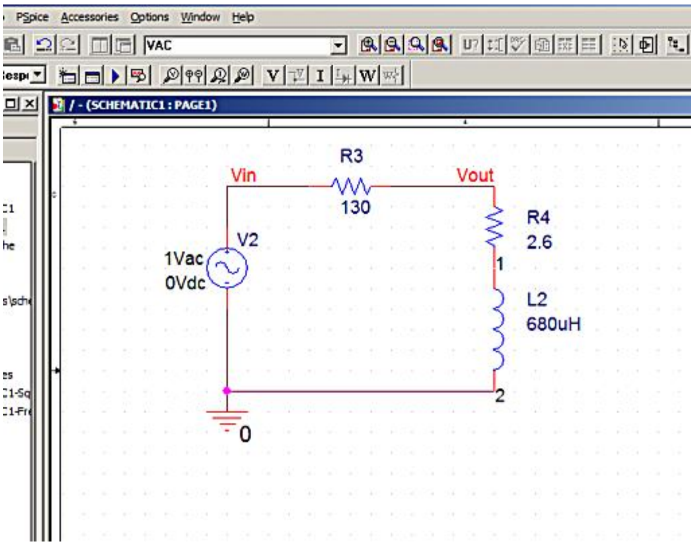
You should use aliases to label your input and output nodes. This makes your nodes easier to find when dealing with simulation output. `V(Vout)` is simpler than finding `V(R1:1)`.

1. Go to **Place ▶ Net Alias**.

2. Enter a name, e.g., `Vout` or `Vin`.

3. Place the label close to a node.

The example below shows a simple circuit with aliases:

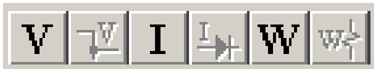


13.3 Simulation

The following sections detail the procedures used in setting up and running simulations.

13.3.1 DC Bias

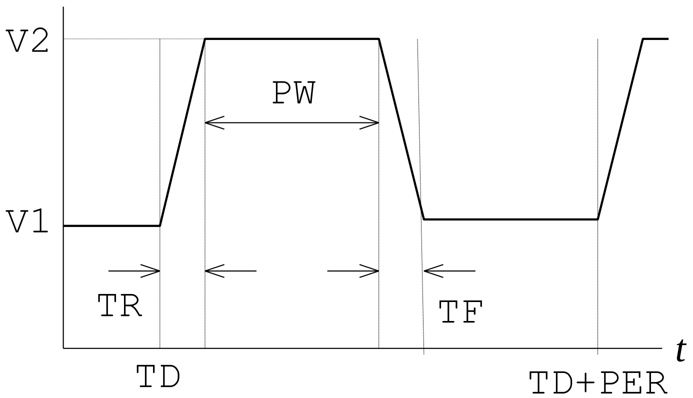
The response of the circuit to DC sources is always calculated. To display DC bias voltages, currents and power on your circuit after you run the simulation, click on the Enable Bias Display buttons:



13.3.2 Time-Domain (Transient) Simulations

For sinusoids, use `VSIN` for your voltage source instead of `VAC` (`VOFF` is the DC offset, `VAMPL` is the amplitude, and `FREQ` is the frequency).

For square and triangular waves, use `VPULSE`:



Parameter	Meaning
`V1`	Initial value (V)
`V2`	Pulsed value (V)
`TD`	Delay time (s)
`TR`	Rise time (s)
`TF`	Fall time (s)
`PW`	Pulse width (s)
`PER`	Period (s)

We normally set the delay time, `TD`, to zero.

Square Wave

A square wave is the `VPULSE` function in the limit of `TR=TF=0` and `PW=0.5*PER` (`PER` is the period of the wave). This limiting case, however, causes numerical difficulties in calculations. In any case, we can never make such a square function in practice. In reality, square waves have very small `TR` and `TF`. Typically, we use a symmetric function, i.e., we set `TR=TF` and `PW=0.5*PER-TR`. Thus, for a given frequency we can set up the square function if we choose `TR`. If we choose `TR` too large, the function does not look like a square wave. If we choose `TR` too small, the program will take a long time to simulate the circuit and for `TR` smaller than a certain value, the simulation will not converge numerically. A good choice for `TR` is to set it to be 1% of the `PER` (a period): `TR=TF=0.01*PER`, `PW=0.49*PER`. This usually results in a nice signal without a huge amount of computational need. Note that `TR` does not have to be exactly 1% of `PER`. You can choose nice round numbers for `TR`, `TF`, and `PW`.

Triangular Wave

A triangular wave is the `VPULSE` function in the limit of `TR=TF=0.5*PER` and `PW=0` (convince yourself that this is the case). As before, the limiting case of `PW=0` causes numerical difficulties in calculations. So we have to choose `PW` to be a reasonably small value. A good choice for `PW` is to set it at 1% of the `PER` (period): `PW=0.01*PER`, `TR=TF=0.495*PER`. This usually results in a nice signal without a huge amount of computational need. Again, note that `PW` does not have to be exactly 1% of `PER`. You can choose nice round numbers for `TR`, `TF`, and `PW`.

Step

A step can be created by setting up a square wave with a period that is much longer than any time constants in the circuit (this is what we do in the lab).

13.10

Simulation Settings

1. Go to **PSpice ▶ Edit Simulation Profile**.

2. Select the "Time Domain (Transient)" Analysis type.

3. Enter a “Run to time”: so that a few periods will be displayed. Remember that the period (seconds) = 1/frequency (Hz).

   Example: If you are using a 1 kHz sine wave, it has a 1/1 kHz = 1 ms period, so use a “Run to time” of 5 ms for 5 periods.

4. Set the “Maximum step size” to be much smaller than the period.

   Example: For a 1 kHz sine wave, it has a 1 ms period, so set a “Maximum step size” of approximately 0.01ms. This works out to 100 data points per period.

5. If you don't set the “Maximum step size”, PSpice may choose one which is too big, making your waveforms look angular and ugly (because it plots straight lines between data points).

13.3.3 AC Sweep / Noise Simulations

Use `VAC` for your voltage source.

Simulation Settings

1. Go to **PSpice ▶ Edit Simulation Profile**.
2. Select the "AC Sweep / Noise" Analysis type.
3. Select the frequency range of interest. Don't start frequency sweeps at 0!
4. Set the "Points/Decade" to be at least 20.

Bode Plots

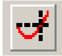
1. Use a logarithmic *x*-axis for the frequency.
2. The magnitude should be measured in decibels. Use the PSpice `DB()` function to convert to decibels.

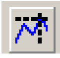
   Example: Use `DB(V(Vout)/V(Vin))`, assuming you have labelled your output and input nodes with `Vout` and `Vin` aliases.

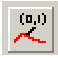
3. Remember you also need a phase response (unless instructed otherwise). Use the PSpice `P()` function to get the phase angle.

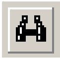
   Example: Use `P(V(Vout)/V(Vin))`, assuming you have labelled your output and input nodes with `Vout` and `Vin` aliases.

4. Be sure to mark the cutoff frequency or other points of interest on your Bode plots (on both magnitude AND phase graphs). Cutoff frequency is defined as 3dB below the peak of the magnitude response (NOT always at -3dB).

   a. Click the "Toggle Cursor" button  (or go through the menu, **Trace ▶ Cursor ▶ Display**). You will now be able to move the cursor along your plot.

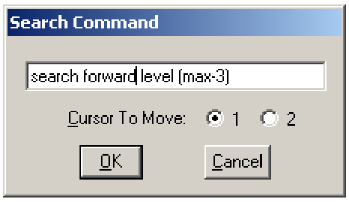
   b. Click the "Cursor Max" button  to find the highest point (or go through the menu, **Trace ▶ Cursor ▶ Max**).

   c. Click the "Mark Label" button  to label that point (or go through the menu, **Plot ▶ Label ▶ Mark**).

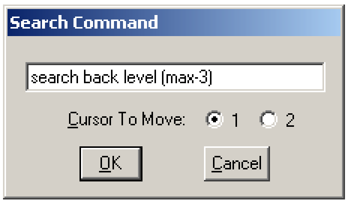
   d. Click the "Cursor Search" button  (or go through the menu, **Trace ▶ Cursor ▶ Search Commands…**).

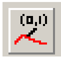
   e. Select 1 for "Cursor To Move" to search along the *y*-axis of trace 1.

   f. Enter "search forward level (max-3)" (don't enter the quotation marks) to move the cursor to the right, to the point which is 3 below the max.



g. Or enter "search back level (max-3)" (don't enter the quotation marks) to move the cursor to the left.



h. Click the "Mark Label" button  to label that cutoff point.

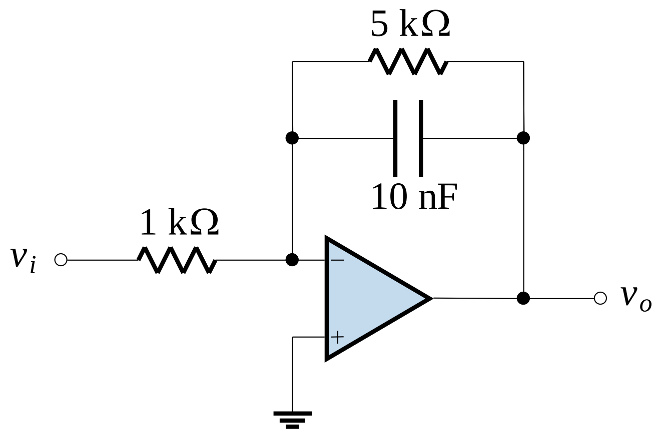
- Unclick the “Toggle Cursor” button to disable the cursor so you can move the label.
- Double-click on the label to edit the text (to add units, or to name the point).

i. It may help to increase the width of the traces in the plot.

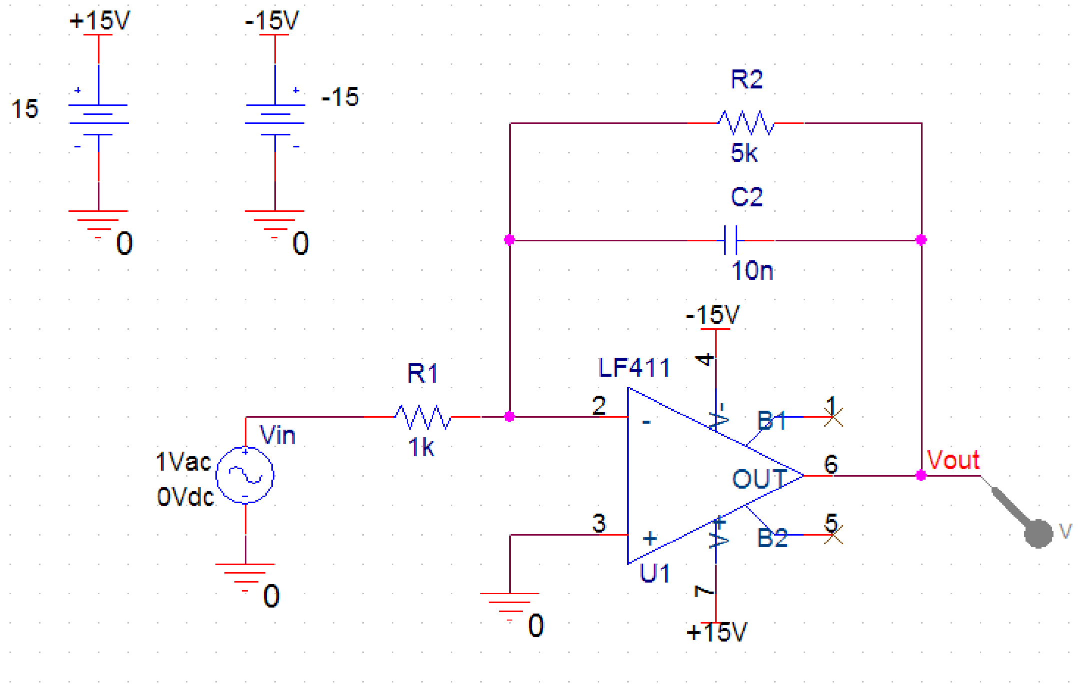
- Right click on a trace. Make sure the selection list has Information, Properties, Cursor 1, and Cursor 2 (if it lists Settings and Properties, you clicked on the background, not on the trace).
- Select Properties.
- You can change the width and other properties of that trace.

EXAMPLE 13.1 Simulation of a Circuit's Magnitude Response

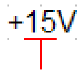
We wish to simulate the following circuit to determine its magnitude response:



We enter the following schematic into OrCAD.

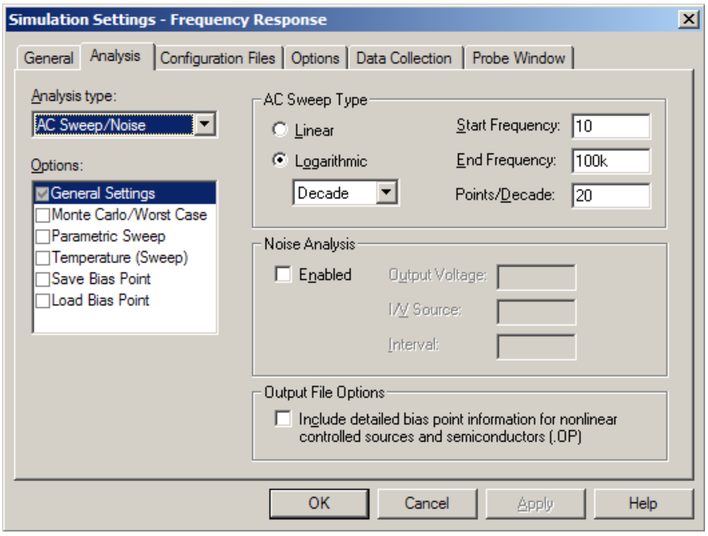


Note the use of a real op-amp – the LF411. You can search for it in the parts library by **Place ▶ Part…** then **Part Search…**. The op-amp requires DC voltage sources on its power pins.

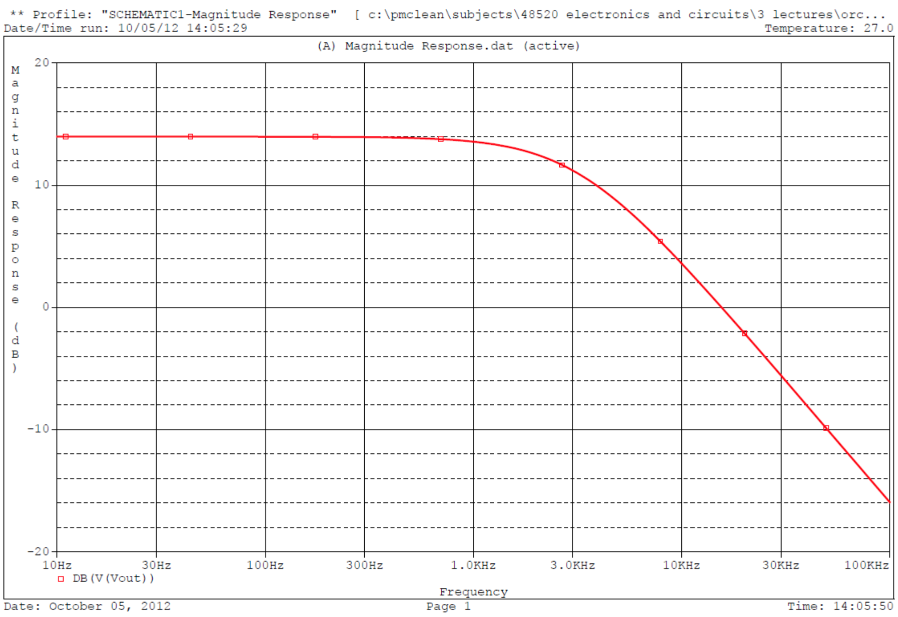
In this schematic, we have used the power symbol 

which simply “ties” nets together with a common name – in this case the name is “+15V”. One “tie” is placed on the DC voltage source, the other on the op-amp power pin. This avoids cluttering wires on the schematic.

We set up a frequency response simulation using AC Sweep as shown below:



for which the output is:

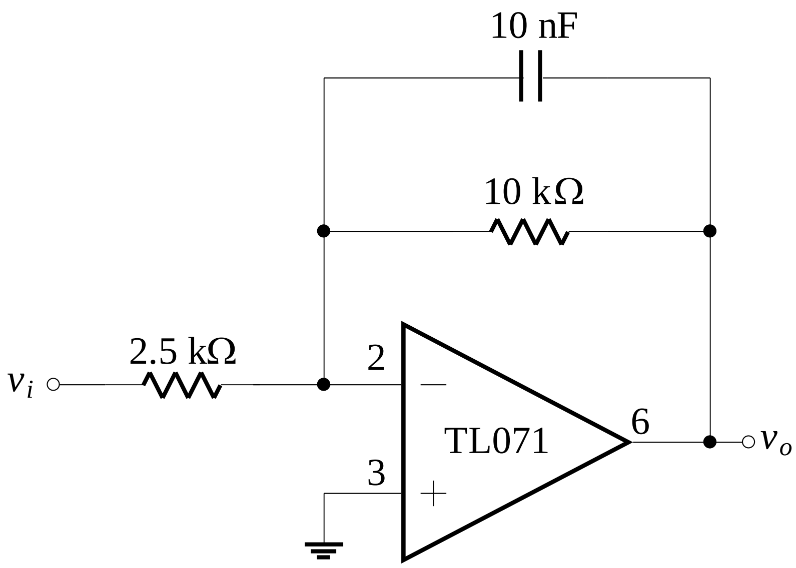


13.16

Exercises

1.

Simulate the step response and frequency response of the following circuit:



Use the following axes for graphing:

Response	**X**	**Y**
Step	0 s to 1 ms	-5 V to 0 V
Magnitude	10 Hz to 100 kHz	-30 dB to 20 dB
Phase	10 Hz to 100 kHz	90° to 180°

14 The Sinusoidal Steady-State Response

Contents

14.2

Introduction

In the analysis of resistive circuits of arbitrary complexity, we are able to employ many different circuit analysis techniques to determine the response – nodal analysis, mesh analysis, superposition, source transformations, Thévenin’s and Norton’s theorems.

Sometimes one method is sufficient, but more often we find it convenient to combine several methods to obtain the response in the most direct manner.

We now want to extend these techniques to the analysis of circuits in the sinusoidal steady-state. We have already seen that impedances combine in the same manner as do resistances. We have seen that KVL and KCL are obeyed by phasors, and we also have an Ohm-like law for the passive elements, $\mathbf{V} = \mathbf{ZI}$. We can therefore extend our resistive circuit analysis techniques to the frequency-domain to determine the phasor response, and therefore the sinusoidal steady-state response.

14.1 Analysis using Phasors

Phasors can only be used for sinusoidal steady-state analysis. They **cannot** be used to determine transients in a circuit.

Phasor analysis is a *transform method* of analysis. In phasor analysis we transform a problem from the time-domain to the frequency-domain. To find a response in the frequency-domain, we solve equations using complex algebra. Once the response is found, we transform the solutions back to the time-domain. This is illustrated conceptually below:

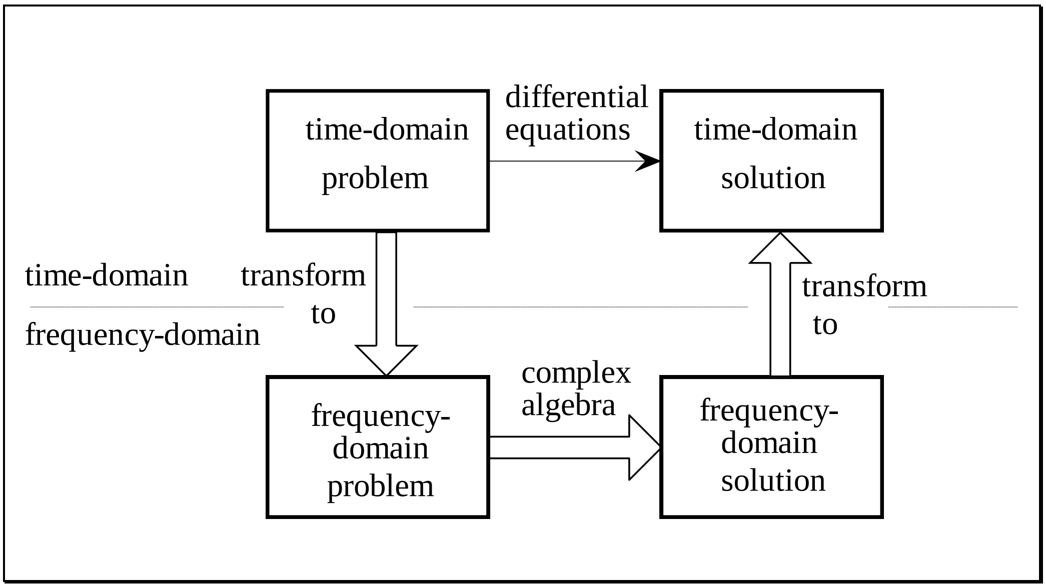


Figure 14.1

Transform methods are a common method of analysis in branches of engineering, and you will be introduced to more powerful transform methods in more advanced subjects.

14.2 Nodal Analysis

As an example of nodal analysis, consider the circuit shown below:

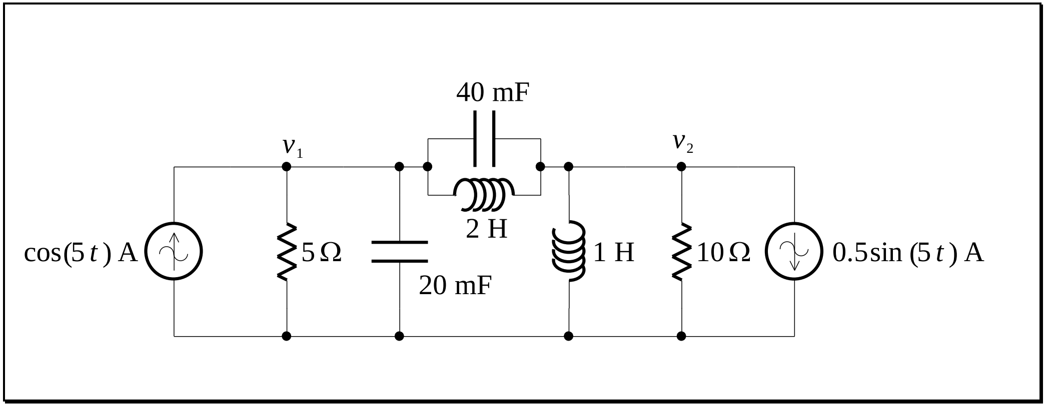


Figure 14.2

Noting from the sources that $\omega = 5\,\text{rads}^{-1}$, we draw the frequency-domain circuit and assign nodal voltages $\mathbf{V}_1$ and $\mathbf{V}_2$:

Nodal analysis in the frequency-domain

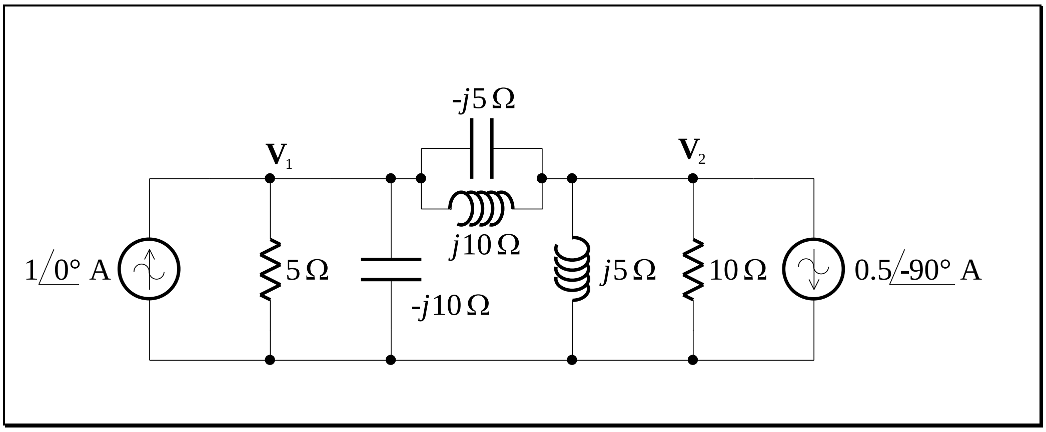


Figure 14.3

Each passive element is specified by its impedance, which has been determined by knowing the frequency of the sources (which are the same) and the element values. Two current sources are given as phasors, and phasor node voltages $\mathbf{V}_1$ and $\mathbf{V}_2$ are indicated.

At the left node, we apply KCL and $\mathbf{I} = \mathbf{V}/\mathbf{Z}$:

$$\frac{\mathbf{V}_1}{5} + \frac{\mathbf{V}_1}{-j10} + \frac{\mathbf{V}_1 - \mathbf{V}_2}{-j5} + \frac{\mathbf{V}_1 - \mathbf{V}_2}{j10} = 1 + j0 \tag{14.1}$$

At the right node:

$$\frac{\mathbf{V}_2-\mathbf{V}_1}{-j5}+\frac{\mathbf{V}_2-\mathbf{V}_1}{j10}+\frac{\mathbf{V}_2}{j5}+\frac{\mathbf{V}_2}{10}=-(-j0.5) \tag{14.2}$$

Combining terms we have the two equations:

$$\begin{aligned}(0.2+j0.2)\mathbf{V}_1-j0.1\mathbf{V}_2&=1\\ -j0.1\mathbf{V}_1+(0.1-j0.1)\mathbf{V}_2&=j0.5\end{aligned} \tag{14.3}$$

Using Cramer's Rule to solve, we obtain:

$$\begin{aligned}\mathbf{V}_1&=\frac{\begin{vmatrix}1 & -j0.1\\ j0.5 & (0.1-j0.1)\end{vmatrix}}{\begin{vmatrix}(0.2+j0.2) & -j0.1\\ -j0.1 & (0.1-j0.1)\end{vmatrix}}=\frac{0.1-j0.1-0.05}{0.02-j0.02+j0.02+0.02+0.01}\\ &=\frac{0.05-j0.1}{0.05}=1-j2\ \text{V}\\ \mathbf{V}_2&=\frac{\begin{vmatrix}(0.2+j0.2) & 1\\ -j0.1 & j0.5\end{vmatrix}}{0.05}=\frac{-0.1+j0.1+j0.1}{0.05}=-2+j4\ \text{V}\end{aligned} \tag{14.4}$$

The time-domain solutions are best obtained by representing the phasors in polar form:

$$\begin{aligned}\mathbf{V}_1&=\sqrt{5}\angle-63.4^\circ\ \text{V}\\ \mathbf{V}_2&=2\sqrt{5}\angle 116.6^\circ\ \text{V}\end{aligned} \tag{14.5}$$

and passing to the time-domain:

$$\begin{aligned}v_1(t)&=\sqrt{5}\cos(5t-63.4^\circ)\ \text{V}\\ v_2(t)&=2\sqrt{5}\cos(5t+116.6^\circ)\ \text{V}\end{aligned} \tag{14.6}$$

Note how simple phasor analysis is compared to the work involved if we stayed in the time-domain solving differential equations!

14.3 Mesh Analysis

As an example of mesh analysis, consider the circuit shown below:

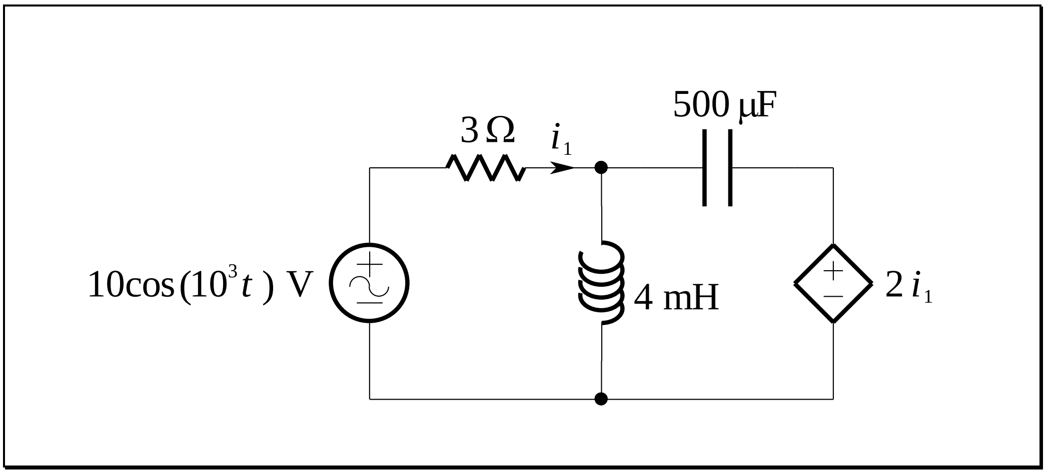


Figure 14.4

Noting from the left source that $\omega = 10^3$ rads^{-1}, we draw the frequency-domain circuit and assign mesh currents $\mathbf{I}_1$ and $\mathbf{I}_2$:

Mesh analysis in the frequency-domain

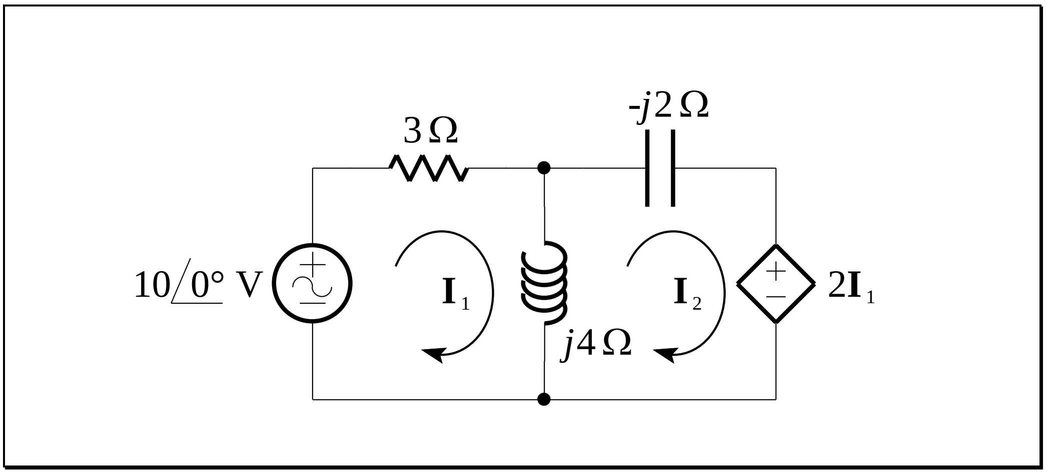


Figure 14.5

Around mesh 1:

$$3\mathbf{I}_1 + j4(\mathbf{I}_1 - \mathbf{I}_2) = 10\angle 0^\circ \qquad (14.7)$$

while mesh 2 leads to:

$$j4(\mathbf{I}_2 - \mathbf{I}_1) - j2\mathbf{I}_2 + 2\mathbf{I}_1 = 0 \qquad (14.8)$$

Combining terms we have the two equations:

$$\begin{aligned}(3+j4)\mathbf{I}_1 - j4\mathbf{I}_2 &= 10 \\ (2-j4)\mathbf{I}_1 + j2\mathbf{I}_2 &= 0\end{aligned} \tag{14.9}$$

Solving:

$$\begin{aligned}\mathbf{I}_1 &= \frac{14+j8}{13} = 1.240\angle 29.7^\circ \ \text{A} \\ \mathbf{I}_2 &= \frac{20+j30}{13} = 2.774\angle 56.3^\circ \ \text{A}\end{aligned} \tag{14.10}$$

or:

$$\begin{aligned}i_1(t) &= 1.240\cos\left(10^3 t + 29.7^\circ\right) \ \text{A} \\ i_2(t) &= 2.774\cos\left(10^3 t + 56.3^\circ\right) \ \text{A}\end{aligned} \tag{14.11}$$

The solution above could be checked by working entirely in the time-domain, but it would be quite an undertaking!

14.4 Superposition

Linear circuits are those that consist of any of the following: idealised linear passive circuit elements (R, L and C), ideal independent voltage and current sources and linearly dependent voltage and current sources. Such circuits are amenable to the superposition principle.

We can analyse linear circuits with phasors and the principle of superposition. (You may remember that linearity and superposition were invoked when we combined real and imaginary sources to obtain a complex source).

Let's look again at the circuit of Figure 14.3, redrawn below with each pair of parallel impedances replaced by a single equivalent impedance (for example, 5 and $-j10$ in parallel yield $4-j2$):

Superposition in the frequency-domain

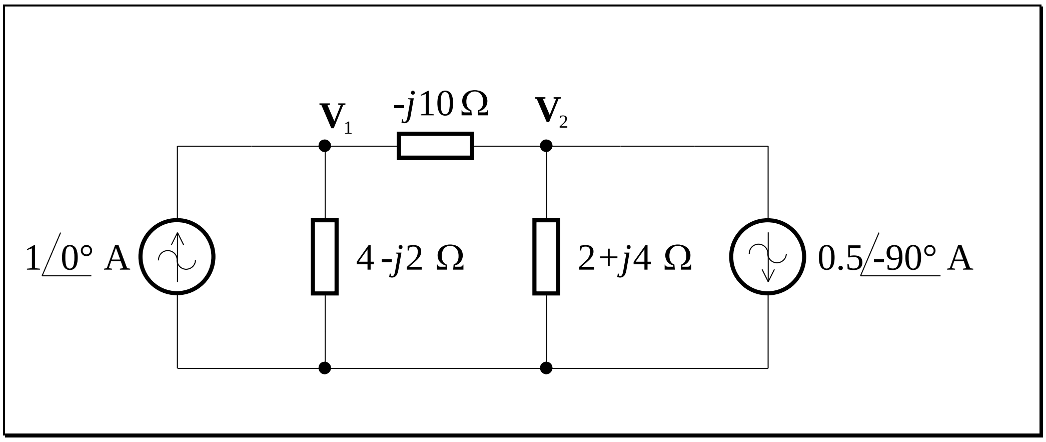


Figure 14.6

To find $\mathbf{V}_1$ we first activate only the left source and find the partial response:

$$\mathbf{V}_{1L} = 1\angle 0^\circ \frac{(4-j2)(-j10+2+j4)}{4-j2-j10+2+j4} = \frac{-4-j28}{6-j8} = 2-j2 \tag{14.12}$$

With only the right source active, current division helps us to obtain:

$$\mathbf{V}_{1R} = (-0.5\angle -90^\circ)\left(\frac{2+j4}{4-j2-j10+2+j4}\right)(4-j2) = \frac{-6+j8}{6-j8} = -1 \tag{14.13}$$

Summing, we get:

$$\mathbf{V}_1 = 2 - j2 - 1 = 1 - j2 \text{ V} \tag{14.14}$$

which agrees with our previous result.

14.5 Thévenin's Theorem

We will use the same circuit to see whether Thévenin's Theorem can help us:

Thévenin's Theorem in the frequency-domain

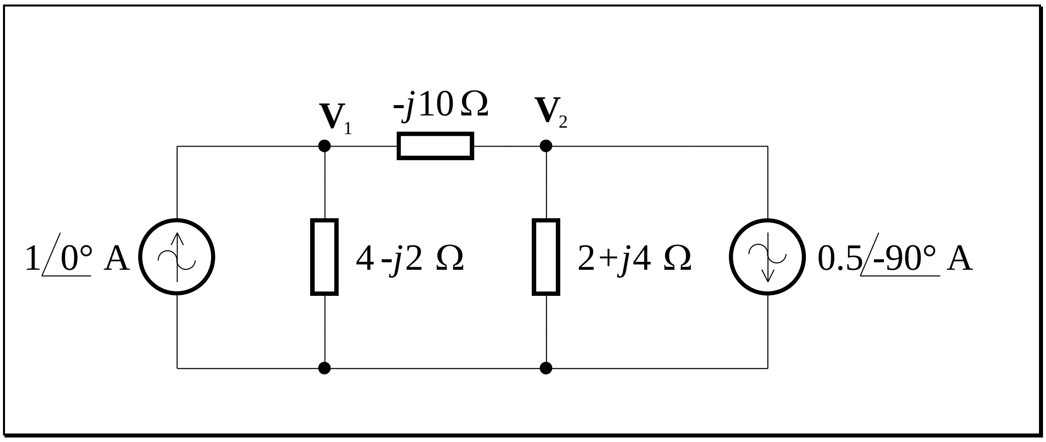


Figure 14.7

Suppose we determine the Thévenin equivalent faced by the $-j10\,\Omega$ impedance. The open circuit voltage (+ reference to the left) is:

$$\begin{aligned}\mathbf{V}_{oc} &= (1\angle 0^\circ)(4 - j2) + (0.5\angle -90^\circ)(2 + j4) \\ &= 4 - j2 + 2 - j1 = 6 - j3\end{aligned} \tag{14.15}$$

The impedance of the inactive circuit, as viewed from the load terminals, is simply the sum of the two remaining impedances (because the current sources are set to zero – open circuits). Hence:

$$\mathbf{Z}_{th} = 6 + j2 \tag{14.16}$$

Thus, when we reconnect the circuit, the current directed from node 1 toward node 2 through the $-j10\,\Omega$ load is:

$$\mathbf{I}_{12} = \frac{6 - j3}{6 + j2 - j10} = 0.6 + j0.3 \tag{14.17}$$

Subtracting this from the left source current, the downward current through the $4 - j2\,\Omega$ branch is found:

$$\mathbf{I}_1 = 1 - (0.6 + j0.3) = 0.4 - j0.3 \tag{14.18}$$

and , thus:

$$\mathbf{V}_1 = \mathbf{Z}_1\mathbf{I}_1 = (4 - j2)(0.4 - j0.3) = 1 - j2\ \text{V} \tag{14.19}$$

14.6 Norton's Theorem

Again using the same circuit, if our chief interest is in $\mathbf{V}_1$ we could use Norton's Theorem on the three right elements:

Norton's Theorem in the frequency-domain

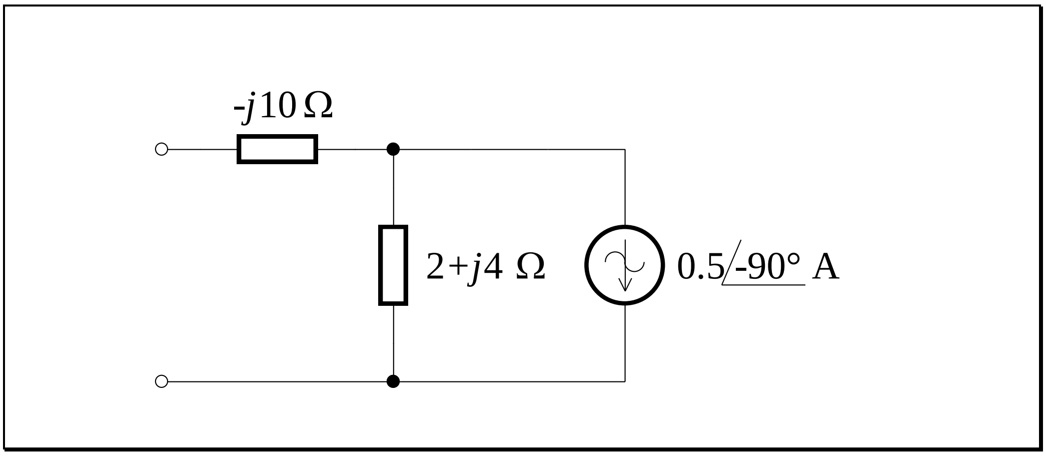


Figure 14.8

The short circuit current is obtained using current division:

$$\mathbf{I}_{sc} = \frac{2+j4}{2+j4-j10}\left(-0.5\angle-90°\right)$$
$$= \frac{-2+j}{2-j6} = -\frac{1+j}{4} \text{ A} \tag{14.20}$$

and the Norton impedance (equal to the Thévenin impedance) is simply:

$$\mathbf{Z}_{th} = 2 - j6 \tag{14.21}$$

We thus need to analyse the circuit:

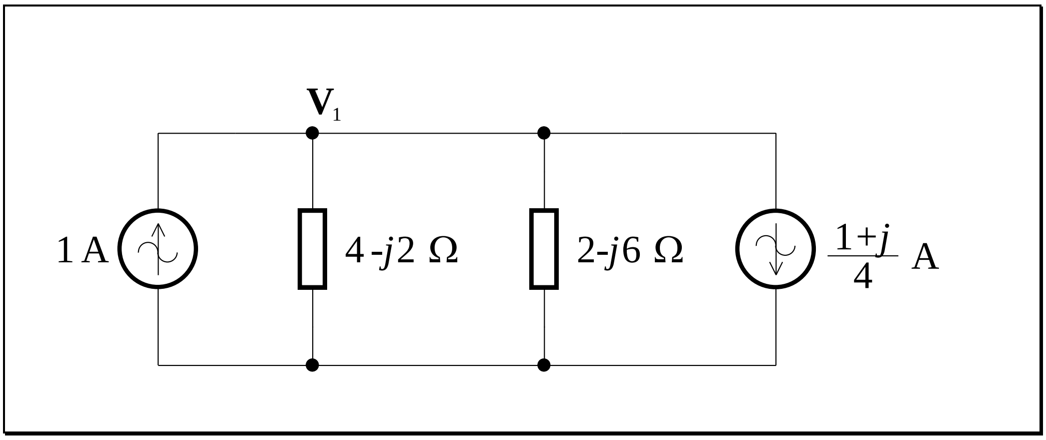


Figure 14.9

The voltage $\mathbf{V}_1$ is therefore:

$$\mathbf{V}_1 = \frac{(4-j2)(2-j6)}{(4-j2+2-j6)}(1-(0.25+j0.25))$$
$$= \frac{-4-j28}{6-j8}(0.75-j0.25) = (2-j2)(0.75-j0.25) = 1-j2 \text{ V} \tag{14.22}$$

It should now be clear that all methods available for linear circuit analysis can be applied to the frequency-domain. The slight additional complexity that is apparent now arises from the necessity of using complex numbers and not from any more involved theoretical considerations.

14.7 Phasor Diagrams

The phasor diagram is a sketch in the complex plane of the phasor voltages and currents throughout a specific circuit. It provides a graphical method for solving problems which may be used to check more exact analytical methods.

A phasor diagram is a graphical sketch of phasors in the complex plane

Since phasor voltages and currents are complex numbers, they may be identified as points in a complex plane. For example, the phasor voltage $\mathbf{V}_1 = 6 + j8 = 10\angle 53.1°$ is identified on the complex voltage plane shown below:

A simple phasor diagram

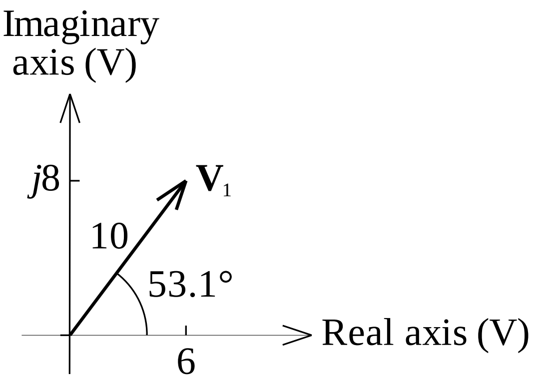


Figure 14.10

The axes are the real voltage axis and the imaginary voltage axis. The voltage $\mathbf{V}_1$ is located by an arrow drawn from the origin. Since addition and subtraction are particularly easy to perform and display on a complex plane, it is apparent that phasors may be easily added and subtracted in a phasor diagram. Multiplication and addition result in a change in magnitude and the addition and subtraction of angles.

14.14

Figure 14.11 shows the sum of $\mathbf{V}_1$ and a second phasor voltage $\mathbf{V}_2 = 3 - j4 = 5\angle -53.1°$:

Phasor diagram showing addition

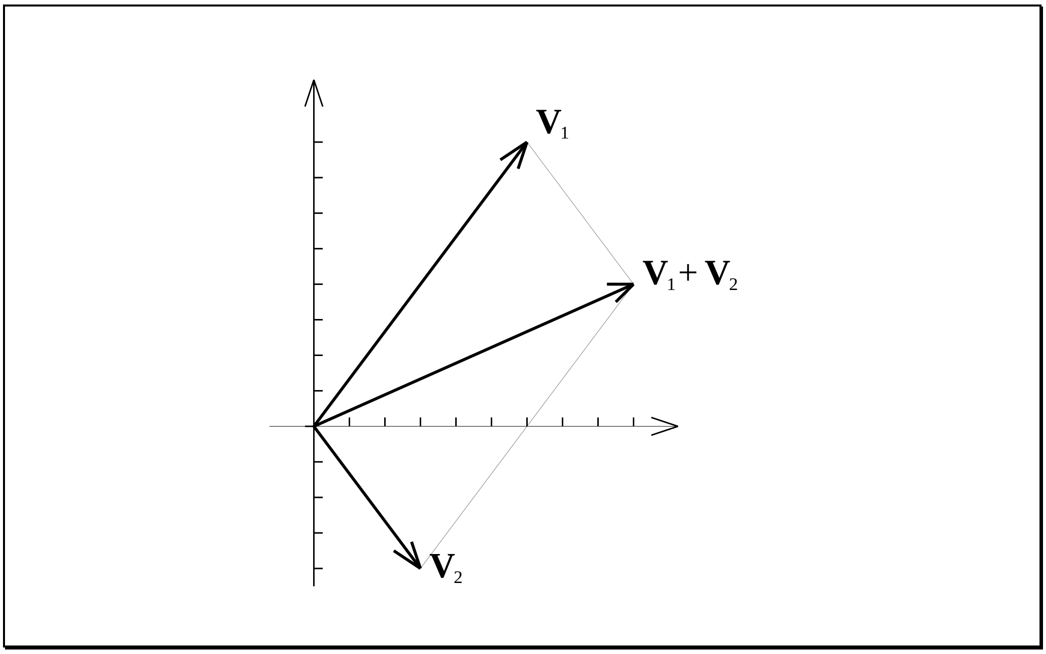


Figure 14.11

Figure 14.12 shows the current $\mathbf{I}_1$, which is the product of $\mathbf{V}_1$ and the admittance $\mathbf{Y} = 1 + j1$:

Phasor diagram showing multiplication

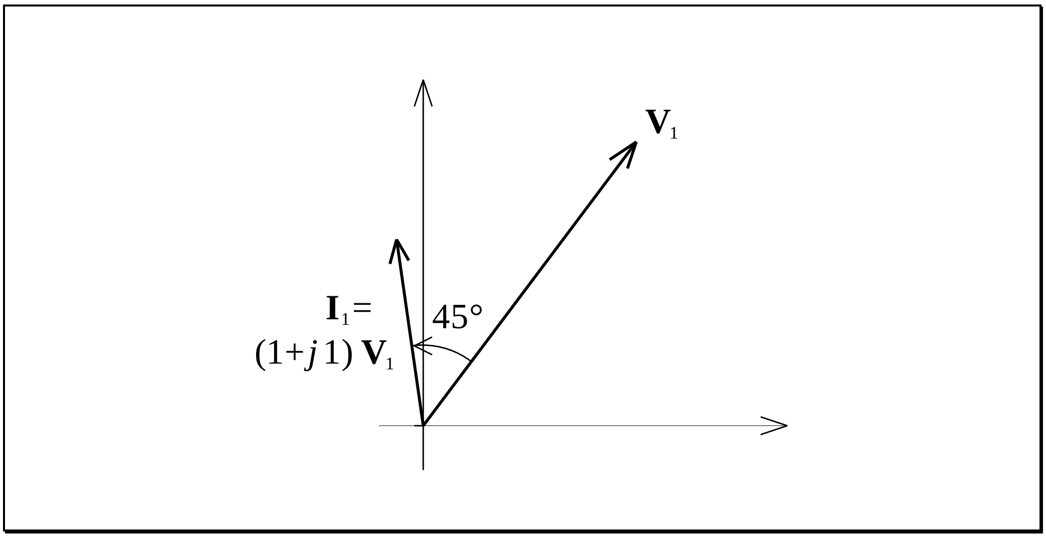


Figure 14.12

This last phasor diagram shows both current and voltage phasors on the same complex plane – it is understood that each will have its own amplitude scale, but a common angle scale.

The phasor diagram can also show the connection between the frequency-domain and the time-domain. For example, let us show the phasor $\mathbf{V} = V_m \angle \theta$ on the phasor diagram, as in (a) below:

Phasor diagram showing transformation to the time-domain by rotating the phasor and taking the real part

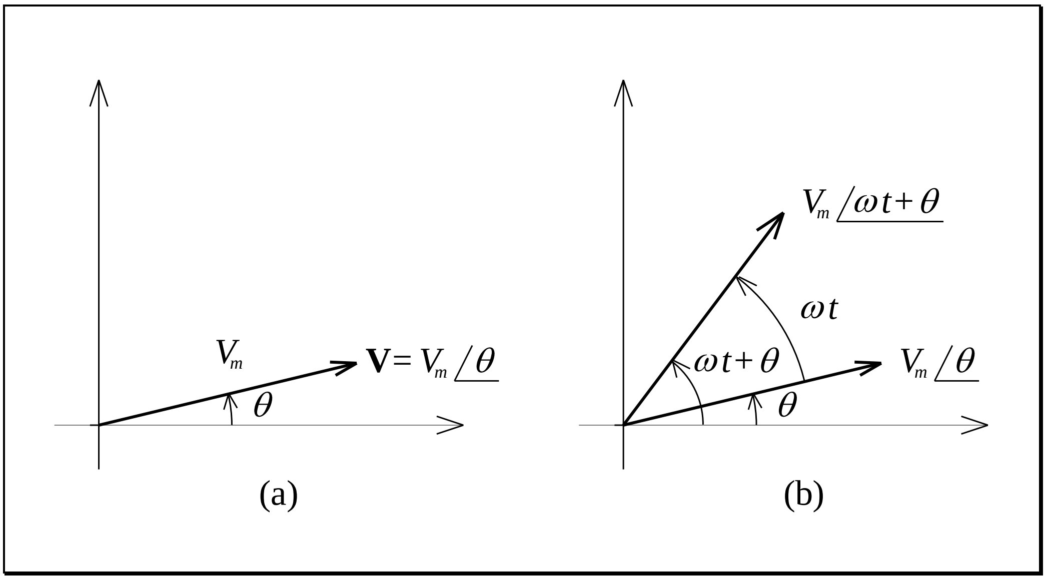


Figure 14.13

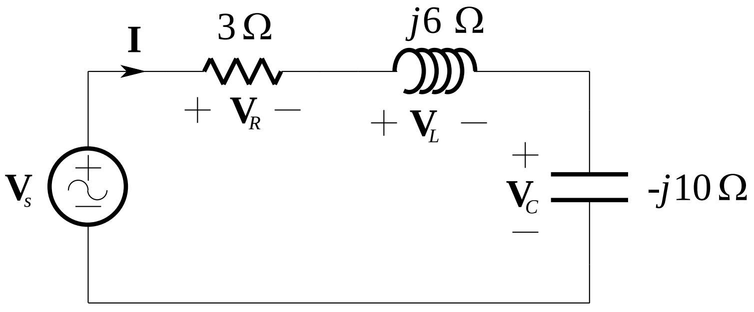
In order to transform $\mathbf{V}$ to the time-domain, we first need to multiply by $e^{j\omega t}$. We now have the complex voltage $\mathbf{V} = V_m e^{j\theta} e^{j\omega t} = V_m \angle \omega t + \theta$. This voltage may be interpreted as a phasor which possesses a phase angle that increases linearly with time. On a phasor diagram it therefore represents a rotating line segment, the instantaneous position being ωt rad ahead (counterclockwise) of $V_m \angle \theta$. Both $V_m \angle \theta$ and $V_m \angle \omega t + \theta$ are shown on the phasor diagram in (b).

The transformation to the time-domain is completed by taking the real part of $V_m \angle \omega t + \theta$, which is the projection of the phasor onto the real axis. It is helpful to think of the arrow representing the phasor $\mathbf{V}$ on the phasor diagram as the snapshot, taken at $t = 0$, of a rotating arrow whose projection onto the real axis is the instantaneous voltage $v(t)$.

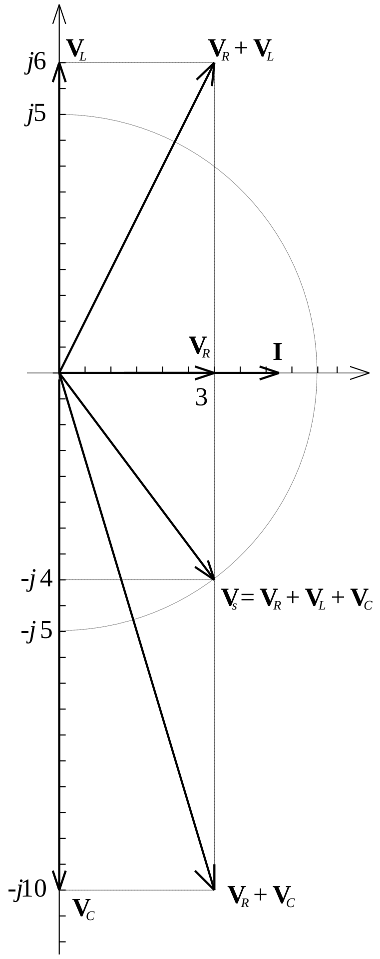
See the "Phasors" PC program

EXAMPLE 14.1 Phasor Diagram of a Series *RLC* Circuit

The series *RLC* circuit shown below has several different voltages associated with it, but only a single current:

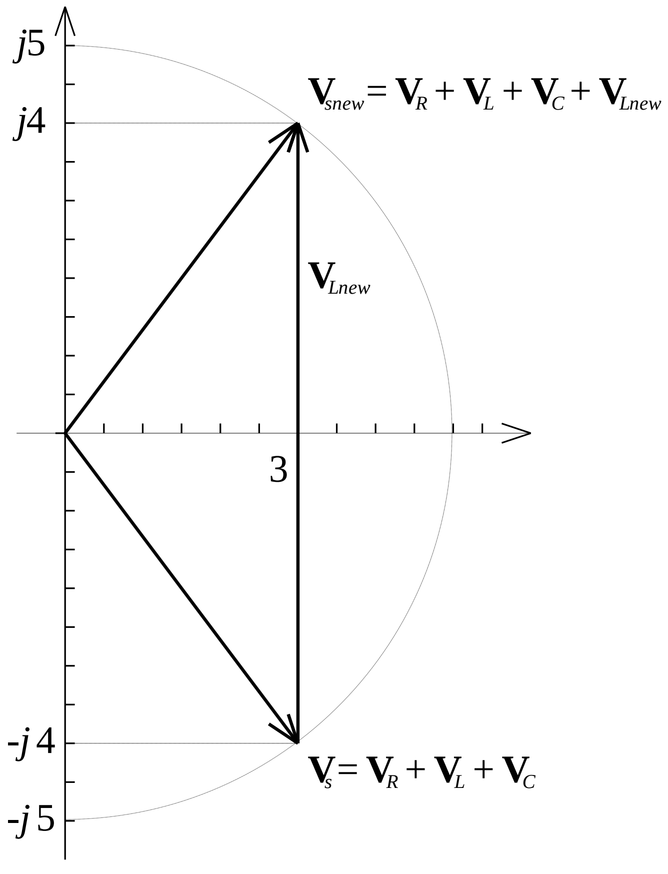


The phasor diagram is constructed most easily by employing the single current as the *reference* phasor – all other phasors with have their angles measured with respect to the reference. Let us arbitrarily set $\mathbf{I} = I_m \angle 0°$ and place it along the real axis of the phasor diagram, as shown below:



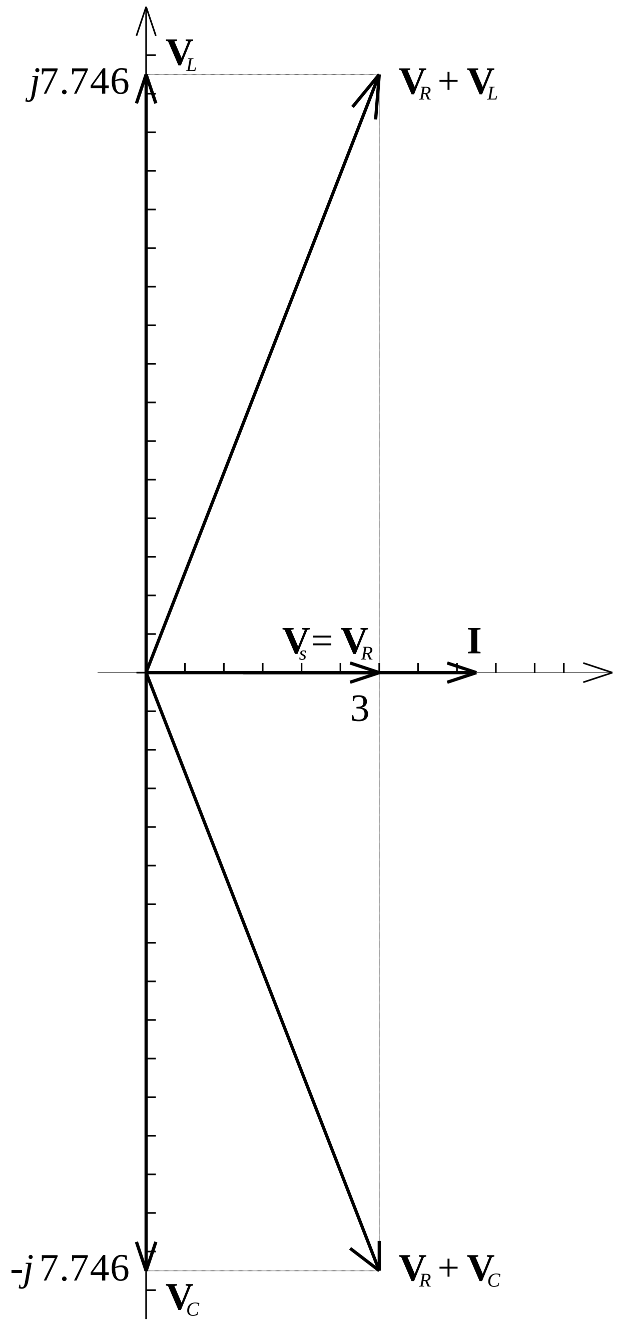
The resistor, inductor and capacitor voltages may next be calculated and placed on the diagram, where the $90°$ phase relationships stand out clearly. The sum of these three voltages is the source voltage for this circuit. The total voltage across the resistance and inductance or resistance and capacitance or inductance and capacitance is easily obtained from the phasor diagram.

We can design using the phasor diagram quite easily instead of embarking on complex algebraic manipulation. For example, suppose we would like to determine a single extra passive element that can be added in series with the circuit so that the magnitude of the current does not change. This additional circuit element will contribute to an additional voltage drop, but we still must have KVL satisfied so that the total voltage drop magnitude equals the source voltage magnitude. Therefore, the addition of the voltage drop due to the new element must keep the source voltage on a circle of radius $|\mathbf{V}_s|$. From the phasor diagram, we can see that we can only add an inductor with an impedance $\mathbf{Z}_{Lnew} = j8\,\Omega$, so that the additional voltage drop still brings us onto the circle of radius $|\mathbf{V}_s|$:



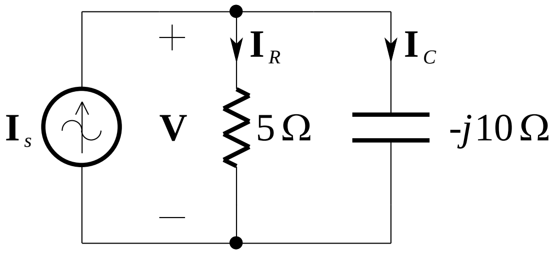
14.18

As another example of the use of the phasor diagram, consider the original circuit again. We know that an increase in frequency will cause the voltage across the inductor to increase, whilst simultaneously decreasing the voltage across the capacitor (although not linearly). In fact, if we increase the frequency by 29.1%, the inductor voltage and capacitor voltage will exactly cancel one another, and we have a condition known as resonance. In this case the supply voltage and current are precisely in phase, and the circuit appears resistive:

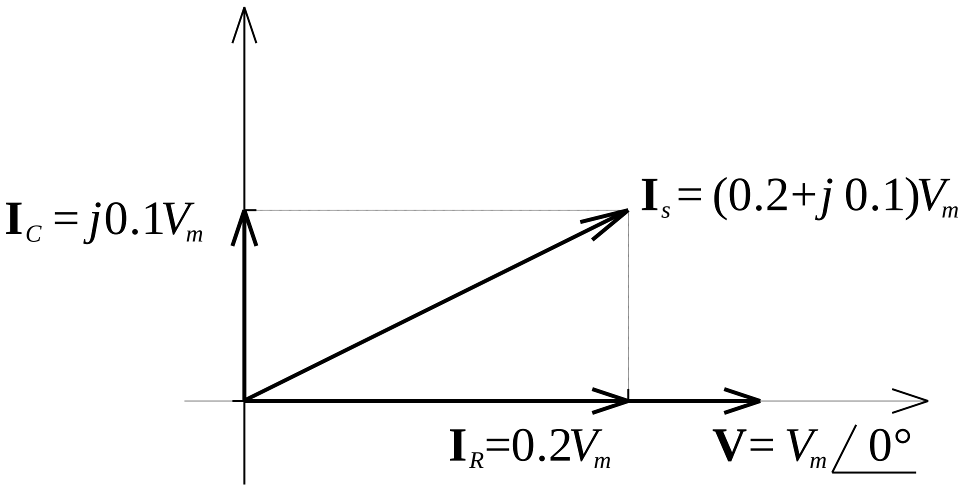


EXAMPLE 14.2 Phasor Diagram of a Parallel *RC* Circuit

The figure below shows a simple parallel circuit in which it is logical to use the single voltage between the two nodes as a reference phasor:

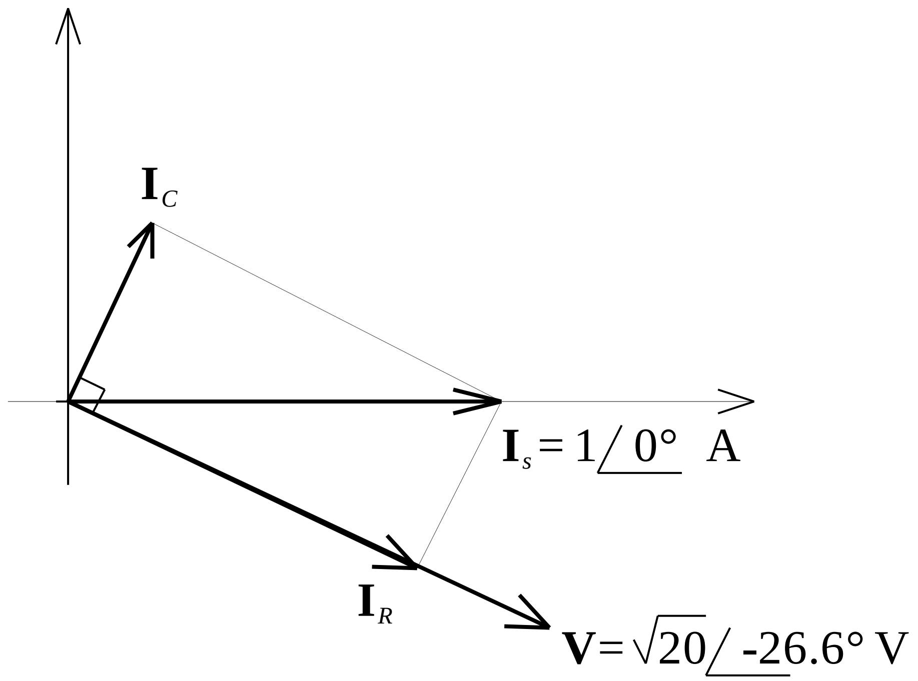


Let us arbitrarily set $\mathbf{V} = V_m \angle 0°$ and place it along the real axis of the phasor diagram. The resistor current is in phase with this voltage, $\mathbf{I}_R = 0.2 V_m$ A, and the capacitor current leads the reference voltage by $90°$, $\mathbf{I}_C = j0.1 V_m$ A. After these two currents are added to the phasor diagram, shown below, they may be summed to obtain the source current. The result is $\mathbf{I}_s = (0.2 + j0.1) V_m$ A.



14.20

If the source current were specified initially as, for example, $\mathbf{I}_s = 1\angle 0°\ \text{A}$, and the node voltage is not initially known, it is still convenient to begin construction of the phasor diagram by assuming, say $\mathbf{V} = 1\angle 0°$. The source current, as a result of the assumed node voltage, is now $\mathbf{I}_s = 0.2 + j0.1\ \text{A}$. The true source current is $1\angle 0°\ \text{A}$, however, and thus the true node voltage is greater by the factor $1/(0.2 + j0.1)$; the true node voltage is therefore $4 - j2\ \text{V}$. The assumed voltage leads to a phasor diagram which differs from the true phasor diagram by a change of scale (the assumed diagram is smaller by a factor of $1/\sqrt{20}$) and an angular rotation (the assumed diagram is rotated clockwise through $26.6°$). The true phasor diagram in this case is shown below:



Phasor diagrams are usually very simple to construct, and most sinusoidal steady-state analyses will be more meaningful if such a diagram is included.

14.8 Power in the Sinusoidal Steady-State

Consider a single sinusoidal source supplying a network as shown below:

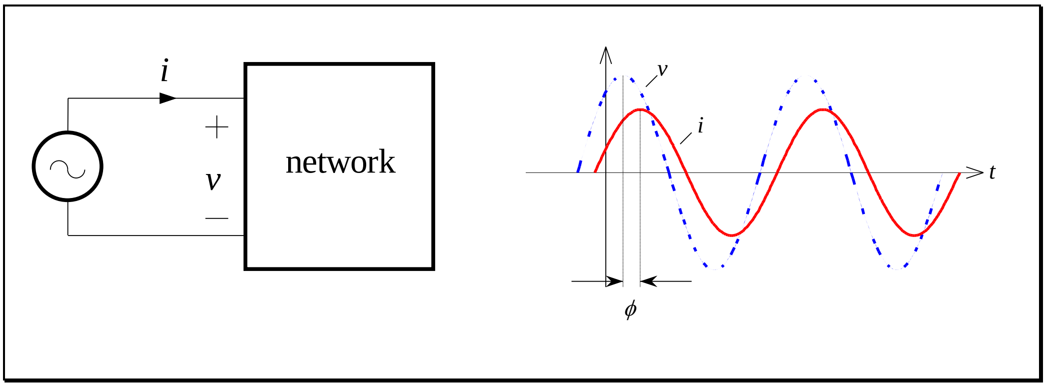


Figure 14.14

If the network contains reactive components, then in general there will be a phase shift between the voltage and current. Let:

$$\begin{aligned} v &= V_m \cos(\omega t + \alpha) \\ i &= I_m \cos(\omega t + \beta) \end{aligned} \tag{14.23}$$

and define:

$$\phi = \alpha - \beta \tag{14.24}$$

Thus ϕ is the angle by which the voltage leads the current.

14.8.1 Instantaneous Power

The instantaneous power delivered to the network is:

$$\begin{aligned} p &= vi \\ &= V_m I_m \cos(\omega t + \alpha)\cos(\omega t + \beta) \\ &= \frac{V_m I_m}{2}\left[\cos\phi + \cos(2\omega t + \alpha + \beta)\right] \end{aligned} \tag{14.25}$$

Notice that the first term is a constant, and the second term oscillates with time at double the supply frequency.

14.8.2 Average Power

Average power is just the average value of the instantaneous power. We define this average in the normal way (the "mean value theorem") as:

$$P = \lim_{T \to \infty} \frac{1}{T} \int_{-T/2}^{T/2} p(t)dt \tag{14.26}$$

If the instantaneous power is periodic with period T_0, we have the special case:

$$P = \frac{1}{T_0} \int_{-T_0/2}^{T_0/2} p(t)dt \tag{14.27}$$

That is, for a periodic instantaneous power, we can integrate over one period, and divide by the period. A graph of the instantaneous power in a network operating in the sinusoidal steady-state, Eq. (14.25), is shown below:

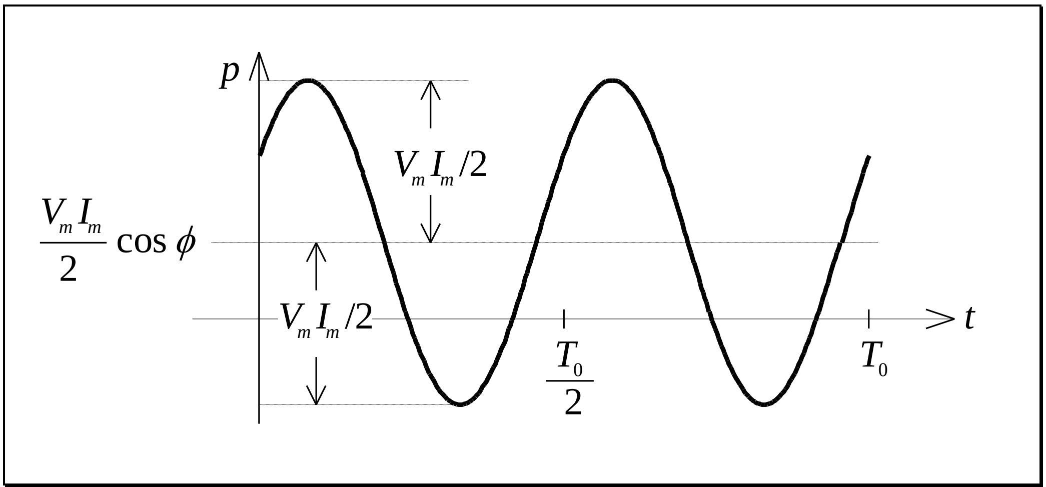


Figure 14.15

From this graph it is easily seen that the *average* power is the constant part of the instantaneous power (the oscillating part averages to zero) and we have:

P is the *average* value of *p*

$$P = \frac{V_m I_m}{2} \cos\phi \quad \text{(W)} \tag{14.28}$$

EXAMPLE 14.3 Instantaneous Power and Average Power

A graph of the instantaneous power is shown below for $\mathbf{V} = 4\angle 0°$ and $\mathbf{I} = 2\angle -60°$:

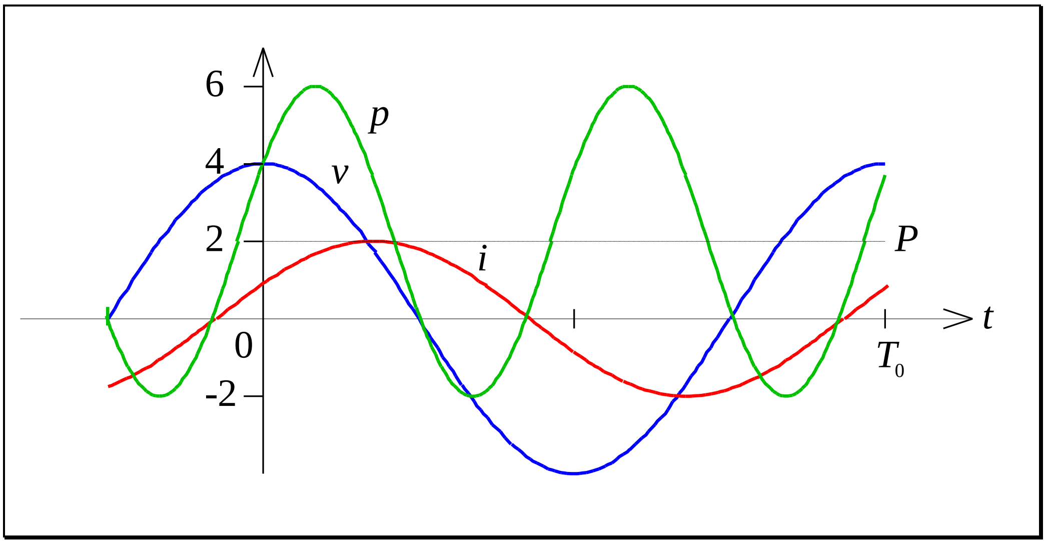


Figure 14.16

Note that, on occasion, the power *delivered* to the network is negative, which implies that the network is actually *sourcing* power back to the voltage supply.

The average power is calculated to be:

$$P = \frac{V_m I_m}{2}\cos\phi = \tfrac{1}{2}(4)(2)\cos 60° = 2\text{ W}$$

Both the 2 W average power and its period, one-half the period of either the current or the voltage, are evident in the graph. The zero value of the instantaneous power at each instant when either the voltage or current is zero is also apparent.

14.8.3 Root-Mean-Square (RMS) Values

It is customary, when dealing with AC power, to refer to voltage and current using a root-mean-square, or RMS value. As we shall see, this leads to some simplification for many power formulas. The concept of an RMS value for a voltage or current comes about by considering the average power dissipated in a resistor when it carries a periodic current:

$$P = \frac{1}{T_0}\int_{-T_0/2}^{T_0/2} Ri^2 dt = RI_{RMS}^2 \tag{14.29}$$

That is, the RMS value of any periodic current is equal to the value of the direct current which delivers the same average power. Removing *R* from the above formula, we thus have:

RMS value defined

$$I_{RMS} = \sqrt{\frac{1}{T_0}\int_{-T_0/2}^{T_0/2} i^2 dt} \tag{14.30}$$

The operation involved in finding this value is the *root* of the *mean* of the *square*, hence the name root-mean-square value, or RMS value for short. A similar expression is obtained for voltage, V_{RMS} (or for any other signal for that matter).

14.8.4 RMS Value of a Sinusoid

A sinusoid is the most important special case of a periodic signal. Consider a sinusoidal current given by:

$$i(t) = I_m \cos(\omega t + \phi) \tag{14.31}$$

The easiest way to find its RMS value is by performing the mean-square operations in Eq. (14.30) graphically. For the arbitrary sinusoid given, we can graph the square of the current, $i^2(t) = I_m^2 \cos^2(\omega t + \phi)$, as shown below:

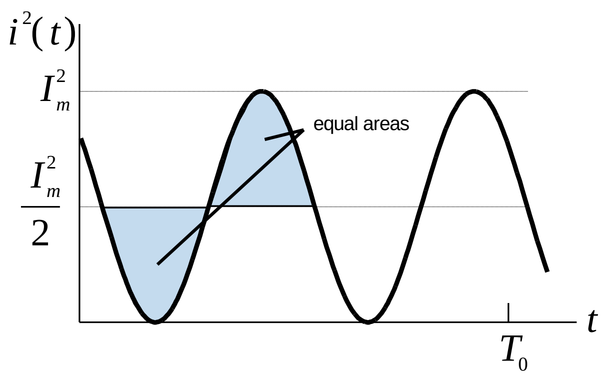


Figure 14.17

Note that in drawing the graph we don't really need to know the identity $\cos^2(\theta) = (1 + \cos 2\theta)/2$ – all we need to know is that if we start off with a sinusoid uniformly oscillating between I_m and $-I_m$, then after squaring we obtain a sinusoid that oscillates (at twice the frequency) uniformly between I_m^2 and 0. We can now see that the average value of the resulting waveform is $I_m^2/2$, because there are equal areas above and below this value. This is the *mean* of the *square*, and so we now just take the *root* and get:

$$I_{RMS} = \frac{I_m}{\sqrt{2}} \tag{14.32}$$

The RMS value of an arbitrary sinusoid

Note that the RMS value depends on the magnitude of the sinusoid only – the frequency and phase is irrelevant!

In the power industry, it is tacitly assumed that values of voltage and current will be measured using their RMS value. For example, in Australia the electricity delivered to your home has a frequency of 50 Hz and an RMS value of 230 V. This means the voltage available at a general power outlet is a 50 Hz sinusoid with a peak value of approximately 325 V.

It should be noted that this formula can only be applied to a sinusoid – for other waveforms, you will obtain a different ratio between the peak and the RMS value. For example, the RMS value of a triangle waveform is $I_m/\sqrt{3}$ whilst for a square wave it is simply I_m.

14.8.5 Phasors and RMS Values

We defined a phasor corresponding to:

$$i(t) = I_m \cos(\omega t + \beta) \tag{14.33}$$

as:

$$\mathbf{I} = I_m \angle \beta \tag{14.34}$$

We could have just as easily defined it to be:

Phasor magnitudes can be defined as RMS values

$$\mathbf{I} = \frac{I_m}{\sqrt{2}} \angle \beta = I_{RMS} \angle \beta \tag{14.35}$$

If we use this definition, then all relationships involving phasors, such as $\mathbf{V} = \mathbf{ZI}$, KCL, KVL, etc. must also use this definition. When working with power and machines, it is customary to use the RMS value for the phasor magnitude. In other fields, such as telecommunications and electronics, we use the amplitude for the phasor magnitude. You need to be aware of this usage.

14.8.6 Average Power Using RMS Values

Returning to the formula for average power, we can now rewrite it using RMS values. We have:

$$P = V_{RMS} I_{RMS} \cos\phi \quad \text{(W)} \tag{14.36}$$

The average power P using RMS values of voltage and current

14.8.7 Apparent Power

The average power in a DC network is simply $P = VI$. In sinusoidal steady-state analysis, we define:

$$\text{apparent power} = V_{RMS} I_{RMS} \quad \text{(VA)} \tag{14.37}$$

Apparent power defined

Dimensionally, average power and apparent power have the same units, since $\cos\phi$ is dimensionless. However, to avoid confusion, the term volt-amperes, or VA, is applied to apparent power.

14.8.8 Power Factor

The ratio of average power to the apparent power is called the *power factor*, symbolized by PF:

$$\text{PF} = \frac{P}{V_{RMS} I_{RMS}} \tag{14.38}$$

Power factor defined

In the sinusoidal case, the power factor is also equal to:

$$\text{PF} = \cos\phi \tag{14.39}$$

We usually refer to *leading* PF or *lagging* PF when referring to loads to resolve the ambiguity in taking the "cos". The terms leading and lagging refer to the *phase of the current with respect to the voltage*. Thus, inductive loads have a lagging power factor, capacitive loads have a leading power factor.

14.8.9 Complex Power

Using RMS phasors of $v = V_m \cos(\omega t + \alpha)$ and $i = I_m \cos(\omega t + \beta)$:

$$\begin{aligned} \mathbf{V} &= \frac{V_m}{\sqrt{2}} \angle\alpha = V_{RMS}\angle\alpha = V_{RMS} e^{j\alpha} \\ \mathbf{I} &= \frac{I_m}{\sqrt{2}} \angle\beta = I_{RMS}\angle\beta = I_{RMS} e^{j\beta} \end{aligned} \tag{14.40}$$

we know that the average power is:

$$P = V_{RMS} I_{RMS} \cos(\alpha - \beta) \tag{14.41}$$

We can associate the average power with the real part of a complex power:

$$\begin{aligned} P &= \operatorname{Re}\left\{V_{RMS} I_{RMS} e^{j(\alpha-\beta)}\right\} \\ &= \operatorname{Re}\left\{V_{RMS} e^{j\alpha} I_{RMS} e^{-j\beta}\right\} \end{aligned} \tag{14.42}$$

Thus, we can define average power as:

$$P = \operatorname{Re}\left\{\mathbf{V}\mathbf{I}^*\right\} \tag{14.43}$$

Note the use of $\mathbf{I}^*$ because of the way we defined $\phi = \alpha - \beta$. It is an accident of history that ϕ was defined this way, as it just as easily could have been defined as $\phi = \beta - \alpha$.

We therefore define *complex power* as:

Complex power defined

$$\mathbf{S} = \mathbf{V}\mathbf{I}^* \quad \text{(complex VA)} \tag{14.44}$$

In polar form the complex power is:

$$\mathbf{S} = \mathbf{VI}^* = V_{RMS} I_{RMS} \angle \phi \tag{14.45}$$

so we can now see that the apparent power is:

$$|\mathbf{S}| = |\mathbf{V}||\mathbf{I}| = V_{RMS} I_{RMS} \tag{14.46}$$

Written in rectangular form, the complex power is:

$$\mathbf{S} = V_{RMS} I_{RMS} (\cos \phi + j \sin \phi) = P + jQ \tag{14.47}$$

14.8.10 Reactive Power

From the rectangular form, we can see that the average power is also the "real" power. It can also be seen that the "imaginary" power, which we call the *reactive power*, is:

Reactive power defined

$$Q = V_{RMS} I_{RMS} \sin \phi \quad \text{(var)} \tag{14.48}$$

It has the same dimensions as the real power P, the complex power $\mathbf{S}$ and the apparent power $|\mathbf{S}|$. In order to avoid confusion, the unit of Q is defined as the volt-ampere-reactive, or var for short.

Reactive power is a measure of the energy flow rate into or out of the *reactive* components of a network. It is positive for inductive loads, and negative for capacitive loads.

The physical interpretation of reactive power causes a lot of confusion. Even though it is the imaginary component of complex power, it has a physical and real interpretation, and must be generated by a power system. (A voltage phasor, such as $\mathbf{V} = 3 + j4$, has an imaginary component of 4 which contributes to the amplitude and phase of the real voltage sinusoid – so we expect Q to also be physically real).

14.30

To see how Q manifests itself in the real world, we return to the formula for instantaneous power. Using $\cos(A-B)=\cos A\cos B+\sin A\sin B$, and noting that $\beta=\alpha-\phi$, we get:

$$\begin{aligned} p &= V_{RMS}I_{RMS}\left[\cos\phi+\cos(2\omega t+\alpha+\beta)\right] \\ &= V_{RMS}I_{RMS}\left[\cos\phi+\cos(2\omega t+2\alpha-\phi)\right] \\ &= V_{RMS}I_{RMS}\left[\cos\phi+\cos(2\omega t+2\alpha)\cos\phi+\sin(2\omega t+2\alpha)\sin\phi\right] \\ &= V_{RMS}I_{RMS}\cos\phi\left[1+\cos(2\omega t+2\alpha)\right]+V_{RMS}I_{RMS}\sin\phi\sin(2\omega t+2\alpha) \end{aligned} \tag{14.49}$$

Then:

$$p=P\left[1+\cos(2\omega t+2\alpha)\right]+Q\sin(2\omega t+2\alpha) \tag{14.50}$$

The instantaneous power associated with the real and reactive power components is shown below:

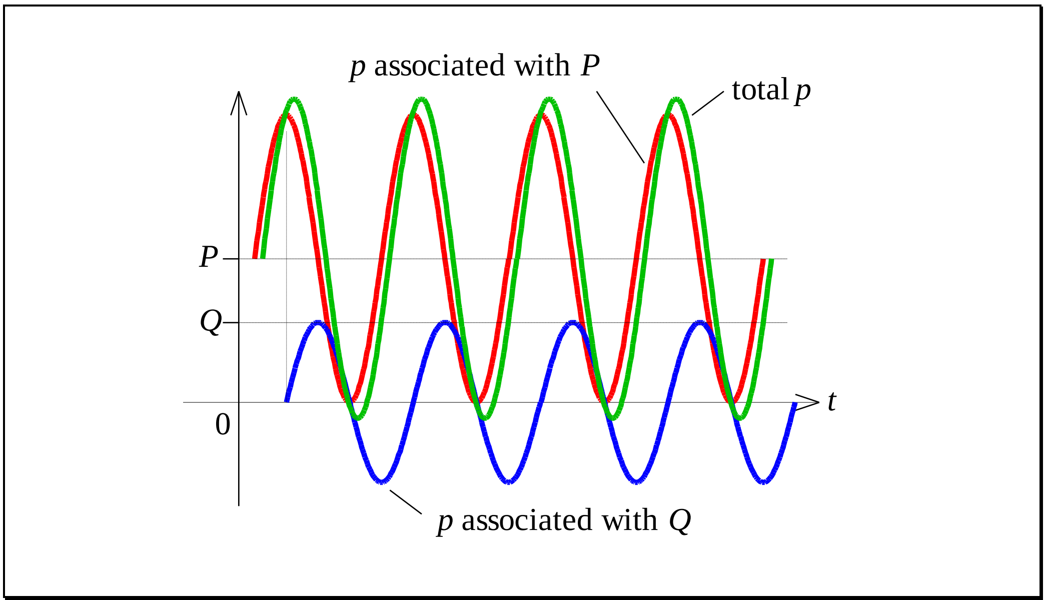


Figure 14.18

Note that the instantaneous power associated with P follows a "cos", and the instantaneous power associated with Q follows a "sin". Thus, the two waveforms are 90° apart and are said to be in "quadrature". You can see that reactive power does not transfer energy – instantaneous power is both delivered to, and received from, the network in a cyclic fashion, with an average of zero. In contrast, real power does transfer energy – instantaneous power is always delivered to the network in a cyclic fashion, but it has a non-zero average.

14.8.11 Summary of Power in AC Circuits

In summary we have:

$$\mathbf{S} = \mathbf{V}\mathbf{I}^* = |\mathbf{S}|\angle\phi = P + jQ \tag{14.51}$$

where:

$$|\mathbf{S}| = |\mathbf{V}||\mathbf{I}| = \text{apparent power} \quad \text{(VA)}$$

$$P = |\mathbf{S}|\cos\phi = \text{real power} \quad \text{(W)}$$

$$Q = |\mathbf{S}|\sin\phi = \text{reactive power} \quad \text{(var)} \tag{14.52}$$

ϕ = angle by which the current **I** lags the voltage **V**

$$|\mathbf{S}|^2 = P^2 + Q^2$$

Components of complex power

Real power is also known as active power and average power

These relationships are illustrated below:

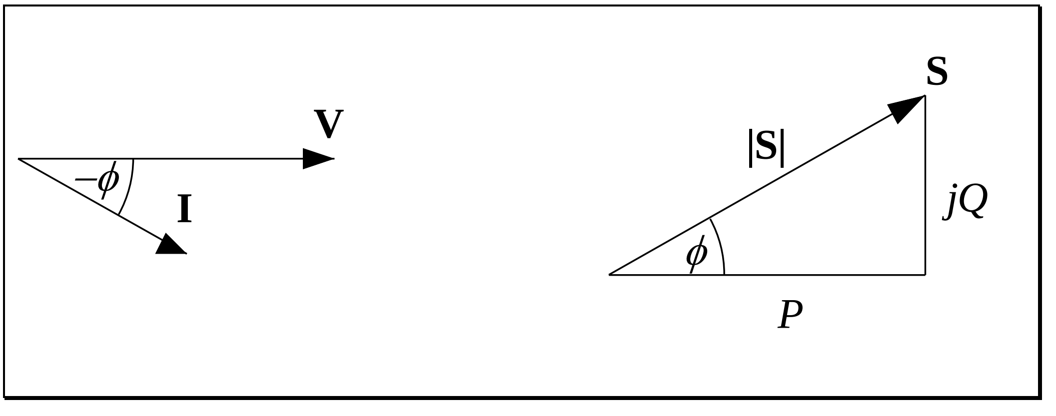


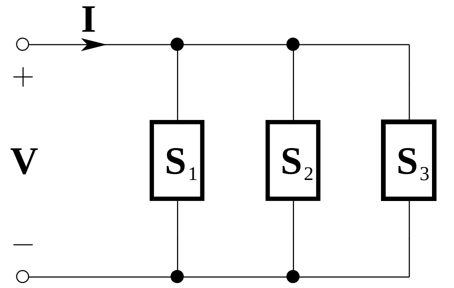
Figure 14.19

The diagram on the right is called a "power triangle". Note that P and Q, are the real and imaginary parts of the complex power **S**.

It can be shown that the total complex power $\mathbf{S} = P + jQ$ consumed by a network is the sum of the complex powers consumed by all the component parts of the network. This conservation property is not true of the apparent power $|\mathbf{S}|$.

EXAMPLE 14.4 Conservation of Complex Power

Consider three loads connected in parallel across a 230 V (RMS) 50 Hz line as shown below:



Load 1 absorbs 10 kW and 7.5 kvar. Load 2 absorbs 3.84 kW at 0.96 PF leading. Load 3 absorbs 5 kW at unity power factor. Find the overall power factor.

The first load is given in rectangular form:

$$\mathbf{S}_1 = 10 + j7.5\,\text{kVA}$$

The complex power supplied to the second load must have a real part of 3.84 kW and an angle (refer to the power triangle) of $\cos^{-1}(0.96) = -16.26°$. Hence,

$$\mathbf{S}_2 = \frac{P}{\cos\phi}\angle\phi = \frac{3.84}{0.96}\angle -16.26° = 4\angle -16.26° = 3.84 - j1.12\,\text{kVA}$$

The third load is simply:

$$\mathbf{S}_3 = 5 + j0\,\text{kVA}$$

The total complex power is:

$$\mathbf{S}_{total} = \mathbf{S}_1 + \mathbf{S}_2 + \mathbf{S}_3 = 10 + j7.5 + 3.84 - j1.12 + 5 + j0 = 18.84 + j6.38\,\text{kVA}$$

Thus, the combined load is operating at a power factor equal to:

$$\text{PF} = \frac{P}{|\mathbf{S}|} = \frac{18.84}{19.89} = 0.9472 \;\; \text{lagging}$$

The magnitude of the line current drawn by the combined load is:

$$|\mathbf{I}| = \frac{|\mathbf{S}|}{|\mathbf{V}|} = \frac{19890}{230} = 86.48 \text{ A RMS}$$

Electricity supply authorities do all they can to improve the PF of their loads by installing capacitors or special machines called synchronous condensers which supply vars to the system. They also impose tariffs which encourage consumers to correct their PF.

If we now seek to raise the PF to 0.98 lagging, without affecting the existing real power, the total complex power must become:

$$\mathbf{S}_{new} = \frac{18.84}{0.98} \angle \cos^{-1}(0.98) = 19.22 \angle 11.48° = 18.84 + j3.826 \text{ kVA}$$

We would therefore need to add a corrective load of:

$$\mathbf{S}_4 = \mathbf{S}_{new} - \mathbf{S}_{total} = -j2.554 \text{ kVA}$$

Now since:

$$\mathbf{S}_4 = \mathbf{V}\mathbf{I}_4^* = \frac{\mathbf{V}\mathbf{V}^*}{\mathbf{Z}_4^*} = \frac{|\mathbf{V}|^2}{\mathbf{Z}_4^*}$$

then:

$$\mathbf{Z}_4 = \frac{|\mathbf{V}|^2}{\mathbf{S}_4^*} = \frac{230^2}{j2554} = -j20.71\,\Omega$$

Thus, the corrective load is a capacitor of value:

$$C = \frac{1}{\omega X_C} = \frac{1}{2\pi \times 50 \times 20.71} = 153.7\ \mu\text{F}$$

The magnitude of the line current drawn by the new combined load reduces to:

$$|\mathbf{I}| = \frac{|\mathbf{S}|}{|\mathbf{V}|} = \frac{19220}{230} = 83.57 \text{ A}$$

14.9 Summary

- A linear circuit can be converted to the frequency-domain where we use the concept of phasors and impedances, and in particular the branch phasor relationship $\mathbf{V} = \mathbf{ZI}$. The circuit is then amenable to normal circuit analysis techniques: nodal analysis, mesh analysis, superposition, source transformations, Thévenin's theorem, Norton's theorem, etc. Time-domain responses are obtained by transforming phasor responses back to the time-domain.

- A phasor diagram is a sketch in the complex plane of the phasor voltages and currents throughout a circuit and is a useful graphical tool to illustrate, analyse and design the sinusoidal steady-state response of the circuit.

- The RMS value of a sinusoid $x(t) = X_m \cos(\omega t + \phi)$ is $X_{RMS} = X_m / \sqrt{2}$.

- There are many power terms in AC circuits:

  - instantaneous power, $p(t) = v(t)i(t)$ (W)
  - average power, $P = V_{RMS} I_{RMS} \cos\phi$ (W)
  - reactive power, $Q = V_{RMS} I_{RMS} \sin\phi$, (var)
  - complex power, $\mathbf{S} = \mathbf{VI}^* = P + jQ$, (VA)
  - apparent power, $|\mathbf{S}| = |\mathbf{V}||\mathbf{I}|$, (VA)
  - power factor, $\mathrm{PF} = \cos\phi$

- The average power delivered to the *resistive* component of a load is nonzero. The average power delivered to the *reactive* component of a load is zero.

- Capacitors are commonly used to improve the PF of industrial loads to minimize the reactive power and current required from the electricity utility.

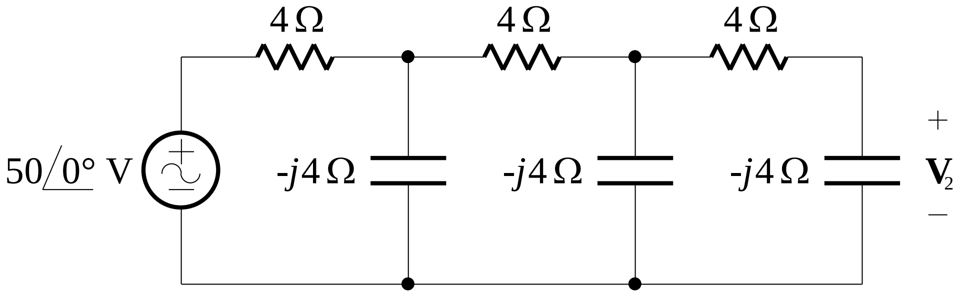
14.10 References

Hayt, W. & Kemmerly, J.: *Engineering Circuit Analysis*, 3rd *Ed*., McGraw-Hill, 1984.

Exercises

1.

Consider the circuit shown below:

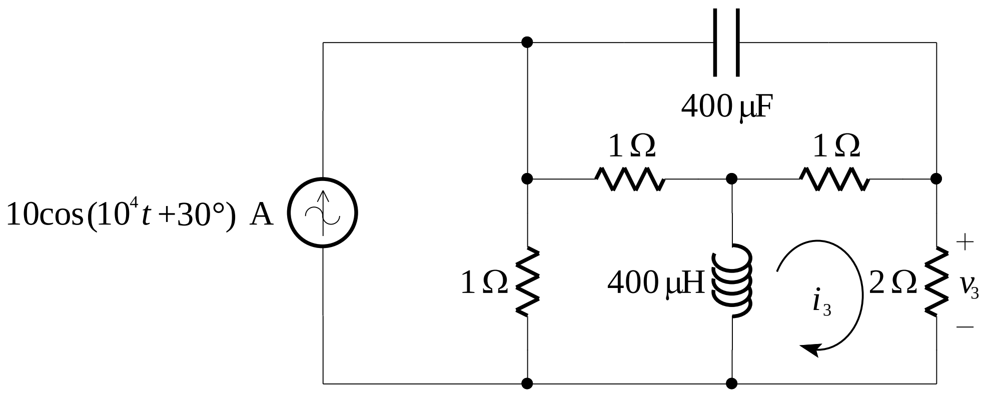


(a) Find $\mathbf{V}_2$.

(b) To what identical value should each of the $4\ \Omega$ resistors be changed so that $\mathbf{V}_2$ is $180°$ out of phase with the source voltage?

2.

Consider the circuit shown below:



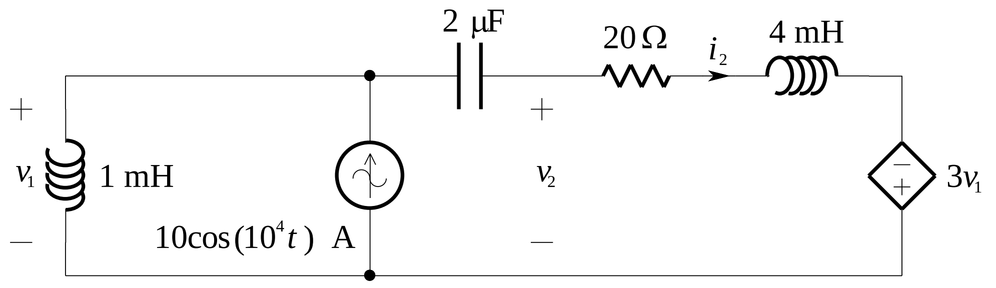
(a) Find $v_3(t)$ in the steady-state by using nodal analysis.

(b) Find $i_3(t)$ in the steady-state by using mesh analysis

14.36

3.

Consider the circuit shown below:

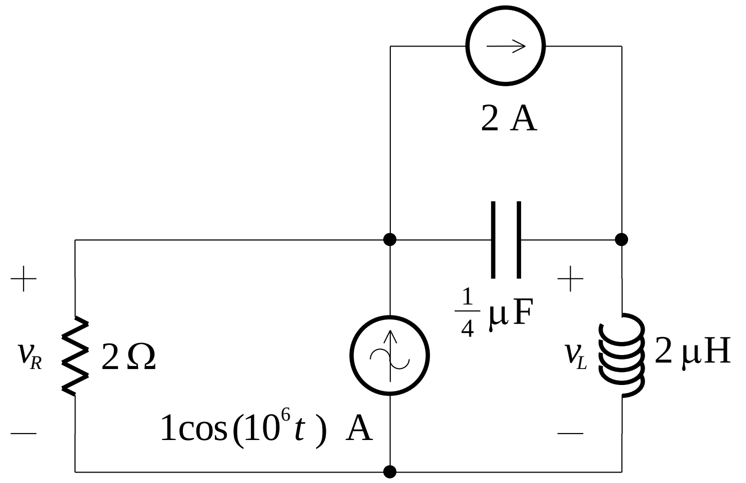


(a) Find $i_2(t)$.

(b) Change the control voltage from $3v_1$ to $3v_2$ and again find $i_2(t)$.

4.

Consider the circuit shown below:

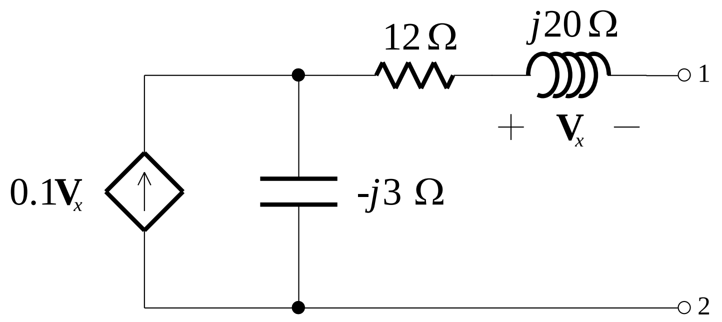


(a) Find $v_R(t)$.

(b) Find $v_L(t)$.

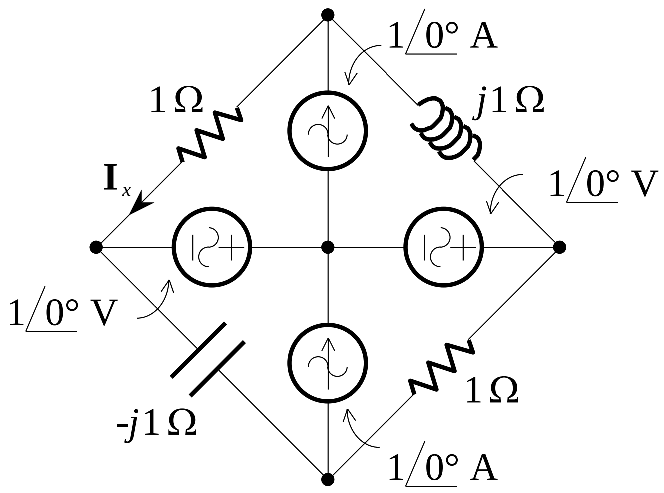
5.

Find the output impedance of the network shown below:



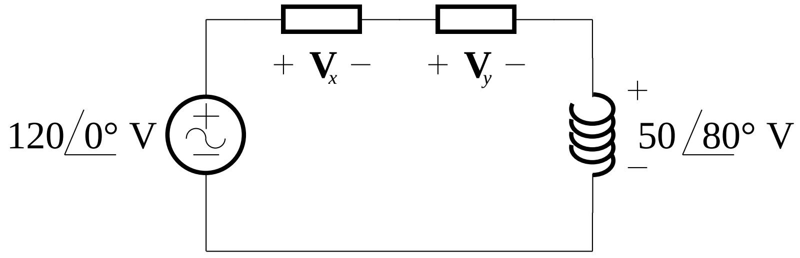
6.

Use superposition to find $\mathbf{I}_x$ in the circuit below:



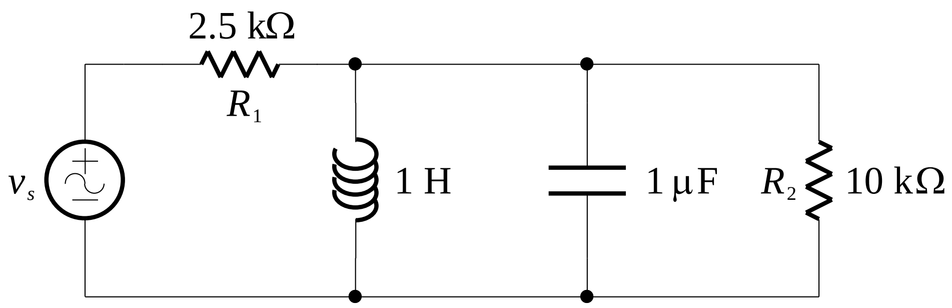
7.

For the circuit below, find the phase angles of $\mathbf{V}_x$ and $\mathbf{V}_y$ graphically if $|\mathbf{V}_x| = 90\ \text{V}$ and $|\mathbf{V}_y| = 150\ \text{V}$.



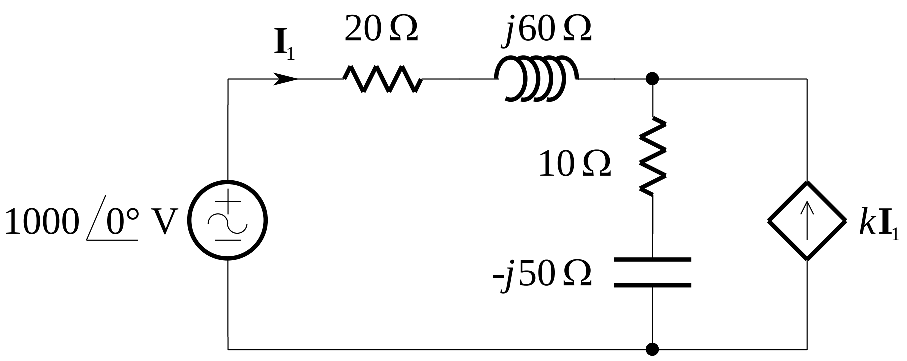
8.

If $v_s = 20\cos(1000t + 30°)$ V in the circuit below, find the power being absorbed by each passive element at $t = 0$.



9.

Determine the average power delivered to each resistor in the network shown below if: (a) $k = 0$ (b) $k = 1$

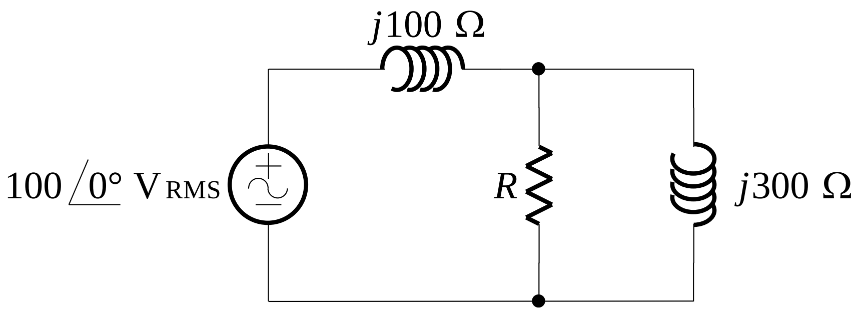


10.

The series combination of a 1000 Ω resistor and a 2 H inductor must not absorb more than 100 mW of power at any instant. Assuming a sinusoidal current with $\omega = 400\ \text{rads}^{-1}$, what is the largest RMS current that can be tolerated?

11.

Consider the following circuit:



(a) What value of R will cause the RMS voltages across the inductors to be equal?

(b) What is the value of that RMS voltage?

12.

A composite load consists of three loads connected in parallel.

One draws 100 W at 0.9 lagging PF, another takes 200 W at 0.8 lagging PF, and the third requires 150 W at unity PF. The composite load is supplied by a source $\mathbf{V}_s$ in series with a $10\,\Omega$ resistor. If the loads are all to operate at 110 V RMS, determine:

(a) the RMS current through the source

(b) the PF of the composite load

13.

A load operating at 2300 V RMS draws 25 A RMS at a power factor of 0.815 lagging. Find:

(a) the real power taken by the load

(b) the reactive power

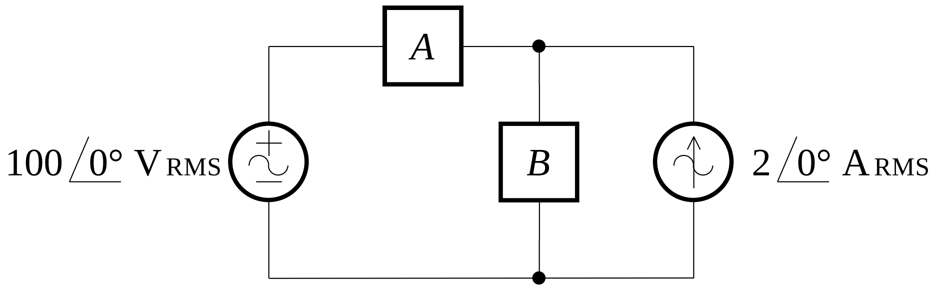
(c) the complex power

(d) the apparent power drawn by the load

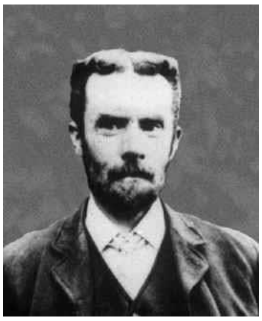
(e) the impedance of the load

14.

In the circuit shown below, load *A* receives $\mathbf{S}_A = 80 + j40\,\text{VA}$, while load *B* absorbs $\mathbf{S}_B = 100 - j200\,\text{VA}$. Find the complex power supplied by each source.



Oliver Heaviside (1850-1925)



The mid-Victorian age was a time when the divide between the rich and the poor was immense (and almost insurmountable), a time of unimaginable disease and lack of sanitation, a time of steam engines belching forth a steady rain of coal dust, a time of horses clattering along cobblestoned streets, a time when social services were the fantasy of utopian dreamers. It was into this smelly, noisy, unhealthy and class-conscious world that Oliver Heaviside was born the son of a poor man on 18 May, 1850.

A lucky marriage made Charles Wheatstone (of Wheatstone Bridge fame) Heaviside's uncle. This enabled Heaviside to be reasonably well educated, and at the age of sixteen he obtained his first (and last) job as a telegraph operator with the Danish-Norwegian-English Telegraph Company. It was during this job that he developed an interest in the physical operation of the telegraph cable. At that time, telegraph cable theory was in a state left by Professor William Thomson (later Lord Kelvin) – a diffusion theory modelling the passage of electricity through a cable with the same mathematics that describes heat flow.

I remember my first look at the great treatise of Maxwell's…I saw that it was great, greater and greatest, with prodigious possibilities in its power. – Oliver Heaviside

By the early 1870's, Heaviside was contributing technical papers to various publications – he had taught himself calculus, differential equations, solid geometry and partial differential equations. But the greatest impact on Heaviside was Maxwell's treatise on electricity and magnetism – Heaviside was swept up by its power.

In 1874 Heaviside resigned from his job as telegraph operator and went back to live with his parents. He was to live off his parents, and other relatives, for the rest of his life. He dedicated his life to writing technical papers on telegraphy and electrical theory – much of his work forms the basis of modern circuit theory and field theory.

In 1876 he published a paper entitled *On the extra current* which made it clear that Heaviside (a 26-year-old unemployed nobody) was a brilliant talent. He had extended the mathematical understanding of telegraphy far beyond

Thomson's submarine cable theory. It showed that inductance was needed to permit finite-velocity wave propagation, and would be the key to solving the problems of long distance telephony. Unfortunately, although Heaviside's paper was correct, it was also unreadable by all except a few – this was a trait of Heaviside that would last all his life, and led to his eventual isolation from the "academic world". In 1878, he wrote a paper *On electromagnets, etc.* which introduced the expressions for the AC impedances of resistors, capacitors and inductors. In 1879, his paper *On the theory of faults* showed that by "faulting" a long telegraph line with an inductance, it would actually improve the signalling rate of the line – thus was born the idea of "inductive loading", which allowed transcontinental telegraphy and long-distance telephony to be developed in the USA.

Now all has been blended into one theory, the main equations of which can be written on a page of a pocket notebook. That we have got so far is due in the first place to Maxwell, and next to him to Heaviside and Hertz. – H.A. Lorentz

When Maxwell died in 1879 he left his electromagnetic theory as twenty equations in twenty variables! It was Heaviside (and independently, Hertz) who recast the equations in modern form, using a symmetrical vector calculus notation (also championed by Josiah Willard Gibbs (1839-1903)). From these equations, he was able to solve an enormous amount of problems involving field theory, as well as contributing to the *ideas* behind field theory, such as energy being carried by fields, and not electric charges.

Rigorous mathematics is narrow, physical mathematics bold and broad. – Oliver Heaviside

A major portion of Heaviside's work was devoted to "operational calculus".[1] This caused a controversy with the mathematicians of the day because although it seemed to solve physical problems, it's mathematical rigor was not at all clear. His knowledge of the physics of problems guided him correctly in many instances to the development of suitable mathematical processes. In 1887 Heaviside introduced the concept of a *resistance operator*, which in modern terms would be called *impedance*, and Heaviside introduced the symbol Z for it. He let p be equal to time-differentiation, and thus the resistance operator for an inductor would be written as pL. He would then treat p just like an algebraic

[1] The Ukrainian Mikhail Egorovich Vashchenko-Zakharchenko published *The Symbolic Calculus and its Application to the Integration of Linear Differential Equations* in 1862. Heaviside independently invented (and applied) his *own* version of the operational calculus.

quantity, and solve for voltage and current in terms of a power series in p. In other words, Heaviside's operators allowed the reduction of the *differential* equations of a physical system to equivalent *algebraic* equations.

Heaviside was fond of using the unit-step as an input to electrical circuits, especially since it was a very practical matter to send such pulses down a telegraph line. The unit-step was even called the Heaviside step, and given the symbol $H(t)$, but Heaviside simply used the notation **1**. He was tantalizingly close to discovering the impulse by stating "… $p \cdot \mathbf{1}$ means a function of t which is wholly concentrated at the moment $t = 0$, of total amount 1. It is an impulsive function, so to speak…[it] involves only ordinary ideas of differentiation and integration pushed to their limit."

Paul Dirac derived the modern notion of the impulse, when he used it in 1927, at age 25, in a paper on quantum mechanics. He did his undergraduate work in electrical engineering and was both familiar with all of Heaviside's work and a great admirer of his.

Heaviside also played a role in the debate raging at the end of the 19th century about the age of the Earth, with obvious implications for Darwin's theory of evolution. In 1862 Thomson wrote his famous paper *On the secular cooling of the Earth*, in which he imagined the Earth to be a uniformly heated ball of molten rock, modelled as a semi-infinite mass. Based on experimentally derived thermal conductivity of rock, sand and sandstone, he then mathematically allowed the globe to cool according to the physical law of thermodynamics embedded in Fourier's famous partial differential equation for heat flow. The resulting age of the Earth (100 million years) fell short of that needed by Darwin's theory, and also went against geologic and palaeontologic evidence. John Perry (a professor of mechanical engineering) redid Thomson's analysis using discontinuous diffusivity, and arrived at approximate results that could (based on the conductivity and specific heat of marble and quartz) put the age of the Earth into the billions of years. But Heaviside, using his operational calculus, was able to solve the diffusion equation for a finite spherical Earth. We now know that such a simple model is based on faulty premises – radioactive decay within the Earth maintains the thermal gradient without a continual cooling of the planet. But the power of Heaviside's methods to solve remarkably complex problems became readily apparent.

The practice of eliminating the physics by reducing a problem to a purely mathematical exercise should be avoided as much as possible. The physics should be carried on right through, to give life and reality to the problem, and to obtain the great assistance which the physics gives to the mathematics. – Oliver Heaviside, *Collected Works*, Vol II, p.4

Throughout his "career", Heaviside released 3 volumes of work entitled *Electromagnetic Theory*, which was really just a collection of his papers.

Heaviside shunned all honours, brushing aside his honorary doctorate from the University of Göttingen and even refusing to accept the medal associated with his election as a Fellow of the Royal Society, in 1891.

In 1902, Heaviside wrote an article for the *Encyclopedia Britannica* entitled *The theory of electric telegraphy*. Apart from developing the wave propagation theory of telegraphy, he extended his essay to include "wireless" telegraphy, and explained how the remarkable success of Marconi transmitting from Ireland to Newfoundland might be due to the presence of a permanently conducting upper layer in the atmosphere. This supposed layer was referred to as the "Heaviside layer", which was directly detected by Edward Appleton and M.A.F. Barnett in the mid-1920s. Today we merely call it the "ionosphere".

Heaviside spent much of his life being bitter at those who didn't recognise his genius – he had disdain for those that could not accept his mathematics without formal proof, and he felt betrayed and cheated by the scientific community who often ignored his results or used them later without recognising his prior work. It was with much bitterness that he eventually retired and lived out the rest of his life in Torquay on a government pension. He withdrew from public and private life, and was taunted by "insolently rude imbeciles". Objects were thrown at his windows and doors and numerous practical tricks were played on him.

Heaviside should be remembered for his vectors, his field theory analyses, his brilliant discovery of the distortionless circuit, his pioneering applied mathematics, and for his wit and humor. – P.J. Nahin

Today, the historical obscurity of Heaviside's work is evident in the fact that his vector analysis and vector formulation of Maxwell's theory have become "basic knowledge". His operational calculus was made obsolete with the 1937 publication of a book by the German mathematician Gustav Doetsch – it showed how, with the Laplace transform, Heaviside's operators could be replaced with a mathematically rigorous and systematic method.

The last five years of Heaviside's life, with both hearing and sight failing, were years of great privation and mystery. He died on 3rd February, 1925.

References

Nahin, P.: *Oliver Heaviside: Sage in Solitude*, IEEE Press, 1988.

15 Amplifier Characteristics

Contents

15.2

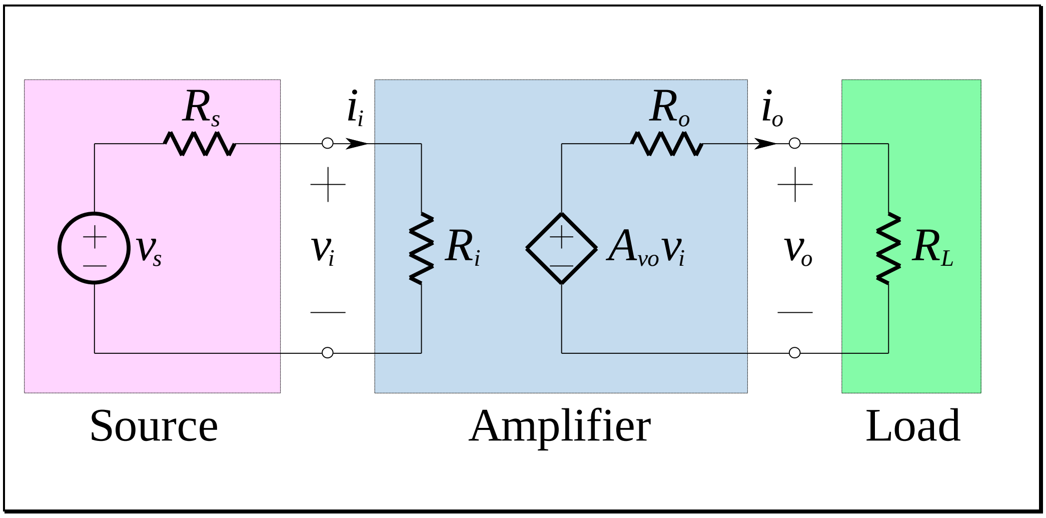
Introduction

Amplifiers are designed to have certain performance characteristics that meet the requirements of a specific application. For example, an amplifier may be required to amplify a very small voltage from a source that has a relatively large internal resistance, such as an antenna. The amplifier would need to have a very large input impedance, and a large gain over a wide band of frequencies. An audio amplifier may be required to amplify signals in the audio range with minimal distortion and provide a low impedance output to drive a speaker.

Since each amplifier application is different, we need to study the various characteristics of amplifiers that will enable us to intelligently design or choose a suitable amplifier. The characteristics we will look at are: voltage gain, current gain, power gain, power efficiency, input impedance, output impedance, frequency response, linear waveform distortion, pulse response and harmonic distortion.

15.1 Amplifier Performance

The model of a voltage amplifier is shown below, together with an attached source and load:



A voltage amplifier with source and load attached

Figure 15.1

The input resistance, R_i, of the amplifier is the equivalent resistance seen when looking into the input terminals. Later we will generalise this concept so that the amplifier has an input impedance, $\mathbf{Z}_i$.

The open-circuit voltage gain of the amplifier, A_{vo}, is the value of the VCCS in the amplifier model (if the load is an open circuit, there is no voltage drop across the output resistance, and then $v_o = A_{vo}v_i$).

The output resistance, R_o, of the amplifier models the fact that when the load draws a current there is a reduction in the voltage at the output terminals.

15.1.1 Voltage Gain

The voltage gain of the *loaded* amplifier is:

$$A_v = \frac{v_o}{v_i} = A_{vo}\frac{R_L}{R_L + R_o} \tag{15.1}$$

and the overall gain between the load and source is:

$$\begin{aligned} G_v &= \frac{v_o}{v_s} = \frac{R_i}{R_i + R_s} A_v \\ &= \frac{R_i}{R_i + R_s} A_{vo} \frac{R_L}{R_L + R_o} \\ &= \left\{\begin{matrix}\text{voltage divider} \\ \text{due to source}\end{matrix}\right\}\left\{\begin{matrix}\text{open-circuit} \\ \text{amplifier gain}\end{matrix}\right\}\left\{\begin{matrix}\text{voltage divider} \\ \text{due to load}\end{matrix}\right\} \end{aligned} \tag{15.2}$$

15.1.2 Current Gain

The *current gain* A_i of a *loaded* amplifier is the ratio of the output current to the input current:

$$A_i = \frac{i_o}{i_i} = \frac{v_o/R_L}{v_i/R_i} = A_v \frac{R_i}{R_L} \tag{15.3}$$

15.1.3 Power Gain

The power delivered to the input terminals by the signal source is called the input power P_i, and the power delivered to the load by the amplifier is the output power P_o. The *power gain* A_P of an amplifier is the ratio of the output power to the input power:

$$A_P = \frac{v_o i_o}{v_i i_i} = A_v A_i \tag{15.4}$$

15.2 Cascaded Amplifiers

When we cascade amplifiers, we need to take into account "loading" effects. Loading refers to the fact that the input impedance of the second amplifier in the cascade will appear as a load impedance to the first amplifier, thus changing the gain.

Cascade connection of two amplifiers

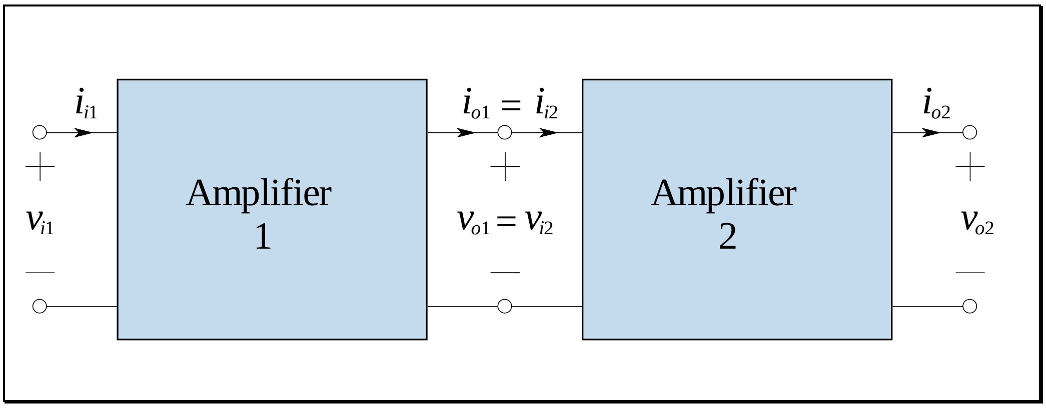


Figure 15.2

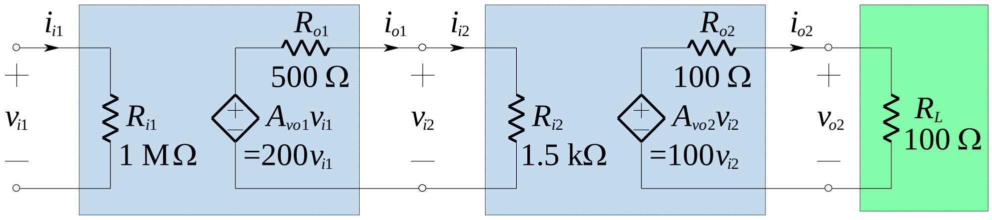
In the figure above, two voltage amplifiers are cascaded. The overall voltage gain is determined be considering the gains of each amplifier, *under loaded conditions*. Since $v_{i2} = v_{o1}$, we have for the overall gain:

$$A_v = \frac{v_{o2}}{v_{i1}} = \frac{v_{o1}}{v_{i1}} \cdot \frac{v_{o2}}{v_{o1}} = \frac{v_{o1}}{v_{i1}} \cdot \frac{v_{o2}}{v_{i2}} = A_{v1} A_{v2} \qquad (15.5)$$

In this equation A_{v1} and A_{v2} represent the voltage gains of the individual stages of the cascade, the values of which are calculated using the specific loading conditions of the cascade.

EXAMPLE 15.1 Cascaded Amplifiers

Consider the cascaded amplifiers below:



The voltage gain of the first amplifier, which is loaded by the input resistance of the second amplifier, is:

$$A_{v1} = A_{vo1} \frac{R_{i2}}{R_{i2} + R_{o1}} = 200 \cdot \frac{1500}{2000} = 150$$

The voltage gain of the second amplifier, which has a load resistance attached, is:

$$A_{v2} = A_{vo2} \frac{R_L}{R_L + R_{o2}} = 100 \cdot \frac{100}{200} = 50$$

Thus, the overall voltage gain of the cascaded amplifiers is:

$$A_v = A_{v1} A_{v2} = 150 \cdot 50 = 7500$$

The current gain of the first amplifier is:

$$A_{i1} = A_{v1} \frac{R_{i1}}{R_{i2}} = 150 \cdot \frac{10^6}{1500} = 10^5$$

and the current gain of the second amplifier is:

$$A_{i2} = A_{v2} \frac{R_{i2}}{R_L} = 50 \cdot \frac{1500}{100} = 750$$

Thus, the overall current gain of the cascaded amplifiers is:

$$A_i = A_{i1} A_{i2} = 10^5 \cdot 750 = 7.5 \times 10^7$$

We can derive an equivalent circuit, or model, of the two cascaded amplifiers, by considering the input resistance, output resistance, and overall open-circuit voltage gain. Looking in from the left, we have:

$$R_i = R_{i1} = 1\,\mathrm{M\Omega}$$

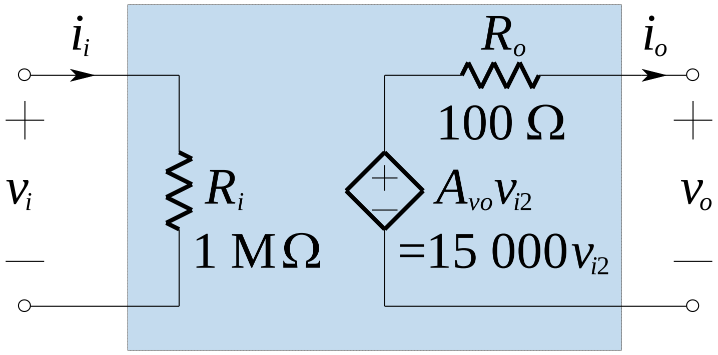
whilst looking in from the right, we have:

$$R_o = R_{o2} = 100\ \Omega$$

The overall *open-circuit* voltage gain of the cascaded amplifiers (i.e. with the first amplifier being loaded by the input resistance of the second amplifier, but with the second amplifier with no load) is:

$$A_{vo} = A_{v1}A_{vo2} = 150 \cdot 100 = 1.5 \times 10^4$$

We then have a simplified model for the cascaded amplifiers of:



15.3 Power Supplies and Efficiency

Power is supplied to the internal circuitry of amplifiers from power supplies. There are two voltage sources to a bipolar amplifier and the total power supplied is the sum of the powers supplied by each voltage source.

The power supply delivers power to the amplifier from several DC voltage sources

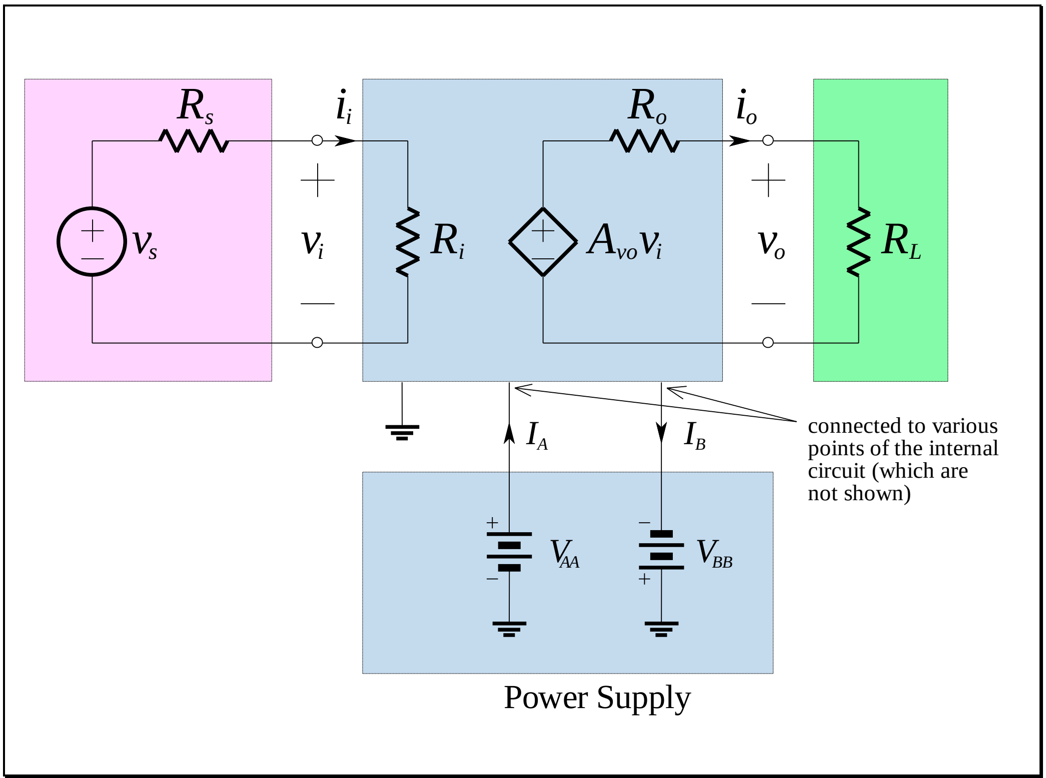


Figure 15.3

In the figure above, the total power supplied to the amplifier is:

$$P_s = V_{AA}I_A + V_{BB}I_B \tag{15.6}$$

It is quite typical for amplifiers to take power without a zero volt reference or common – op-amps are a prime example. In these cases the designer of the internal amplifier circuit has arranged for the output to be exactly half way between the voltage supplies under zero input conditions. If the bipolar supplies are symmetric (i.e. they are equal in magnitude), then the output will be 0 V for no input, with respect to the power supply common. Note that for an op-amp, which doesn't have a GND pin, the single-ended voltage output is normally taken with respect to the power supply common, and symmetric supplies are assumed.

Since amplifiers typically provide power gain, this power must be taken from the power supplies for conservation of energy to hold.

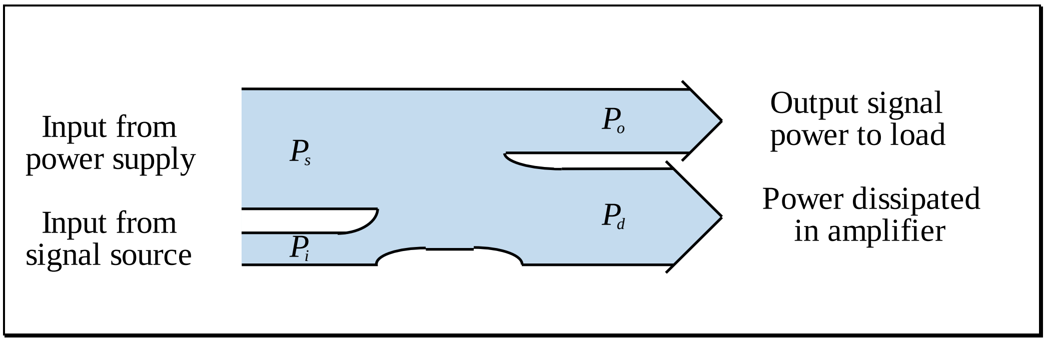


Illustration of power flow

Figure 15.4

The principle of Conservation of Energy requires that:

$$
\begin{aligned}
P_i + P_s &= P_o + P_d \\
\begin{Bmatrix} \text{power} \\ \text{supplied} \end{Bmatrix} &= \begin{Bmatrix} \text{power} \\ \text{absorbed} \end{Bmatrix}
\end{aligned}
\qquad (15.7)
$$

15.3.1 Efficiency

The efficiency η of an amplifier is the proportion of power supplied that is converted into output power:

$$\eta = \frac{P_o}{P_s} \times 100\% \qquad (15.8)$$

Note that P_i is often insignificant compared to the power supply power, so it does not appear in the above equation.

15.4 Amplifier Models

As far as the external terminal characteristics are concerned, the type of circuit model used to represent the amplifier is irrelevant. We used this idea to convert from a practical voltage source to a practical current source via a source transformation. For an amplifier, which is modelled with dependent sources, there are four types of equivalent model – each with a different dependent source. Some models are more suited to particular amplifiers or analysis techniques, so it is important to be aware of the different models and be able to convert between them.

15.4.1 Voltage Amplifier

We have already seen the model for a voltage amplifier – it is a practical model for devices such as the op-amp (since the op-amp is designed to amplify voltage) as well as for circuits built from op-amps that are designed to amplify voltage, such as the inverting and noninverting op-amp configurations.

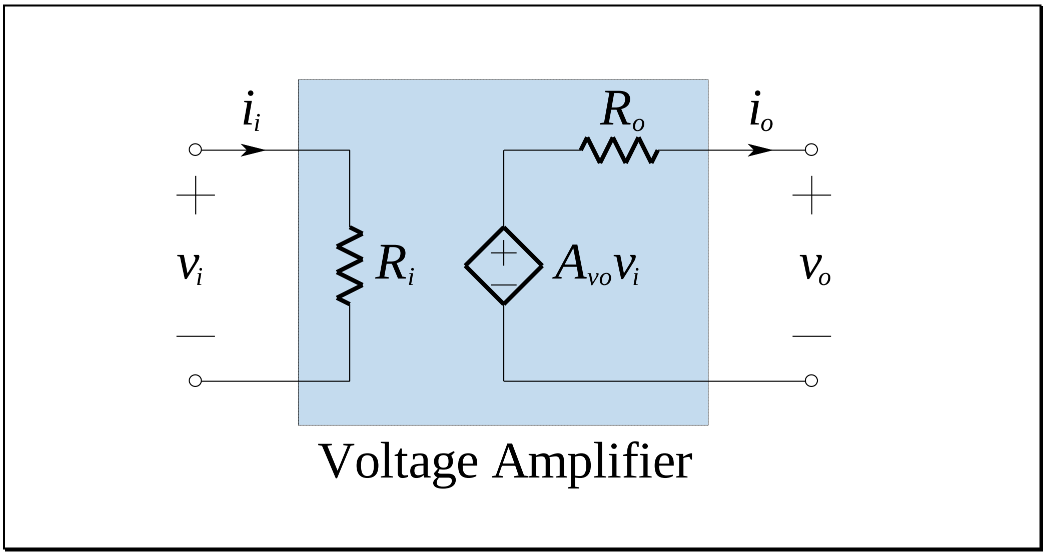


Figure 15.5

Note that the output side of the amplifier has a series connection – it is like the Thévenin equivalent of a linear circuit, except it uses a dependent source to model the amplifier behaviour. If we do KVL at the output we get:

$$v_o = A_{vo}v_i - R_o i_o \tag{15.9}$$

15.4.2 Current Amplifier

An alternative but equivalent model of the voltage amplifier is the *current amplifier*, shown below:

Current amplifier model

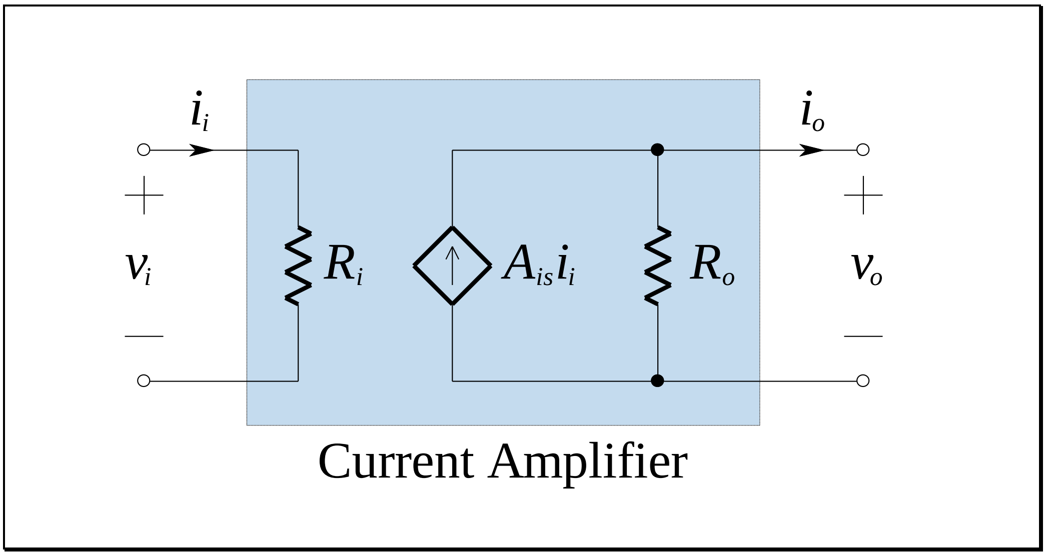


Figure 15.6

Note that the output side of the amplifier has a parallel connection – it is like the Norton equivalent of a linear circuit, except it uses a dependent source to model the amplifier behaviour. If we do KCL at the output we get:

$$i_o = A_{is}i_i - v_o/R_o \tag{15.10}$$

Rearranging in terms of the output voltage, and noting that $i_i = v_i/R_i$, we have:

$$v_o = (A_{is}R_o/R_i)v_i - R_o i_o \tag{15.11}$$

Comparison with the voltage amplifier's output equation, Eq. (15.9), reveals that we can convert from one to the other using:

$$A_{vo} = A_{is}\frac{R_o}{R_i} \tag{15.12}$$

The values of the input and output resistors are the same in each model.

15.4.3 Transconductance Amplifier

Another model for an amplifier is the *transconductance amplifier*:

Transconductance amplifier model

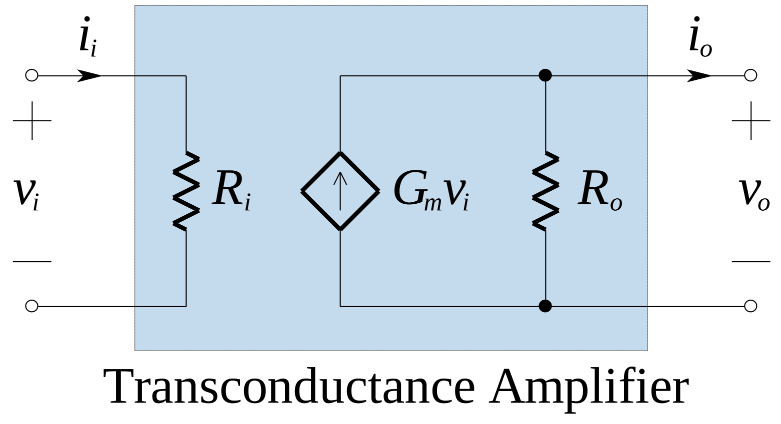


Figure 15.7

Here the controlled source is a voltage-controlled current source. If we do KCL at the output we get:

$$i_o = G_m v_i - v_o / R_o \tag{15.13}$$

Rearranging in terms of the output voltage, we have:

$$v_o = G_m R_o v_i - R_o i_o \tag{15.14}$$

Comparison with the voltage amplifier's output equation reveals that we can convert from one to the other using:

$$A_{vo} = G_m R_o \tag{15.15}$$

The values of the input and output resistors are the same in each model.

Transconductance models are often used in the analysis of amplifier circuits that use a transistor known as a metal-oxide semiconductor field-effect transistor (MOSFET). This type of transistor is controlled by an electric field (hence a voltage) and can behave like a voltage-controlled current source.

15.4.4 Transresistance Amplifier

The fourth type of model for an amplifier is the *transresistance amplifier*:

Transresistance amplifier model

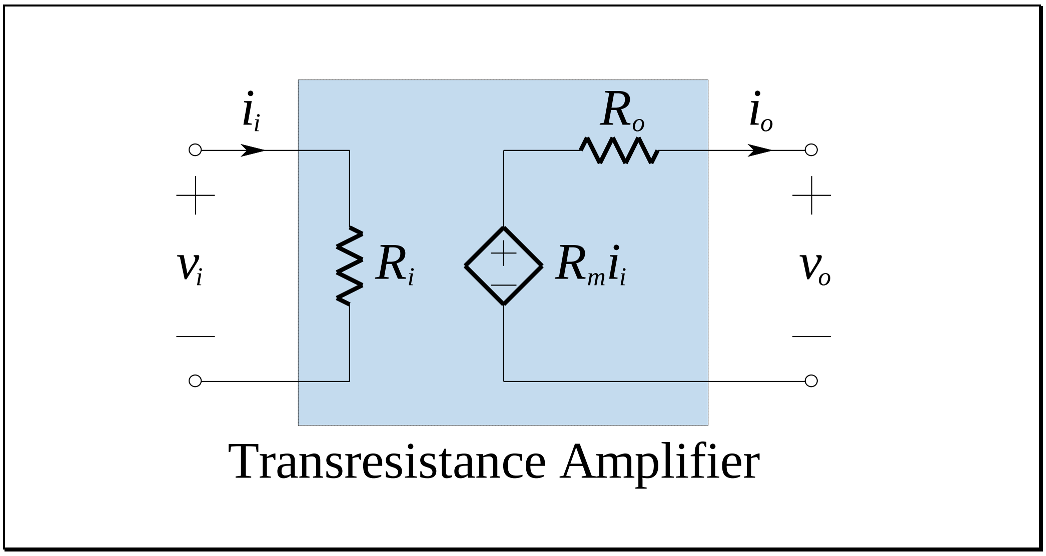


Figure 15.8

Here the controlled source is a current-controlled voltage source. If we do KVL at the output we get:

$$v_o = R_m i_i - R_o i_o \tag{15.16}$$

Noting that $i_i = v_i / R_i$, we have:

$$v_o = \left(R_m / R_i\right) v_i - R_o i_o \tag{15.17}$$

Comparison with the voltage amplifier's output equation reveals that we can convert from one to the other using:

$$A_{vo} = \frac{R_m}{R_i} \tag{15.18}$$

The values of the input and output resistors are the same in each model.

Transresistance amplifiers are often used to convert a current into a voltage, such as when amplifying the output of photo multiplier tubes, accelerometers, photo detectors and other types of sensors.

15.5 Amplifier Impedances

Amplifiers are designed for specific applications which place demands on their characteristics. For example, suppose we wish to build a voltage amplifier for a very weak signal source, such as the output of an electret microphone or a TV antenna:

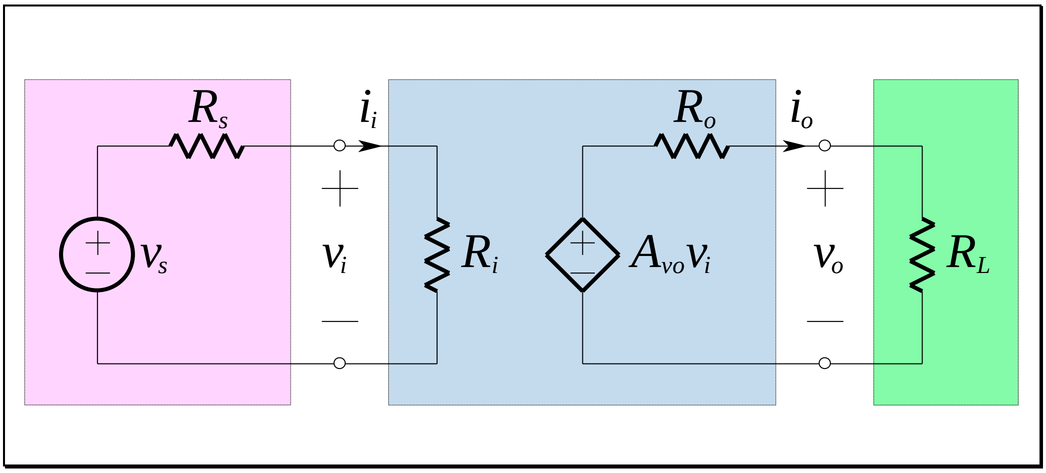


Figure 15.9

In the figure above, the source has been modelled using a Thévenin equivalent circuit. Since the application calls for the source voltage to be amplified, we want to minimise the voltage drop occurring in the source due to the current drawn by the input of the amplifier. The voltage amplifier therefore needs to present a very high impedance to the source.

Since the overall gain between the load and source is:

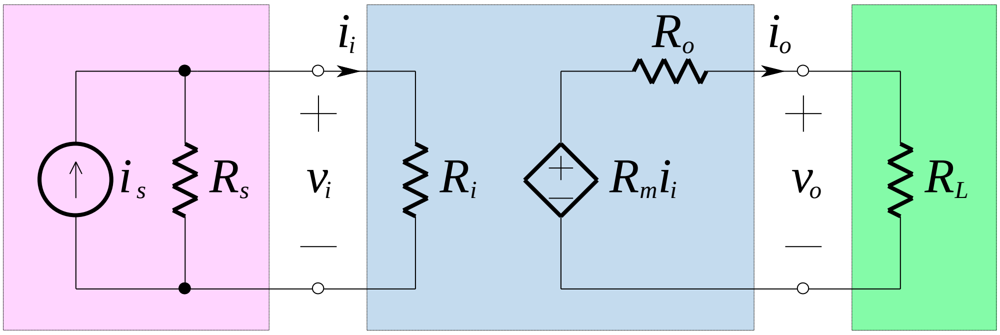
$$G_v = \frac{v_o}{v_s} = \frac{R_i}{R_i + R_s} A_{vo} \frac{R_L}{R_L + R_o} \tag{15.19}$$

we can see that the ideal case would be for the amplifier input impedance to be infinite, and for the output impedance to be zero – we would then achieve the maximum gain $G_v = A_{vo}$.

Conversely, applications that call for the internal current produced by the source to be amplified need to present a low impedance to the source, and as large an impedance as possible to the load.

EXAMPLE 15.2 Transresistance Amplifier

Consider a current source connected to a transresistance amplifier. A model of such a situation is shown below:



The expression for the overall gain is:

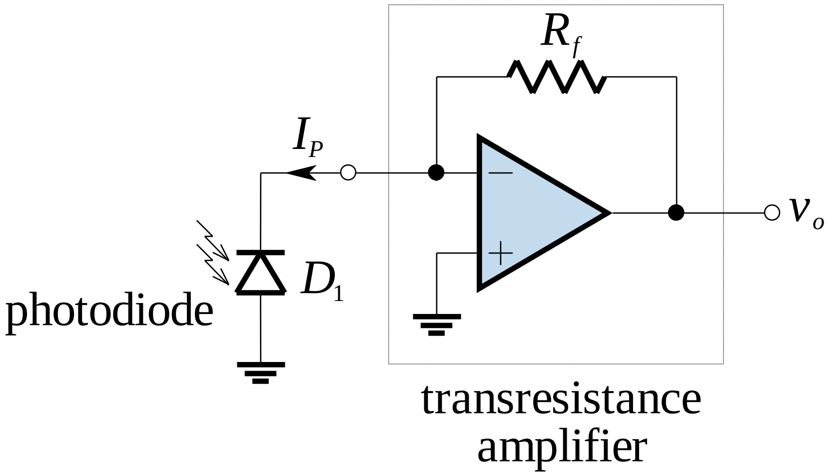
$$\frac{v_o}{i_s} = \frac{R_s}{R_i + R_s} R_m \frac{R_L}{R_L + R_o}$$

$$= \begin{Bmatrix}\text{current divider}\\ \text{due to source}\end{Bmatrix}\begin{Bmatrix}\text{open-circuit}\\ \text{amplifier gain}\end{Bmatrix}\begin{Bmatrix}\text{voltage divider}\\ \text{due to load}\end{Bmatrix}$$

To maximise the gain:

- the input resistance of the amplifier, R_i, should be as **small** as possible (ideally zero).
- the output resistance of the amplifier, R_o, should be as **small** as possible (ideally zero).

A practical transresistance amplifier that achieves these goals is shown below.



In this example the source is a photodiode (which acts like a current source), and the requirement of the amplifier is to convert the low-level current to a voltage. Photodiodes are often used for accurate measurement of light intensity in science and industry. They are also widely used in various medical applications, such as detectors for computed tomography, instruments to analyze medical samples (immunoassay), and pulse oximeters. The photodiode is designed to operate in reverse bias, where the current is almost linearly proportional to the light intensity.

Assuming an ideal op-amp, the output voltage is given by:

$$v_o = R_f I_P$$

Due to the virtual short-circuit across the op-amp input terminals, the input resistance of the amplifier is effectively zero. Since the output is taken from the op-amp output, the output resistance of the amplifier is effectively zero.

Thus, the circuit satisfies the criteria to make a good transresistance amplifier. It is also known as a current-to-voltage converter.

15.6 Frequency Response

If we apply a variable-frequency sinusoidal input signal to an amplifier, we will find that gain is a function of frequency. Moreover, the amplifier affects the phase as well as the amplitude of the sinusoid. For a voltage amplifier, we define the (complex) gain to be the ratio of the output voltage phasor to the input voltage phasor:

$$\mathbf{A}_v = \frac{\mathbf{V}_o}{\mathbf{V}_i} \tag{15.20}$$

Frequency response of an amplifier is the output phasor over the input phasor

EXAMPLE 15.3 Complex Gain

The input voltage to a certain amplifier is:

$$v_i(t) = 0.1\cos(2000\pi t - 30°)$$

and the output voltage is:

$$v_o(t) = 10\cos(2000\pi t + 15°)$$

To find the gain as a complex number, we convert the voltages into their respective phasors:

$$\mathbf{V}_i = 0.1\angle -30° \quad \text{and} \quad \mathbf{V}_o = 10\angle 15°$$

The complex gain is then:

$$\mathbf{A}_v = \frac{\mathbf{V}_o}{\mathbf{V}_i} = \frac{10\angle 15°}{0.1\angle -30°} = 100\angle 45°$$

The meaning of this complex voltage gain is that the output signal is 100 times larger in amplitude than the input signal. Furthermore, the output signal is phase shifted by 45° relative to the input signal.

15.6.1 AC Coupling and Direct Coupling

Some amplifier stages are "coupled" together using capacitors which block DC. Thus, only AC signals are amplified. Such amplifiers are known as *AC coupled* amplifiers. Examples include audio amplifiers and electro-cardiographs.

An AC coupled amplifier

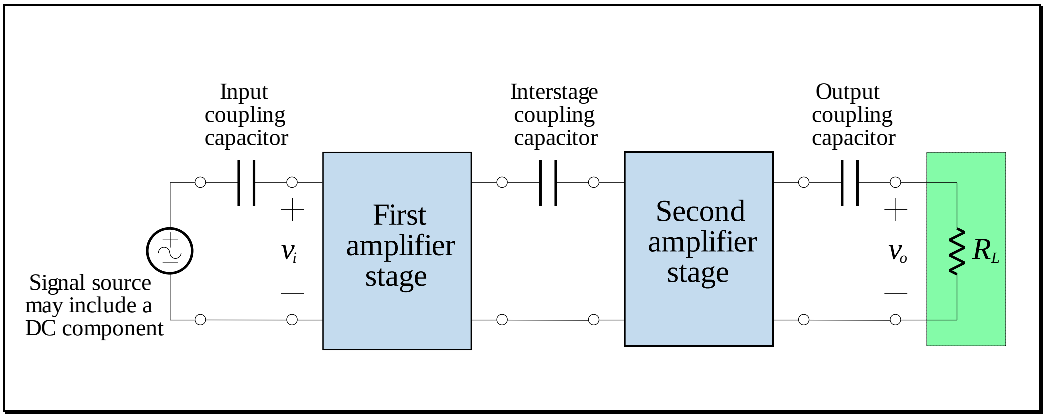


Figure 14.1

The capacitors used to couple the stages together are known as *coupling* capacitors or *DC blocking* capacitors. These coupling capacitors may be external to the amplifier or may be internal to the amplifier. The most common reason why amplifiers are AC coupled is so that the internal DC biasing arrangements of the transistors in the amplifier are undisturbed by any DC component of the input signal, and because the output AC signal is usually superimposed upon an unwanted DC bias signal.

Amplifiers that provide gain all the way down to DC are known as *DC coupled* or *direct coupled* amplifiers. Examples are op-amps and other integrated circuit amplifiers where the coupling capacitors cannot be fabricated in integrated form.

If we plot the *magnitude* of the gain of a typical amplifier versus frequency, a plot such as the one shown below results:

Gain magnitude versus frequency for AC and DC coupled amplifiers

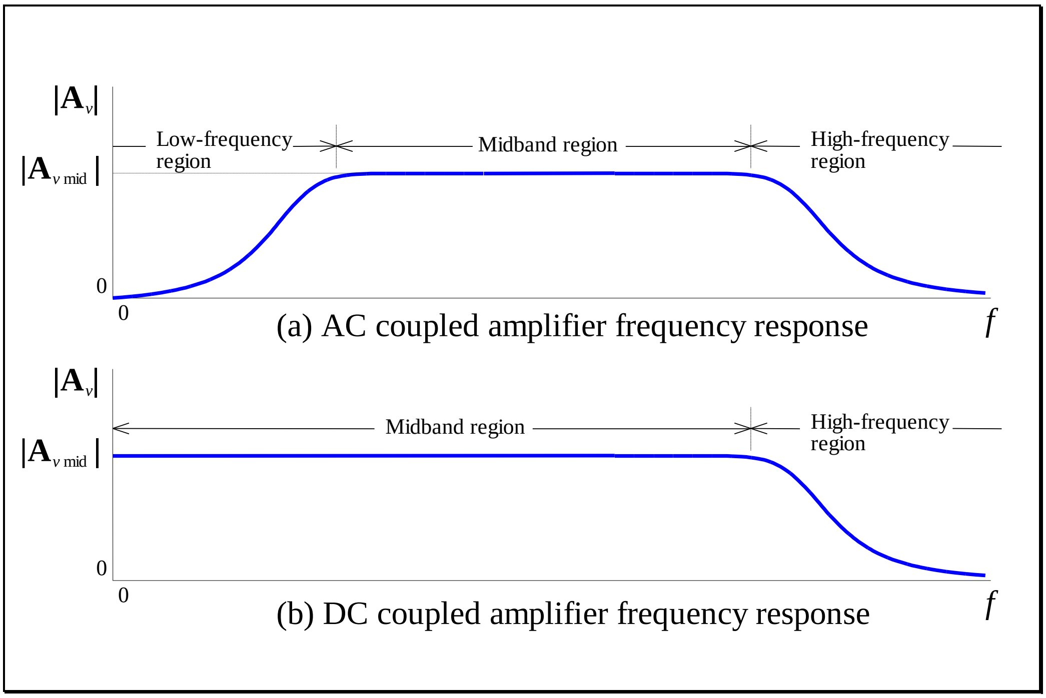


Figure 14.2

The gain magnitude is constant over a wide range of frequencies known as the *midband region*. At high frequencies the gain of an amplifier inevitably falls off due to the intrinsic capacitances of the internal transistors, which appear more and more like short-circuits. For AC coupled amplifiers the gain drops off at low frequencies and is eventually zero at DC due to the coupling capacitors.

15.6.2 Half-Power Frequencies and Bandwidth

We usually specify the approximate useful frequency range of an amplifier by giving the frequencies for which the voltage (or current) gain magnitude is $1/\sqrt{2}$ times the midband gain magnitude. These are known as the *half-power frequencies* because the output power level is half the value of the midband region. In decibels, the half-power frequencies occur when the gain drops by $20\log_{10}\left(1/\sqrt{2}\right) \approx -3.01\,\text{dB}$ compared to the midband gain.

The bandwidth B of an amplifier is the difference between the upper and lower half-power frequencies.

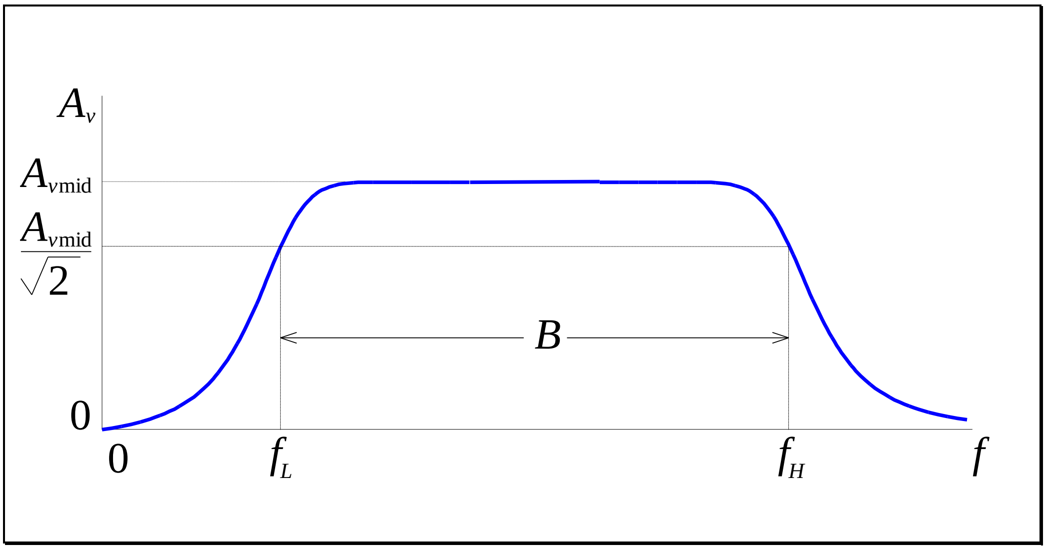


Figure 14.3

In the figure above f_L and f_H are the lower and upper half-power frequencies, respectively. The half-power bandwidth of the amplifier is then given by:

Bandwidth of an amplifier defined

$$B = f_H - f_L \tag{15.21}$$

15.7 Linear Waveform Distortion

Linear waveform distortion will arise in an amplifier if the magnitude response is not perfectly flat (amplitude distortion) or if the phase response is not proportional to frequency (phase distortion). It is termed *linear* waveform distortion because the amplifier is still being a linear device (i.e. obeys the principle of superposition).

15.7.1 Amplitude Distortion

If the gain of an amplifier has a different magnitude for the various frequency components of the input signal, then *amplitude distortion* occurs. Audio systems tend to suffer from amplitude distortion due to the frequency response of the speakers (not the audio amplifier).

15.7.2 Phase Distortion

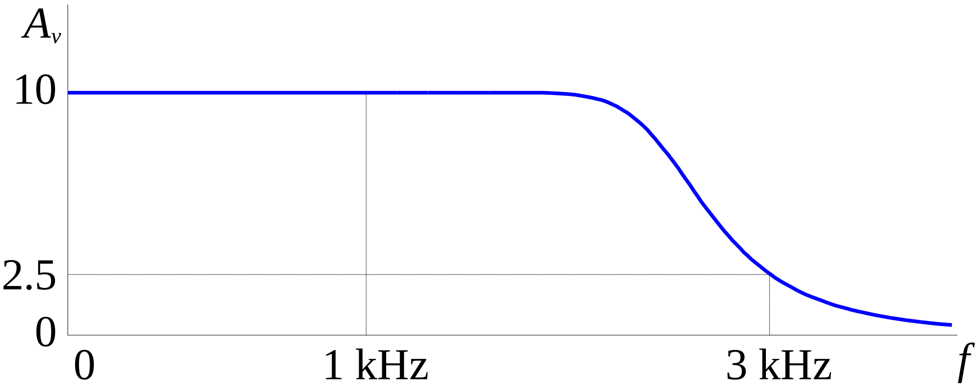
If the phase shift of an amplifier is not proportional to frequency, *phase distortion* occurs, and the output waveform shape is different to the input. On the other hand, if the phase shift of the amplifier is proportional to frequency, the output waveform is a time-shifted version of the input – in this case we do not say that distortion has occurred because the shape of the waveform is unchanged.

EXAMPLE 15.4 Amplitude Distortion

The input voltage to a certain amplifier consists of two frequency components and is given by:

$$v_i(t) = 3\cos(2000\pi t) - 2\cos(6000\pi t)$$

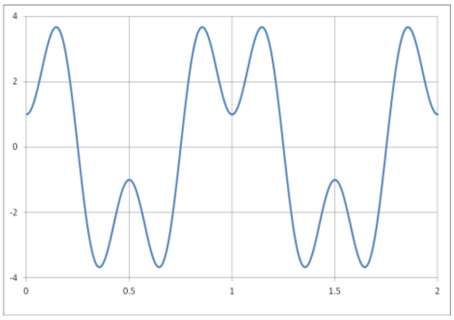
The *magnitude* response of the amplifier is shown below, and the phase response is zero (an ideal case that never occurs in practice):



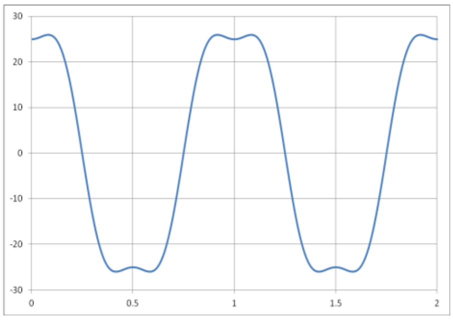
The first term of the input signal (a 1 kHz sinusoid) experiences a gain of 10, whilst the second term of the input signal (a 3 kHz sinusoid) experiences a gain of 2.5. The output signal is then:

$$v_o(t) = 30\cos(2000\pi t) - 5\cos(6000\pi t)$$

Plots of the input and output waveform are shown below:



(a) Input waveform



(b) Output waveform

Notice that the output waveform has a different shape than the input waveform because of amplitude distortion.

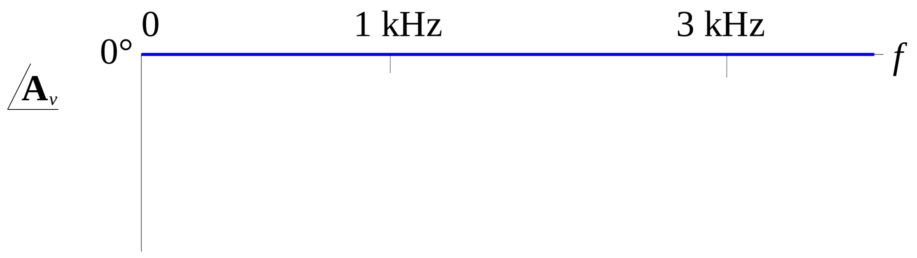
EXAMPLE 15.5 Phase Distortion

Suppose that the input signal given by:

$$v_i(t) = 3\cos(2000\pi t) - \cos(6000\pi t)$$

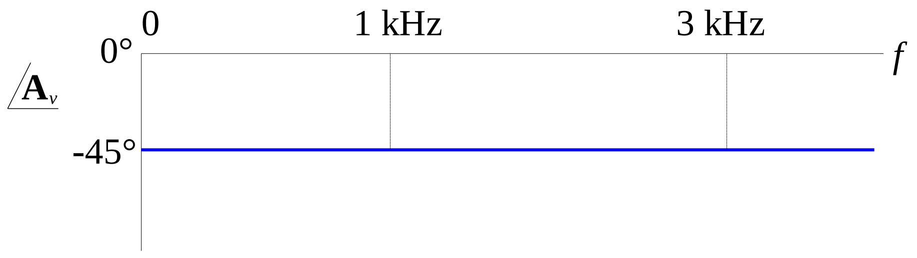
is applied to the inputs of three amplifiers having the *phase* responses shown below:

(a)



(b)

(c)



The amplifiers can be categorised as:

(a) No phase shift

(b) Linear phase versus frequency

(c) Constant phase shift

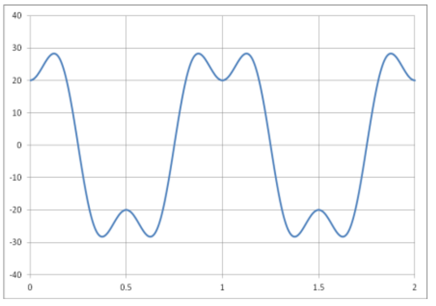
The magnitude response of all three amplifiers is assumed to be 10 for all frequencies.

15.24

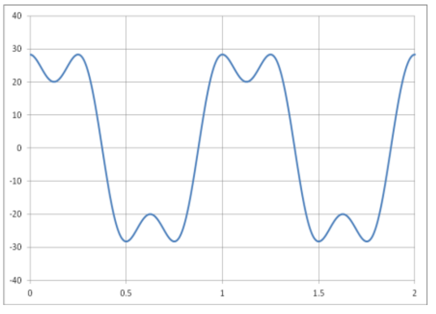
Applying the gain and phase shifts of the individual amplifiers, we find the output signals for the amplifiers to be:

$$v_{oa}(t) = 30\cos(2000\pi t) - 10\cos(6000\pi t)$$
$$v_{ob}(t) = 30\cos(2000\pi t - 45°) - 10\cos(6000\pi t - 135°)$$
$$v_{oc}(t) = 30\cos(2000\pi t - 45°) - 10\cos(6000\pi t - 45°)$$

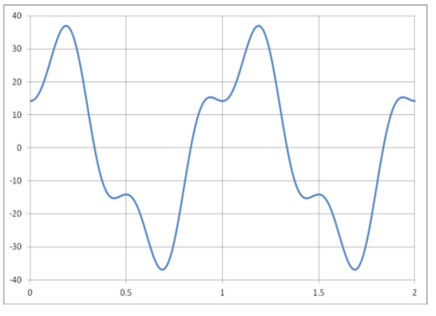
Plots of the output waveforms are shown below:



(a) No phase shift



(b) Linear phase versus frequency

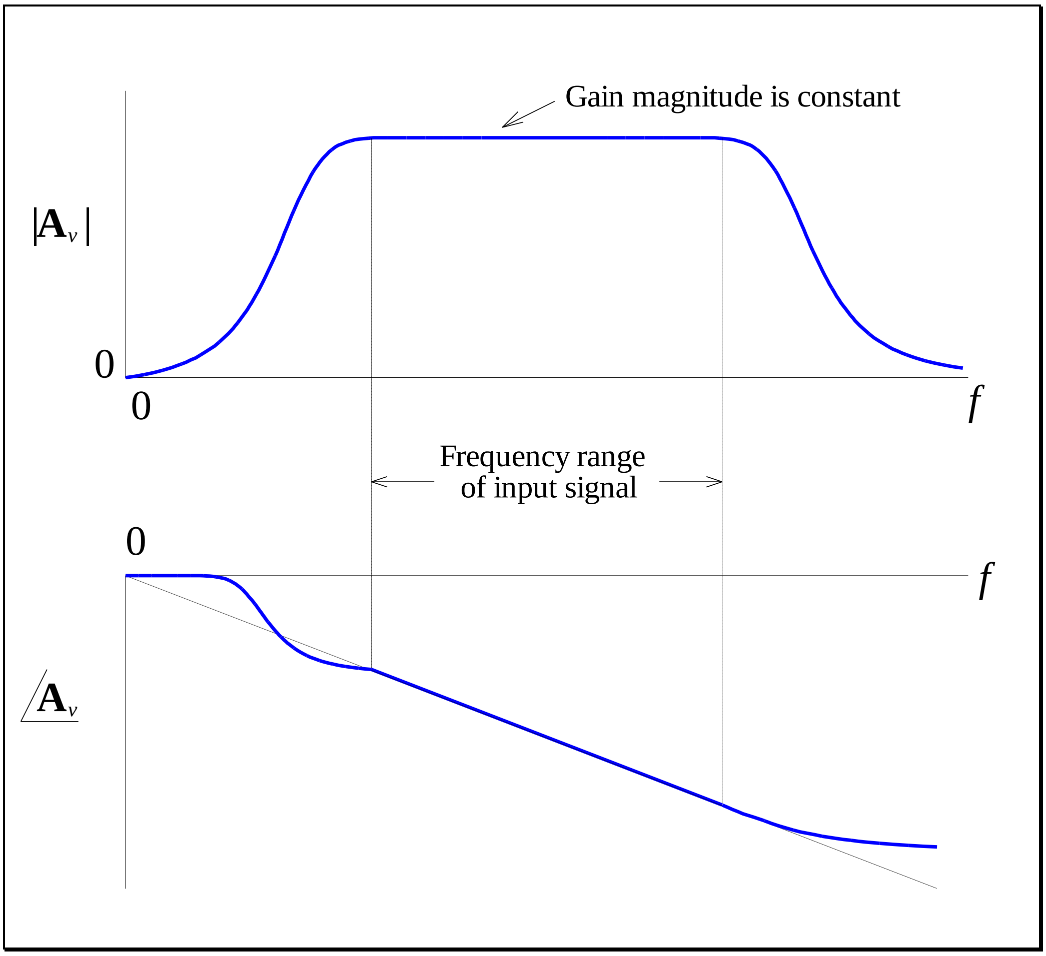


(c) Constant phase shift

Amplifier (a) produces an output waveform identical to the input (just amplified), and amplifier (b) produces an output waveform identical to the input, except for a time delay. Amplifier (c) produces a distorted output waveform.

15.7.3 Distortionless Amplification

To avoid linear waveform distortion, an amplifier should have constant gain magnitude and a phase response that is linear versus frequency for the range of frequencies expected from the input signal. Departure from these requirements outside the frequency range of the input signal does not result in distortion. These requirement for distortionless amplification are shown below:



Requirements on an amplifier's frequency response for distortionless amplification

Figure 14.4

The criteria for distortionless amplification can be generalised to any system that possesses a frequency response. For example, a twisted pair (as commonly found in Cat 5 or Cat 6 Ethernet cabling) can be modelled as a passive lowpass *RC* circuit. With a time constant around $T = RC = 1\,\text{ns}$, the twisted pair is distortionless for frequencies from DC up to around 100 MHz. In this frequency range waveforms will appear at the output undistorted but delayed by about $1\,\text{ns}$. At higher frequencies, we get distortion – this fundamentally limits the rate of "signalling" in digital transmission systems.

15.8 Step Response

Often we need to specify the performance of an amplifier in the time domain. In such cases it is usual to look at the step response of an amplifier. A typical step response is shown below, from which we can define certain quantities:

Time-domain specifications of an amplifier's step response

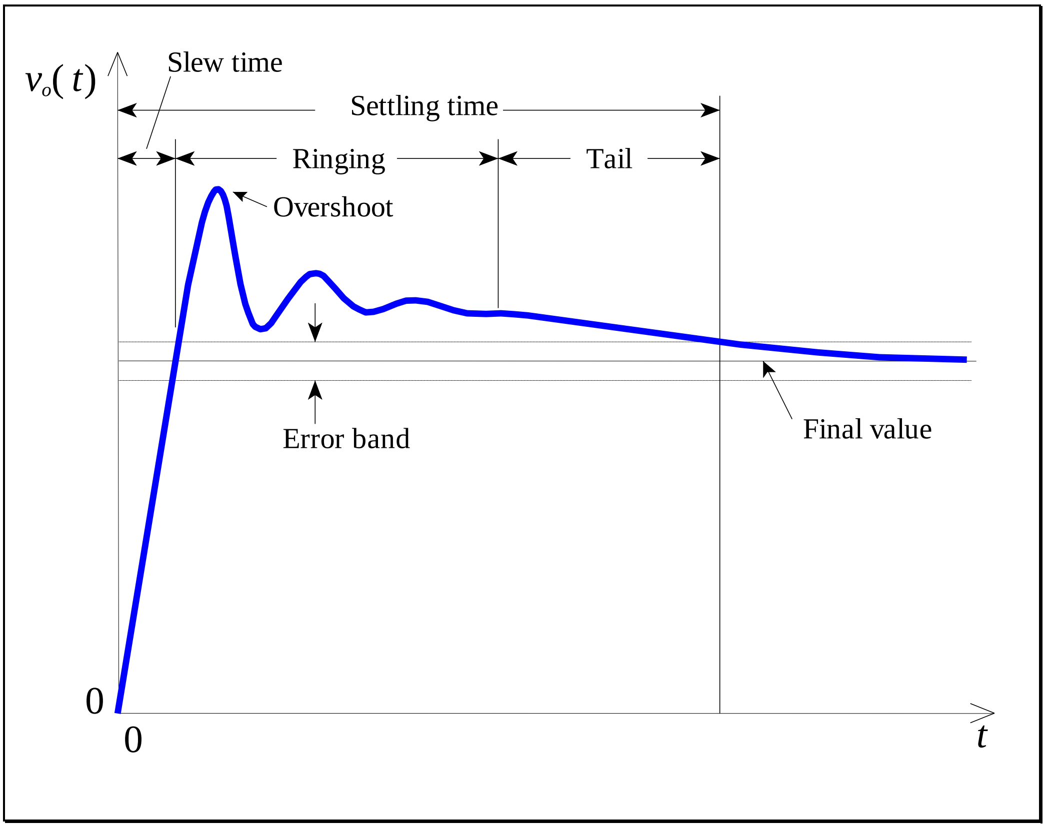


Figure 14.5

The step response displays *overshoot* and *ringing*, and the leading edge is gradual rather than abrupt due to slew rate limiting. The error band is defined as a given percentage of the final value, and the *settling time* is the time it takes for the output to fall within the error band.

If the amplifier is AC coupled, a final value will never be reached and the output response will gradually decay to zero.

15.9 Harmonic Distortion

Real amplifiers have transfer characteristics that depart from straight lines, particularly at large amplitudes. This is shown in the figure below:

Transfer characteristic of a real amplifier showing deviation from the ideal

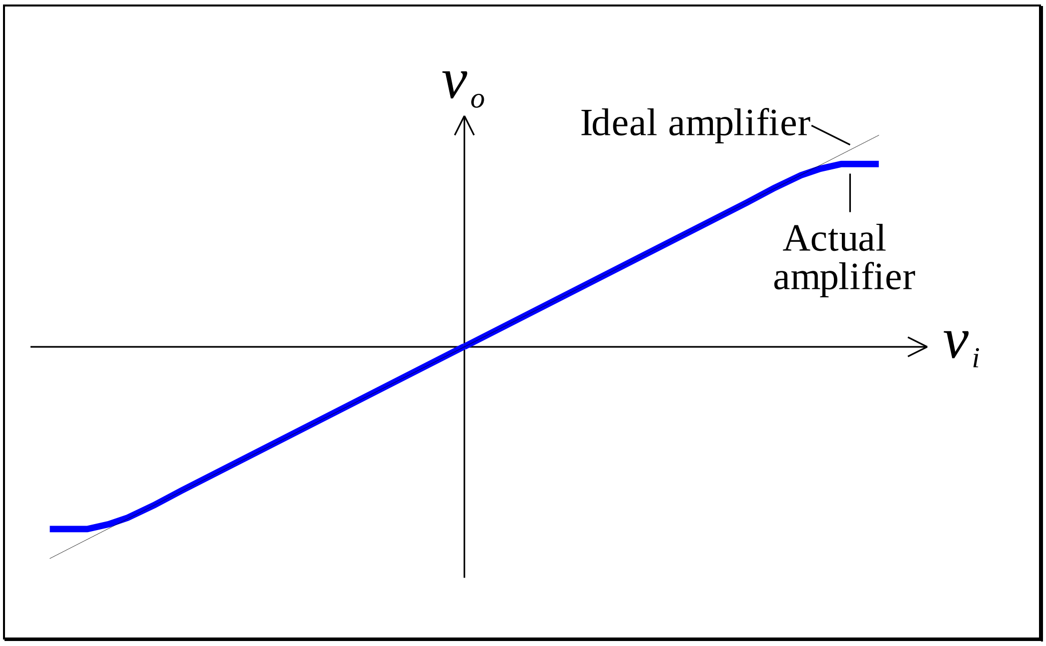


Figure 14.6

Curvature of the transfer characteristic results in an undesirable effect known as *nonlinear distortion*.

The input-output relationship of a nonlinear amplifier can be written as:

$$v_o = A_1 v_i + A_2 v_i^2 + A_3 v_i^3 + A_4 v_i^4 \cdots \tag{15.22}$$

where A_1, A_2, A_3, A_4, and so on, are constants selected so that the equation matches the curvature of the nonlinear transfer characteristic. Usually, the even powered terms in the expression are negligible since most transfer characteristics possess odd symmetry.

Consider the case where the input signal is a sinusoid given by:

$$v_i(t) = V_m \cos(\omega_o t) \tag{15.23}$$

Substituting into Eq. (15.22), applying trigonometric identities for $[\cos(\omega_0 t)]^n$, collecting terms, and defining V_0 to be equal to the sum of all the constant terms, V_1 to be the sum of the coefficients of the terms with frequency ω_0, and so on, we find that:

$$v_o(t) = V_0 + V_1 \cos(\omega_o t) + V_2 \cos(2\omega_o t) + V_3 \cos(3\omega_o t) + \cdots \tag{15.24}$$

The desired output is the $V_1 \cos(\omega_0 t)$ term, which we call the *fundamental* component. The V_0 term represents a shift in the DC level (which does not appear at the load if it is AC coupled). In addition, terms at multiples of the input frequency have resulted from the second and higher power terms of the transfer characteristic. These terms are called *harmonic distortion*. The $2\omega_0$ term is called the *second harmonic*, the $3\omega_0$ term is called the *third harmonic*, and so on.

Harmonic distortion is objectionable in an audio amplifier because it degrades the aesthetic qualities of the sound produced by the loudspeakers.

The n^{th}-*harmonic distortion factor* D_n is defined as the ratio of the amplitude of the n^{th} harmonic to the amplitude of the fundamental:

$$D_n = \frac{V_n}{V_1} \tag{15.25}$$

The *total harmonic distortion* (THD), denoted by *D*, is the ratio of the RMS value of the sum of all the harmonic distortion terms to the RMS value of the fundamental:

Total harmonic distortion defined

$$D = \sqrt{D_2^2 + D_3^2 + D_4^2 + D_5^2 + \cdots} \tag{15.26}$$

We will often find THD expressed as a percentage. A well-designed audio amplifier might have a THD specification of 0.01% at rated power output.

15.10 Summary

- Amplifier performance measures the voltage gain, current gain and power gain under certain source and load conditions.

- In a cascade connection, the output of each amplifier is connected to the input of the next amplifier. Each amplifier presents a load to the preceding stage which must be taken into account when calculating gain.

- Amplifiers are active devices that require power. The efficiency of an amplifier is the percentage of the supply power that is converted into output signal power.

- Several models are useful in characterising amplifiers. They are the voltage amplifier model, the current amplifier model, the transconductance amplifier model and the transresistance amplifier model. The different models are suited to particular amplifiers or configurations.

- The requirements for the **input** impedance of an amplifier are:
  - To sense the open-circuit voltage of a source, an amplifier needs as **large** a $\mathbf{Z}_i$ as possible (ideally ∞).
  - To sense the short-circuit current of a source, an amplifier needs as **small** a $\mathbf{Z}_i$ as possible (ideally 0).

- The requirements for the **output** impedance of an amplifier are:
  - If the load requires constant voltage, an amplifier needs as **small** a $\mathbf{Z}_o$ as possible (ideally 0).
  - If the load requires constant current, an amplifier needs as **large** a $\mathbf{Z}_o$ as possible (ideally ∞).

- The gain of an amplifier is a function of frequency, i.e. the magnitude of the gain and amount of phase shift applied to a single input sinusoid depends on the frequency. Amplifiers can be AC or DC coupled. The bandwidth of an amplifier is a measure of the frequency range for which the gain is greater than half the power of the midband gain.

- Linear distortion can be either amplitude or phase distortion. Amplitude distortion occurs if the gain magnitude is different for various components of the input signal. Phase distortion occurs is the amplifier phase shift is not proportional to frequency. Distortionless amplification requires the amplifier to have constant gain magnitude and a phase response that is linear versus frequency for the range of frequencies expected from the input signal.

- Amplifier step response is characterised by slew time, overshoot, ringing and settling time.

- Nonlinear distortion arises in an amplifier when its transfer characteristic deviates from a straight line. Assuming a sinusoidal input signal, nonlinear distortion causes harmonics to appear in the output. The total harmonic distortion (THD) rating of an amplifier indicates the degree of nonlinear distortion.

15.11 References

Hambley, A.R..: *Electrical Engineering – Principles and Applications*, 5th *Ed*., Pearson, 2011.

16 Frequency Response

Contents

Introduction

The *frequency response* of a circuit, by definition, is the sinusoidal steady-state response of the circuit, as a function of frequency. The frequency response is therefore a description of a circuit's behaviour in the *frequency-domain* – we will use the concepts of voltage and current phasors, and impedance, to facilitate this view.

It turns out that the frequency response of a circuit gives us enough information to *completely characterise*[1] a circuit. Thus, if we know the frequency response of a circuit, then we can determine the circuit's response to any input, not just individual sinusoids.

The frequency-domain view of a circuit turns out to be not only a very powerful mathematical tool, but also a concept that aids engineers to design real systems with real benefits – the fields of telecommunications and automated control rely heavily on 'thinking' and designing in the frequency-domain.

For example, in telecommunications the frequency-domain view leads us to the concept of information being carried by sinusoids – the circuits that generate and receive telecommunication signals can then be said to carry out 'information processing' on these signals. This leads us to the notion of a *filter* – something which retains or rejects a signal based on its frequency. A familiar everyday example of a filter is the circuit that controls the treble and bass in an audio system. The use of circuits to filter signals is one that is fundamental to electrical engineering.

---

[1] A formal link between the frequency response of a circuit and it's time-domain description will be given in later subjects, where the ideas and mathematics related to the Fourier and Laplace transforms are elucidated.

16.1 Frequency Response Function

If a linear time-invariant (LTI) circuit has a sinusoidal input $x(t) = A\cos(\omega_0 t)$, then the steady-state response is also a sinusoid of the same frequency, but with different amplitude and phase, and is given by $y(t) = B\cos(\omega_0 t + \theta)$. Using phasors, the sinusoidal input can be represented by $\mathbf{X} = A\angle 0^0$ and the output response by $\mathbf{Y} = B\angle\theta$:

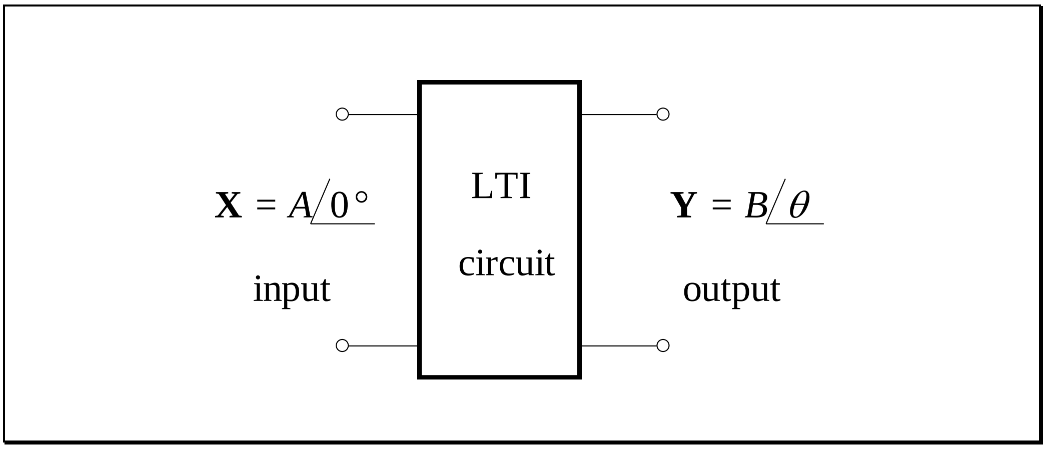


Figure 16.1

We now introduce the *frequency response* function, denoted by $\mathbf{T}(j\omega)$, which is defined to be the ratio of the output phasor to the input phasor of the circuit:

$$\mathbf{T}(j\omega) = \frac{\mathbf{Y}}{\mathbf{X}} \tag{16.1}$$

Note that $\mathbf{T}(j\omega)$ is, in general, a complex function of the complex variable $j\omega$. Thus, like all complex numbers, $\mathbf{T}(j\omega)$ has both a magnitude and an angle – both of which are functions of $j\omega$.

We call $|\mathbf{T}(j\omega)|$ the *magnitude response* of the circuit, and we call $\angle\mathbf{T}(j\omega)$ the *phase response* of the circuit. Thus, the system is completely specified if we know both the magnitude response and the phase response.

The definition above is precisely how we determine the frequency response experimentally – we input a sinusoid and, in the steady-state, measure the magnitude and phase change at the output.

16.2 Frequency Response Representation

The complex function $\mathbf{T}(j\omega)$ can be written using a complex exponential in terms of magnitude and phase:

$$\mathbf{T}(j\omega) = |\mathbf{T}(j\omega)| e^{j\theta(\omega)} \tag{16.2}$$

which is normally written in polar coordinates:

The frequency response in terms of magnitude and phase

$$\mathbf{T}(j\omega) = |\mathbf{T}(j\omega)| \angle \theta(\omega) \tag{16.3}$$

We plot the magnitude and phase of $\mathbf{T}(j\omega)$ as a function of ω or f. We use both linear and logarithmic scales. The phase function is usually plotted in degrees.

If the logarithm (base 10) of the *magnitude* is multiplied by 20, then we have the *gain* of the frequency response in decibels (dB):

The magnitude of the frequency response in dB

$$A(\omega) = 20 \log_{10} |\mathbf{T}(j\omega)| \text{ dB} \tag{16.4}$$

A *negative* gain in decibels is referred to as *attenuation*. For example, -3 dB gain is the same as 3 dB attenuation.

16.3 Determining the Frequency Response from Circuit Analysis

We can get the frequency response of a circuit by undertaking frequency-domain analysis using phasors. In this case, the input sinusoid has a general frequency ω and we are interested in how the output varies as a function of the frequency.

Let us analyse the following simple circuit:

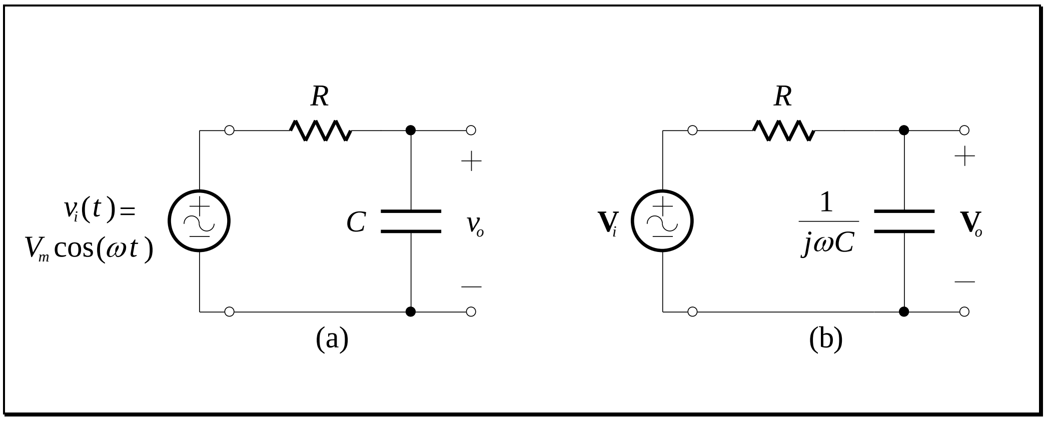


Figure 16.2

The circuit shown in (a) is the real physical circuit, whilst that shown in (b) is its frequency-domain representation. The forcing function is regarded as the input to the circuit, and the voltage across the capacitor is regarded as the output response. We can find the forced response of this circuit using the voltage divider rule:

$$\begin{aligned} \mathbf{V}_o &= \frac{1/j\omega C}{R + 1/j\omega C}\mathbf{V}_i \\ &= \frac{1/RC}{j\omega + 1/RC}\mathbf{V}_i \end{aligned} \qquad (16.5)$$

The frequency response is then the ratio:

$$\mathbf{T}(j\omega) = \frac{\mathbf{V}_o}{\mathbf{V}_i} = \frac{1/RC}{1/RC + j\omega} \qquad (16.6)$$

EXAMPLE 16.1 Frequency Response of an *RC* Circuit

For the *RC* circuit, let $\omega_0 = 1/RC$ so that the frequency response can be written as:

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

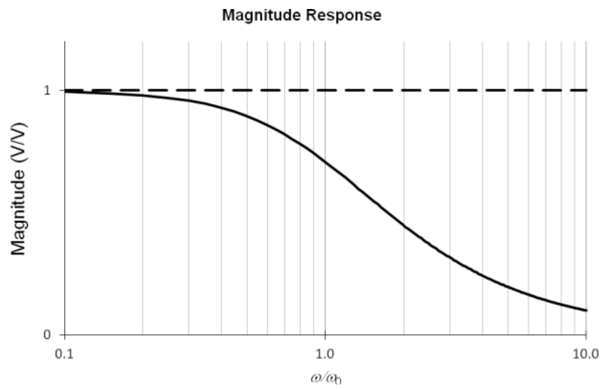
The magnitude function is found directly as:

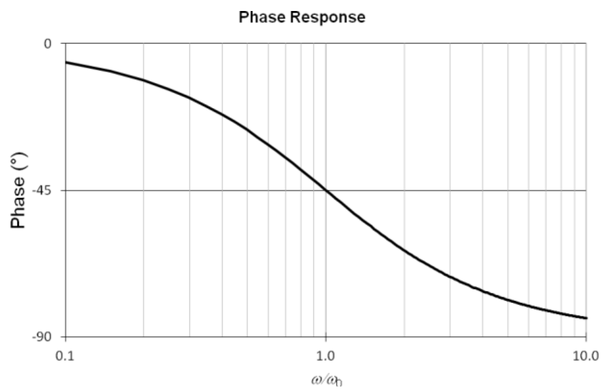
$$\left|\mathbf{T}(j\omega)\right| = \frac{1}{\sqrt{1 + (\omega/\omega_0)^2}}$$

The phase is:

$$\theta(\omega) = -\tan^{-1}\left(\frac{\omega}{\omega_0}\right)$$

These are graphed below, using a normalised log scale for ω:





16.4 Magnitude Responses

A magnitude response is the magnitude of the frequency response, plotted against the frequency of the input. Magnitude responses can be classified according to their particular properties. To look at these properties, we will use linear magnitude versus linear frequency plots. For the simple first-order *RC* circuit that you are so familiar with, the magnitude function has three frequencies of special interest corresponding to these values of $|\mathbf{T}(j\omega)|$:

The magnitude response is the magnitude of the transfer function in the sinusoidal steady state

$$\begin{aligned} |\mathbf{T}(j0)| &= 0 \\ |\mathbf{T}(j\omega_0)| &= \frac{1}{\sqrt{2}} \approx 0.707 \\ |\mathbf{T}(j\infty)| &\to 1 \end{aligned} \tag{16.7}$$

The frequency ω_0 is known as the *half-power frequency*. The plot below shows the complete magnitude response, $|\mathbf{T}(j\omega)|$, as a function of ω, and the circuit that produces it:

A simple lowpass filter

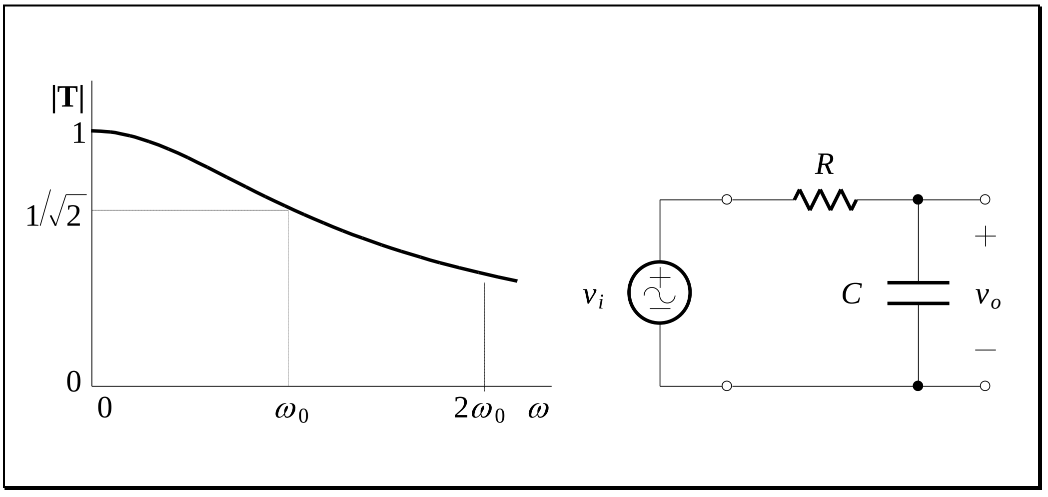


Figure 16.3

An idealisation of the response in Figure 16.3, known as a *brick wall*, and the circuit that produces it are shown below:

An ideal lowpass filter

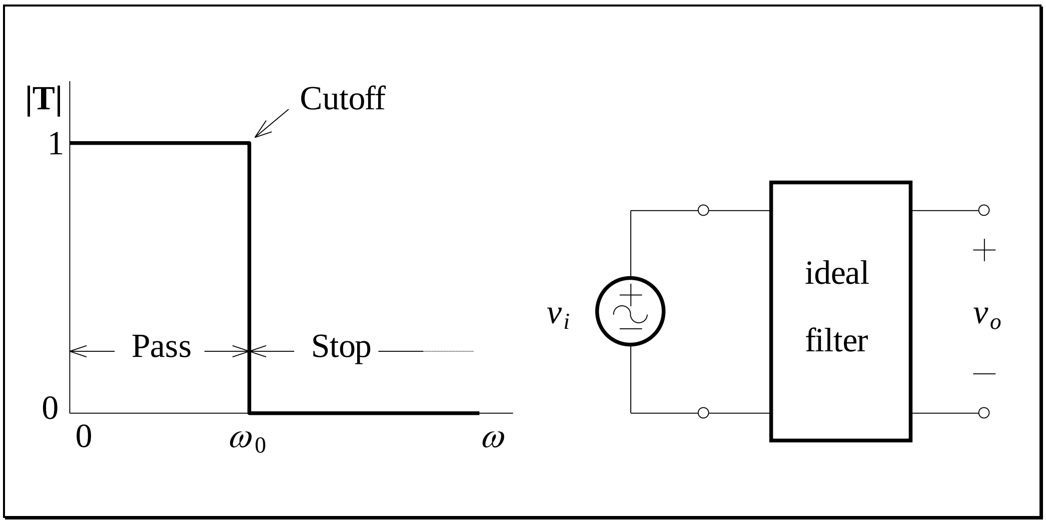


Figure 16.4

Pass and stop bands defined

For the ideal filter, the output voltage remains fixed in amplitude until a critical frequency is reached, called the *cutoff frequency*, ω_0. At that frequency, and for all higher frequencies, the output is zero. The range of frequencies with output is called the *passband*; the range with no output is called the *stopband*. The obvious classification of the filter is a *lowpass* filter.

Even though the response shown in the plot of Figure 16.3 differs from the ideal, it is still known as a lowpass filter, and, by convention, the half-power frequency is taken as the cutoff frequency.

If the positions of the resistor and capacitor in the circuit of Figure 16.2 are interchanged, then the resulting circuit is:

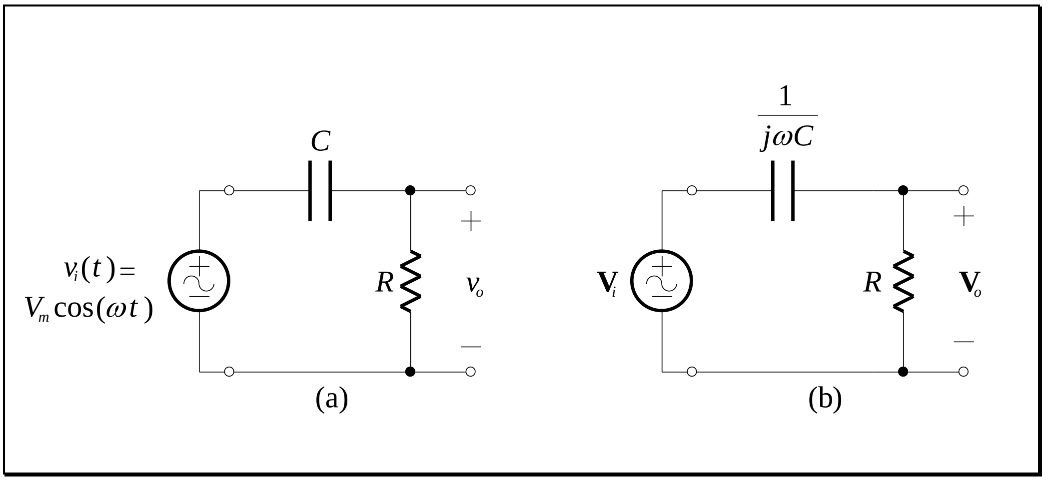


Figure 16.5

Show that the frequency response is:

$$\mathbf{T}(j\omega) = \frac{j\omega}{1/RC + j\omega} \tag{16.8}$$

Letting $\omega_0 = 1/RC$ again, we can write it in the standard form:

$$\mathbf{T}(j\omega) = \frac{j\omega/\omega_0}{1 + j\omega/\omega_0} \tag{16.9}$$

The magnitude function of this equation, at the three frequencies of special interest, is:

$$\begin{aligned} |\mathbf{T}(j0)| &= 0 \\ |\mathbf{T}(j\omega_0)| &= \frac{1}{\sqrt{2}} \approx 0.707 \\ |\mathbf{T}(j\infty)| &\to 1 \end{aligned} \tag{16.10}$$

16.10

A simple highpass filter

The plot below shows the complete magnitude response, $\mathbf{T}(j\omega)$, as a function of ω, and the circuit that produces it:

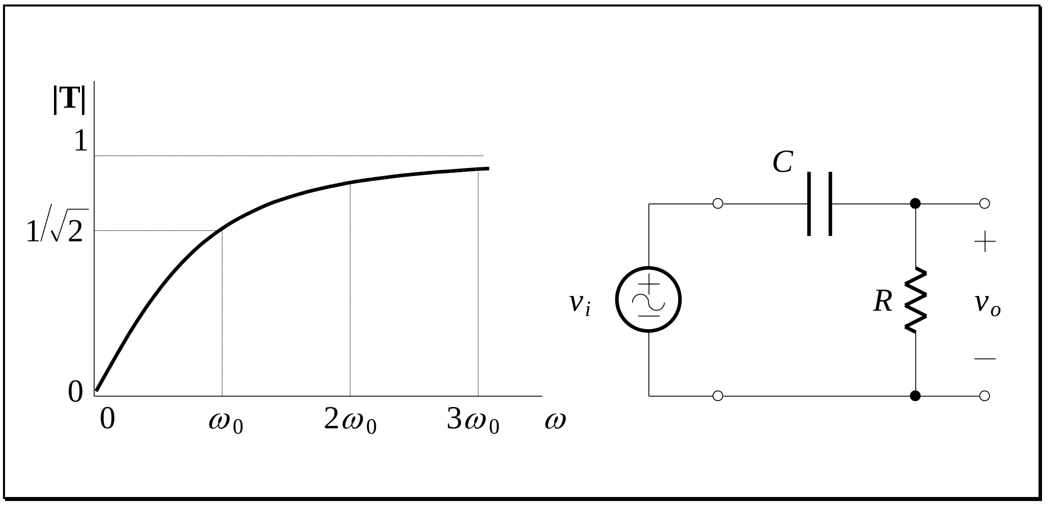


Figure 16.6

This filter is classified as a *highpass filter*. The ideal brick wall highpass filter is shown below:

An ideal highpass filter

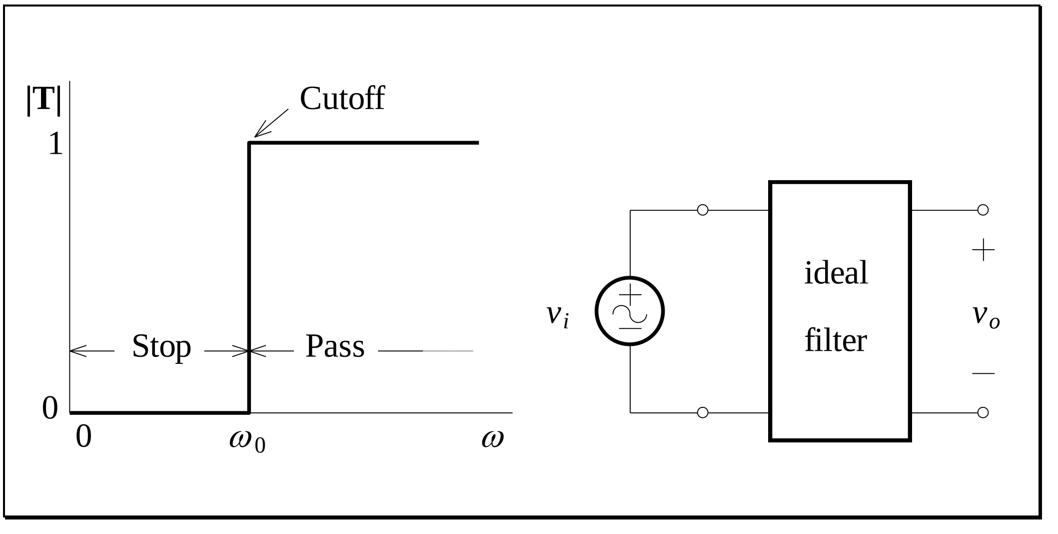


Figure 16.7

The cutoff frequency is ω_0, as it was for the lowpass filter.

16.5 Phase Responses

Like magnitude responses, phase responses are only meaningful when we look at sinusoidal steady-state signals. From Eq. (16.1), a frequency response can be expressed in polar form as:

$$\mathbf{T}(j\omega) = \frac{\mathbf{V}_o}{\mathbf{V}_i} = \frac{|\mathbf{V}_o|\angle\theta}{|\mathbf{V}_i|\angle 0} = |\mathbf{T}|\angle\theta \tag{16.11}$$

Phase response is obtained in the sinusoidal steady-state

where the input is taken as reference (zero phase).

We use the sign of the phase angle to classify circuits. Those giving positive θ are known as *lead* circuits, those giving negative θ as *lag* circuits.

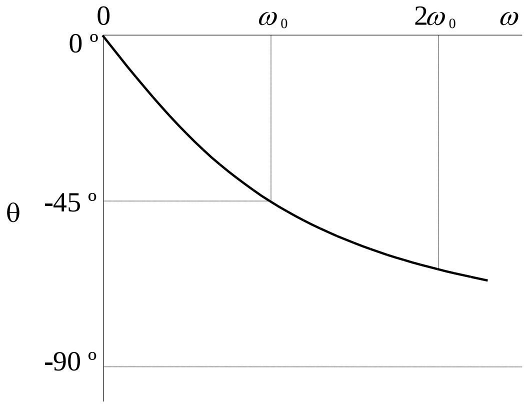
Lead and lag circuits defined

EXAMPLE 16.2 Lagging Phase Response

For the simple *RC* circuit of Figure 16.3, we have already seen that:

$$\theta = -\tan^{-1}(\omega/\omega_0)$$

Since θ is negative for all ω, the circuit is a lag circuit. When $\omega = \omega_0$, $\theta = -\tan^{-1}(1) = -45°$. A complete plot of the phase response is shown below:

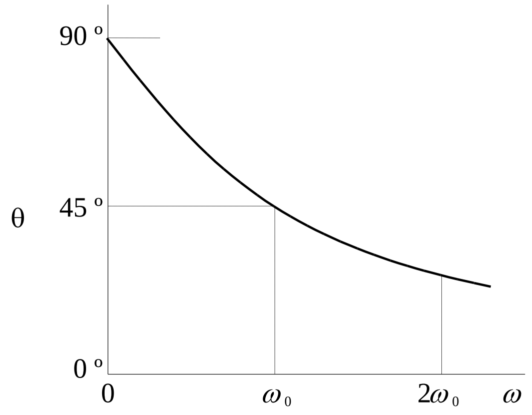


EXAMPLE 16.3 Leading Phase Response

For the simple *RC* circuit of , we can show that the phase is given by:

$$\theta = 90° - \tan^{-1}(\omega/\omega_0)$$

The phase response is shown below:



The angle θ is positive for all ω, and so the circuit is a lead circuit.

16.6 Determining the Frequency Response Experimentally

Experimentally, we can apply a sinusoid of a certain frequency to a circuit, and measure the steady-state output. The output will be a sinusoid of the same frequency, but with a different amplitude and phase. If we graph the magnitude change versus frequency and the phase change versus frequency, we have an experimentally derived frequency response. For example, the frequency response for an op-amp circuit derived experimentally is shown below:

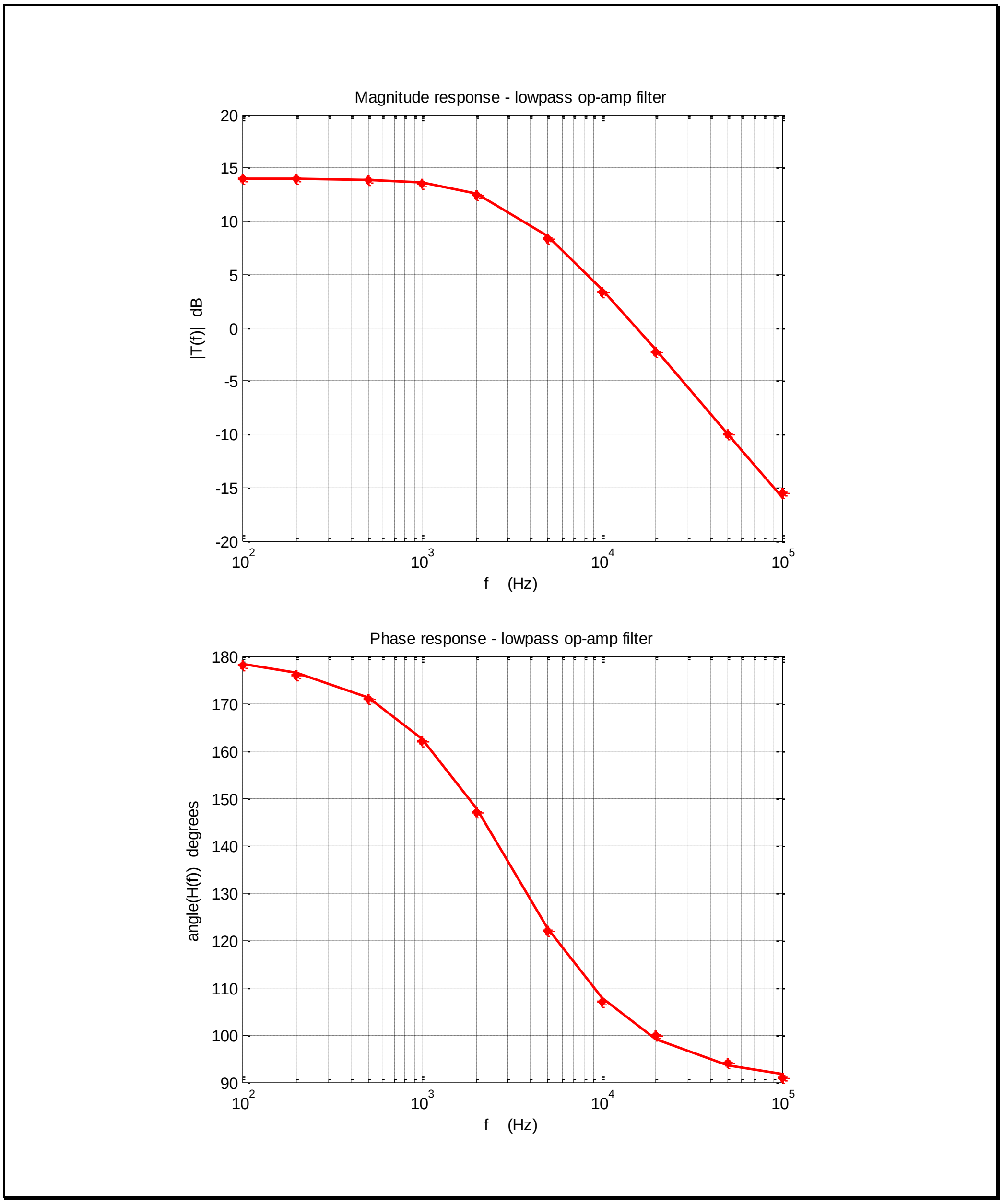


Figure 16.8

16.7 Bode Plots

The advantages of using Bode plots

Bode[*] plots are plots of the magnitude function $A(\omega) = 20\log|\mathbf{T}(j\omega)|$ and the phase function $\theta(\omega)$, where the scale of the frequency variable (usually ω) is logarithmic. The use of logarithmic scales has several desirable properties:

- we can *approximate* a frequency response with straight lines. This is called an *approximate* Bode plot.
- the shape of a Bode plot is preserved if we decide to scale the frequency – this makes design easy.
- we add and subtract individual factors in a frequency response, rather than multiplying and dividing.
- the slope of all asymptotic lines is $\pm 20n\ \text{dB/decade}$ in a magnitude plot, and $\pm n45°\text{/decade}$ in a phase plot, where *n* is any integer.
- by examining a few features of a Bode plot, we can readily determine the frequency response function (for simple systems).

We normally don't deal with equations when drawing Bode plots – we rely on our knowledge of the asymptotic approximations for the handful of factors that go to make up a frequency response.

[*] Dr. Hendrik Bode grew up in Urbana, Illinois, USA, where his name is pronounced *boh-dee*. Purists insist on the original Dutch *boh-dah*. No one uses *bohd*.

EXAMPLE 16.4 Bode Plot of an *RC* Circuit's Frequency Response

For the simple lowpass *RC* circuit, let $\omega_0 = 1/RC$ so that the frequency response can be written as:

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

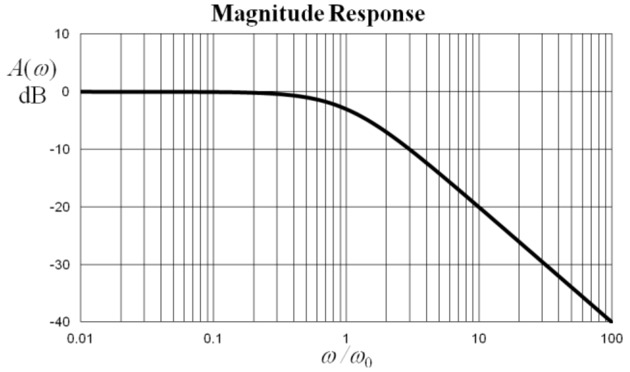
The magnitude function is found directly as:

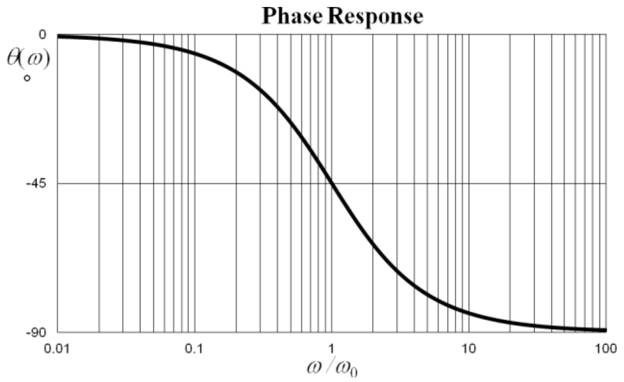
$$A(\omega) = 20\log_{10}|\mathbf{T}(j\omega)| = 20\log_{10}\left(\frac{1}{\sqrt{1+(\omega/\omega_0)^2}}\right)$$
$$= -10\log_{10}\left(1+(\omega/\omega_0)^2\right)$$

The phase is:

$$\theta(\omega) = -\tan^{-1}\left(\frac{\omega}{\omega_0}\right)$$

The Bode plots are graphed below, using a normalised log scale for ω:





16.7.1 Bode Plot Factors

Bode plot factors are additive if the magnitude scale is logarithmic

The primary advantage of a logarithmic scale for Bode magnitude plots is the conversion of multiplicative factors into additive factors by virtue of the definition of the logarithm. The phase plots are additive by definition of the multiplication of complex numbers. For example, if we have a frequency response function of the form:

$$\mathbf{T}(j\omega) = \frac{\mathbf{T}_1\mathbf{T}_2}{\mathbf{T}_3\mathbf{T}_4} \tag{16.12}$$

then clearly:

$$\begin{aligned} A(\omega) &= 20\log_{10}|\mathbf{T}_1| + 20\log_{10}|\mathbf{T}_2| \\ &\quad - 20\log_{10}|\mathbf{T}_3| - 20\log_{10}|\mathbf{T}_4| \\ &= A_1 + A_2 - A_3 - A_4 \end{aligned} \tag{16.13}$$

and:

$$\begin{aligned} \theta(\omega) &= \angle\mathbf{T}_1 + \angle\mathbf{T}_2 - \angle\mathbf{T}_3 - \angle\mathbf{T}_4 \\ &= \theta_1 + \theta_2 - \theta_3 - \theta_4 \end{aligned} \tag{16.14}$$

There are four different kinds of factors that may occur in a frequency response function:

The four factors that can occur in a frequency response function

Name	Factor
Constant gain	K
Pole (or zero) at the origin	$j\omega$
Pole (or zero) on the real axis	$1+j\omega/\omega_0$
Complex conjugate poles (or zeros)	$1+\left(1/Q_0\omega_0\right)j\omega+\left(j\omega/\omega_0\right)^2$

We can determine the logarithmic magnitude plot and phase plot for these four factors and then utilize them to obtain a Bode plot for any general form of a frequency response function. Typically, the curves for each factor are obtained and then added together graphically to obtain the curve for the complete frequency response function.

16.7.2 Approximating Bode Plots

Approximate responses can be easily drawn for the individual frequency response factors

The hand drawing of the individual Bode plot frequency response factors can be simplified by using linear approximations to the exact curves.

Approximate Magnitude Response

Consider a frequency response factor $1/(1+j\omega/\omega_0)$. The exact magnitude response is given by:

$$A(\omega) = -10\log_{10}\left(1+(\omega/\omega_0)^2\right) \tag{16.15}$$

For very small frequencies, such that $\omega << \omega_0$, we can say:

$$A(\omega) \approx -10\log_{10}(1) = 0\text{ dB} \tag{16.16}$$

For very large frequencies, such that $\omega >> \omega_0$, we can say:

$$A(\omega) \approx -20\log_{10}(\omega/\omega_0)\text{ dB} \tag{16.17}$$

Thus, on a set of axes where the horizontal axis is $\log_{10}\omega$, the "asymptotic" curves for the magnitude response are straight lines as shown below:

Exact and approximate magnitude response for a "real pole" factor

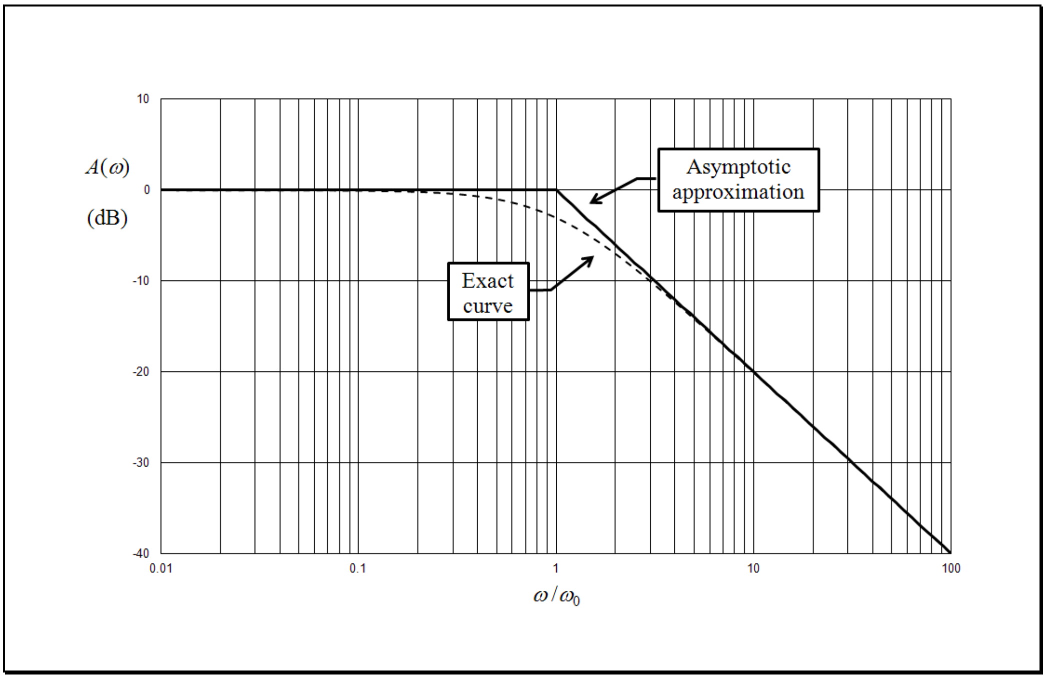


Figure 16.9

The low and high frequency asymptotes meet at the frequency ω_0, which is often called the *break frequency* or *corner frequency*. The slope of the high frequency asymptote can be ascertained from Eq. (16.17).

Break frequency and corner frequency defined

An interval of two frequencies with a ratio equal to ten, such as from ω_1 to $\omega_2 = 10\omega_1$, is called a *decade*. The difference between the logarithmic gains for $\omega >> \omega_0$, over a decade of frequency is approximated by:

$$\begin{aligned} A(\omega_2) - A(\omega_1) &\approx -20\log_{10}(\omega_2/\omega_0) \\ &\quad -(-20\log_{10}(\omega_1/\omega_0)) \\ &= -20\log_{10}(\omega_2/\omega_1) \\ &= -20\log_{10}(10) \\ &= -20\text{ dB} \end{aligned} \tag{16.18}$$

The slope of all asymptotic lines on a magnitude plot is a multiple of ±20 dB/decade…

That is, the slope of the high frequency asymptotic line for this frequency response factor is -20 dB/decade.

A frequency interval from ω_1 to $\omega_2 = 2\omega_1$ is often used and is called an *octave*. The difference between the logarithmic gains for $\omega >> \omega_0$, over an octave is approximated by:

$$\begin{aligned} A(\omega_2) - A(\omega_1) &\approx = -20\log_{10}(\omega_2/\omega_1) \\ &= -20\log_{10}(2) \\ &= -6.021\text{ dB} \end{aligned} \tag{16.19}$$

…or ±6 dB/octave

Therefore, the slope of the high frequency asymptote can be specified as either -6 dB/octave or -20 dB/decade.

Note that the *actual* gain at the break frequency $\omega = \omega_0$ is -3 dB, so ω_0 is also sometimes referred to as “the -3 dB frequency”.

Approximate Phase Response

Consider a frequency response factor $1/(1+j\omega/\omega_0)$. The exact phase response is given by:

$$\theta(\omega) = -\tan^{-1}\left(\frac{\omega}{\omega_0}\right) \tag{16.20}$$

Instead of graphing this nonlinear function, we often resort to a piece-wise linear approximation, as shown below:

Exact and approximate phase response for a "real pole" factor

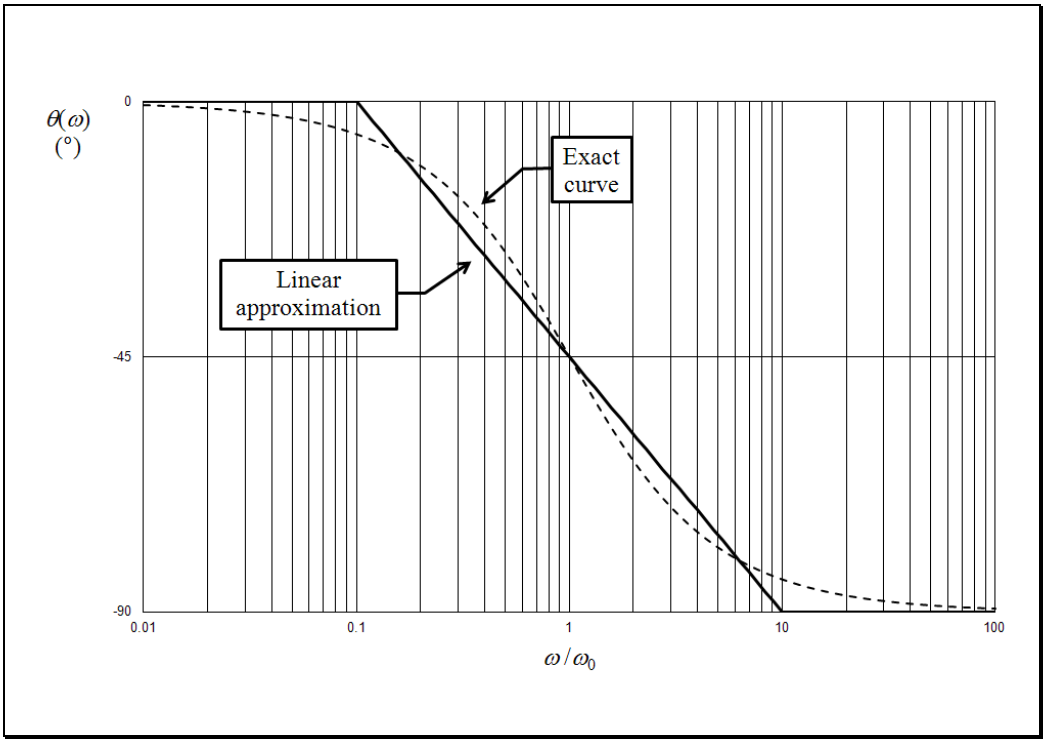


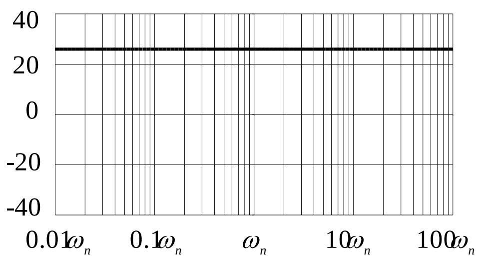
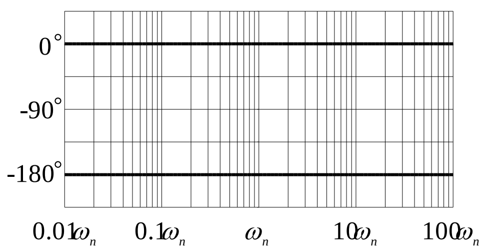
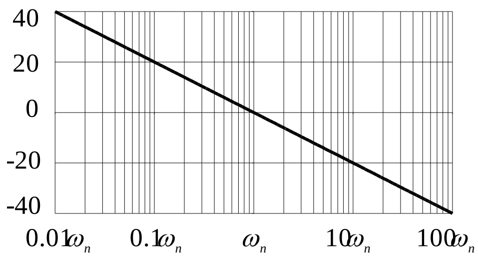
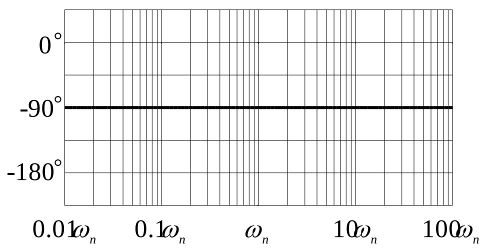
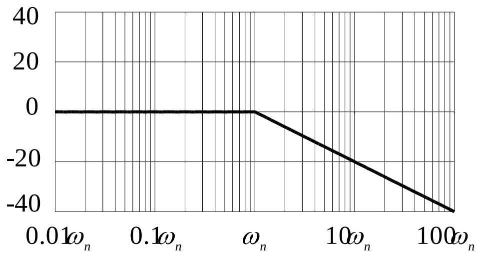
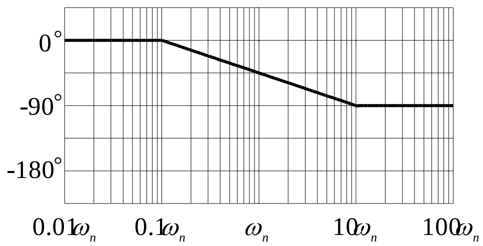
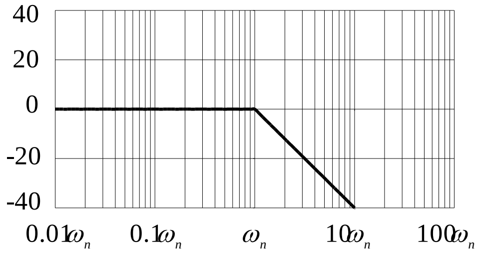
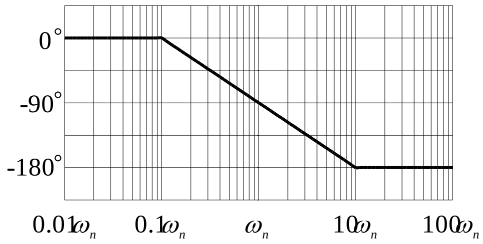
Figure 16.10

The slope of all asymptotic lines on a phase plot is a multiple of ±45°/decade

The piece-wise linear approximation passes through the correct phase of -45° at the break frequency, and is within 6° of the actual phase curve for all frequencies. As can be seen from the graph, the slope of the line passing through the -45° point is -45°/decade. This line only continues a decade above and a decade below the break frequency. For $\omega \ll \omega_0$ the approximating line is flat at 0°, whilst for $\omega \gg \omega_0$ the approximating line is flat at -90°.

16.8 Approximate Bode Plot Frequency Response Factors

The table below gives the four frequency response factors and their corresponding magnitude asymptotic plots and phase linear approximations:

<table>
<tr><th>Magnitude
Asymptote
A, dB</th><th>Phase
Linear Approximation
θ, °</th><th>Frequency Response Factor</th></tr>
<tr><td>
</td><td>
</td><td>K</td></tr>
<tr><td>
</td><td>
</td><td>$\dfrac{1}{(j\omega/\omega_0)}$</td></tr>
<tr><td>
</td><td>
</td><td>$\dfrac{1}{(1+j\omega/\omega_0)}$</td></tr>
<tr><td>
</td><td>
</td><td>$\dfrac{1}{\left(1-\dfrac{\omega^2}{\omega_0^2}+\dfrac{j\omega}{Q_0\omega_0}\right)}$</td></tr>
</table>

The corresponding numerator factors are obtained by "mirroring" the above plots about the 0 dB line and 0° line.

16.9 Summary

- The *frequency response* $\mathbf{T}(j\omega)$ is a complex function and can be written using a complex exponential in terms of magnitude and phase:

$$\mathbf{T}(j\omega)=|\mathbf{T}(j\omega)|e^{j\theta(\omega)}$$

- The *magnitude response* is a graph of $|\mathbf{T}(j\omega)|$ versus frequency.

- The *phase response* is a graph of $\theta(j\omega)$ versus frequency.

- A Bode plot is a graph of the magnitude response, $A(\omega)=20\log_{10}|\mathbf{T}(j\omega)|$, in dB, and the phase response, $\theta(j\omega)$, on a logarithmic frequency scale. The advantage of such a representation is that when circuits are cascaded their Bode plots are simply added.

- Approximate Bode plot factors can be used to analyse and design simple circuits.

16.10 References

Kamen, E. & Heck, B.: *Fundamentals of Signals and Systems using MATLAB®*, Prentice-Hall, 1997.

Exercises

1.

With respect to a reference frequency $f_0 = 20\text{ Hz}$, find the frequency which is (a) 2 decades above f_0 and (b) 3 octaves below f_0.

2.

Express the following magnitude ratios in dB: (a) 1, (b) 40, (c) 0.5

17 First-Order Op-Amp Filters

Contents

17.2

Introduction

Filters are essential to all modern electronic systems

Filters are essential to electrical engineering. They are used in *all* modern electronic systems. In communications, filters are essential for the generation and detection of analog and digital signals, whether via cable, optic fibre, air or satellite. In instrumentation, filters are essential in "cleaning up" noisy signals, or to recover some "special" part of a complicated signal. In control, feedback through a filter is used to achieve a desired response. In power, filters are used to inject high frequency signals on the power line for control purposes, or for removing harmonic components of a current. In machines, filters are used to suppress the generation of harmonics, or for controlling switching transients. The design of filters is therefore a useful skill to possess.

Filters can be of two types: analog and digital. In this subject, we will concentrate on analog filters. There are two reasons for this: analog filters are necessary components in "digital" systems, and analog filter theory serves as a precursor to digital filter design. The analog filters we will be looking at will also be of two types: passive and active. Active filters represent the most common, and use electronic components (such as op-amps) for their implementation. This is opposed to passive filters, which use the ordinary circuit elements: resistors, capacitors, inductors.

A filter is a circuit that implements a specific frequency response

Filter has the commonly accepted meaning of something retained, something rejected. For us, a filter is very simple: it is an electric circuit designed to implement a specific frequency response. Given a filter, obtaining the frequency response is just a matter of applying circuit theory. This is analysis. The choice of a frequency response and the choice of an implementation for a filter, however, are never unique. This is called design.

17.1 Bilinear Frequency Responses

Filter design and analysis is predominantly carried out in the frequency-domain. The circuits we design and analyse will be assumed to be operating with sinusoidal sources and be in the steady-state. This means we can use phasors and complex numbers to describe a circuit's response.

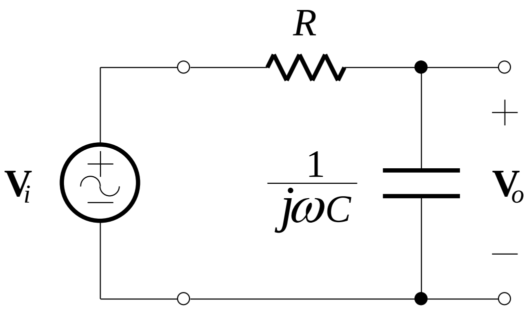
Recall that we define a circuit's frequency response as the ratio of the output voltage phasor to the input voltage phasor, as a function of frequency:

$$\mathbf{T}(j\omega)=\frac{\mathbf{V}_o}{\mathbf{V}_i} \tag{17.1}$$

Frequency response defined

EXAMPLE 17.1 Frequency Response of an *RC* Circuit

Consider a simple *RC* circuit:



The frequency response is:

$$\mathbf{T}(j\omega)=\frac{\mathbf{V}_o}{\mathbf{V}_i}=\frac{1/j\omega C}{R+1/j\omega C}=\frac{1}{1+j\omega RC}=\frac{1}{1+j\omega/\omega_0}, \qquad \omega_0=\frac{1}{RC}$$

This particular frequency response is one member of a larger class of frequency response functions.

The variable of the frequency response, $\mathbf{T}(j\omega)$, is $j\omega$. A first-order polynomial in this variable would be written as $\mathbf{P}(j\omega) = j\omega a_1 + a_2$ since the power of $j\omega$ is one. A first-order polynomial in real numbers could be written as $y = mx + b$, which is the equation of a straight line. Thus a first-order polynomial such as $j\omega a_1 + a_2$ is also known as a linear polynomial. When a frequency response is the quotient of two linear polynomials it is said to be *bilinear*. Thus a bilinear frequency response is of the form:

Bilinear frequency response defined

$$\mathbf{T}(j\omega) = \frac{j\omega a_1 + a_2}{j\omega b_1 + b_2} \tag{17.2}$$

where the *a* and *b* constants are real numbers. If Eq. (17.2) is written in a "standard form":

and rewritten in terms of poles and zeros

$$\mathbf{T}(j\omega) = \frac{a_2}{b_2}\frac{1 + j\omega a_1/a_2}{1 + j\omega b_1/b_2} = K\frac{1 + j\omega/z}{1 + j\omega/p} \tag{17.3}$$

then "-*z*" is the *zero* of $\mathbf{T}(j\omega)$ and "-*p*" is the *pole* of $\mathbf{T}(j\omega)$ (the reason for these names will become apparent later). For real circuits, *p* will always be a positive real number, while *z* may be either a positive or negative real number. The constant *K* depends on the circuit used to create the bilinear frequency response – for active circuits (e.g. those with op-amps) it can be positive or negative.

Thus for the *RC* circuit in the example, if we were to write the frequency response in the standard form, we would identify:

$$\begin{aligned} K &= 1 \\ z &= \infty \\ p &= \omega_0 = 1/RC \end{aligned} \tag{17.4}$$

17.1.1 Bilinear Magnitude Response

For the bilinear frequency response function:

$$\mathbf{T}(j\omega) = K \frac{1 + j\omega/z}{1 + j\omega/p} \tag{17.5}$$

The magnitude response is the magnitude of the frequency response

the magnitude response is:

$$\left|\mathbf{T}(j\omega)\right| = |K| \frac{|1 + j\omega/z|}{|1 + j\omega/p|} \tag{17.6}$$

Thus, the magnitude response is made up of the following three factors:

$$T_k = |K|$$

$$T_z = |1 + j\omega/z|$$

$$T_p = \frac{1}{|1 + j\omega/p|} \tag{17.7}$$

The first factor, $|K|$, is flat with frequency – it is just a real number.

17.6

The second and third factors look like:

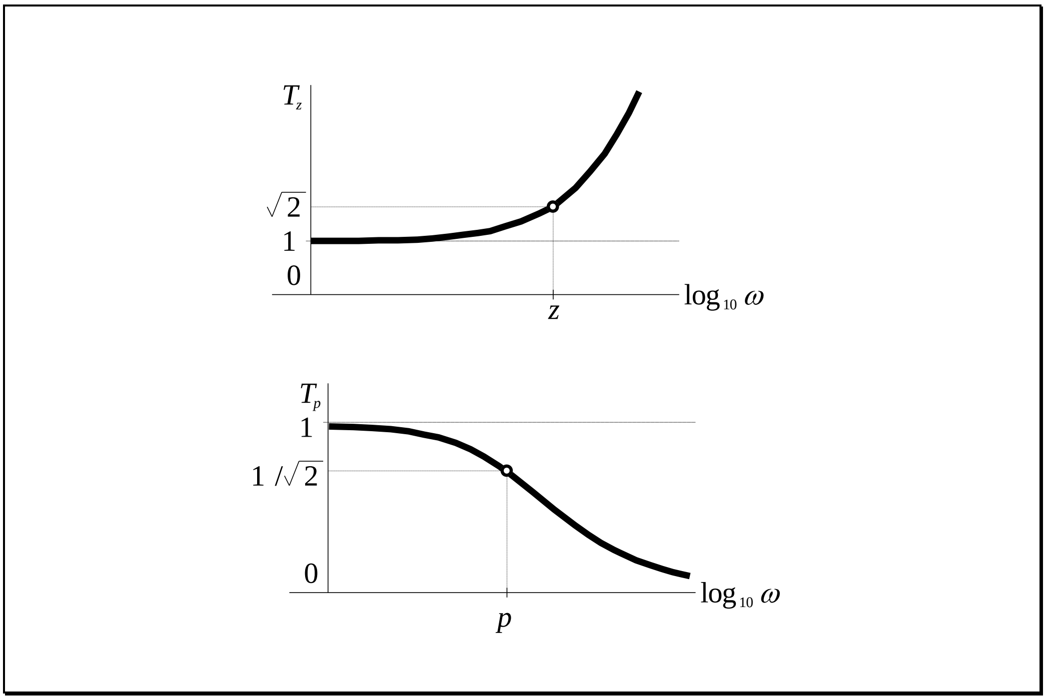


Figure 17.1

If all three factors are combined then the magnitude response takes on one of two shapes, depending on whether $p > z$ or not:

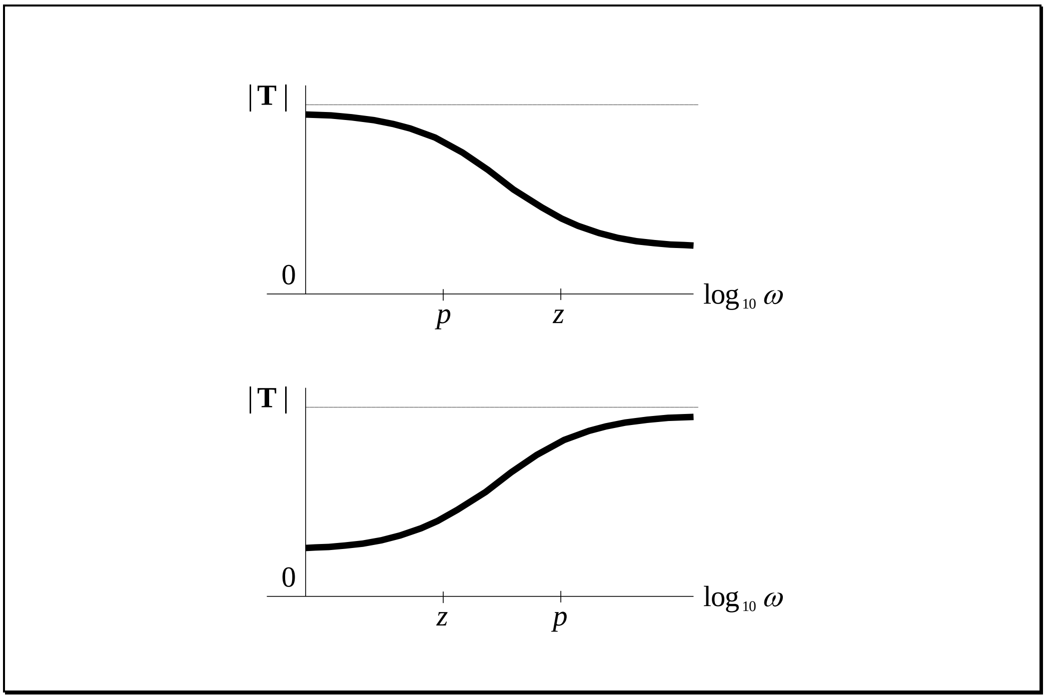


Figure 17.2

Thus, we can use the bilinear frequency response to create either a lowpass response ($p < z$) or a highpass response ($p > z$).

17.1.2 Bilinear Phase Response

For the bilinear frequency response function:

$$\mathbf{T}(j\omega) = K \frac{1 + j\omega/z}{1 + j\omega/p} \tag{17.8}$$

the phase is:

$$\theta = \angle K + \tan^{-1}\left(\frac{\omega}{z}\right) - \tan^{-1}\left(\frac{\omega}{p}\right) \tag{17.9}$$

The phase of the bilinear frequency response

Thus, the phase response is made up of the following three factors:

$$\begin{aligned} \theta_k &= \angle K \\ \theta_z &= \tan^{-1}\left(\frac{\omega}{z}\right) \\ \theta_p &= -\tan^{-1}\left(\frac{\omega}{p}\right) \end{aligned} \tag{17.10}$$

The first factor, $\angle K$, is flat with frequency – it is just a real number. If K is positive, its phase is $0°$, and if negative it is $180°$.

The second and third factors look like:

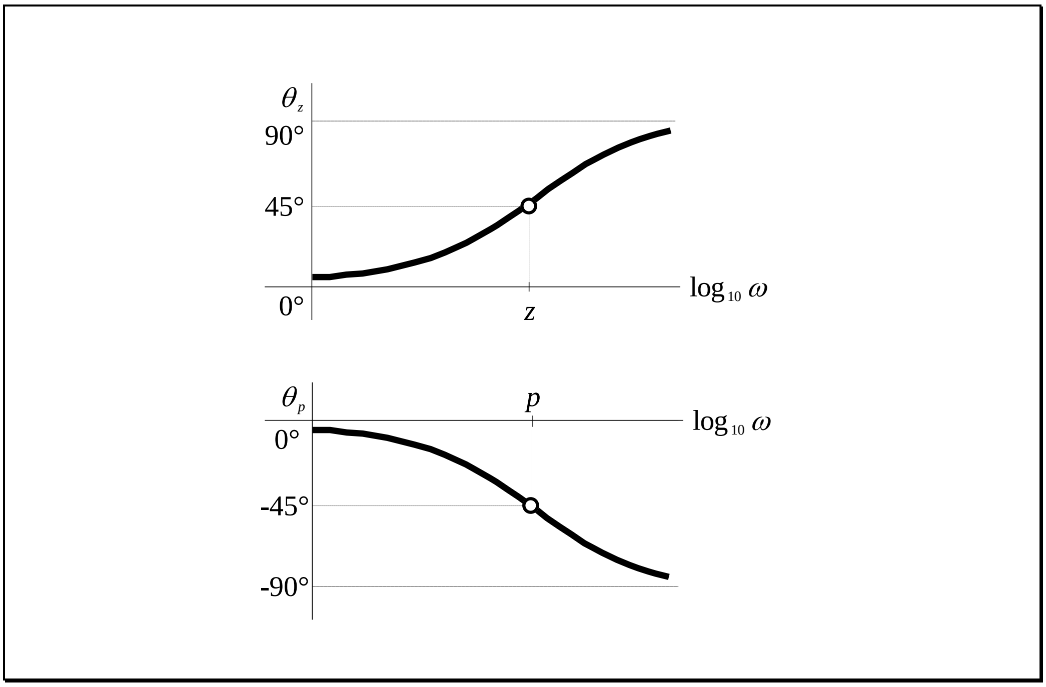


Figure 17.3

Assuming $\angle K = 0°$, If all three factors are combined then the phase response takes on one of two shapes, depending on whether $p > z$ or not:

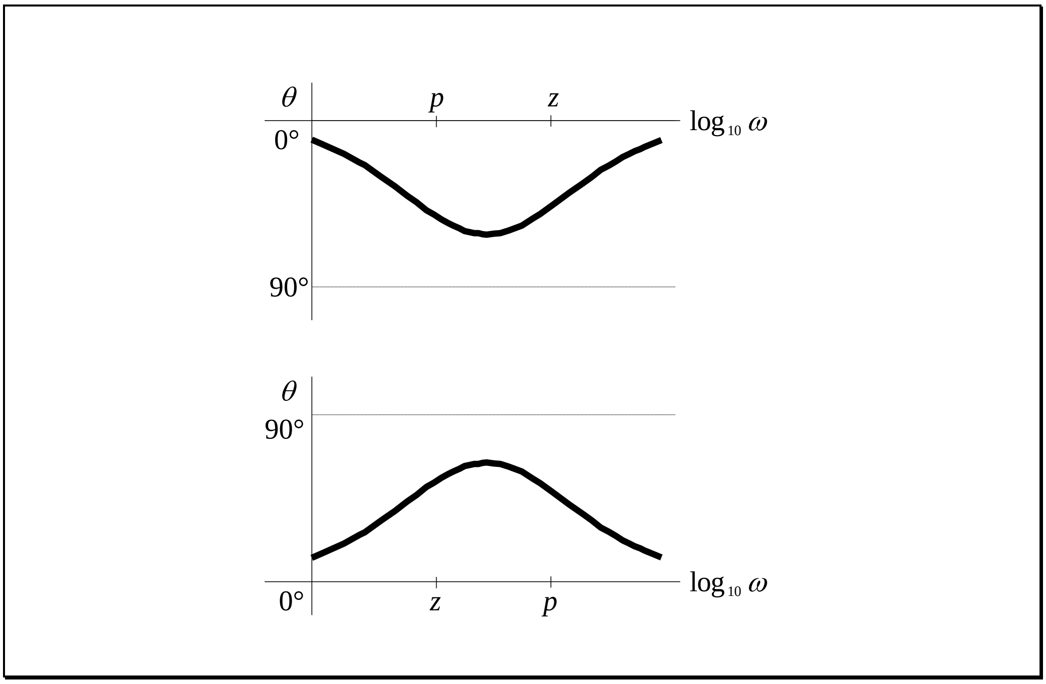
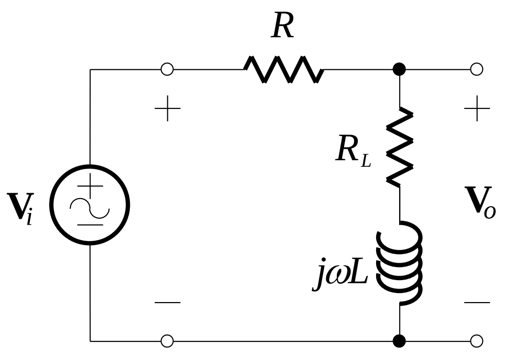


Figure 17.4

Thus, we can use the bilinear frequency response to create either a lag circuit ($p < z$) or a lead circuit ($p > z$).

EXAMPLE 17.2 Phase Response of an *RL* Circuit

Consider the *RL* circuit shown below:



Using the voltage divider rule, we obtain:

$$\mathbf{T}(j\omega) = \frac{R_L + j\omega L}{R + R_L + j\omega L}$$

If we write this in the standard form:

$$\mathbf{T}(j\omega) = K \frac{1 + j\omega/z}{1 + j\omega/p}$$

then:

$$K = \frac{R_L}{R + R_L}, \quad z = \frac{R_L}{L} \quad \text{and} \quad p = \frac{R + R_L}{L}$$

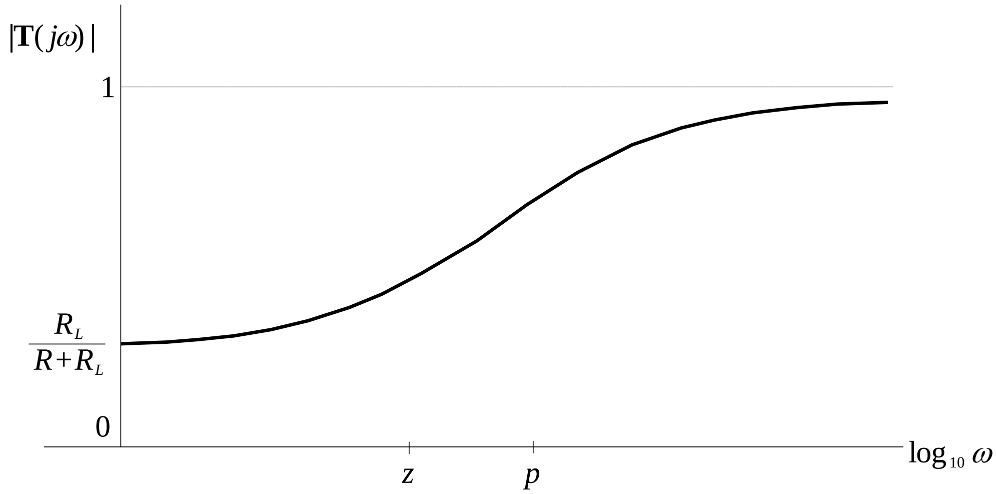
so that $p > z$. We can see that:

$$|\mathbf{T}(j0)| = K = \frac{R_L}{R + R_L}$$

$$|\mathbf{T}(j\infty)| = K \frac{1/z}{1/p} = 1$$

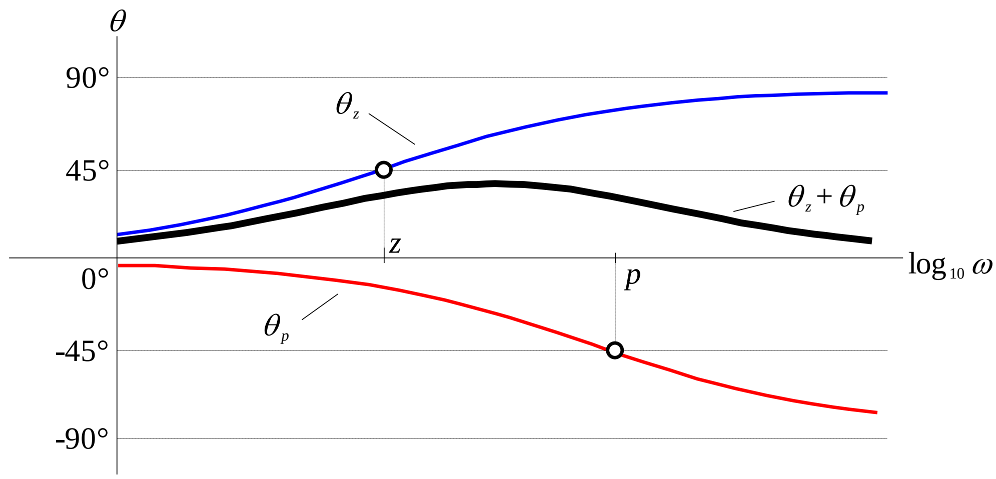
17.10

The magnitude response is:



This does not approximate an ideal brickwall filter very well, but it is still known as a highpass filter.

From Eq. (17.9) we see that θ is characterized by the sum of three angles, the first due to the constant $\angle K = 0°$, the second a function of the zero numerator term, θ_z, and the third a function of the pole denominator term, θ_p. For $p > z$ the phase function θ_z reaches $+45°$ at a low frequency, while θ_p reaches $-45°$ for a higher frequency. Therefore, the phase response looks like:



Thus, the circuit provides phase lead.

17.1.3 Summary of Bilinear Frequency Responses

We can summarize the magnitude and phase responses of $\mathbf{T}(j\omega)$ for various values of z and p, where K is assumed to be positive, in the table below:

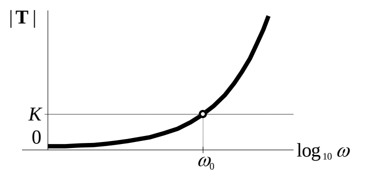
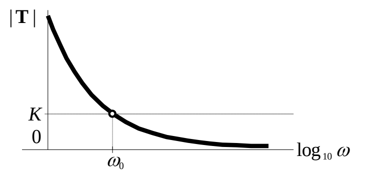
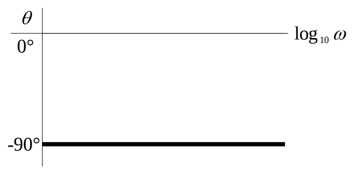
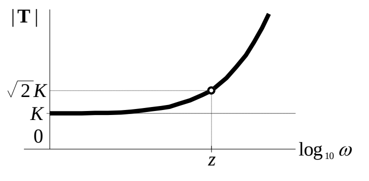
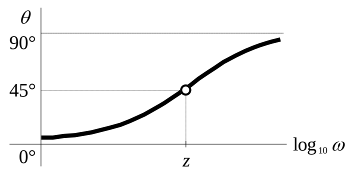
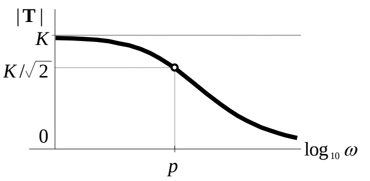
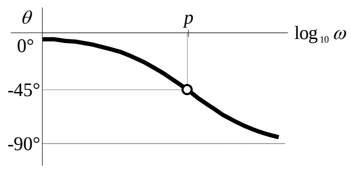
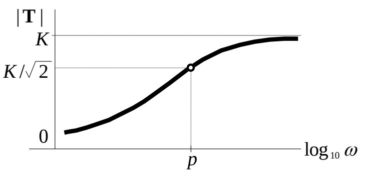
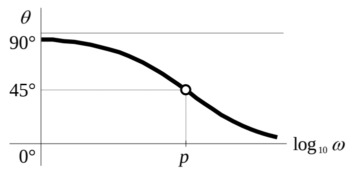
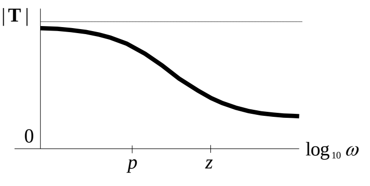
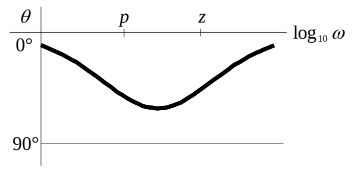
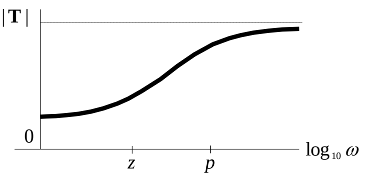
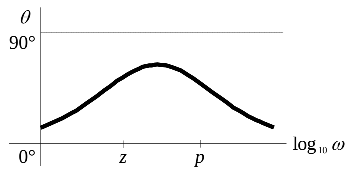
$\mathbf{T}(j\omega)$	Magnitude Response	Phase Response
1 $K\, j\omega/\omega_0$		
2 $\dfrac{K}{j\omega/\omega_0}$		
3 $K(1+j\omega/z)$		
4 $K\dfrac{1}{1+j\omega/p}$		
5 $K\dfrac{j\omega/p}{1+j\omega/p}$		
6 $K\dfrac{1+j\omega/z}{1+j\omega/p}$ $z>p$		
7 $K\dfrac{1+j\omega/z}{1+j\omega/p}$ $p>z$		

Table 17.1 – Summary of Bilinear Frequency Responses

17.2 Frequency and Magnitude Scaling

In filter design, it is common practice to *normalise* equations so that they have the same form. For example, we have seen the bilinear frequency response:

$$\mathbf{T}(j\omega)=\frac{1/RC}{1/RC+j\omega} \tag{17.11}$$

expressed as:

$$\mathbf{T}(j\omega)=\frac{1}{1+j\omega/\omega_0} \tag{17.12}$$

Normalising the cutoff frequency means setting it to 1

We *normalise* the equation by setting $\omega_0 = 1$:

$$\mathbf{T}(j\omega)=\frac{1}{1+j\omega} \tag{17.13}$$

Every equation in filter design will be normalised so that $\omega_0 = 1$. This is helpful, since every equation will be able to be compared on the same base.

The difficulty that now arises is *denormalising* the resulting equations, values or circuit designs.

17.2.1 Frequency Scaling (Denormalising)

Frequency scaling, or denormalising, means we want to change ω_0 from 1 back to its original value. To do this, we must change all frequency dependent terms in the frequency response, which also means frequency dependent elements in a circuit. Furthermore, the frequency scaling should not affect the magnitude of any impedance in the frequency response.

and denormalising means setting the frequency back to its original value

To scale the frequency by a factor k_f, for a capacitor, we must have:

$$|\mathbf{Z}_C| = \frac{1}{\omega C} = \frac{1}{(k_f \omega)(1/k_f)C} = \frac{1}{(k_f \omega)C_{\text{new}}} \tag{17.14}$$

We must decrease the capacitance by the amount $1/k_f$, while increasing the frequency by the amount k_f if the magnitude of the impedance is to remain the same.

Frequency scaling must keep the magnitude of the impedance the same

For an inductor, we must have:

$$|\mathbf{Z}_L| = \omega L = (k_f \omega)\frac{1}{k_f} L = (k_f \omega)L_{\text{new}} \tag{17.15}$$

Therefore, new element values may be expressed in terms of old values as follows:

$$R_{\text{new}} = R_{\text{old}} \tag{17.16}$$

$$L_{\text{new}} = \frac{1}{k_f} L_{\text{old}} \tag{17.17}$$

$$C_{\text{new}} = \frac{1}{k_f} C_{\text{old}} \tag{17.18}$$

The frequency scaling equations

17.2.2 Magnitude Scaling

Magnitude scale to get realistic element values

Since a frequency response is always a ratio, if we increase the impedances in the numerator and denominator by the same amount, it changes nothing. We do this to obtain realistic values for the circuit elements in the implementation.

If the impedance magnitudes are normally:

$$\mathbf{Z}_R = R, \quad |\mathbf{Z}_L| = \omega L, \quad |\mathbf{Z}_C| = \frac{1}{\omega C} \tag{17.19}$$

then after magnitude scaling with a constant k_m they will be:

$$\begin{aligned} k_m \mathbf{Z}_R &= k_m R, \\ k_m |\mathbf{Z}_L| &= k_m \omega L, \\ k_m |\mathbf{Z}_C| &= \frac{1}{\omega C / k_m} \end{aligned} \tag{17.20}$$

Therefore, new element values may be expressed in terms of old values as follows:

The magnitude scaling equations

$$R_{\text{new}} = k_m R_{\text{old}} \tag{17.21}$$

$$L_{\text{new}} = k_m L_{\text{old}} \tag{17.22}$$

$$C_{\text{new}} = \frac{1}{k_m} C_{\text{old}} \tag{17.23}$$

An easy-to-remember rule in scaling *R*'s and *C*'s in electronic circuits is that *RC* products should stay the same. For example, for the lowpass *RC* filter, the cutoff frequency is given by $\omega_0 = 1/RC$. Therefore, if the resistor value goes up, then the capacitor value goes down by the same factor – and vice versa.

17.3 Cascading Circuits

How can we create circuits with higher than first-order frequency responses by "cascading" first-order circuits? Consider the following circuit:

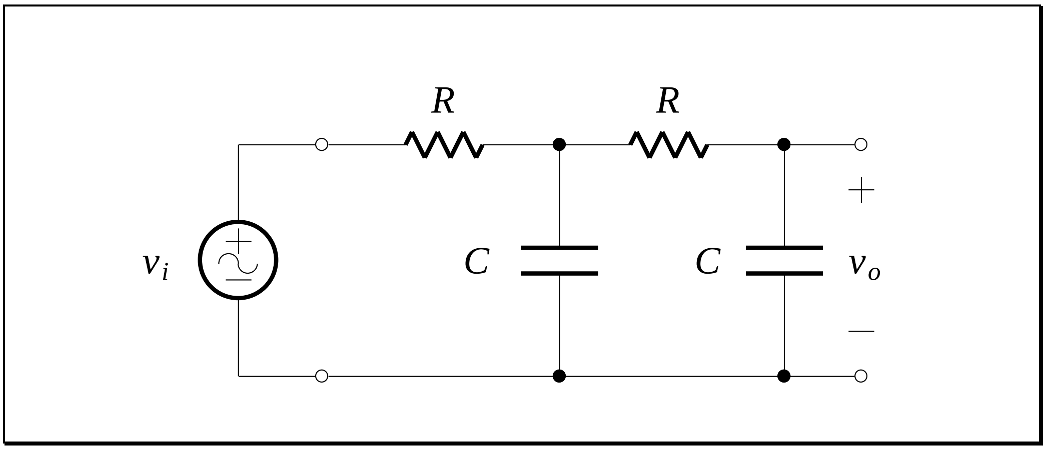


Figure 17.5

Show that the frequency response for the above circuit is:

$$\frac{\mathbf{V}_o}{\mathbf{V}_i} = \frac{(1/RC)^2}{(1/RC)^2 - \omega^2 + j\omega(3/RC)} \tag{17.24}$$

Compare with the following circuit:

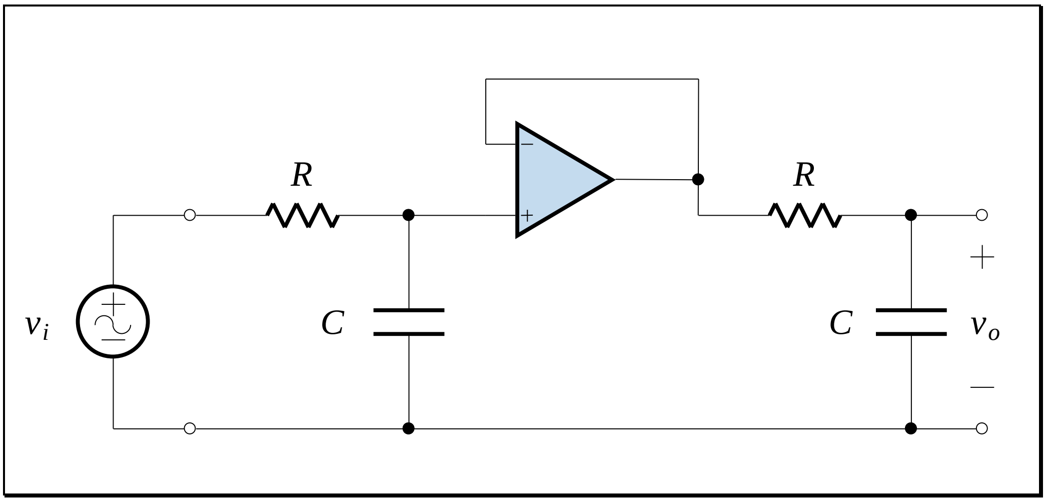


Figure 17.6

which has the frequency response:

$$\frac{\mathbf{V}_o}{\mathbf{V}_i} = \left(\frac{1/RC}{1/RC + j\omega}\right)\left(\frac{1/RC}{1/RC + j\omega}\right)$$
$$= \frac{(1/RC)^2}{(1/RC)^2 - \omega^2 + j\omega(2/RC)} \tag{17.25}$$

Cascading buffered circuits is highly desirable from a design perspective

Therefore, when we cascade circuits, if the "output" of each circuit presents a low impedance to the next stage, so that each successive circuit does not "load" the previous circuit, then we can simply multiply the frequency responses of the individual circuits to achieve an overall frequency response. Op-amp circuits of both the inverting and non-inverting type are ideal for cascading.

17.4 Inverting Bilinear Op-Amp Circuit

We seek an implementation of the bilinear frequency response that can be cascaded. The frequency response of the inverting op-amp circuit is:

$$\mathbf{T}(j\omega) = -\frac{\mathbf{Z}_2}{\mathbf{Z}_1} \tag{17.26}$$

Therefore, we would like to have:

The inverting op-amp circuit is one way to implement a bilinear frequency response

$$\mathbf{T} = -\frac{\mathbf{Z}_2}{\mathbf{Z}_1} = -K\frac{1 + j\omega/z}{1 + j\omega/p} \tag{17.27}$$

The specifications of the design problem are the values K, z and p. These may be found from a Bode plot – the break frequencies and the gain at some frequency – or obtained in any other way. The solution to the design problem involves finding a circuit and the values of the elements in that circuit. Since we are using an active device – the op-amp – inductors are excluded from our circuits. Therefore, we want to find values for the R's and the C's. Once found, these values can be adjusted by any necessary frequency scaling, and then by magnitude scaling to obtain convenient element values.

For the general bilinear frequency response, we can make the following identification:

$$K\frac{1+j\omega/z}{1+j\omega/p}=\frac{K/(1+j\omega/p)}{1/(1+j\omega/z)}=\frac{\mathbf{Z}_2}{\mathbf{Z}_1} \tag{17.28}$$

Therefore:

$$\mathbf{Z}_2=\frac{1}{1/K+j\omega/pK}=\frac{1}{1/R_2+j\omega C_2} \tag{17.29}$$

$$\mathbf{Z}_1=\frac{1}{1+j\omega/z}=\frac{1}{1/R_1+j\omega C_1} \tag{17.30}$$

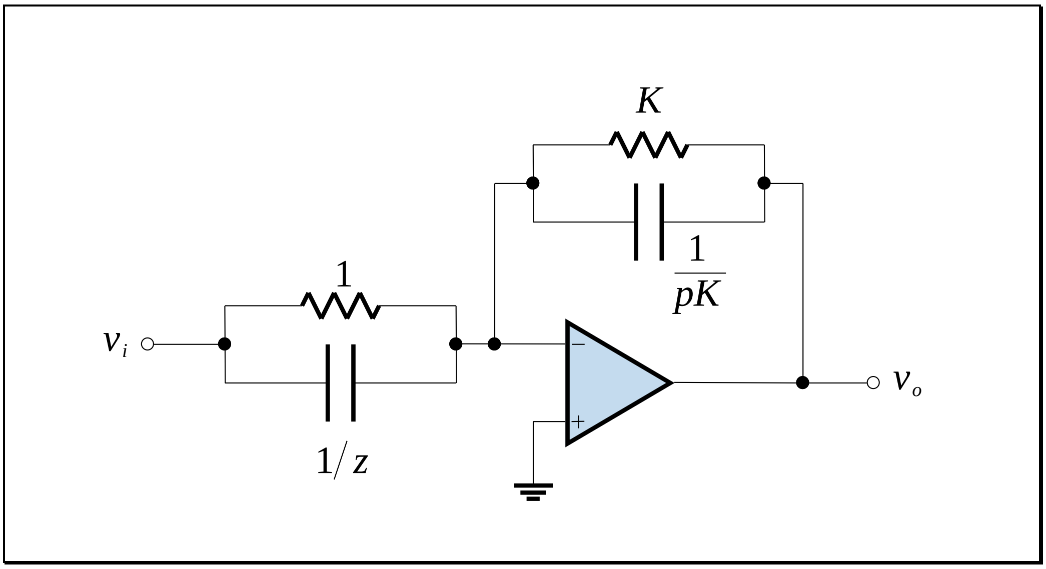
The impedances for the inverting op-amp circuit to implement a bilinear frequency response

The design equations become:

$$R_1=1,\quad C_1=\frac{1}{z},\quad R_2=K,\quad C_2=\frac{1}{pK} \tag{17.31}$$

Element values for the inverting op-amp circuit in terms of the pole, zero and gain

One realisation of the bilinear frequency response is then:



An inverting op-amp circuit that implements a bilinear frequency response

Figure 17.7

17.5 Inverting Op-Amp Circuits

The approach we took in obtaining a circuit to implement the bilinear frequency response can be applied to other forms of $\mathbf{T}(j\omega)$ to give the entries in the table below:

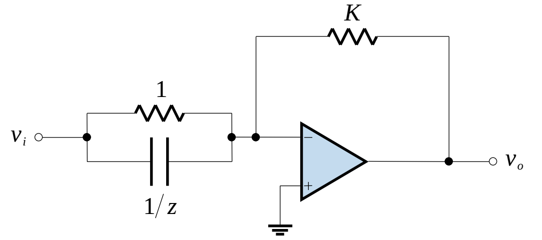
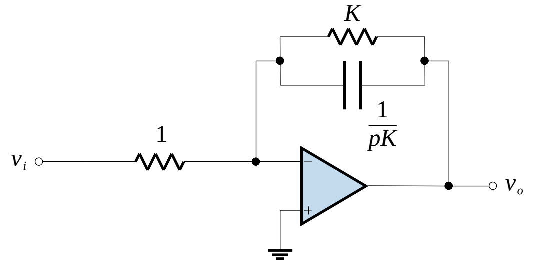
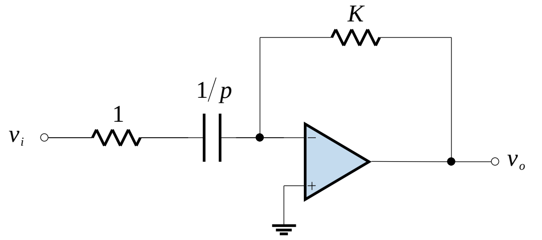
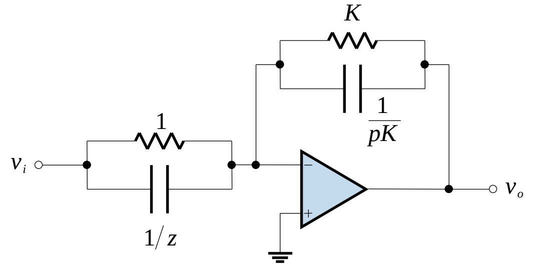
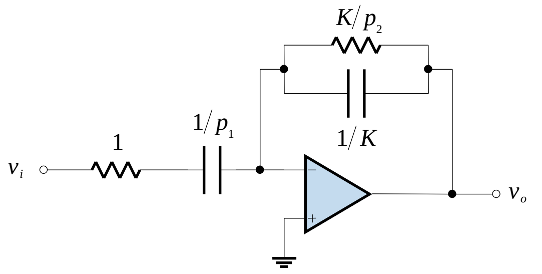
Frequency Response $\mathbf{T}(j\omega)$	**Circuit**
1 $-K(1+j\omega/z)$	
2 $-K\dfrac{1}{1+j\omega/p}$	
3 $-K\dfrac{j\omega/p}{1+j\omega/p}$	
4 $-K\dfrac{1+j\omega/z}{1+j\omega/p}$	
5 $-K\dfrac{j\omega/p_1p_2}{(1+j\omega/p_1)(1+j\omega/p_2)}$	

Table 17.2 – Inverting Op-Amp Circuits

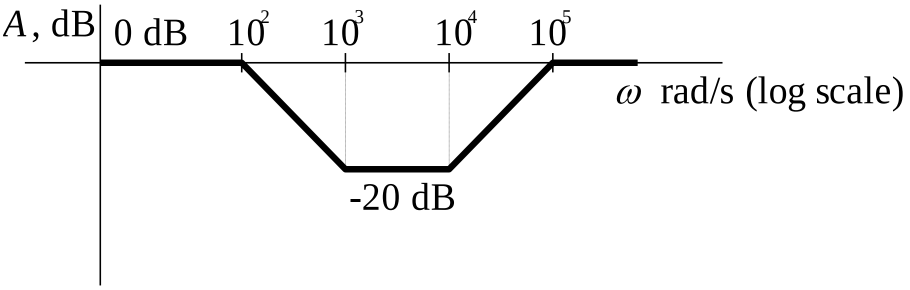
The last entry in the table illustrates an important point. If we use a series *RC* connection for $\mathbf{Z}_1$ but a parallel *RC* connection for $\mathbf{Z}_2$, then the frequency response becomes one of second-order. So the manner in which the capacitors are connected in the circuit determines the order of the circuit.

17.6 Cascade Design

We can make use of cascaded modules, each of first-order, to satisfy specifications that are more complicated than the bilinear function.

EXAMPLE 17.3 Cascade Design of a Bandstop Filter

The asymptotic Bode plot shown below is for a bandstop filter:



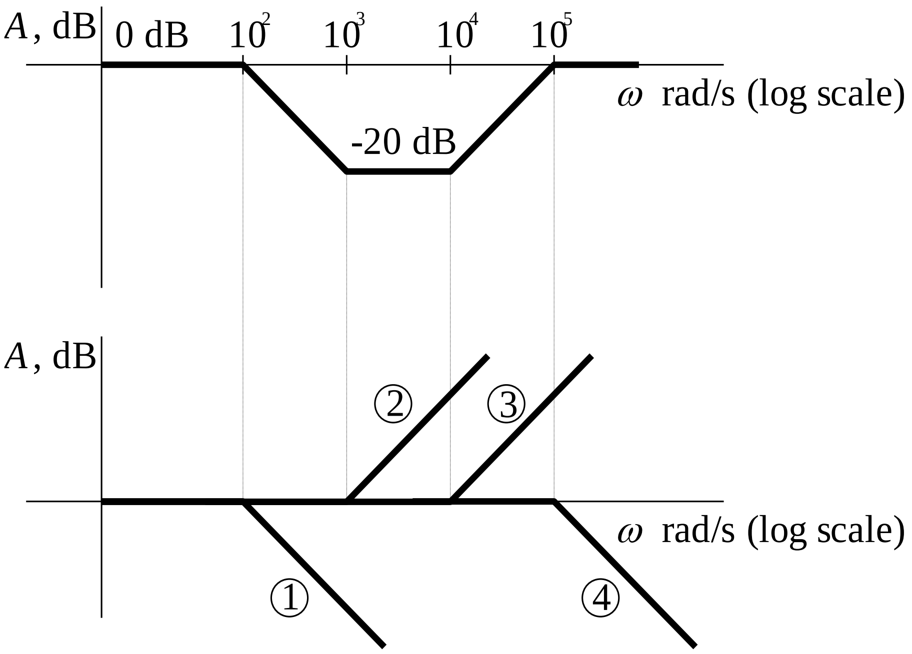
A bandstop filter

The gain at low and high frequencies is unchanged, but 20 dB of attenuation is provided in an intermediate frequency range. We wish to design a filter to these specifications and the additional requirement that all capacitors have the value $C = 10\ \text{nF}$.

17.20

The composite plot may be decomposed into four first-order factors as shown below:

Decomposing a Bode plot into first-order factors



Those marked (1) and (4) represent pole factors, while those marked (2) and (3) are zero factors.

From the break frequencies given, we have:

$$\mathbf{T}(j\omega)=\frac{(2)(3)}{(1)(4)}=\frac{\left(1+j\omega/10^3\right)\left(1+j\omega/10^4\right)}{\left(1+j\omega/10^2\right)\left(1+j\omega/10^5\right)}$$

We next write $\mathbf{T}(j\omega)$ as a product of bilinear functions:

The frequency response as a cascade of bilinear functions

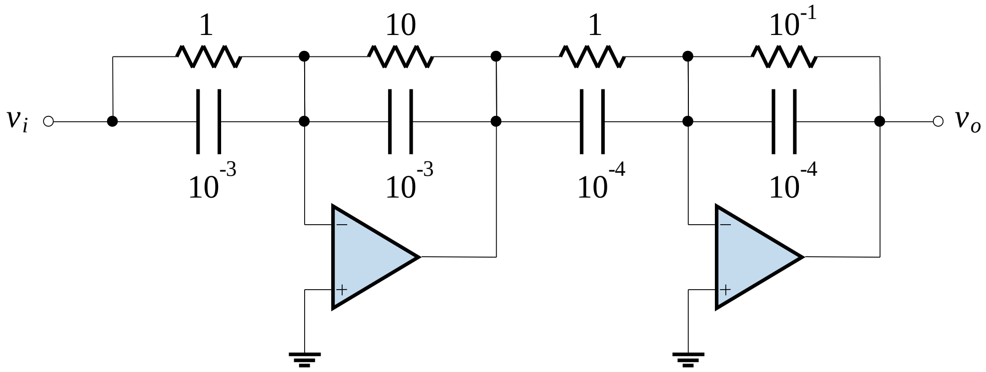
$$\mathbf{T}(j\omega)=\mathbf{T}_1(j\omega)\mathbf{T}_2(j\omega)$$

For a circuit realisation of $\mathbf{T}_1$ and $\mathbf{T}_2$ we decide to use the inverting op-amp circuit 4 in Table 17.2. Since one of the design requirements was that all capacitor values are equal, we need to choose $K=z/p$ in each circuit. This leads to:

$$\begin{aligned}\mathbf{T}(j\omega)&=\mathbf{T}_1(j\omega)\mathbf{T}_2(j\omega)\\&=10\frac{1+j\omega/10^3}{1+j\omega/10^2}\times 0.1\frac{1+j\omega/10^4}{1+j\omega/10^5}\end{aligned}$$

Note that, in this design, the DC gains $K = z/p$ of each circuit multiply to give us the overall DC gain we required (of 1). In general though, we might need to add an additional gain / attenuation stage to the cascade to "meet spec".

Using the formulas for element values given in Table 17.2, we obtain the realisation shown below:



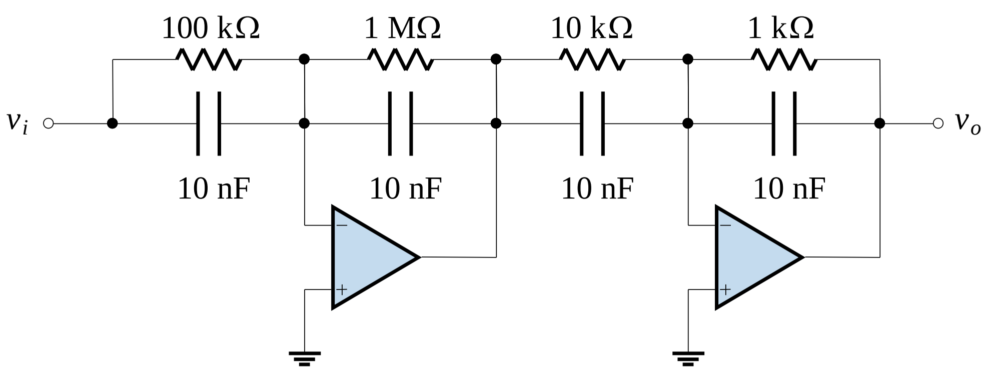
A realisation of the specifications

Frequency scaling is not required since we have worked directly with specified frequencies (we did not normalize the initial Bode plot). The magnitude scaling of the circuit is accomplished with the equations:

$$C_{\text{new}} = \frac{1}{k_m} C_{\text{old}} \quad \text{and} \quad R_{\text{new}} = k_m R_{\text{old}}$$

Magnitude scaling is required to get realistic element values

Since the capacitors are to have the value 10 nF, this means $k_m = 10^5$ for the first circuit and $k_m = 10^4$ for the second circuit. The element values that result are shown below and the design is complete:

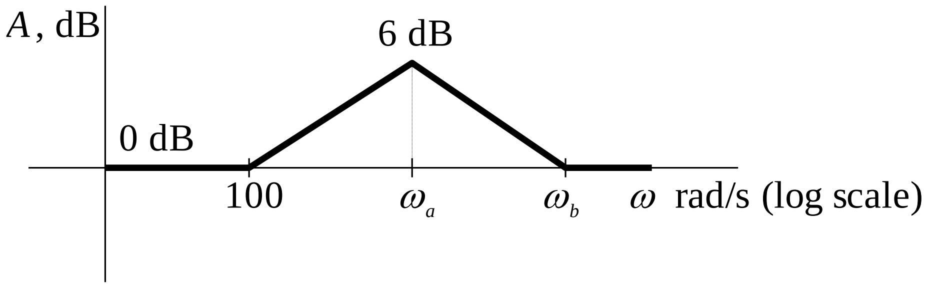


A realistic implementation of the specifications

A band-enhancement filter

EXAMPLE 17.4 Cascade Design of a Band-Enhancement Filter

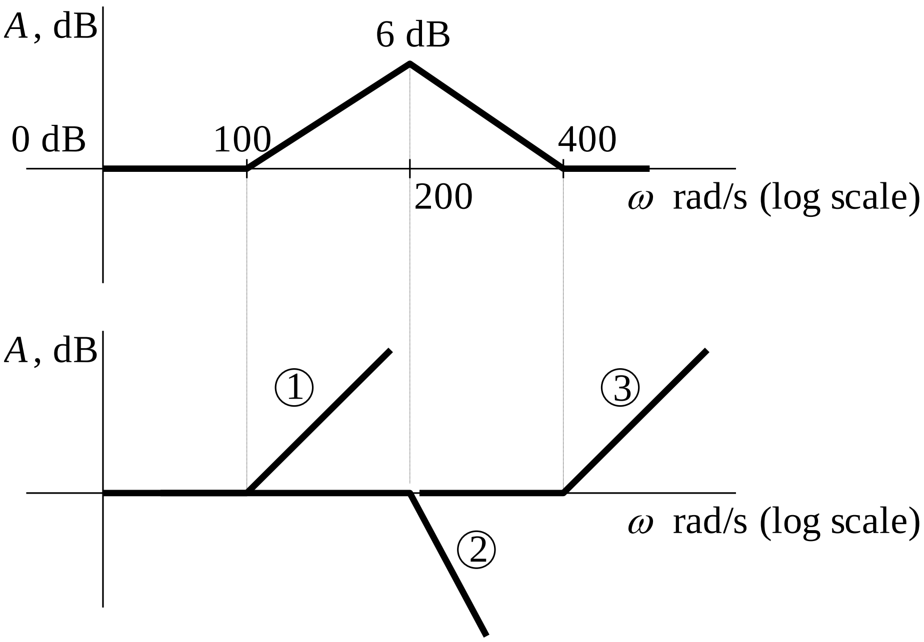
The asymptotic Bode plot shown below describes a band-enhancement filter:



We wish to provide additional gain over a narrow band of frequencies, leaving the gain at higher and lower frequencies unchanged. We wish to design a filter to these specifications and the additional requirement that all capacitors have the value $C = 10\text{ nF}$.

The maximum gain of the asymptotic plot is 6 dB. An asymptotic plot increases at 20 dB/decade, or equivalently, at 6 dB/octave, so that $\omega_a = 200\text{ rads}^{-1}$. Since the plot returns to 0 dB, ω_b must be one octave greater than ω_a, or 400 rads^{-1}.

The four first-order factors, which make up the composite Bode plot, are shown below:



Decomposing a Bode plot into first-order factors

As frequency increases, the first break frequency identifies a zero factor (1), then a double pole factor (2), followed by another zero factor (3).

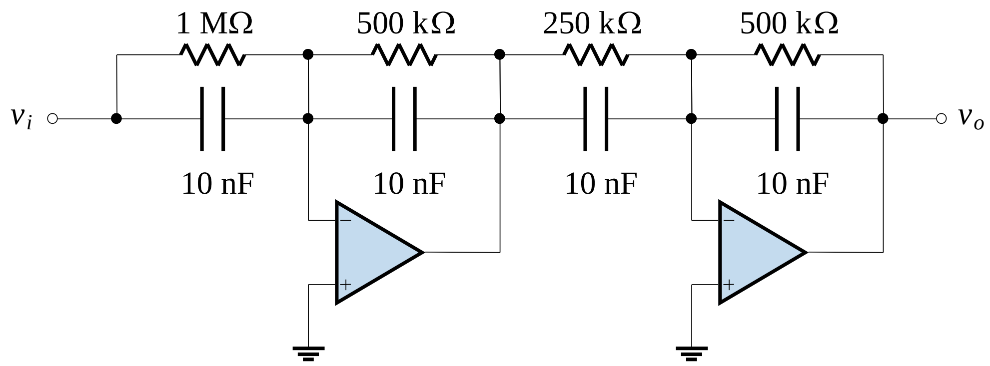
From this information, we construct:

$$\mathbf{T}(j\omega) = \frac{(1)(3)}{(2)} = \frac{(1 + j\omega/100)(1 + j\omega/400)}{(1 + j\omega/200)^2}$$

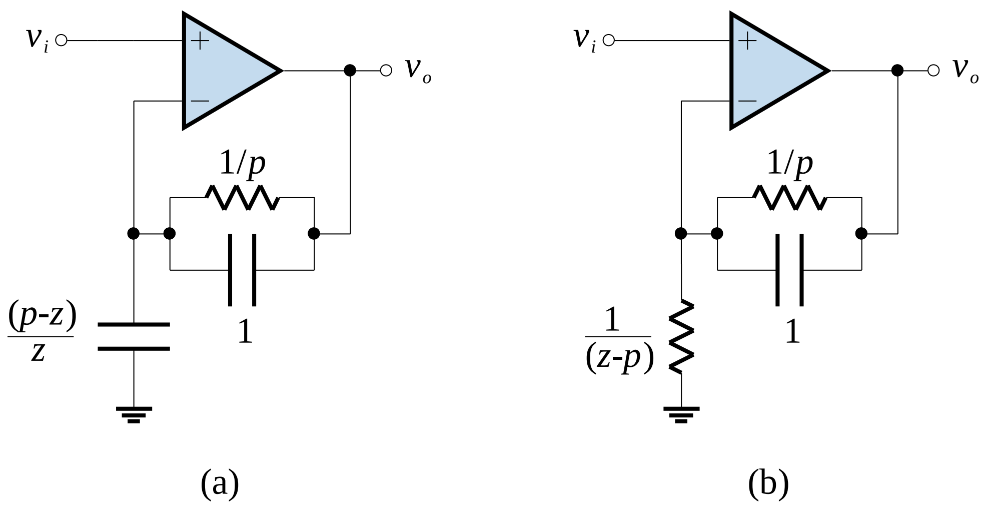
If we make use of the same strategy that was used in the previous example, we write $\mathbf{T}(j\omega)$ as the product of two bilinear functions:

$$\begin{aligned}\mathbf{T}(j\omega) &= \mathbf{T}_1(j\omega)\mathbf{T}_2(j\omega) \\ &= 0.5\frac{1 + j\omega/100}{1 + j\omega/200} \times 2\frac{1 + j\omega/400}{1 + j\omega/200}\end{aligned}$$

and then obtain the practical realisation shown below:



In design there are always many possibilities. Suppose that we decide to try a design for this frequency response using non-inverting op-amp configurations. Two circuits, and their frequency responses, are shown below:

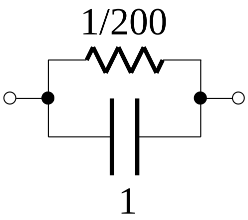
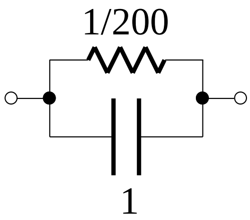
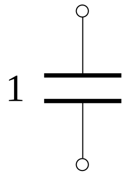
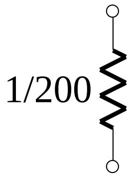


$$\mathbf{T}_a(j\omega) = \frac{1 + j\omega/z}{1 + j\omega/p}, \quad p > z$$

$$\mathbf{T}_b(j\omega) = K\frac{1 + j\omega/z}{1 + j\omega/p}, \quad K = z/p, \quad z > p$$

You should derive these frequency responses to confirm this.

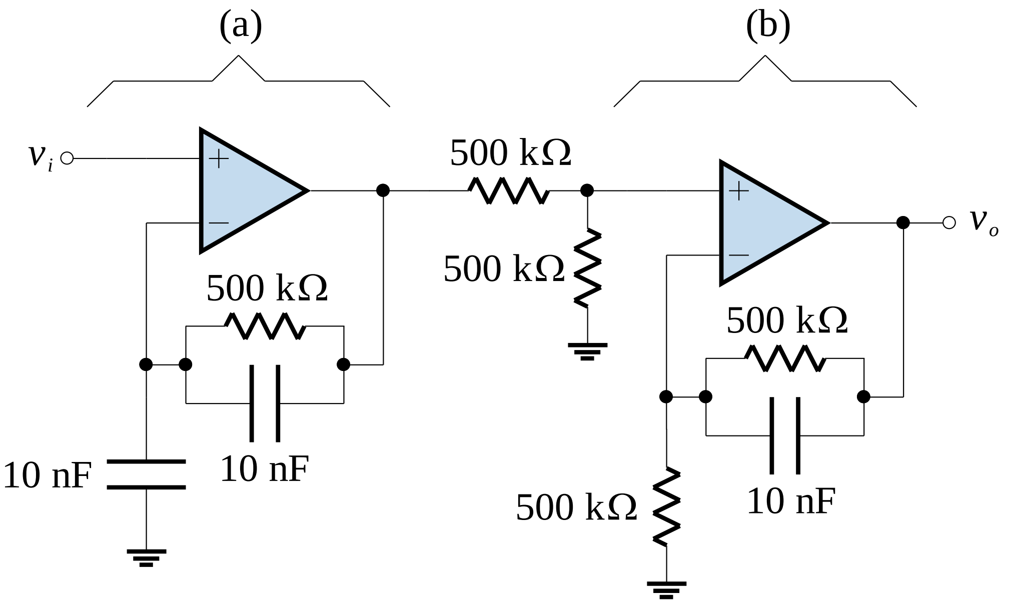
The sequence of design steps to accomplish our specification is shown in the table below:

Item	*Module 1*	*Module 2*
Frequency response	$= \dfrac{1 + j\omega/100}{1 + j\omega/200}$	$\dfrac{1 + j\omega/400}{1 + j\omega/200}$
Choice of circuit	(a)	(b)
$\mathbf{Z}_2$ element values		
$\mathbf{Z}_1$ element value		

These circuits are good choices for this particular design because all *R*'s have the same value and all *C*'s the value of 1. Frequency scaling is not required and magnitude scaling is done to obtain the required *C* values of 10 nF by the choice of $k_m = 10^8$. Note that module 2, by virtue of $K = z/p$, produces a DC gain of 2. Therefore, a voltage divider is needed to reduce the gain by ½ to meet specifications exactly.

17.26

The final circuit is shown below:



Now the problem remaining for the designer is to decide whether to use the inverting implementation or the non-inverting implementation, or whether to find other designs before a final selection is made. One advantage that the circuit above has over the inverting implementation is that it effectively has an infinite input impedance (which may or may not be important – it depends on the application).

17.7 Summary

- A bilinear frequency response can be put in the “standard form”:

$$\mathbf{T}(j\omega) = K \frac{1 + j\omega/z}{1 + j\omega/p}$$

  where z is termed a *zero*, and p is a *pole*. Zeros make the magnitude response increase with frequency, whereas poles make the magnitude response decrease with frequency. Zeros provide phase lead, whereas poles provide phase lag.

- It is common in both the algebraic description of a system, and in its circuit representation, to *normalise* ω_0 to 1. We can then *frequency scale* inductors and capacitors to set a new ω_0. We *magnitude scale* circuit elements to get realistic values.

- A circuit can be cascaded when it does not “load” the output of the previous stage. Circuits in which the output is taken from an op-amp output terminal are ideal for cascading, since the op-amp effectively acts as an ideal source.

- Simple op-amp circuits, of both the inverting and non-inverting variety, can be used to implement bilinear frequency responses.

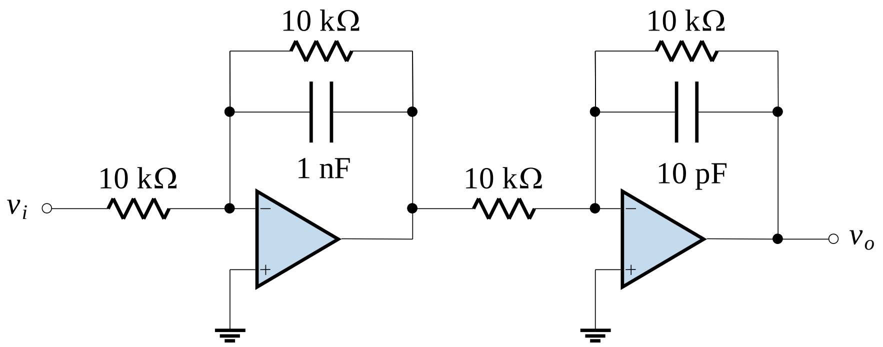
17.8 References

Van Valkenburg, M. E.: *Analog Filter Design*, Holt-Saunders, Tokyo, 1982.

Exercises

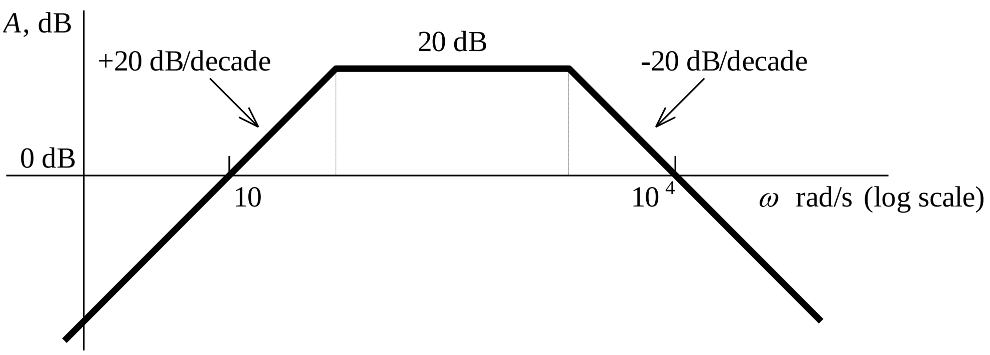
1.

For the circuit shown below, prepare the asymptotic Bode plot for the magnitude of $\mathbf{T}(j\omega)$. Carefully identify all slopes and low and high frequency asymptotes.



2.

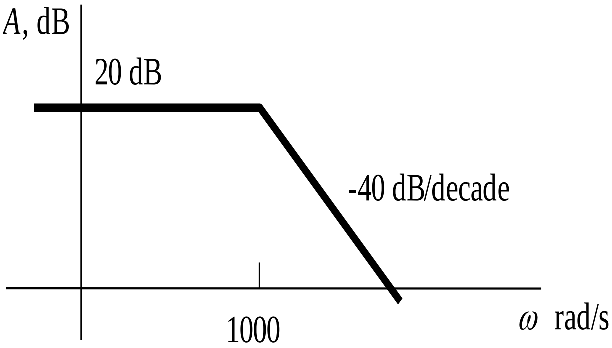
Design an *RC* op-amp filter to realise the bandpass response shown below.



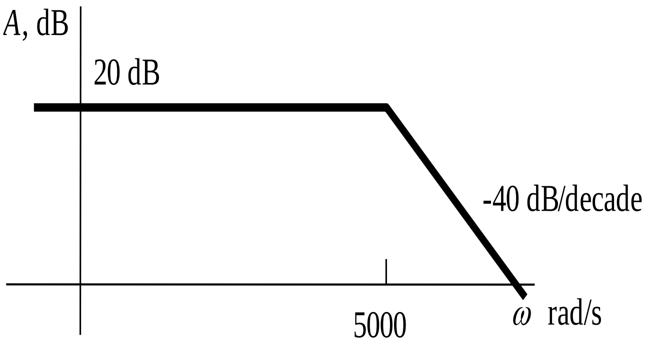
Use a minimum number of op-amps in your design, and scale so that the elements are in a practical range.

3.

The asymptotic Bode plot shown below represents a lowpass filter-amplifier with a break frequency of $\omega_0 = 1000 \text{ rads}^{-1}$.

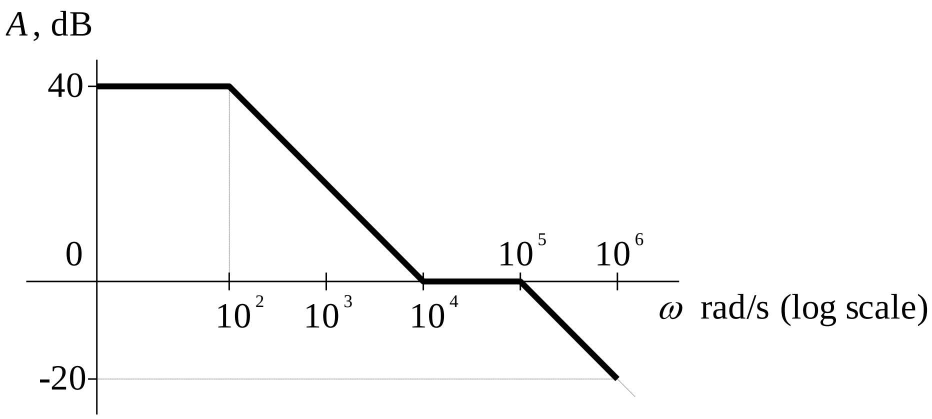


Design a circuit to be connected in cascade with the amplifier such that the break frequency is extended to $\omega_0 = 5000 \text{ rads}^{-1}$:



4.

An asymptotic Bode plot is shown below for a desired magnitude response. Design an amplifier-filter using a minimum number of op-amps.



18 The Second-Order Step Response

Contents

18.2

Introduction

Circuits that contain resistors, inductors and capacitors are known as *RLC* circuits. We will examine the simplest *RLC* circuits – the parallel *RLC* circuit where the resistor, inductor and capacitor are in parallel, and the series *RLC* circuit in which they are connected in series.

Many circuits can be reduced to one of these equivalent circuits. For example, transistor amplifiers in analog radios, crystal oscillators driving the clocks of digital circuits and microprocessors, and power system "power factor correction" circuits, can each be reduced to a parallel *RLC* circuit.

The inclusion of both inductance and capacitance in a circuit leads to at least a second-order system – that is, one that is characterized by a linear differential equation including a second-order derivative, or by a set of two first-order differential equations. We shall see that such systems are much more complicated than first-order systems – there will be two arbitrary constants to find, we will need initial conditions for variables and their derivatives, and there are three different functional forms of the natural response that depend upon element values.

However, there is a side benefit to studying second-order systems: we will see a unifying underlying structure emerge for all linear circuits, regardless of their complexity. We will also discover a way to exploit this underlying structure.

18.1 Solution of the Homogeneous Linear Differential Equation

We have seen that any linear homogeneous differential equation with constant coefficients of order *n* is an equation that can be written:

$$a_n \frac{d^n y}{dt^n} + a_{n-1} \frac{d^{n-1} y}{dt^{n-1}} + \cdots + a_1 \frac{dy}{dt} + a_0 y = 0 \tag{18.1}$$

or more simply using the *D* operator as:

$$f(D)y = 0 \tag{18.2}$$

We have seen that the solution to such an equation was found by using the property that:

$$f(D)e^{st} = f(s)e^{st} \tag{18.3}$$

then letting:

$$f(s) = 0 \tag{18.4}$$

which leads to $y = e^{st}$ as a solution. To find the general solution, we find the *n* roots of the n^{th}-order characteristic equation $f(s) = 0$, and then form a linear sum of each individual solution.

18.1.1 Distinct Real Roots

The characteristic equation is of degree *n*. Let its roots be s_1, s_2, …, s_n. If these roots are all real and distinct, then the *n* solutions:

$$y_1 = e^{s_1 t},\ y_2 = e^{s_2 t},\ \ldots,\ y_n = e^{s_n t} \tag{18.5}$$

are linearly independent and the general solution can be written at once. It is:

$$y = c_1 e^{s_1 t} + c_2 e^{s_2 t} + \cdots + c_n e^{s_n t} \tag{18.6}$$

in which c_1, c_2, …, c_n are arbitrary constants.

EXAMPLE 18.1 Solution of a Second-Order ODE with Distinct Real Roots

Solve the equation:

$$\frac{d^2 y}{dt^2} + 7\frac{dy}{dt} + 6y = 0$$

First write the characteristic equation:

$$\begin{aligned} s^2 + 7s + 6 &= 0 \\ (s+1)(s+6) &= 0 \end{aligned}$$

whose roots are $s = -1, -6$. Then the general solution is seen to be:

$$y = c_1 e^{-t} + c_2 e^{-6t}$$

18.1.2 Repeated Real Roots

Suppose that $f(s)=0$ has repeated roots. Then Eq. (18.6) does not yield the general solution. To see this, let the characteristic equation have three equal roots $s_1=b$, $s_2=b$ and $s_3=b$. The corresponding part of the solution yielded by Eq. (18.6) is:

$$\begin{aligned} y &= c_1 e^{bt} + c_2 e^{bt} + c_3 e^{bt} \\ &= (c_1+c_2+c_3)e^{bt} \\ &= c_4 e^{bt} \end{aligned} \tag{18.7}$$

Thus, corresponding to the three roots under consideration, this method yields only one solution.

What is needed is a method for obtaining m linearly independent solutions corresponding to m equal roots of the characteristic equation. Suppose that the characteristic equation $f(s)=0$ has the m equal roots:

$$s_1 = s_2 = \cdots = s_m = b \tag{18.8}$$

Then the operator $f(D)$ must have a factor $(D-b)^m$. We wish to find m linearly independent y's for which:

$$(D-b)^m y = 0 \tag{18.9}$$

Turning to Eq. (8.16) in Topic 8, and writing $s=b$, we find that:

$$(D-b)^m \left(t^k e^{bt}\right) = 0, \quad k = 0, 1, \ldots, (m-1) \tag{18.10}$$

The functions $y_k = t^k e^{bt}$ where $k = 0, 1, \ldots, (m-1)$ are linearly independent because, aside from the common factor e^{bt}, they contain only the respective powers $t^0, t^1, t^2, \ldots, t^{m-1}$.

The general solution of Eq. (18.9) is therefore:

$$y = c_1 e^{bt} + c_2 t e^{bt} + \cdots + c_m t^{m-1} e^{bt} \tag{18.11}$$

Furthermore, if $f(D)$ contains the factor $(D-b)^m$, then the equation $f(D)y = 0$ can be written $g(D)(D-b)^m y = 0$ where $g(D)$ contains all the factors of $f(D)$ except $(D-b)^m$. Then any solution of $(D-b)^m y = 0$ is also a solution of $g(D)(D-b)^m y = 0$ and therefore of $f(D)y = 0$.

18.1.3 Only Real Roots

Now we are in a position to write the solution of $f(D)y = 0$ whenever the characteristic equation has only real roots. Each root of the characteristic equation is either distinct from all the other roots or it is one of a set of equal roots.

Corresponding to a root s_i distinct from all others, there is the solution:

$$y_i = c_i e^{s_i t} \tag{18.12}$$

and corresponding to m equal roots s_1, s_2, …, s_m, each equal to b, there are solutions:

$$c_1 e^{bt}, \; c_2 t e^{bt}, \ldots, \; c_m t^{m-1} e^{bt} \tag{18.13}$$

EXAMPLE 18.2 Solution of an ODE with Repeated Real Roots

Solve the equation:

$$\left(D^4 + 7D^3 + 18D^2 + 20D + 8\right)y = 0$$

The characteristic equation:

$$s^4 + 7s^3 + 18s^2 + 20s + 8 = 0$$

has the roots $s = -1, -2, -2, -2$. Then the general solution is:

$$y = c_1 e^{-t} + c_2 e^{-2t} + c_3 t e^{-2t} + c_4 t^2 e^{-2t}$$

or:

$$y = c_1 e^{-t} + \left(c_2 + c_3 t + c_4 t^2\right)e^{-2t}$$

EXAMPLE 18.3 Solution of an ODE with Repeated Real Roots

Solve the equation:

$$\frac{d^4 y}{dt^4} + 2\frac{d^3 y}{dt^3} + \frac{d^2 y}{dt^2} = 0$$

The characteristic equation is:

$$s^4 + 2s^3 + s^2 = 0$$

with roots $s = 0, 0, -1, -1$. Hence the desired solution is:

$$y = c_1 + c_2 t + c_3 e^{-t} + c_4 t e^{-t}$$

18.1.4 Distinct Complex Roots

In all real systems, the describing differential equation $f(D)y=0$ has a characteristic equation $f(s)=0$ that has real coefficients. We know from algebra that if the characteristic equation has any complex roots then those roots must occur in conjugate pairs. Thus, if:

$$s_1 = a + jb \tag{18.14}$$

is a root of the equation $f(s)=0$ with *a* and *b* real and $b \neq 0$, then:

$$s_2 = a - jb \tag{18.15}$$

is also a root of $f(s)=0$. According to the preceding section, $f(D)y=0$ is satisfied by:

$$y = c_1 e^{(a+jb)t} + c_2 e^{(a-jb)t} \tag{18.16}$$

where for *y* to be real, we must have $c_1 = c_2^*$. Since *t* is real along with *a* and *b*, Euler's identity gives us the result:

$$y = c_1 e^{at}(\cos bt + j\sin bt) + c_2 e^{at}(\cos bt - j\sin bt) \tag{18.17}$$

This may be written:

$$y = (c_1 + c_2)e^{at}\cos bt + j(c_1 - c_2)e^{at}\sin bt \tag{18.18}$$

Finally, let $c_1 + c_2 = 2\operatorname{Re}\{c_1\} = c_3$, and $j(c_1 - c_2) = -2\operatorname{Im}\{c_1\} = c_4$, where c_3 and c_4 are *real* arbitrary constants. Then $f(D)y=0$ is seen to have the solution:

$$\boxed{y = c_3 e^{at}\cos bt + c_4 e^{at}\sin bt} \tag{18.19}$$

corresponding to the two roots $s_1 = a + jb$ and $s_2 = a - jb$ $(b \neq 0)$ of the characteristic equation.

EXAMPLE 18.4 Solution of an ODE with Distinct Complex Roots

Solve the equation:

$$\left(D^3+5D^2+17D+13\right)y=0$$

For the characteristic equation:

$$s^3+5s^2+17s+13=0$$

one root, $s_1=-1$, is easily found. When the factor $(s+1)$ is removed, it is seen that the other two roots are solutions of the quadratic equation:

$$s^2+4s+13=0$$

Those roots are found to be $s_2=-2+j3$ and $s_3=-2-j3$. The characteristic equation has the roots $s=-1,-2\pm j3$. Hence the general solution of the differential equation is:

$$y=c_1e^{-t}+c_2e^{-2t}\cos 3t+c_3e^{-2t}\sin 3t$$

18.1.5 Repeated Complex Roots

Repeated complex roots lead to solutions analogous to those brought in by repeated real roots. For instance, if the roots $s = a \pm jb$ occur three times, then the corresponding six linearly independent solutions of the differential equation are those appearing in the expression:

$$\left(c_1 + c_2 t + c_3 t^2\right)e^{at} \cos bt + \left(c_4 + c_5 t + c_6 t^2\right)e^{at} \sin bt \tag{18.20}$$

EXAMPLE 18.5 Solution of an ODE with Repeated Complex Roots

Solve the equation:

$$\left(D^4 + 8D^2 + 16\right)y = 0$$

The characteristic equation $s^4 + 8s^2 + 16 = 0$ may be written:

$$\left(s^2 + 4\right)^2 = 0$$

so its roots are seen to be $s = \pm j2, \pm j2$. The roots, $s_1 = j2$ and $s_2 = -j2$ occur twice each. Thinking of $j2$ as $0 + j2$ and recalling that $e^{0t} = 1$, we write the solution of the differential equation as:

$$y = \left(c_1 + c_2 t\right)\cos 2t + \left(c_3 + c_4 t\right)\sin 2t$$

18.2 The Source-Free Parallel *RLC* Circuit

A suitable model for portions of many practical circuits is given by the parallel *RLC* circuit:

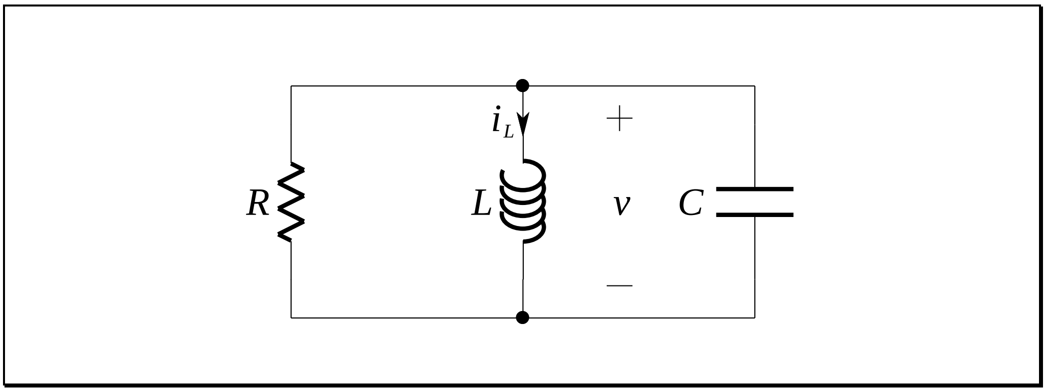


Figure 18.1

We will assume that energy may be stored initially in both the inductor and the capacitor, and thus nonzero values of both inductor current and capacitor voltage are initially present. With reference to the circuit above, we may then write the single nodal equation:

$$\frac{v}{R}+\frac{1}{L}\int_{t_0}^{t} v\,dt+i_L(t_0)+C\frac{dv}{dt}=0 \qquad (18.21)$$

When both sides are differentiated once with respect to time and divided by C the result is the linear second-order homogeneous differential equation:

$$\frac{d^2v}{dt^2}+\frac{1}{RC}\frac{dv}{dt}+\frac{1}{LC}v=0 \qquad (18.22)$$

We must solve this equation subject to the initial conditions:

$$\begin{aligned} i_L(0^+)&=I_0 \\ v(0^+)&=V_0 \end{aligned} \qquad (18.23)$$

With our theory of the D operator behind us, we now embark on solving Eq. (18.22). We write the characteristic equation:

$$f(s) = s^2 + \frac{1}{RC}s + \frac{1}{LC} = 0 \tag{18.24}$$

and identify the two roots:

$$\begin{aligned} s_1 &= -\frac{1}{2RC} + \sqrt{\left(\frac{1}{2RC}\right)^2 - \frac{1}{LC}} \\ s_2 &= -\frac{1}{2RC} - \sqrt{\left(\frac{1}{2RC}\right)^2 - \frac{1}{LC}} \end{aligned} \tag{18.25}$$

These roots can be either real and distinct, real and repeated, or complex and distinct depending on the values of R, L and C.

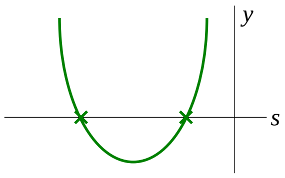
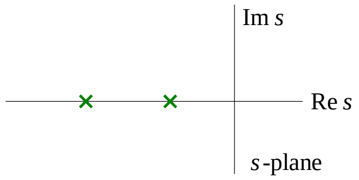
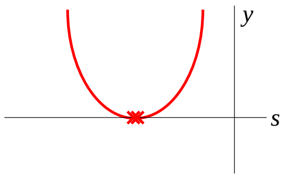
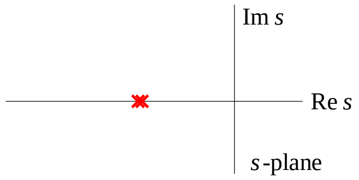
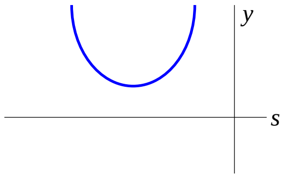
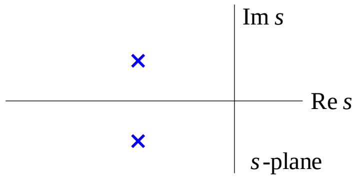
Case	**Graph of** $y = f(s)$	**Solution of** $f(s) = 0$
Real and distinct		
Real and repeated		
Complex and distinct		

Table 18.1 – Three types of roots to the 2nd-order characteristic equation

Therefore, there will be three different forms of the natural response, corresponding to each of the cases in Table 18.1.

It will be helpful to make some simplifying substitutions into Eq. (18.25) for the sake of conceptual clarity. Let us define the *undamped natural frequency*:

$$\omega_0 = \frac{1}{\sqrt{LC}} \tag{18.26}$$

We will also define the *exponential damping coefficient*:

$$\alpha = \frac{1}{2RC} \quad \text{(parallel)} \tag{18.27}$$

This description is used because α is a measure of how rapidly the natural response decays or damps out to its steady final value (usually zero).

Finally, s, s_1 and s_2 are called *complex frequencies* and will form the basis for some of our later work.

Thus, our characteristic equation becomes:

$$s^2 + 2\alpha s + \omega_0^2 = 0 \tag{18.28}$$

and has the two roots:

$$\begin{aligned} s_1 &= -\alpha + \sqrt{\alpha^2 - \omega_0^2} \\ s_2 &= -\alpha - \sqrt{\alpha^2 - \omega_0^2} \end{aligned} \tag{18.29}$$

It is now apparent that the nature of the response depends upon the relative magnitudes of α and ω_0. The square root appearing in the expressions for s_1 and s_2 will be real when $\alpha > \omega_0$, zero when $\alpha = \omega_0$, and imaginary when $\alpha < \omega_0$. Each of these cases will be considered separately.

18.3 The Overdamped Parallel *RLC* Circuit

When $\alpha > \omega_0$, both s_1 and s_2 will be real and distinct:

$$s_1 = -\alpha + \sqrt{\alpha^2 - \omega_0^2}$$
$$s_2 = -\alpha - \sqrt{\alpha^2 - \omega_0^2} \tag{18.30}$$

Thus, the natural response is of the form:

$$\boxed{v_n = K_1 e^{s_1 t} + K_2 e^{s_2 t}} \tag{18.31}$$

We also know that:

$$\alpha > \sqrt{\alpha^2 - \omega_0^2} \tag{18.32}$$

and therefore:

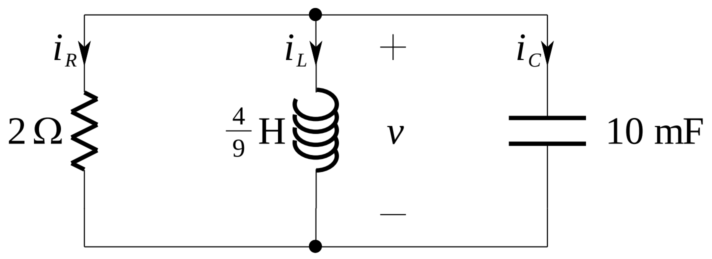
$$\left(-\alpha - \sqrt{\alpha^2 - \omega_0^2}\right) < \left(-\alpha + \sqrt{\alpha^2 - \omega_0^2}\right) < 0$$
$$s_2 < s_1 < 0 \tag{18.33}$$

This shows that both s_1 and s_2 are *negative* real numbers. Thus, the response $v(t)$ is the algebraic sum of two decreasing exponential terms, both of which approach zero as time increases without limit. In fact, since the absolute value of s_2 is larger than that of s_1, the term containing s_2 has the more rapid rate of decrease.

It only remains to find the arbitrary constants K_1 and K_2 using the initial conditions, and we have the solution.

EXAMPLE 18.6 The Overdamped Parallel *RLC* Circuit

Consider the circuit:



in which $v(0)=0$ and $i_L(0)=-8$. We may easily determine the values of the several parameters:

$$\alpha = 25 \quad \omega_0 = 15$$
$$s_1 = -5 \quad s_2 = -45$$

and since $\alpha > \omega_0$ immediately write the general form of the natural response:

$$v(t) = K_1 e^{-5t} + K_2 e^{-45t}$$

Only evaluation of the two constants K_1 and K_2 remains. Using the initial value of $v(t)$:

$$v(0) = 0$$

and therefore:

$$0 = K_1 + K_2$$

A second relation between K_1 and K_2 must be obtained by taking the derivative of $v(t)$ with respect to time, determining the initial condition of this derivative through the use of the remaining initial condition $i_L(0)=-8$, and equating the results.

Taking the derivative of both sides of the response, we get:

$$\frac{dv}{dt} = -5K_1 e^{-5t} - 45K_2 e^{-45t}$$

Evaluating the derivative at $t = 0$:

$$\left.\frac{dv}{dt}\right|_{t=0} = -5K_1 - 45K_2$$

We next pause to consider how the initial value of the derivative can be found numerically. This next step is always suggested by the derivative itself, dv/dt suggests capacitor current, for:

$$i_C = C\frac{dv}{dt}$$

Thus:

$$\left.\frac{dv}{dt}\right|_{t=0} = \frac{i_C(0)}{C} = \frac{-i_L(0) - i_R(0)}{C} = \frac{-i_L(0) - v(0)/R}{C} = 800\ \text{Vs}^{-1}$$

We thus have our second equation:

$$800 = -5K_1 - 45K_2$$

and we can solve our two equations in K_1 and K_2 simultaneously to get $K_1 = 20$ and $K_2 = -20$. Thus, the final numerical solution for the natural response is:

$$v(t) = 20\left(e^{-5t} - e^{-45t}\right)$$

We can interpret this result. We note that $v(t)$ is zero at $t = 0$, as required. We also interpret the first exponential term as having a time constant of $1/5\,\text{s}$ and the other exponential, a time constant of $1/45\,\text{s}$. Each starts with unity amplitude, but the second decays more rapidly – $v(t)$ is thus never negative. We thus have a response curve which is zero at $t = 0$, zero at $t = \infty$, and is never negative. Since it is not everywhere zero, it must have at least one maximum, and this can be found easily.

We differentiate the response:

$$\frac{dv}{dt} = 20\left(-5e^{-5t} + 45e^{-45t}\right)$$

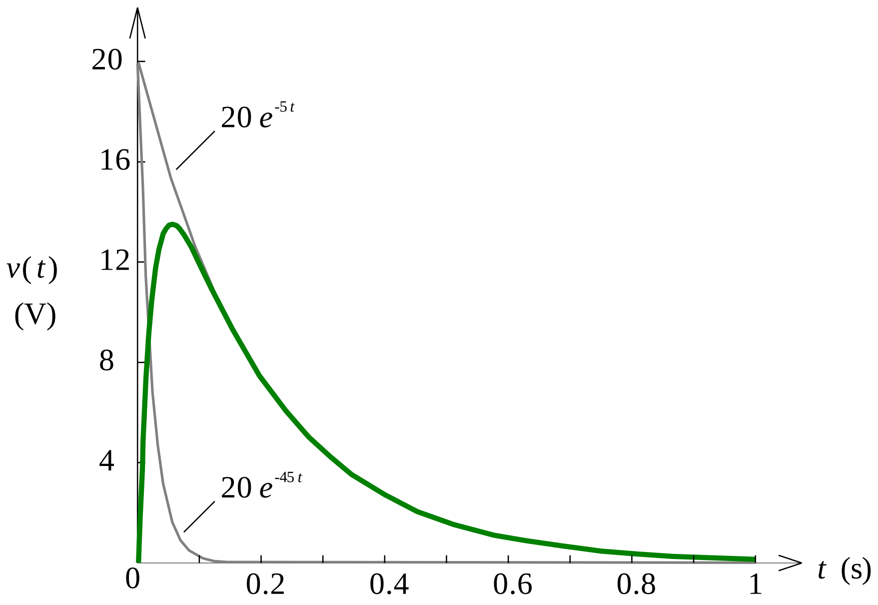
Then set the derivative to zero to determine the time t_p at which the voltage reaches its peak:

$$\begin{aligned} 0 &= -5e^{-5t_p} + 45e^{-45t_p} \\ e^{40t_p} &= 45/5 \\ t_p &= \frac{\ln 9}{40} \\ &= 54.93 \text{ ms} \end{aligned}$$

and obtain:

$$v(t_p) = 13.51 \text{ V}$$

A reasonable sketch of the response may be obtained by plotting the two exponential terms $20e^{-5t}$ and $20e^{-45t}$, taking their difference, and noting the peak value obtained above:



You can see that the dominant term is $20e^{-5t}$ for large t since the other term has effectively decayed to zero.

18.4 The Critically Damped Parallel *RLC* Circuit

The very special case of $\alpha = \omega_0$ is known as *critical damping*. Any one of the three elements *R*, *L* or *C* may be changed to obtain critical damping – however it is usual to select *R* to obtain critical damping in a circuit, and thus leave ω_0 unchanged.

For critical damping, we have real repeated roots :

$$\begin{aligned} s_1 &= -\alpha = -\omega_0 \\ s_2 &= -\alpha = -\omega_0 \end{aligned} \tag{18.34}$$

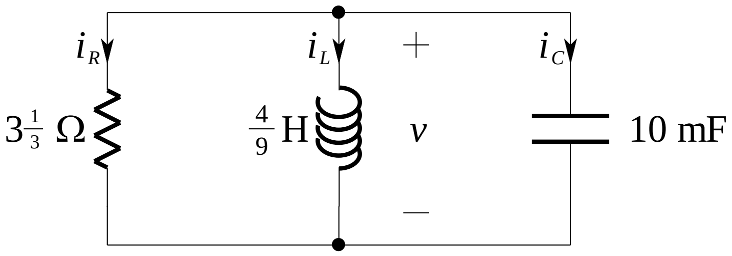
Thus, the natural response is of the form:

$$v_n = (K_1 + K_2 t)e^{-\alpha t} \tag{18.35}$$

In a real physical system it is impossible to obtain the exact conditions necessary for critical damping, since even an infinitesimally small change in the circuit's conditions will cause α to differ from ω_0. However, it can always be a design goal to obtain a critically damped response.

EXAMPLE 18.7 The Critically Damped Parallel *RLC* Circuit

We will use the same circuit as before as an example, but this time set $R = 3\frac{1}{3}\,\Omega$ to obtain critical damping:



The initial conditions are again $v(0)=0$ and $i_L(0)=-8$. In this case:

$$\alpha = \omega_0 = 15$$
$$s_1 = s_2 = -15$$

and since $\alpha = \omega_0$ immediately write the general form of the natural response:

$$v(t) = (K_1 + K_2 t)e^{-15t}$$

We establish the values of K_1 and K_2 by first imposing the initial condition on $v(t)$ itself, $v(0)=0$. Thus, $K_1 = 0$. This simple result occurs in this example because the initial value of the response was selected as zero.

The second initial condition must be applied to the derivative dv/dt just as in the overdamped case. We therefore differentiate, remembering that $K_1 = 0$:

$$\frac{dv}{dt} = K_2 t(-15)e^{-15t} + K_2 e^{-15t}$$

Evaluating the derivative at $t = 0$:

$$\left.\frac{dv}{dt}\right|_{t=0} = K_2$$

We next express the derivative in terms of the initial capacitor current:

$$\left.\frac{dv}{dt}\right|_{t=0} = \frac{i_C(0)}{C} = \frac{-i_L(0) - i_R(0)}{C} = \frac{-i_L(0) - v(0)/R}{C} = 800 \text{ Vs}^{-1}$$

and thus:

$$K_2 = 800$$

The natural response is therefore:

$$v(t) = 800te^{-15t}$$

Before plotting the response, we try to anticipate its form by qualitative reasoning. We note that $v(t)$ is zero at $t = 0$, as required. It is not immediately apparent that the response also approaches zero as *t* becomes infinitely large.

Using l'Hôpital's rule:

$$\lim_{t\to\infty} v(t) = 800 \lim_{t\to\infty} \frac{t}{e^{15t}} = 800 \lim_{t\to\infty} \frac{1}{15e^{15t}} = 0$$

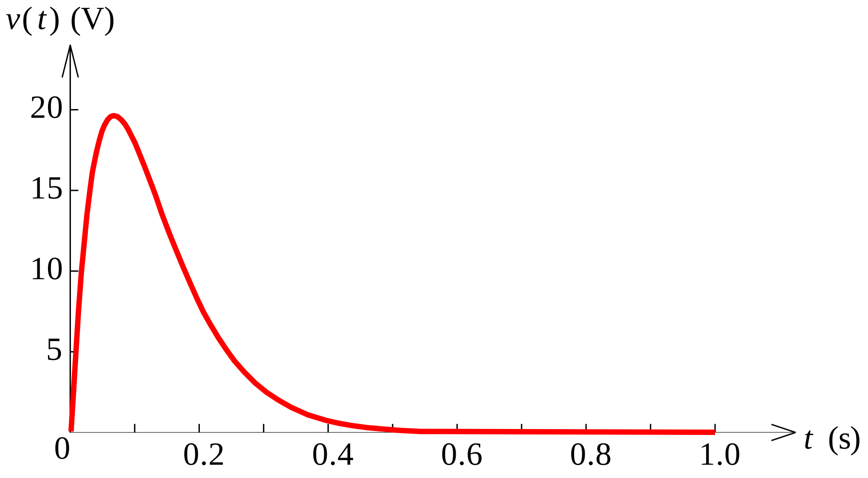
and once again we have a response which begins and ends at zero and has positive values at all other times.

A maximum value v_p again occurs at time t_p:

$$t_p = 1/15 \text{ s} \qquad \text{and} \qquad v_p = 19.62 \text{ V}$$

This maximum is larger than that obtained in the overdamped case and is a result of the smaller losses that occur in the larger resistor. The time of the maximum response is slightly larger than it was with overdamping.

The natural response curve for critical damping is shown below:



18.5 The Underdamped Parallel *RLC* Circuit

If we further increase the resistance R from the value we had at critical damping, whilst leaving L and C unchanged, the damping coefficient α decreases while ω_0 remains constant. We thus have $\alpha < \omega_0$, and the roots of the characteristic equation become:

$$\begin{aligned} s_1 &= -\alpha + \sqrt{\alpha^2 - \omega_0^2} = -\alpha + j\sqrt{\omega_0^2 - \alpha^2} \\ s_2 &= -\alpha - \sqrt{\alpha^2 - \omega_0^2} = -\alpha - j\sqrt{\omega_0^2 - \alpha^2} \end{aligned} \tag{18.36}$$

We now take the new square root, which is real for the underdamped case, and call it ω_d, the *damped natural frequency*:

$$\omega_d = \sqrt{\omega_0^2 - \alpha^2} \tag{18.37}$$

Thus, the roots are distinct complex conjugates and are located at:

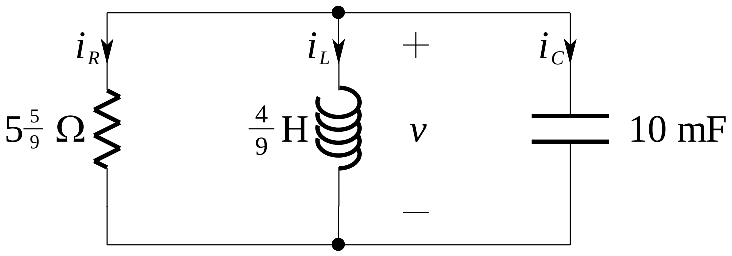
$$\begin{aligned} s_1 &= -\alpha + j\omega_d \\ s_2 &= -\alpha - j\omega_d \end{aligned} \tag{18.38}$$

Thus, the natural response is of the form:

$$v_n = e^{-\alpha t}\left(K_1 \cos \omega_d t + K_2 \sin \omega_d t\right) \tag{18.39}$$

EXAMPLE 18.8 The Underdamped Parallel *RLC* Circuit

We will use the same circuit as before as an example, but this time set $R = 5\frac{5}{9}\,\Omega$, L and C are unchanged:



Again we have $v(0) = 0$ and $i_L(0) = -8$. In this case:

$$\alpha = \frac{1}{2RC} = 9$$

$$\omega_0 = \frac{1}{\sqrt{LC}} = 15$$

and since $\alpha < \omega_0$ we identify:

$$\omega_d = \sqrt{\omega_0^2 - \alpha^2} = 12$$

and immediately write the general form of the natural response:

$$v(t) = e^{-9t}\left(K_1 \cos 12t + K_2 \sin 12t\right)$$

The determination of the two constants proceeds as before. Again, if $v(0) = 0$ then $K_1 = 0$. Hence:

$$v(t) = K_2 e^{-9t} \sin 12t$$

The derivative is:

$$\frac{dv}{dt} = 12K_2 e^{-9t} \cos 12t - 9K_2 e^{-9t} \sin 12t$$

and at $t = 0$:

$$\left.\frac{dv}{dt}\right|_{t=0} = 12K_2 = \frac{-i_L(0)}{C} = 800$$

The natural response is therefore:

$$v(t) = 200/3e^{-9t}\sin 12t$$

Notice that, as before, $v(t)$ is zero at $t = 0$, as required. The response also has a final value of zero because the exponential term vanishes for large values of *t*.

As *t* increases from zero through small positive values, $v(t)$ increases because the exponential term remains unable to damp the increase due to the sinusoidal term. But at a time t_p, the exponential function begins to decrease more rapidly than $\sin 12t$ is increasing, so $v(t)$ reaches a maximum v_p and begins to decrease. We should note that t_p is not the value of *t* for which $\sin 12t$ is a maximum, but must occur before $\sin 12t$ reaches its maximum value.

When $t = \pi/\omega_d = \pi/12$, $v(t) = 0$. For the interval $\pi/\omega_d < t < 2\pi/\omega_d$, the response is negative, becoming zero again at $t = 2\pi/\omega_d$.

Thus $v(t)$ is an *oscillatory* function of time and crosses the time axis an infinite number of times at $t = n\pi/\omega_d$ where *n* is any positive integer.

The oscillatory nature of the response becomes more noticeable as α decreases. If $\alpha = 0$, which corresponds to an infinitely large resistance, then $v(t)$ is an undamped sinusoid which oscillates with constant amplitude. We have merely assumed an initial energy in the circuit and have not provided any means to dissipate this energy. It is transferred from its initial location in the inductor to the capacitor, then returns to the inductor, and so on, forever.

Differentiation of the natural response locates the first maximum of $v(t)$,

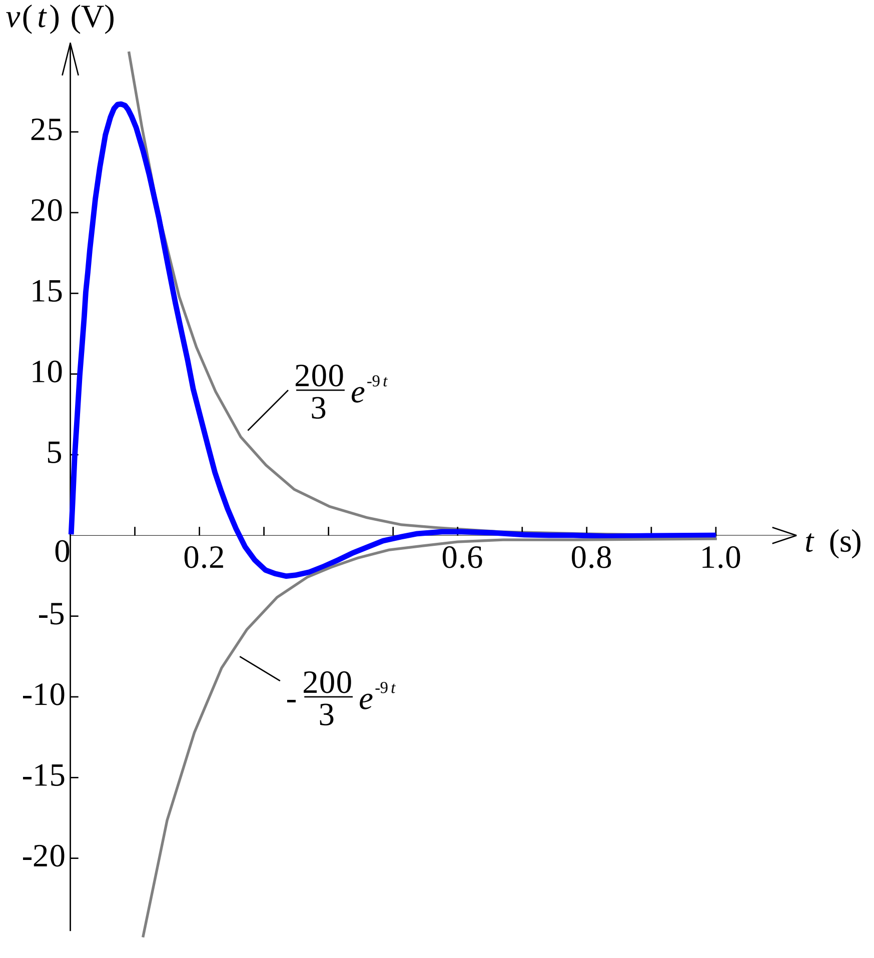
$$t_{p1} = 77.27 \text{ ms} \quad \text{and} \quad v_{p1} = 26.60 \text{ V}$$

the succeeding minimum,

$$t_{p2} = 339.1 \text{ ms} \quad \text{and} \quad v_{p2} = -2.522 \text{ V}$$

and so on.

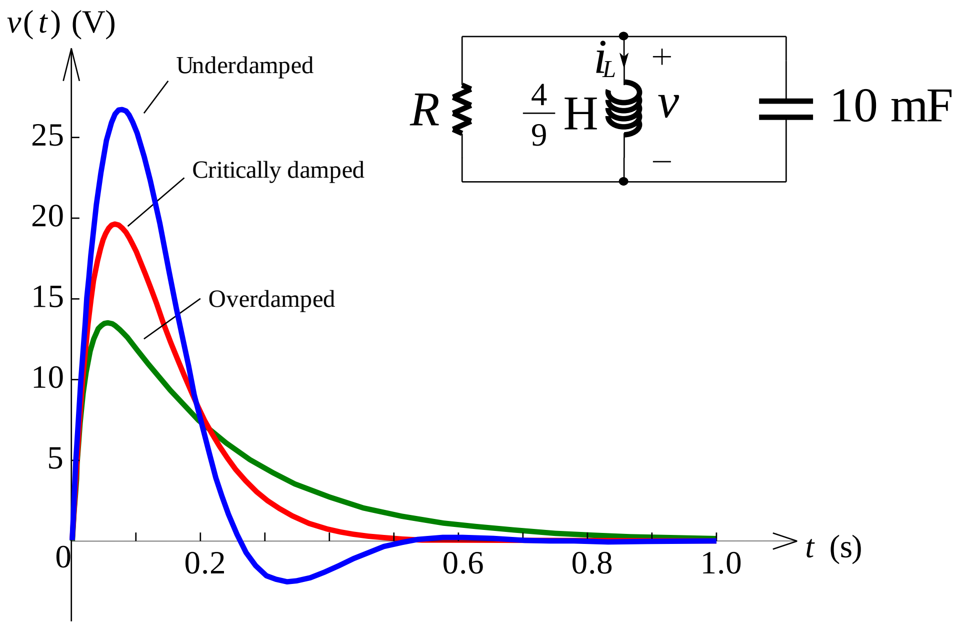
The natural response curve for the underdamped case is shown below:



Notice that the "envelope" of the damped sinusoid is given by $Ke^{-\alpha t}$.

18.6 Response Comparison

The overdamped, critically damped and underdamped responses for the example circuit are shown on the same graph below:



The table below shows the possibilities and names associated with the second-order natural response.

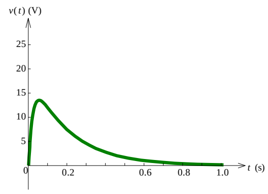
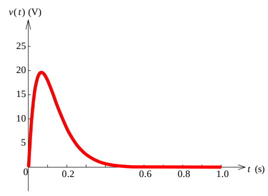
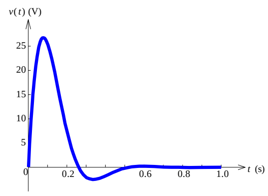
Condition	Criteria	Natural Response	Example
Overdamped	$\alpha > \omega_0$	$f_n = K_1 e^{s_1 t} + K_2 e^{s_2 t}$ $s_{1,2} = -\alpha \pm \sqrt{\alpha^2 - \omega_0^2}$	
Critically damped	$\alpha = \omega_0$	$f_n = (K_1 + K_2 t) e^{-\alpha t}$ $s_{1,2} = -\alpha$	
Underdamped	$\alpha < \omega_0$	$f_n = e^{-\alpha t}(K_1 \cos \omega_d t + K_2 \sin \omega_d t)$ $s_{1,2} = -\alpha \pm j\omega_d$ $\omega_d = \sqrt{\omega_0^2 - \alpha^2}$	

Table 18.2 – Second-Order Natural Responses

18.7 The Source-Free Series *RLC* Circuit

The series *RLC* circuit is the dual of the parallel *RLC* circuit, and this makes the analysis fairly simple. The figure below shows the source-free series *RLC* circuit:

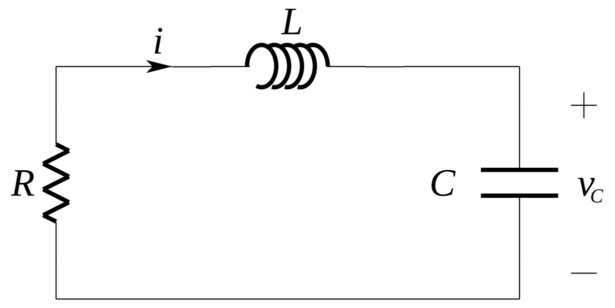


Figure 18.2

KVL around the circuit gives:

$$Ri + L\frac{di}{dt} + \frac{1}{C}\int_{t_0}^{t} i dt + v_C(t_0) = 0 \tag{18.40}$$

When both sides are differentiated once with respect to time and divided by *L* the result is the linear second-order homogeneous differential equation:

$$\frac{d^2 i}{dt^2} + \frac{R}{L}\frac{di}{dt} + \frac{1}{LC} i = 0 \tag{18.41}$$

This is the dual of Eq. (18.22). Thus, if we define:

$$\alpha = \frac{R}{2L} \qquad \text{(series)} \tag{18.42}$$

then we get the same characteristic equation as for the parallel *RLC* circuit, Eq. (18.28). It is now apparent that our discussion of the parallel *RLC* circuit is directly applicable to the series *RLC* circuit.

18.8 Complete Response of the *RLC* Circuit

Consider a series *RLC* circuit with a DC source that is connected by a switch that closes at $t = 0$:

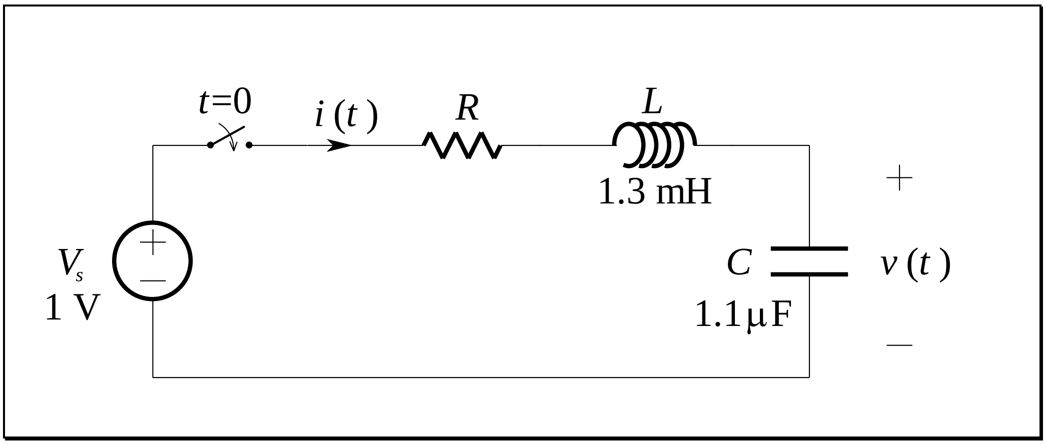


Figure 18.3

Assuming zero initial conditions, we would like to calculate the capacitor voltage $v(t)$ at $t = 120\ \mu\text{s}$ for the following values of resistance *R*:

1. $R = 330\ \Omega$

2. $R = 68.7552\ \Omega$

3. $R = 33\ \Omega$

18.8.1 Forced Response

First, we find the forced response. Since we have a DC source, we can find this part of the solution by replacing the inductance by a short circuit and the capacitance by an open circuit, i.e. we analyse the circuit under DC conditions. This is shown below:

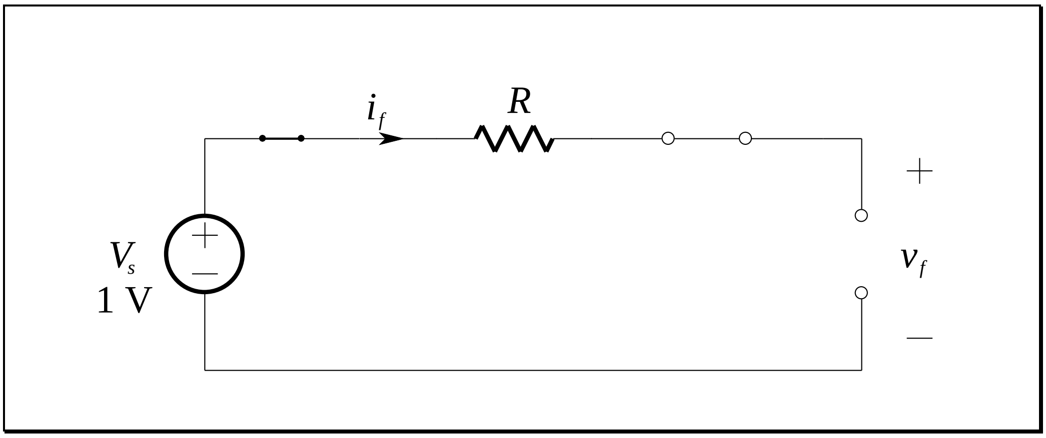


Figure 18.4

The current is zero, the voltage drop across the resistance is zero, and the voltage across the capacitance (an open circuit) is equal to the DC source voltage. Therefore, the forced response is:

$$v_f = V_s = 1\text{ V} \tag{18.43}$$

Notice that in this circuit the forced response for $v(t)$ is the same for all three values of resistance.

18.8.2 Natural Response

First, we can write an expression for the current in terms of the voltage across the capacitance:

$$i = C\frac{dv}{dt} \tag{18.44}$$

Then, we write a KVL equation for the circuit:

$$L\frac{di}{dt} + Ri + v = V_s u(t) \tag{18.45}$$

Using Eq. (18.44) to substitute for i, we get, for $t > 0$:

$$LC\frac{d^2v}{dt^2} + RC\frac{dv}{dt} + v = V_s \tag{18.46}$$

Dividing through by *LC*, we have:

$$\frac{d^2v}{dt^2} + \frac{R}{L}\frac{dv}{dt} + \frac{1}{LC}v = \frac{1}{LC}V_s \tag{18.47}$$

In *D* operator notation, the equation is:

$$\left(D^2 + \frac{R}{L}D + \frac{1}{LC}\right)v = \frac{1}{LC}V_s \tag{18.48}$$

Therefore, the characteristic equation is:

$$s^2 + \frac{R}{L}s + \frac{1}{LC} = 0 \tag{18.49}$$

If we let:

$$\alpha = \frac{R}{2L} \quad \text{and} \quad \omega_0 = \frac{1}{\sqrt{LC}} \tag{18.50}$$

then the characteristic equation can be written as:

$$s^2 + 2\alpha s + \omega_0^2 = 0 \tag{18.51}$$

which has general solutions:

$$\begin{aligned} s_1 &= -\alpha + \sqrt{\alpha^2 - \omega_0^2} \\ s_2 &= -\alpha - \sqrt{\alpha^2 - \omega_0^2} \end{aligned} \tag{18.52}$$

Next, we find the natural response and complete response for each value of *R*. For all three cases we have:

$$\omega_0 = \frac{1}{\sqrt{LC}} = 26444 \tag{18.53}$$

Case I – Overdamped $(R = 330\ \Omega)$

In this case, we get:

$$\alpha = \frac{R}{2L} = 126923 \tag{18.54}$$

Since we have $\alpha > \omega_0$, this is the *overdamped* case. The roots of the characteristic equation are given by:

$$\begin{aligned} s_1 &= -\alpha + \sqrt{\alpha^2 - \omega_0^2} \\ &= -126923 + \sqrt{126923^2 - 26444^2} \\ &= -2785 \end{aligned} \tag{18.55}$$

and

$$\begin{aligned} s_2 &= -\alpha - \sqrt{\alpha^2 - \omega_0^2} \\ &= -251060 \end{aligned} \tag{18.56}$$

For the overdamped case, the natural response has the form:

$$v_n = K_1 e^{s_1 t} + K_2 e^{s_2 t} \tag{18.57}$$

Adding the forced response given by Eq. (18.43) to the natural response, we obtain the complete response:

$$v(t) = 1 + K_1 e^{s_1 t} + K_2 e^{s_2 t} \tag{18.58}$$

Now, we must find values of K_1 and K_2 so the solution matches the known initial conditions in the circuit. It was given that the initial voltage on the capacitance is zero, hence $v(0)=0$. Evaluating Eq. (18.58) at $t=0$, we obtain:

$$0=1+K_1+K_2 \tag{18.59}$$

Furthermore, the initial current was given as $i(0)=0$. Since the current through the capacitance is given by:

$$i=C\frac{dv}{dt} \tag{18.60}$$

we conclude that:

$$\frac{dv}{dt}=0 \tag{18.61}$$

Taking the derivative of Eq. (18.58) and evaluating at $t=0$, we have:

$$s_1K_1+s_2K_2=0 \tag{18.62}$$

Now, we can solve Eqs. (18.59) and (18.62) for the values of K_1 and K_2. The results are $K_1=-1.0112$ and $K_2=0.01122$. Substituting these values into Eq. (18.58), we have the solution:

$$\boxed{v(t)=1-1.0112\,e^{-2785t}+0.01122\,e^{-251060t}} \tag{18.63}$$

Evaluating this expression at $t=120\ \mu\text{s}$, we get:

$$v(120\mu\text{s})=0.27607\ \text{V} \tag{18.64}$$

18.34

Case II – Critically Damped $(R = 68.7552\ \Omega)$

In this case, we get:

$$\alpha = \frac{R}{2L} = 26444 \tag{18.65}$$

Since $\alpha = \omega_0$, this is the *critically damped* case. The roots of the characteristic equation are given by:

$$s_1 = s_2 = -\alpha + \sqrt{\alpha^2 - \omega_0^2} = -\alpha = -26444 \tag{18.66}$$

The natural response has the form of:

$$v_n = (K_1 + K_2 t)e^{s_1 t} \tag{18.67}$$

Adding the forced response to the natural response, we find that:

$$v(t) = 1 + (K_1 + K_2 t)e^{s_1 t} \tag{18.68}$$

As in Case I, the initial conditions require $v(0) = 0$ and $dv(0)/dt = 0$. Thus, substituting $t = 0$ into Eq. (18.68), we get:

$$0 = 1 + K_1 \tag{18.69}$$

Differentiating Eq. (18.68) and substituting $t = 0$ yields:

$$s_1 K_1 + K_2 = 0 \tag{18.70}$$

Solving Eqs. (18.69) and (18.70) yields $K_1 = -1$ and $K_2 = -26444$. Thus the complete response is:

$$\boxed{v(t) = 1 - e^{-26444t} - 26444\, te^{-26444t}} \tag{18.71}$$

Evaluating this expression at $t = 120\ \mu\text{s}$, we get:

$$v(120\mu\text{s}) = 0.82528\ \text{V} \tag{18.72}$$

Case III – Underdamped $(R = 33\,\Omega)$

For this value of resistance, we have:

$$\alpha = \frac{R}{2L} = 12692 \tag{18.73}$$

Since $\alpha < \omega_0$, this is the *underdamped* case. Using:

$$\omega_d = \sqrt{\omega_0^2 - \alpha^2} \tag{18.74}$$

we compute the damped natural frequency:

$$\omega_d = 23199 \tag{18.75}$$

The natural response has the form:

$$v_n = e^{-\alpha t}\left(K_1 \cos\omega_d t + K_2 \sin\omega_d t\right) \tag{18.76}$$

Adding the forced response found earlier to the natural response, we obtain the complete response:

$$v(t) = 1 + e^{-\alpha t}\left(K_1 \cos\omega_d t + K_2 \sin\omega_d t\right) \tag{18.77}$$

As in the previous cases, the initial conditions are $v(0) = 0$ and $dv(0)/dt = 0$. Evaluating Eq. (18.77) at $t = 0$, we obtain:

$$0 = 1 + K_1 \tag{18.78}$$

Differentiating Eq. (18.77), we get:

$$\begin{aligned}\frac{dv(t)}{dt} &= e^{-\alpha t}\left(-\omega_d K_1 \sin\omega_d t + \omega_d K_2 \cos\omega_d t\right)\\ &\quad - \alpha e^{-\alpha t}\left(K_1 \cos\omega_d t + K_2 \sin\omega_d t\right)\end{aligned} \tag{18.79}$$

Evaluating at $t=0$, we have:

$$\omega_d K_2 - \alpha K_1 = 0 \tag{18.80}$$

Solving Eqs. (18.78) and (18.80), we obtain $K_1 = -1$ and $K_2 = -0.5471$. Thus, the complete solution is:

$$\begin{aligned} v(t) &= 1 - e^{-12692t}\cos 23199t \\ &\quad - 0.5471 e^{-12692t}\sin 23199t \end{aligned} \tag{18.81}$$

Evaluating this expression at $t = 120\ \mu\text{s}$, we get:

$$v(120\mu\text{s}) = 1.16248\ \text{V} \tag{18.82}$$

For this underdamped case, the response will look like:

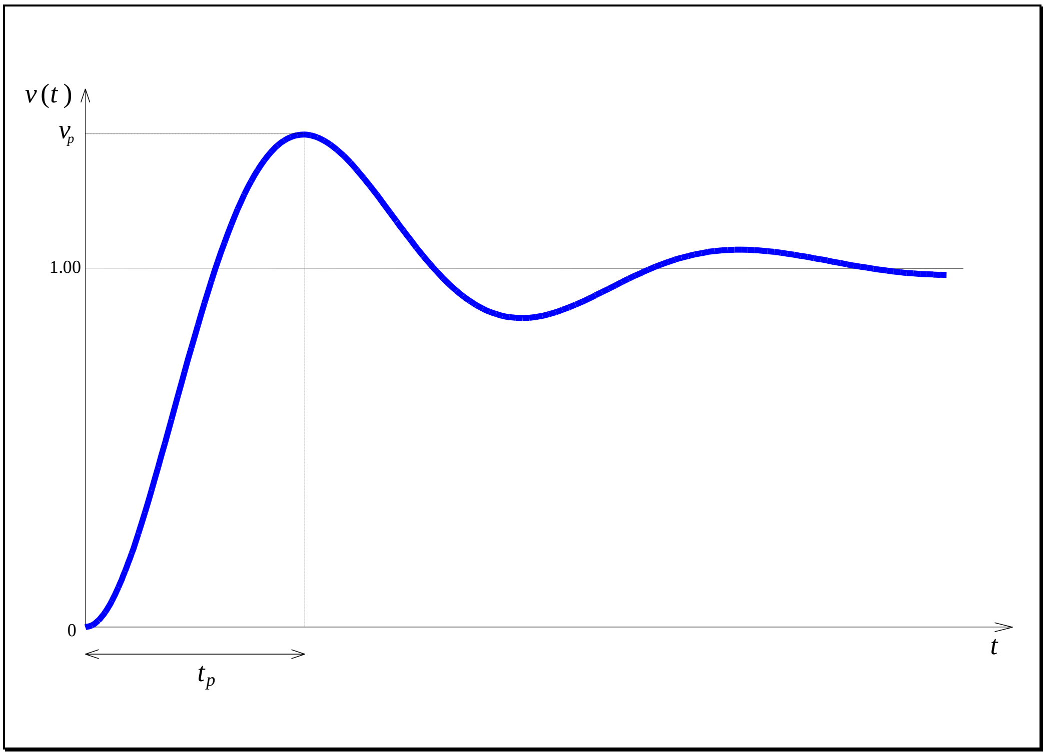


Figure 18.5

18.8.3 Maximum Value and Peak Time

We can evaluate the maximum value, and the time to reach the maximum value (called the peak time), by observing that the derivative of $v(t)$ is zero at relative maxima and minima.

We already found the derivative of $v(t)$:

$$\begin{aligned}\frac{dv(t)}{dt} &= -\alpha K_1 e^{-\alpha t}\cos\omega_d t - \omega_d K_1 e^{-\alpha t}\sin\omega_d t \\ &\quad - \alpha K_2 e^{-\alpha t}\sin\omega_d t + \omega_d K_2 e^{-\alpha t}\cos\omega_d t\end{aligned} \tag{18.83}$$

Grouping terms, this can be written:

$$\frac{dv(t)}{dt} = e^{-\alpha t}\left[(\omega_d K_2 - \alpha K_1)\cos\omega_d t - (\omega_d K_1 + \alpha K_2)\sin\omega_d t\right] \tag{18.84}$$

By applying the $dv(0)/dt = 0$ initial condition, we can see that:

$$\omega_d K_2 - \alpha K_1 = 0 \tag{18.85}$$

Therefore, the derivative is:

$$\frac{dv(t)}{dt} = -e^{-\alpha t}(\omega_d K_1 + \alpha K_2)\sin\omega_d t \tag{18.86}$$

Equating this to zero, we get:

$$\sin\omega_d t = 0 \tag{18.87}$$

Therefore, the times of relative maxima and minima are:

$$t_m = \frac{n\pi}{\omega_d} \tag{18.88}$$

The case of $n=1$ corresponds to the peak time. Thus:

$$t_p = \frac{\pi}{\omega_d} \tag{18.89}$$

Substituting this value into the output expression:

$$v(t) = 1 + e^{-\alpha t}\left(K_1 \cos \omega_d t + K_2 \sin \omega_d t\right) \tag{18.90}$$

we get:

$$v_p = v(t_p) = 1 - K_1 e^{-\alpha\pi/\omega_d} \tag{18.91}$$

Putting values of this particular case into Eqs. (18.89) and (18.91), we get:

$$t_p = \frac{\pi}{23199} = 0.13542 \text{ ms} \tag{18.92}$$

and:

$$v_p = 1 + e^{-\frac{12692\pi}{23199}} = 1.17929 \text{ V} \tag{18.93}$$

Thus, the output exhibits a "peak overshoot" of about 18%.

It is important to note that the formula obtained for the peak time, Eq. (18.89), is **only** valid for this particular case of zero initial conditions. In the general case it can be shown that the times of relative maxima and minima for an underdamped response are given by:

$$t_m = \frac{1}{\omega_d} \tan^{-1}\left(\frac{\omega_d K_2 - \alpha K_1}{\omega_d K_1 + \alpha K_2}\right) \tag{18.94}$$

from which Eq. (18.89) is a special case. *Derive this general formula*.

18.9 Summary

- For the *RLC* circuit, we define the undamped natural frequency:

$$\omega_0 = \frac{1}{\sqrt{LC}}$$

- For the *RLC* circuit, we define the exponential damping coefficient:

$$\alpha = \frac{1}{2RC} \quad \text{parallel}$$

$$\alpha = \frac{R}{2L} \quad \text{series}$$

- The *RLC* circuit exhibits three different forms for the natural response:

Overdamped ($\alpha > \omega_0$ **)**

The natural response is of the form:

$$f_n = K_1 e^{s_1 t} + K_2 e^{s_2 t}$$

Critically Damped ($\alpha = \omega_0$ **)**

The natural response is of the form:

$$f_n = (K_1 + K_2 t)e^{-\alpha t}$$

Underdamped ($\alpha < \omega_0$ **)**

The natural response is of the form:

$$f_n = e^{-\alpha t}(K_1 \cos \omega_d t + K_2 \sin \omega_d t), \qquad \omega_d = \sqrt{\omega_0^2 - \alpha^2}$$

- The complete response is the sum of the forced response and the natural response.

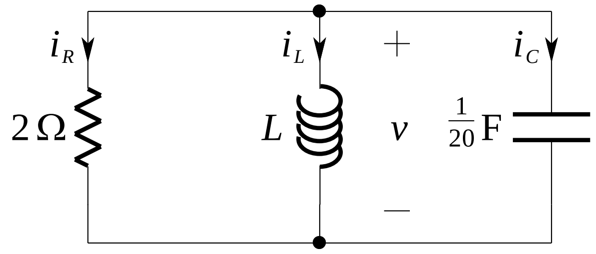
18.10 References

Hayt, W. & Kemmerly, J.: *Engineering Circuit Analysis*, 3rd *Ed.*, McGraw-Hill, 1984.

Exercises

1.

Consider the circuit shown below:

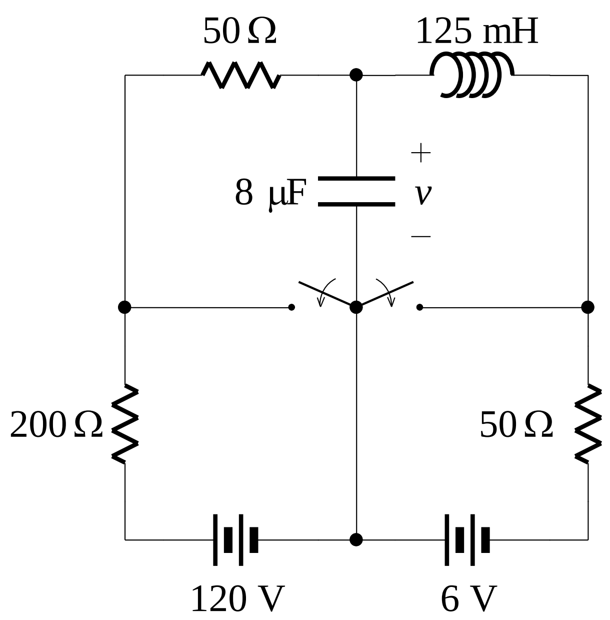


Let $L = 1.25$ H and determine $v(t)$ if $v(0^+) = 100$ V and:

(a) $i_C(0^+) = 20$ A   (b) $i_L(0^+) = 20$ A

2.

Consider the circuit shown below:



Both switches close at $t = 0$ after having been open for a very long time.

(a) Find $v(t)$.

(b) Determine the maximum and minimum values of $v(t)$.

3.

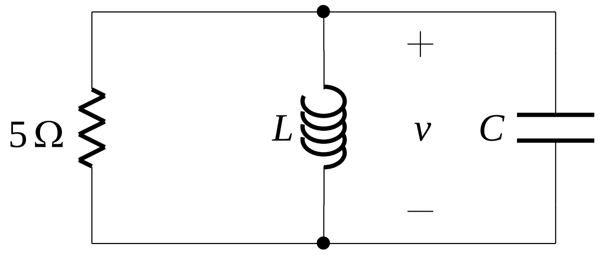
The capacitor voltage in a parallel *RLC* circuit that is critically damped is given by $v(t)=1000e^{-500t}(t-0.01)$ V. If the energy stored in the capacitor is 2 mJ at $t=0$, find:

(a) R.

(b) The initial energy stored in the inductor.

4.

Consider the circuit shown below:



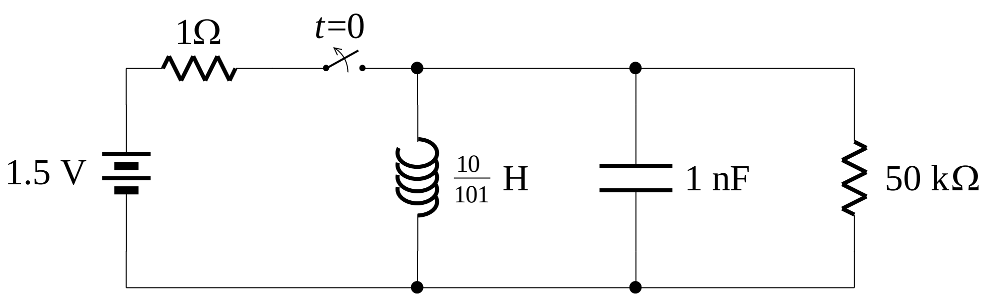
The voltage is given by $v(t)=e^{-7t}(20\cos 24t+5\sin 24t)$ V for $t\geq 0$. Find:

(a) L and C.

(b) The initial energy stored in the circuit.

5.

Consider the circuit shown below:

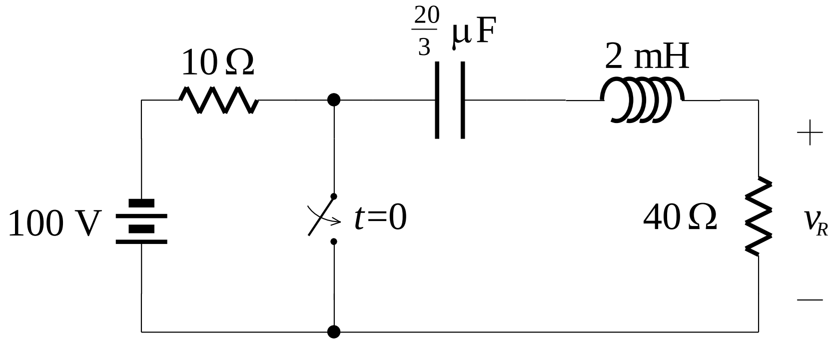


What is the maximum voltage magnitude present across the switch after $t=0$? (Note that it is much safer to solve this problem analytically than to do so experimentally.)

18.42

6.

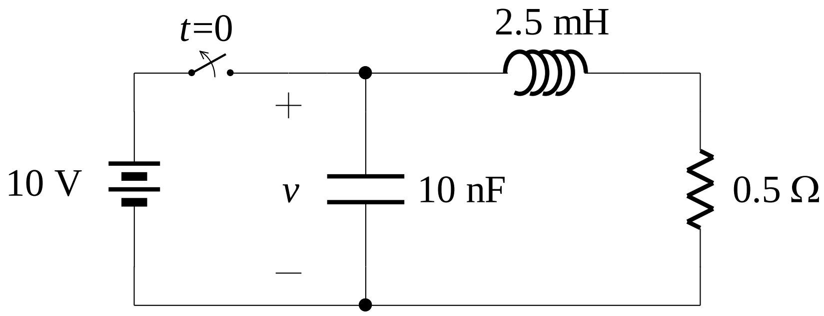
Consider the circuit shown below:



The switch closes at $t=0$. Find $v_R(t)$.

7.

Consider the circuit shown below:



The switch has been closed for hours. It is opened at $t=0$. Show that a 10 V battery can create a high voltage by finding v at $t=2.5\pi$ μs .

8.

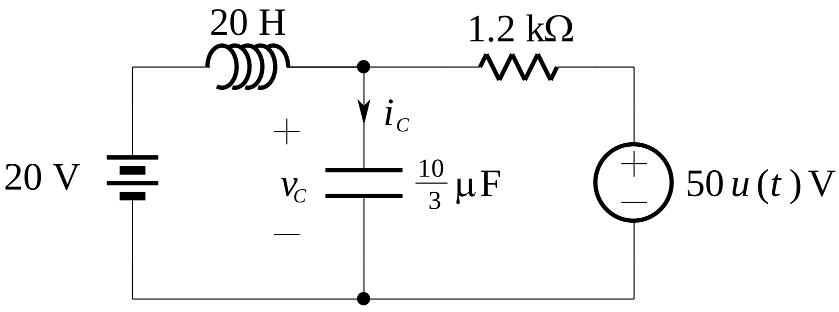
A 2.5 H inductor, a $4\,\Omega$ resistor, and a 25 mF capacitor are in parallel. An 80 V battery is then placed in series with the inductor.

(a) After a long time has passed, find the energy stored in the inductor and in the capacitor.

(b) The battery voltage drops suddenly to 40 V at $t=0$. Find the energy stored in the inductor and in the capacitor 0.25 s later.

9.

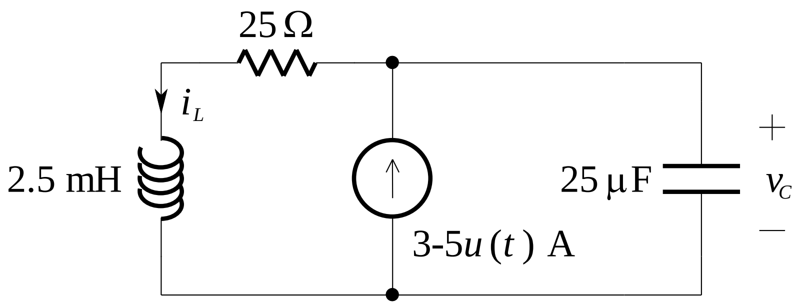
Consider the circuit shown below:



Find $v_C(t)$ and $i_C(t)$.

10.

Consider the circuit shown below:



Find:

(a) $v_C(t)$

(b) $i_L(t)$

11.

A 5 mH inductor, a $50\,\mu\text{F}$ capacitor, and a $25\,\Omega$ resistor are in series with a voltage source $v_s(t)$. The source voltage is zero prior to $t=0$. At $t=0$ it jumps to 75 V, at $t=1\,\text{ms}$ it drops to zero, at $t=2\,\text{ms}$ it again jumps to 75 V, and it continues in this periodic manner thereafter. Find the source current at:

(a) $t=0$ (b) $t=1\,\text{ms}$ (c) $t=2\,\text{ms}$

William Thomson (Lord Kelvin) (1824-1907)

William Thomson was probably the first true electrical engineer. His engineering was firmly founded on a solid bedrock of mathematics. He invented, experimented, advanced the state-of-the-art, was entrepreneurial, was a businessman, had a multi-disciplinary approach to problems, held office in the professional body of his day (the Royal Society), published papers, gave lectures to lay people, strived for an understanding of basic physical principles and exploited that knowledge for the benefit of mankind.

William Thomson was born in Belfast, Ireland. His father was a professor of engineering. When Thomson was 8 years old his father was appointed to the chair of mathematics at the University of Glasgow. By age 10, William Thomson was attending Glasgow University. He studied astronomy, chemistry and natural philosophy (physics, heat, electricity and magnetism). Prizes in Greek, logic (philosophy), mathematics, astronomy and physics marked his progress. In 1840 he read Fourier's *The Analytical Theory of Heat* and wrote:

> *…I had become filled with the utmost admiration for the splendour and poetry of Fourier… I took Fourier out of the University Library; and in a fortnight I had mastered it - gone right through it.*

At the time, lecturers at Glasgow University took a strong interest in the approach of the French mathematicians towards physical science, such as Lagrange, Laplace, Legendre, Fresnel and Fourier. In 1840 Thomson also read Laplace's *Mécanique Céleste* and visited Paris.

In 1841 Thomson entered Cambridge and in the same year he published a paper on *Fourier's expansions of functions in trigonometrical series*. A more important paper *On the uniform motion of heat and its connection with the mathematical theory of electricity* was published in 1842.

The examinations in Cambridge were fiercely competitive exercises in problem solving against the clock. The best candidates were trained as for an athletics contest. Thomson (like Maxwell later) came second. A day before he left Cambridge, his coach gave him two copies of Green's *Essay on the*

Application of Mathematical Analysis to the Theories of Electricity and Magnetism.

After graduating, he moved to Paris on the advice of his father and because of his interest in the French approach to mathematical physics. Thomson began trying to bring together the ideas of Faraday, Coulomb and Poisson on electrical theory. He began to try and unify the ideas of “action-at-a-distance”, the properties of the “ether” and ideas behind an “electrical fluid”. He also became aware of Carnot’s view of heat.

In 1846, at the age of twenty two, he returned to Glasgow on a wave of testimonials from, among others, De Morgan, Cayley, Hamilton, Boole, Sylvester, Stokes and Liouville, to take up the post of professor of natural philosophy. In 1847-49 he collaborated with Stokes on hydrodynamic studies, which Thomson applied to electrical and atomic theory. In electricity Thomson provided the link between Faraday and Maxwell. He was able to mathematise Faraday’s laws and to show the formal analogy between problems in heat and electricity. Thus the work of Fourier on heat immediately gave rise to theorems on electricity and the work of Green on potential theory immediately gave rise to theorems on heat flow. Similarly, methods used to deal with linear and rotational displacements in elastic solids could be applied to give results on electricity and magnetism. The ideas developed by Thomson were later taken up by Maxwell in his new theory of electromagnetism.

Thomson’s other major contribution to fundamental physics was his combination of the almost forgotten work of Carnot with the work of Joule on the conservation of energy to lay the foundations of thermodynamics. The thermodynamical studies of Thomson led him to propose an absolute temperature scale in 1848 (The Kelvin absolute temperature scale, as it is now known, was defined much later after conservation of energy was better understood).

The Age of the Earth

In the first decades of the nineteenth century geological evidence for great changes in the past began to build up. Large areas of land had once been under water, mountain ranges had been thrown up from lowlands and the evidence of fossils showed the past existence of species with no living counterparts. Lyell, in his *Principles of Geology*, sought to explain these changes "by causes now in operation". According to his theory, processes – such as slow erosion by wind and water; gradual deposition of sediment by rivers; and the cumulative effect of earthquakes and volcanic action – combined over very long periods of time to produce the vast changes recorded in the Earth's surface. Lyell's so-called 'uniformitarian' theory demanded that the age of the Earth be measured in terms of hundreds of millions and probably in terms of billions of years. Lyell was able to account for the disappearance of species in the geological record but not for the appearance of new species. A solution to this problem was provided by Charles Darwin (and Wallace) with his theory of evolution by natural selection. Darwin's theory also required vast periods of time for operation. For natural selection to operate, the age of the Earth had to be measured in many hundreds of millions of years.

Such demands for vast amounts of time run counter to the laws of thermodynamics. Every day the sun radiates immense amounts of energy. By the law of conservation of energy there must be some source of this energy. Thomson, as one of the founders of thermodynamics, was fascinated by this problem. Chemical processes (such as the burning of coal) are totally insufficient as a source of energy and Thomson was forced to conclude that gravitational potential energy was turned into heat as the sun contracted. On this assumption his calculations showed that the Sun (and therefore the Earth) was around 100 million years old.

However, Thomson's most compelling argument concerned the Earth rather than the Sun. It is well known that the temperature of the Earth increases with depth and

> *this implies a continual loss of heat from the interior, by conduction outwards through or into the upper crust. Hence, since the upper crust does not become hotter from year to year there must be a…loss of heat from the whole earth. It is possible that no cooling may result from this loss of heat but only an exhaustion of potential energy which in this case could scarcely be other than chemical.*

Since there is no reasonable mechanism to keep a chemical reaction going at a steady pace for millions of years, Thomson concluded "…that the earth is merely a warm chemically inert body cooling". Thomson was led to believe that the Earth was a solid body and that it had solidified at a more or less uniform temperature. Taking the best available measurements of the conductivity of the Earth and the rate of temperature change near the surface, he arrived at an estimate of 100 million years as the age of the Earth (confirming his calculations of the Sun's age).

The problems posed to Darwin's theory of evolution became serious as Thomson's arguments sank in. In the fifth edition of *The Origin of Species*, Darwin attempted to adjust to the new time scale by allowing greater scope for evolution by processes other than natural selection. Darwin was forced to ask for a suspension of judgment of his theory and in the final chapter he added

> *With respect to the lapse of time not having been sufficient since our planet was consolidated for the assumed amount of organic change, and this objection, as argued by [Thomson], is probably one of the gravest yet advanced, I can only say, firstly that we do not know at what rate species change as measured by years, and secondly, that many philosophers are not yet willing to admit that we know enough of the constitution of the universe and of the interior of our globe to speculate with safety on its past duration.*

(Darwin, *The Origin of Species*, Sixth Edition, p.409)

The chief weakness of Thomson's arguments was exposed by Huxley

A variant of the adage: "Garbage in equals garbage out".

> *…this seems to be one of the many cases in which the admitted accuracy of mathematical processes is allowed to throw a wholly inadmissible appearance of authority over the results obtained by them. Mathematics may be compared to a mill of exquisite workmanship, which grinds you stuff of any degree of fineness; but nevertheless, what you get out depends on what you put in; and as the grandest mill in the world will not extract wheat-flour from peascods, so pages of formulae will not get a definite result out of loose data.*

(*Quarterly Journal of the Geological Society of London*, Vol. 25, 1869)

However, Thomson's estimates were the best available and for the next thirty years geology took its time from physics, and biology took its time from geology. But Thomson and his followers began to adjust his first estimate down until at the end of the nineteenth century the best physical estimates of the age of the Earth and Sun were about 20 million years whilst the minimum the geologists could allow was closer to Thomson's original 100 million years.

Then in 1904 Rutherford announced that the radioactive decay of radium was accompanied by the release of immense amounts of energy and speculated that this could replace the heat lost from the surface of the Earth.

> *The discovery of the radioactive elements…thus increases the possible limit of the duration of life on this planet, and allows the time claimed by the geologist and biologist for the process of evolution.*

(Rutherford quoted in Burchfield, p.164)

A problem for the geologists was now replaced by a problem for the physicists. The answer was provided by a theory which was just beginning to be gossiped about. Einstein's theory of relativity extended the principle of conservation of energy by taking matter as a form of energy. It is the conversion of matter to heat which maintains the Earth's internal temperature and supplies the energy radiated by the sun. The ratios of lead isotopes in the Earth compared to meteorites now leads geologists to give the Earth an age of about 4.55 billion years.

The Transatlantic Cable

The invention of the electric telegraph in the 1830s led to a network of telegraph wires covering England, western Europe and the more settled parts of the USA. The railroads, spawned by the dual inventions of steam and steel, were also beginning to crisscross those same regions. It was vital for the smooth and safe running of the railroads, as well as the running of empires, to have speedy communication.

Attempts were made to provide underwater links between the various separate systems. The first cable between Britain and France was laid in 1850. The operators found the greatest difficulty in transmitting even a few words. After 12 hours a trawler accidentally caught and cut the cable. A second, more heavily armoured cable was laid and it was a complete success. The short lines worked, but the operators found that signals could not be transmitted along submarine cables as fast as along land lines without becoming confused.

In spite of the record of the longer lines, the American Cyrus J. Fields proposed a telegraph line linking Europe and America. Oceanographic surveys showed that the bottom of the Atlantic was suitable for cable laying. The connection of existing land telegraph lines had produced a telegraph line of the length of the proposed cable through which signals had been passed extremely rapidly. The British government offered a subsidy and money was rapidly raised.

Faraday had predicted signal retardation but he and others like Morse had in mind a model of a submarine cable as a hosepipe which took longer to fill with water (signal) as it got longer. The remedy was thus to use a thin wire (so that less electricity was needed to charge it) and high voltages to push the signal through. Faraday's opinion was shared by the electrical adviser to the project, Dr Whitehouse (a medical doctor).

Thomson's researches had given him a clearer mathematical picture of the problem. The current in a telegraph wire in air is approximately governed by the wave equation. A pulse on such a wire travels at a well defined speed with no change of shape or magnitude with time. Signals can be sent as close together as the transmitter can make them and the receiver distinguish them.

In undersea cables of the type proposed, capacitive effects dominate and the current is approximately governed by the diffusion (i.e. heat) equation. This equation predicts that electric pulses will last for a time that is proportional to the length of the cable squared. If two or more signals are transmitted within this time, the signals will be jumbled at the receiver. In going from submarine cables of 50 km length to cables of length 2400 km, retardation effects are 2500 times worse. Also, increasing the voltage makes this jumbling (called intersymbol interference) worse. Finally, the diffusion equation shows that the wire should have as large a diameter as possible (small resistance).

Whitehouse, whose professional reputation was now involved, denied these conclusions. Even though Thomson was on the board of directors of Field's company, he had no authority over the technical advisers. Moreover the production of the cable was already underway on principles contrary to Thomson's. Testing the cable, Thomson was astonished to find that some sections conducted only half as well as others, even though the manufacturers were supplying copper to the then highest standards of purity.

Realising that the success of the enterprise would depend on a fast, sensitive detector, Thomson set about to invent one. The problem with an ordinary galvanometer is the high inertia of the needle. Thomson came up with the mirror galvanometer in which the pointer is replaced by a beam of light.

In a first attempt in 1857 the cable snapped after 540 km had been laid. In 1858, Europe and America were finally linked by cable. On 16 August it carried a 99-word message of greeting from Queen Victoria to President Buchanan. But that 99-word message took 16½ hours to get through. In vain, Whitehouse tried to get his receiver to work. Only Thomson's galvanometer was sensitive enough to interpret the minute and blurred messages coming through. Whitehouse ordered that a series of huge two thousand volt induction coils be used to try to push the message through faster – after four weeks of this treatment the insulation finally failed; 2500 tons of cable and £350 000 of capital lay useless on the ocean floor.

In 1859 eighteen thousand kilometres of undersea cable had been laid in other parts of the world, and only five thousand kilometres were operating. In 1861 civil war broke out in the United States. By 1864 Field had raised enough capital for a second attempt. The cable was designed in accordance with Thomson's theories. Strict quality control was exercised: the copper was so pure that for the next 50 years 'telegraphist's copper' was the purest available.

Once again the British Government supported the project – the importance of quick communication in controlling an empire was evident to everybody. The new cable was mechanically much stronger but also heavier. Only one ship was large enough to handle it and that was Brunel's *Great Eastern*. She was fives time larger than any other existing ship.

This time there was a competitor. The Western Union Company had decided to build a cable along the overland route across America, Alaska, the Bering Straits, Siberia and Russia to reach Europe the long way round. The commercial success of the cable would therefore depend on the rate at which messages could be transmitted. Thomson had promised the company a rate of 8 or even 12 words a minute. Half a million pounds was being staked on the correctness of the solution of a partial differential equation.

In 1865 the *Great Eastern* laid cable for nine days, but after 2000 km the cable parted. After two weeks of unsuccessfully trying to recover the cable, the expedition left a buoy to mark the spot and sailed for home. Since communication had been perfect up until the final break, Thomson was confident that the cable would do all that was required. The company decided to build and lay a new cable and then go back and complete the old one.

Cable laying for the third attempt started on 12 July 1866 and the cable was landed on the morning of the 27^{th}. On the 28^{th} the cable was open for business and earned £1000. Western Union ordered all work on their project to be stopped at a loss of $3 000 000.

On 1 September after three weeks of effort the old cable was recovered and on 8 September a second perfect cable linked America and Europe. A wave of knighthoods swept over the engineers and directors. The patents which Thomson held made him a wealthy man.

For his work on the transatlantic cable Thomson was created Baron Kelvin of Largs in 1892. The Kelvin is the river which runs through the grounds of Glasgow University and Largs is the town on the Scottish coast where Thomson built his house.

Other Achievements

There are many other factors influencing local tides – such as channel width – which produce phenomena akin to resonance in the tides. One example of this is the narrow Bay of Fundy, between Nova Scotia and New Brunswick, where the tide can be as high as 21m. In contrast, the Mediterranean Sea is almost tideless because it is a broad body of water with a narrow entrance.

Thomson worked on several problems associated with navigation – sounding machines, lighthouse lights, compasses and the prediction of tides. Tides are primarily due to the gravitational effects of the Moon, Sun and Earth on the oceans but their theoretical investigation, even in the simplest case of a single ocean covering a rigid Earth to a uniform depth, is very hard. Even today, the study of only slightly more realistic models is only possible by numerical computer modelling. Thomson recognised that the forces affecting the tides change periodically. He then approximated the height of the tide by a trigonometric polynomial – a Fourier series with a finite number of terms. The coefficients of the polynomial required calculation of the Fourier series coefficients by numerical integration – a task that "…required not less than twenty hours of calculation by skilled arithmeticians." To reduce this labour Thomson designed and built a machine which would trace out the predicted height of the tides for a year in a few minutes, given the Fourier series coefficients.

Michelson (of Michelson-Morley fame) was to build a better machine that used up to 80 Fourier series coefficients. The production of 'blips' at discontinuities by this machine was explained by Gibbs in two letters to *Nature*. These 'blips' are now referred to as the "Gibbs phenomenon".

Thomson also built another machine, called the *harmonic analyser*, to perform the task "which seemed to the Astronomer Royal so complicated and difficult that no machine could master it' of computing the Fourier series coefficients from the record of past heights. This was the first major victory in the struggle "to substitute brass for brain" in calculation.

Thomson introduced many teaching innovations to Glasgow University. He introduced laboratory work into the degree courses, keeping this part of the work distinct from the mathematical side. He encouraged the best students by offering prizes. There were also prizes which Thomson gave to the student that he considered most deserving.

Thomson worked in collaboration with Tait to produce the now famous text *Treatise on Natural Philosophy* which they began working on in the early 1860s. Many volumes were intended but only two were ever written which cover kinematics and dynamics. These became standard texts for many generations of scientists.

In later life he developed a complete range of measurement instruments for physics and electricity. He also established standards for all the quantities in use in physics. In all he published over 300 major technical papers during the 53 years that he held the chair of Natural Philosophy at the University of Glasgow.

During the first half of Thomson's career he seemed incapable of being wrong while during the second half of his career he seemed incapable of being right. This seems too extreme a view, but Thomson's refusal to accept atoms, his opposition to Darwin's theories, his incorrect speculations as to the age of the Earth and the Sun, and his opposition to Rutherford's ideas of radioactivity, certainly put him on the losing side of many arguments later in his career.

William Thomson, Lord Kelvin, died in 1907 at the age of 83. He was buried in Westminster Abbey in London where he lies today, adjacent to Isaac Newton.

References

Burchfield, J.D.: *Lord Kelvin and The Age of the Earth*, Macmillan, 1975.

Körner, T.W.: *Fourier Analysis*, Cambridge University Press, 1988.

Encyclopedia Britannica, 2004.

Morrison, N.: *Introduction to Fourier Analysis*, John Wiley & Sons, Inc., 1994.

Thomson, S.P.: *The Life of Lord Kelvin*, London, 1976.

Kelvin & Tait: *Treatise on Natural Philosophy*, Appendix B.

19 Waveform Generation

Contents

19.2

Introduction

A comparator uses the op-amp in an *open-loop* mode. For a very small input voltage, the output will saturate close to one of the power supply voltages due to the very large gain of the op-amp.

Positive feedback can be applied to a comparator to create hysteresis. This can be used to "clean-up" noisy digital waveforms, amongst other applications, and is an example of a *bistable* circuit (it has two stable states). It can also be used to make an *astable multivibrator*. The output will oscillate at a rate which can be set by a few passive components.

A comparator with hysteresis can also be used to generate simple waveforms such as square waves and triangle waves. With proper filtering, sinusoids can also be generated.

19.1 Open-Loop Comparator

A comparator is an example of a non-linear op-amp circuit. It is a switching device that produces a high or low output, depending on which of the two inputs is larger. A simple comparator can be made from an op-amp with no feedback connection (*open-loop*) as shown in the figure below:

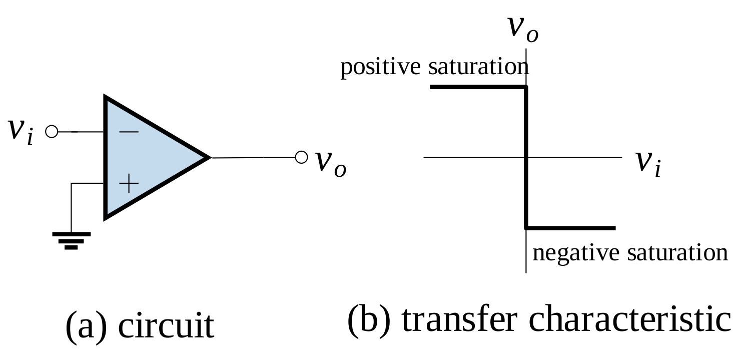


Figure 19.1

Since the open-loop voltage gain of an op-amp is very large, when there is no feedback an input voltage difference of only a few microvolts is sufficient to drive the output voltage to either its maximum (V_{OH}) or to its minimum value (V_{OL}). These values are determined by the op-amp supply voltages and its internal structure; their magnitudes are always slightly lower than that of their respective supply values ($V_{OH} < V_{CC}, |V_{OL}| < |V_{EE}|$).

This feature is used in comparator circuits, when one wishes to know whether a given input is larger or smaller than a reference value. It is especially useful in digital applications, such as in analog to digital converters (ADCs).

In practical applications that require a comparator, an op-amp should not be used. Semiconductor manufacturers produce specific integrated circuit comparators that have a different output stage to op-amps and are specifically designed to optimise operation in "saturation".

19.2 Comparator with Hysteresis (Schmitt Trigger)

The Schmitt trigger shown in Figure 19.2 is an extension of the comparator. The positive feedback and absence of negative feedback ensures that the output will always be at either its highest (V_{OH}) or its lowest (V_{OL}) possible value. The voltage divider formed by R_1 and R_2 sets V_+ at a fraction of the output.

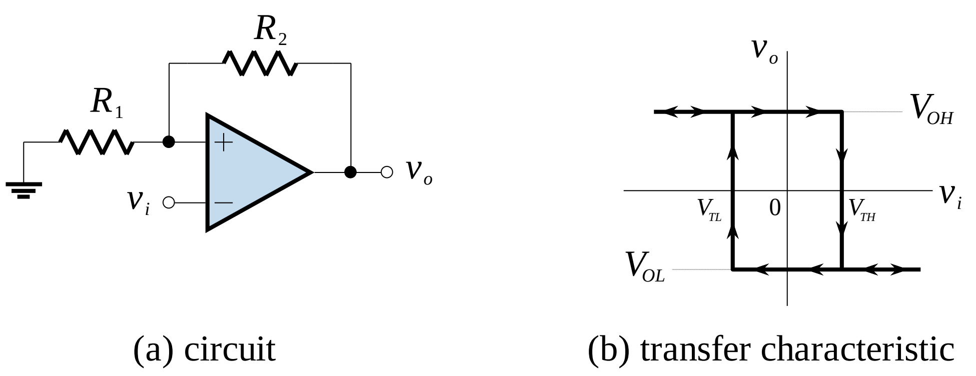


Figure 19.2

If $v_i > V_+$, the output is negative, if $v_i < V_+$ the output is positive. Each time the difference $v_i - V_+$ changes sign, the polarity of the output, and consequently of V_+, changes. No further change is possible until v_i reaches the new reference value V_+. The result is that the output may be at either extreme value (V_{OH} or V_{OL}) for the same value of the input; whether the output is positive or negative is determined by its previous state. The circuit therefore possesses memory. The consequence of this is that the voltage transfer characteristic of a Schmitt trigger follows a different curve, depending on whether the independent variable is increasing or decreasing. This property is called *hysteresis* and is depicted in Figure 19.2. Since the circuit has two stable states, it is also called a *bistable* circuit.

The thresholds for a change of an output state can be calculated as:

$$V_{TL} = V_{OL} \frac{R_1}{R_1 + R_2}$$
$$V_{TH} = V_{OH} \frac{R_1}{R_1 + R_2} \qquad (19.1)$$

It is important to note that in order for the output to change state all that is needed is a short departure of the input voltage above or below the respective threshold. This initiates the regenerative process that results in changing the state.

The figure below shows a *noninverting* Schmitt trigger with an adjustable reference voltage.

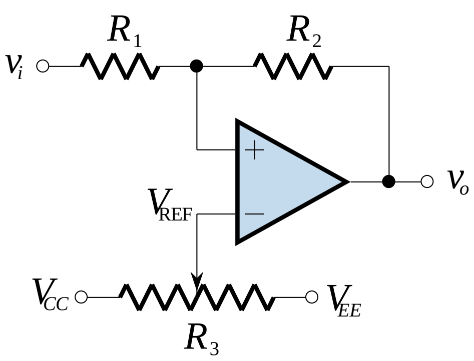


Figure 19.3

Using superposition, we can write the expression for v_+:

$$v_+ = v_i \frac{R_2}{R_1 + R_2} + v_o \frac{R_1}{R_1 + R_2} \qquad (19.2)$$

Let's assume that the circuit is in the positive stable state with $v_o = V_{OH}$. Then, in order to change this state to negative output, we must make $v_+ < V_- = V_{REF}$.

This means we need to apply:

$$v_i < V_{TL} = V_{\mathrm{REF}}\left(1+\frac{R_1}{R_2}\right) - V_{OH}\frac{R_1}{R_2} \tag{19.3}$$

Similarly, to change the state from low to high, the input voltage must satisfy (even for a brief moment) the following inequality:

$$v_i > V_{TH} = V_{\mathrm{REF}}\left(1+\frac{R_1}{R_2}\right) - V_{OL}\frac{R_1}{R_2} \tag{19.4}$$

19.3 Astable Multivibrator (Schmitt Trigger Clock)

When a negative feedback path consisting of a resistor R and a capacitor C is added to the Schmitt trigger in Figure 19.2, the new circuit has no stable state. The output will continuously switch between its two extremes at a rate determined by the time constant $T = RC$. The circuit is shown below:

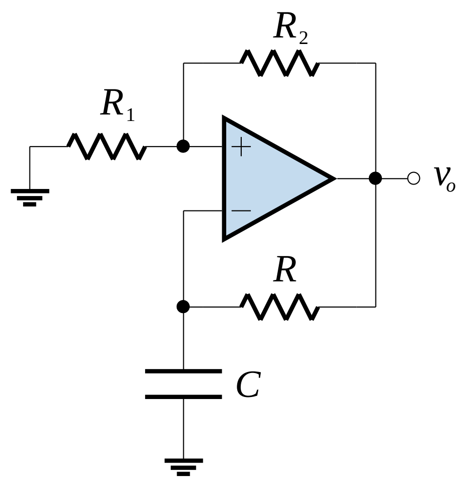


Figure 19.4

Immediately after a transition of the output to either its positive extreme (V_{OH}) or its negative extreme (V_{OL}), the RC network will begin an exponential transition; the capacitor will begin to charge or discharge, depending on its previous state, with its voltage approaching the new value of v_o. When the capacitor voltage v_- passes the value of v_+, which is determined by R_1 and R_2, the op-amp output will suddenly switch to its opposite extreme. The capacitor voltage will then begin to charge in the opposite direction until switching occurs again. The process will be repeated indefinitely, giving a square-wave output without the need for an input voltage source.

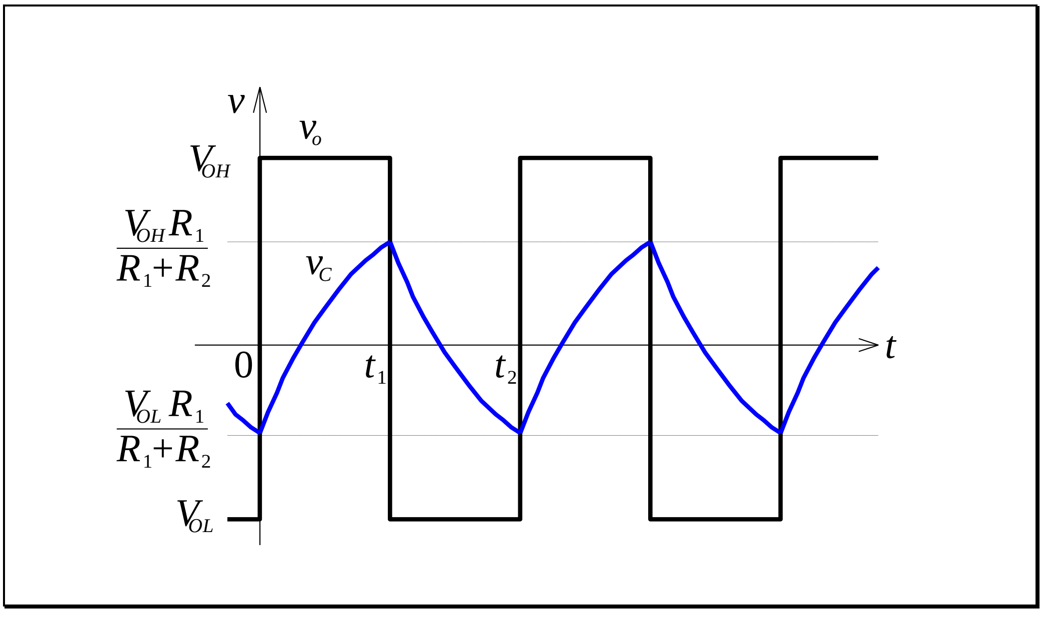


Figure 19.5

Suppose that at $t=0$ the output voltage is V_{OL}, and the capacitor voltage v_- has just fallen below $v_+ = V_{OL}R_1/(R_1+R_2)$. The output will switch from V_{OL} to V_{OH} because $v_+ - v_-$ has just become positive. The capacitor voltage begins to increase, and is given by:

$$v_C(t) = V_{OH} + \left(V_{OL}\frac{R_1}{R_1+R_2} - V_{OH}\right)e^{-\frac{t}{RC}} \qquad t \geq 0 \qquad (19.5)$$

Substitution of $t=0$ shows that the above equation indeed satisfies the initial condition $v_C(0) = V_{OL}R_1/(R_1+R_2)$. When $t \to \infty$, we obtain $\lim_{t\to\infty} v_C(t) = V_{OH}$. So, the capacitor voltage begins to increase toward V_{OH}, reaching $v_+ = V_{OH}R_1/(R_1+R_2)$ at time t_1.

Solving the above equation for this condition, one gets:

$$t_1 = -RC \ln\left[\frac{V_{OH} R_2}{V_{OH}(R_1 + R_2) - V_{OL} R_1}\right] \tag{19.6}$$

At this point $v_+ - v_-$ changes sign and v_C begins to decrease, now governed by the equation:

$$v_C(t) = V_{OL} + \left(V_{OH} \frac{R_1}{R_1 + R_2} - V_{OL}\right) e^{-\frac{t-t_1}{RC}} \qquad t \ge t_1 \tag{19.7}$$

At time t_2, v_C reaches $v_C(t_2) = V_{OL} R_1 / (R_1 + R_2)$. Solving the above equation for this condition, one gets:

$$t_2 - t_1 = -RC \ln\left[\frac{V_{OL} R_2}{V_{OL}(R_1 + R_2) - V_{OH} R_1}\right] \tag{19.8}$$

The period of the output waveform is just $T_0 = t_2$. Therefore we have:

$$\begin{aligned} T_0 &= (t_2 - t_1) + t_1 \\ &= -RC \ln\left[\frac{V_{OL} R_2}{V_{OL}(R_1 + R_2) - V_{OH} R_1}\right] - RC \ln\left[\frac{V_{OH} R_2}{V_{OH}(R_1 + R_2) - V_{OL} R_1}\right] \\ &= -RC \ln\left[\frac{V_{OL} R_2}{V_{OL}(R_1 + R_2) - V_{OH} R_1} \cdot \frac{V_{OH} R_2}{V_{OH}(R_1 + R_2) - V_{OL} R_1}\right] \end{aligned} \tag{19.9}$$

In the *special* case of $R_1 = R_2$ and $V_{OL} = -V_{OH}$, the above equation simplifies to a function of only *R* and *C*:

$$T_0 = RC \ln 9 \approx 2.2RC \tag{19.10}$$

19.4 Waveform Generator

The exponential waveform (across the capacitor) generated in the astable circuit of Figure 19.4 can be changed to triangular by replacing the lowpass *RC* circuit with an integrator (the integrator is, after all, a lowpass circuit with a corner frequency at DC). The integrator causes linear charging and discharging of the capacitor, thus producing a triangular waveform. The resulting circuit is shown below:

Triangular and square waveform generator

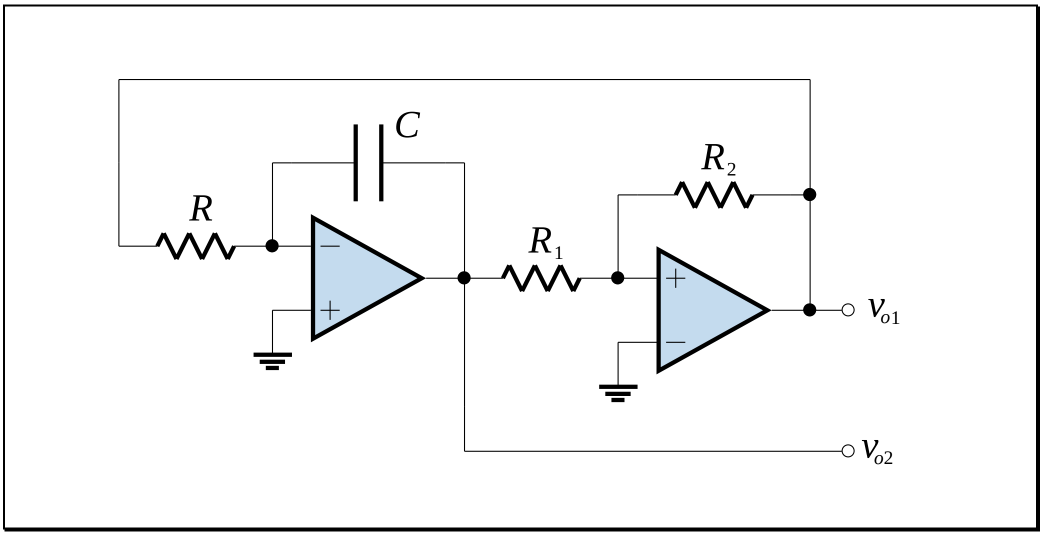


Figure 19.6

This circuit oscillates and generates a square waveform at the output of the noninverting Schmitt trigger, v_{o1}, and a triangular waveform at the output of the inverting integrator, v_{o2}.

19.10

Let the output of the bistable circuit be at V_{OH}. A current equal to V_{OH}/R will go into the resistor R and then on to the capacitor C, causing the output of the integrator to *linearly* decrease with the slope $-V_{OH}/RC$, as shown in the figure below. This will continue until the integrator output reaches the lower threshold, V_{TL}, of the bistable circuit.

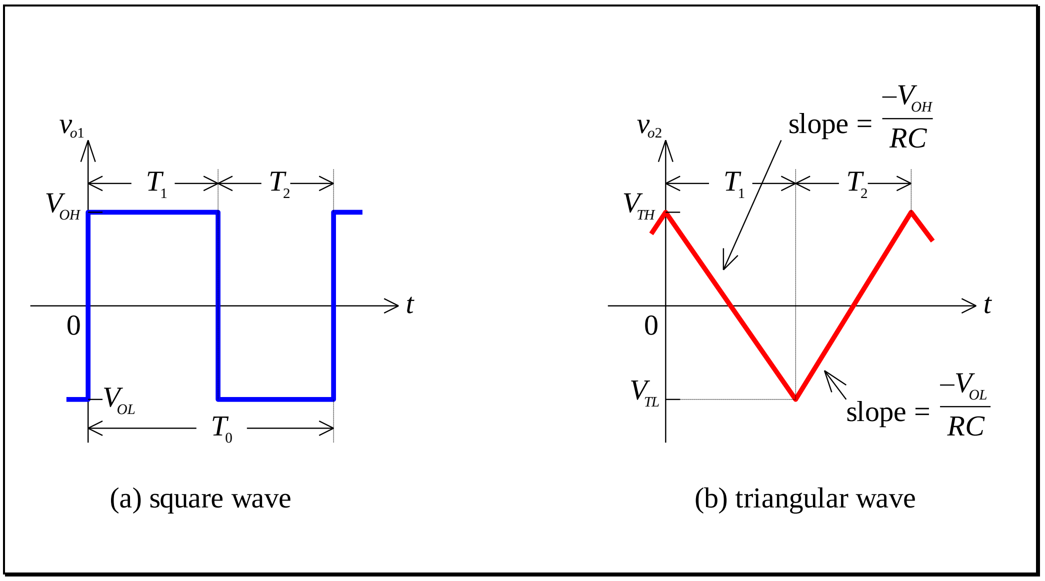


Figure 19.7

At this point the bistable circuit will switch states, its output becoming negative and equal to V_{OL}. At this moment the current through R will reverse direction and its value will become equal to $|V_{OL}|/R$. The output of the integrator will therefore linearly increase with time. This will continue until the integrator output voltage reaches the positive threshold of the Schmitt trigger, V_{TH}. The Schmitt trigger switches states again, starting the new cycle.

From Figure 19.7 it is relatively easy to derive an expression for the period T_0 of the square and triangular waveforms. During the interval T_1 we have:

$$\frac{V_{TH} - V_{TL}}{T_1} = \frac{V_{OH}}{RC} \quad \Rightarrow \quad T_1 = RC\frac{V_{TH} - V_{TL}}{V_{OH}} \qquad (19.11)$$

Similarly, during T_2 we have:

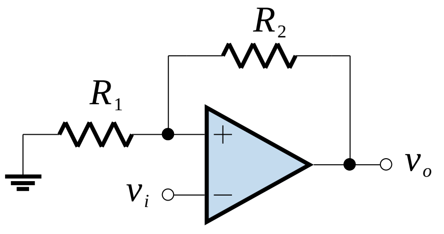
$$\frac{V_{TH} - V_{TL}}{T_2} = \frac{-V_{OL}}{RC} \quad \Rightarrow \quad T_2 = RC\frac{V_{TH} - V_{TL}}{-V_{OL}} \qquad (19.12)$$

Thus, to obtain symmetrical waveforms we need a bistable circuit with $V_{OL} = -V_{OH}$. The oscillation frequency is equal to:

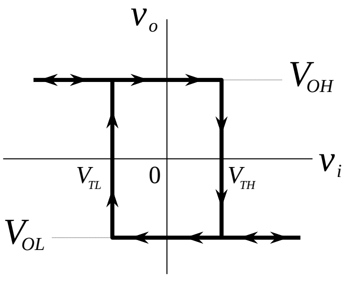
$$f_0 = \frac{1}{T_0} = \frac{1}{T_1 + T_2} = \frac{1}{RC}\frac{V_{OL}V_{OH}}{(V_{TH} - V_{TL})(V_{OL} - V_{OH})} \qquad (19.13)$$

19.5 Summary

- An op-amp comparator with hysteresis is known as a Schmitt trigger:

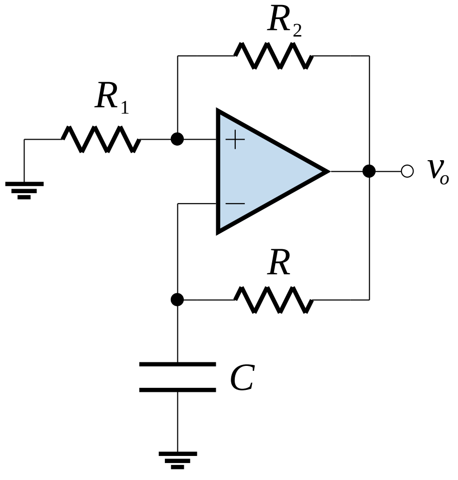


(a) circuit

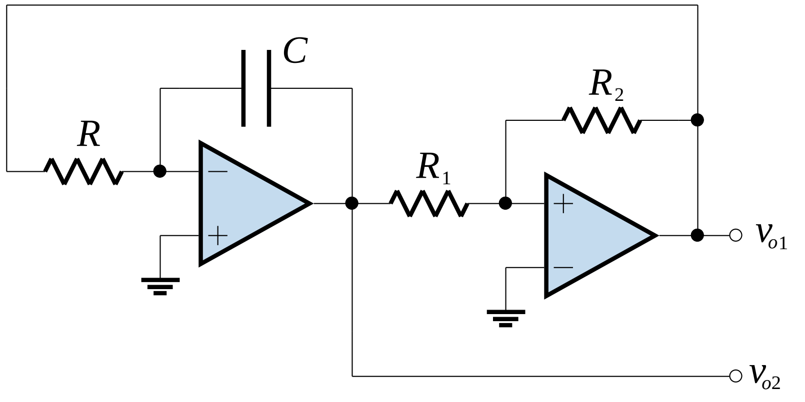


(b) transfer characteristic

- The astable multivibrator is a simple circuit based on the Schmitt trigger that can produce a square wave at low frequencies:



- A waveform generator circuit that utilises a Schmitt trigger and produces both triangular and square waveforms is:



19.6 References

Sedra, A. and Smith, K.: *Microelectronic Circuits*, Saunders College Publishing, New York, 1991.

20 The Second-Order Frequency Response

Contents

20.2

Introduction

The use of both capacitors and inductors in a circuit gives rise to an important phenomenon – the exchange of electric and magnetic stored energy in a sinusoidal fashion between the ideal elements without dissipation. The peak energy stored in the elements reaches a maximum when a sinusoidal forcing function drives the circuit into *resonance* at a frequency close to the natural frequency of the circuit.

The resonance phenomenon occurs in many natural systems, and is the result of systems being able to store energy in different ways. For example, a pendulum, without friction, will oscillate forever once set in motion, and there will be a continual exchange of kinetic energy (the velocity of the pendulum mass) and potential energy (stored by virtue of the position of the pendulum in a gravitational field) as the pendulum oscillates in a sinusoidal fashion.

In an electric circuit the resonant condition can be used to create a highly "selective" circuit in the sense that a narrow band of frequencies of the forcing function will cause the circuit response to be large whilst all other frequencies result in a response which is much smaller. This property of second-order circuits is not exhibited by passive first-order circuits (or cascades of them).

The "width of the magnitude response", or *bandwidth B*, for a second-order circuit will be seen to be related to a quantity called the *quality factor*, Q_0. The quality factor, together with the undamped natural frequency, ω_0, uniquely determine the properties of many second-order systems. These parameters can be determined for a parallel passive *RLC* circuit, an electronic op-amp circuit, a mechanical system, a hydraulic system, etc. Thus, it pays to express second-order frequency response quantities in terms of ω_0 and Q_0 for the sake of uniformity across the disciplines.

20.1 Resonance

A system being driven by a sinusoidal forcing function will produce a sinusoidal steady-state response at the frequency of the driving force. The amplitude of the response *may be* larger than the forcing function when the frequency of the driving force is near a "natural frequency of oscillation" of the system.[1] This dramatic increase in amplitude near a natural frequency is called *resonance*, and we denote the frequency at which it occurs as ω_r, which is called the *resonance frequency* of the system.[2]

The phenomenon of resonance is a familiar one (at least qualitatively). For example, a child using a swing realizes that if the pushes are properly timed, the swing can move with a very large amplitude. The quartz crystal in a computer or watch is made to mechanically vibrate at a resonance frequency which is determined by its "cut" (its shape and size). A further illustration is furnished by the shattering of a crystal wineglass when exposed to a musical tone of the right pitch. The condition of resonance may or may not be desirable, depending on the circumstances.

The reason for large-amplitude oscillations at the resonance frequency is that energy is being transferred to the system under favourable conditions. In fact, for a parallel *RLC* circuit, maximum power will be dissipated by the circuit when the forcing function's voltage and current are in phase – i.e. the circuit appears to be purely resistive. This leads us to a precise definition for a resonance frequency, ω_r.

---

[1] A natural frequency of oscillation only occurs in systems that are second-order or higher. Even then, a large amplitude response will occur only under certain conditions.

[2] This assumes that one natural frequency gives rise to one resonance frequency. Certain topologies of circuit components can give rise to two resonance frequencies even though there is only one natural frequency. Circuits and systems of high-order can have multiple natural frequencies and consequently may have multiple resonance frequencies. Also, even though the amplitude response is very large at a resonance frequency, it is *not* necessarily the maximum amplitude response – it depends on how we define "response".

20.4

In a two-terminal electrical network containing at least one inductor and one capacitor, we define *resonance* to be the condition which exists when the input impedance of the network is purely resistive:

Resonance defined

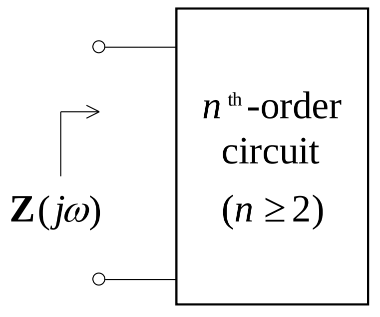


Figure 20.1

The resonant condition can be achieved by adjusting any circuit parameter. We normally devote our attention to the case for which ω is the variable, and since we have denoted the resonance frequency as ω_r, then resonance occurs if:

$$\operatorname{Im}\left[\mathbf{Z}\left(j\omega_r\right)\right] = 0 \qquad (20.1)$$

Thus, a two-terminal circuit is said it be in resonance when the sinusoidal voltage and current at the circuit input terminals are in phase.

This definition also implies that at resonance:

$$\operatorname{Im}\left[\mathbf{Y}\left(j\omega_r\right)\right] = 0 \qquad (20.2)$$

since $\mathbf{Z} = R$ is real at resonance and therefore $\mathbf{Y} = 1/R = G$ is real also.

It should be noted that the resonance condition $\operatorname{Im}\left[\mathbf{Z}\left(j\omega_r\right)\right] = 0$ may not be satisfied at any real positive frequency ω_r. Thus, resonance is a condition which a circuit *may* achieve, but only if its topology and component values allow it.

20.2 Parallel Resonance

We shall apply the definition of resonance to the parallel *RLC* circuit:

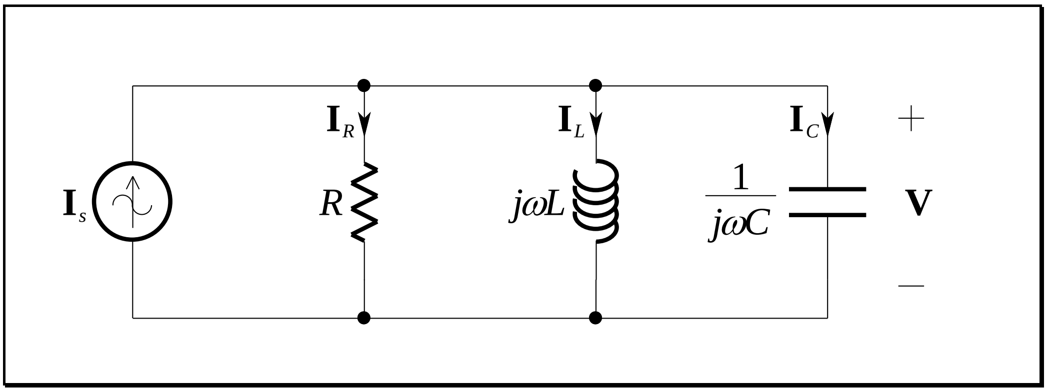


Figure 20.2

The admittance offered to the ideal current source is:

$$\mathbf{Y} = \frac{1}{R} + j\left(\omega C - \frac{1}{\omega L}\right) \tag{20.3}$$

Obviously, if the impedance is to be purely resistive at resonance, then so is the admittance. Thus, resonance occurs when:

$$\omega_r C - \frac{1}{\omega_r L} = 0 \tag{20.4}$$

Hence, the resonance frequency for this simple case is:

$$\omega_r = \frac{1}{\sqrt{LC}} = \omega_0 \tag{20.5}$$

This resonance frequency is identical to the undamped natural frequency that was defined whilst considering the step-response of the parallel *RLC* circuit.

Let us examine the magnitude of the response, the voltage $\mathbf{V}$ as indicated in Figure 20.2. For a constant-amplitude sinusoidal current source input, the response is proportional to the input impedance $\mathbf{Z}$.

The admittance as a function of ω can be written as:

$$\begin{aligned}\mathbf{Y} &= \frac{1}{R} + j\omega C + \frac{1}{j\omega L} \\ &= C\frac{-\omega^2 + j\omega/RC + 1/LC}{j\omega}\end{aligned} \tag{20.6}$$

The response is therefore given by:

$$\mathbf{V} = \mathbf{ZI} = \frac{1}{C}\frac{j\omega}{\left(1/LC - \omega^2\right) + j\omega/RC}\mathbf{I}_s \tag{20.7}$$

The response starts at zero, reaches a maximum value at the resonance frequency, and then drops again to zero as ω becomes infinite:

Typical magnitude and phase response of a second-order bandpass circuit

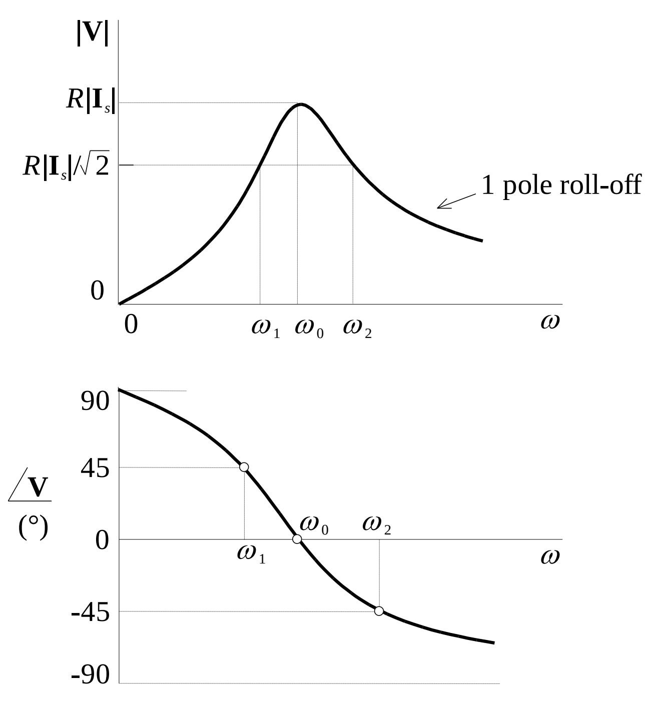


Figure 20.3

The frequency response is referred to as a *bandpass* response, since it passes only "midband" frequencies whilst attenuating low and high frequencies.

To show that the maximum response is $R|\mathbf{I}_s|$, and that this occurs at resonance, we could take the magnitude of Eq. (20.7), differentiate and equate to zero to obtain the frequencies of any relative maxima and minima, and then use these values to obtain the magnitude of the response. However, there is a simpler way.

The admittance contains a constant conductance and a susceptance which has a minimum magnitude (of zero) at resonance. The minimum admittance magnitude therefore occurs at resonance, and it is $1/R$. Hence, the maximum impedance magnitude is R, and it occurs at resonance.

At the resonance frequency, therefore, the voltage across the parallel circuit is simply $R\mathbf{I}_s$, and we can see that the source current goes through the resistor. However, there is also a current in L and C, since there is a voltage across them:

$$\begin{aligned} \mathbf{I}_{Lr} &= \frac{R\mathbf{I}}{j\omega_r L} \\ \mathbf{I}_{Cr} &= j\omega_r CR\mathbf{I} \end{aligned} \tag{20.8}$$

Since $1/\omega_r C = \omega_r L$ at resonance, we find that:

$$\mathbf{I}_{Cr} = -\mathbf{I}_{Lr} \tag{20.9}$$

and therefore $\mathbf{I}_{Lr} + \mathbf{I}_{Cr} = \mathbf{0}$. We thus have a circulating current around the LC part of the circuit which is the cause of the never-ending exchange of energy between the inductor and the capacitor at resonance.

Although the height of the response curve depends only upon the value of R, the width of the curve or the steepness of the sides depends upon the other two element values also. The width of the response curve is most easily expressed when we introduce a very important parameter, the quality factor Q.

20.2.1 Phasor Diagram of the Parallel *RLC* Circuit

It is instructive to illustrate resonance with the aid of phasor diagrams. The figure below shows phasor diagrams, as well as illustrations of the sinusoidal currents, for the parallel *RLC* circuit at three different frequencies:

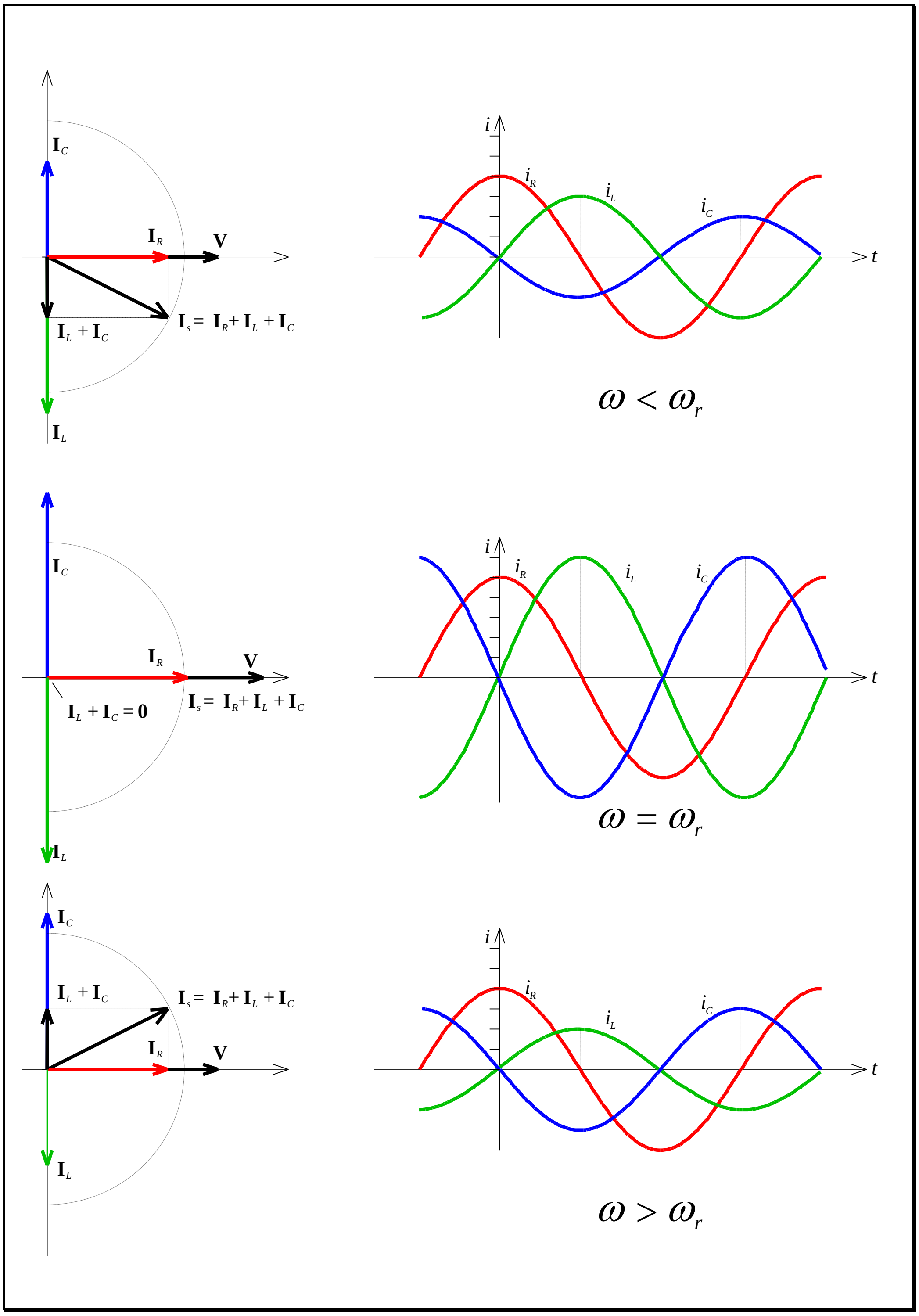


Figure 20.4

20.3 Quality Factor

We define *quality factor*, Q, as:

$$Q = 2\pi \frac{\text{maxmium energy stored}}{\text{total energy lost per period}} \tag{20.10}$$

The constant 2π is put into the definition in order to simplify the more useful expressions for Q that occur in the study of second-order systems.

For the parallel *RLC* circuit, energy is only lost in the resistor. We can therefore express Q in terms of the instantaneous energy associated with each of the reactive elements and the average power dissipated in the resistor:

$$Q = 2\pi \frac{\left[w_L(t) + w_C(t)\right]_{\max}}{P_R T_0} \tag{20.11}$$

We will apply this definition and determine the value of Q at the resonance frequency $\omega_r = \omega_0$, which is denoted by Q_0. We select the forcing function:

$$i(t) = I_m \cos(\omega_0 t) \tag{20.12}$$

and obtain the corresponding voltage at resonance:

$$v(t) = Ri(t) = RI_m \cos(\omega_0 t) \tag{20.13}$$

The instantaneous energy stored in the capacitor is:

$$w_C(t) = \tfrac{1}{2} Cv^2 = \frac{CR^2 I_m^2}{2} \cos^2(\omega_0 t) \tag{20.14}$$

The instantaneous energy stored in the inductor is:

$$w_L(t) = \tfrac{1}{2} L i_L^2 = \tfrac{1}{2} L \left(\frac{1}{L} \int_0^t v\, dt \right)^2$$
$$= \frac{R^2 I_m^2}{2\omega_0^2 L} \sin^2(\omega_0 t) = \frac{C R^2 I_m^2}{2} \sin^2(\omega_0 t) \tag{20.15}$$

The total instantaneous energy stored is therefore constant:

$$w_L(t) + w_C(t) = \frac{C R^2 I_m^2}{2} \tag{20.16}$$

and this constant value must also be the maximum value.

In order to find the energy lost in the resistor in one period, we take the average power absorbed by the resistor:

$$P_R = \tfrac{1}{2} R I_m^2 \tag{20.17}$$

and multiply by one period, obtaining:

$$P_R T_0 = \frac{R I_m^2}{2 f_0} \tag{20.18}$$

We thus find the quality factor at resonance:

$$Q_0 = 2\pi \frac{C R^2 I_m^2 / 2}{R I_m^2 / 2 f_0} = 2\pi f_0 C R = \omega_0 C R \tag{20.19}$$

This equation holds only for the simple parallel *RLC* circuit.

Equivalent expressions for Q_0 which are often useful may be obtained by substitution:

$$Q_0 = \omega_0 CR = \frac{R}{X_{C0}} = \frac{R}{X_{L0}} = R\sqrt{\frac{C}{L}} \tag{20.20}$$

It is apparent that Q_0 is a dimensionless constant which is a function of all three circuit elements in the parallel resonant circuit, and it turns out that it can be evaluated from a knowledge of the natural response, as will be illustrated later.

A useful interpretation of Q_0 is obtained when we inspect the capacitor and inductor currents at resonance:

$$\mathbf{I}_{Cr} = -\mathbf{I}_{Lr} = j\omega_0 CR\mathbf{I} = jQ_0\mathbf{I} \tag{20.21}$$

Each is Q_0 times the source current in magnitude and they are $180°$ out of phase. Thus if we apply 1 mA at the resonance frequency to a parallel resonant circuit with a Q_0 of 50, we find 1 mA in the resistor, and 50 mA in both the inductor and capacitor. A parallel resonant circuit can therefore act as a current amplifier (but not a power amplifier, since it is a passive network).

20.4 Second-Order Circuit Relations

The two parameters α and ω_d were introduced in connection with the natural response of a second-order circuit. These two parameters can be related to the undamped natural frequency, ω_0, and the quality factor at resonance, Q_0.

We have:

$$\alpha = \frac{1}{2RC} = \frac{1}{2\left(Q_0/\omega_0 C\right)C} \tag{20.22}$$

and thus:

$$\alpha = \frac{\omega_0}{2Q_0} \tag{20.23}$$

We also have:

$$\omega_d = \sqrt{\omega_0^2 - \alpha^2} \tag{20.24}$$

and thus:

$$\omega_d = \omega_0\sqrt{1 - \left(\frac{1}{2Q_0}\right)^2} \tag{20.25}$$

When we analyze second-order circuits in the time-domain, we usually resort to using α, ω_0 and ω_d (for an underdamped circuit). When we discuss second-order circuits in the frequency-domain, we usually use ω_0 and Q_0.

20.5 Bandwidth

The "width" of the response curve for the parallel *RLC* circuit can be defined more carefully and related to Q_0.

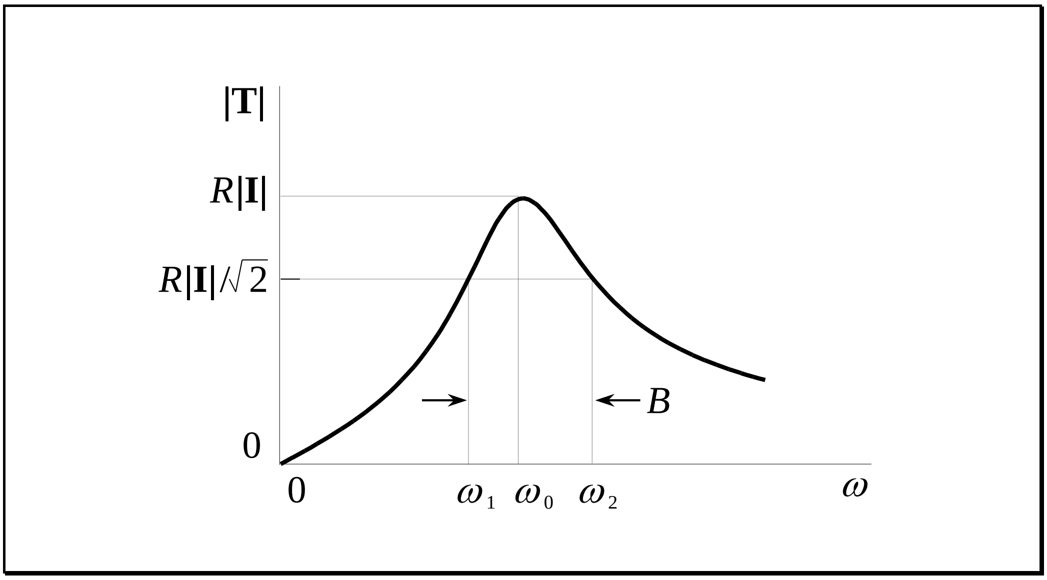


Figure 20.5

The two *half-power frequencies* ω_1 and ω_2 are those frequencies at which the magnitude of the voltage response is $1/\sqrt{2}$ times its maximum value (or -3 dB down from the peak).

We select ω_1 as the *lower half-power frequency* and ω_2 as the *upper half-power frequency*. These names arise from the fact that a voltage which is $1/\sqrt{2}$ times the resonance voltage is equivalent to a squared voltage which is *one-half* the squared voltage (and therefore the power) at resonance.

The *bandwidth* of a resonant circuit is defined as the difference of these two half-power frequencies:

$$B = \omega_2 - \omega_1 \quad \text{rads}^{-1} \tag{20.26}$$

We can also refer to bandwidth as:

$$B = f_2 - f_1 \quad \text{Hz} \tag{20.27}$$

The context of the analysis or design makes the units of bandwidth clear.

The use of *half-power* frequencies in the definition of bandwidth is an arbitrary but widely accepted criterion used by the engineering profession. The concept of bandwidth is used in many other electrical systems and is a very important parameter in the design of filters, amplifiers and electrical systems in general. One should also be aware that bandwidth is only defined for systems with a single peak response – otherwise the definition is ambiguous.

We think of this bandwidth as the "width" of the response curve, even though the curve actually extends from $\omega = 0$ to $\omega = \infty$.

We can express the bandwidth B in terms of ω_0 and Q_0. The admittance of the parallel RLC circuit is:

$$\begin{aligned} \mathbf{Y} &= \frac{1}{R} + j\left(\omega C - \frac{1}{\omega L}\right) \\ &= \frac{1}{R} + j\frac{1}{R}\left(\frac{\omega\omega_0 CR}{\omega_0} - \frac{\omega_0 R}{\omega\omega_0 L}\right) \\ &= \frac{1}{R}\left[1 + jQ_0\left(\frac{\omega}{\omega_0} - \frac{\omega_0}{\omega}\right)\right] \end{aligned} \tag{20.28}$$

The magnitude of the admittance at resonance is $1/R$, and we seek frequencies at which the magnitude reaches $\sqrt{2}/R$ to achieve half-power. This must occur when the imaginary part of the bracketed quantity has a magnitude of unity.

Thus:

$$Q_0\left(\frac{\omega_1}{\omega_0} - \frac{\omega_0}{\omega_1}\right) = -1 \quad \text{and} \quad Q_0\left(\frac{\omega_2}{\omega_0} - \frac{\omega_0}{\omega_2}\right) = 1 \tag{20.29}$$

Solving, we have:

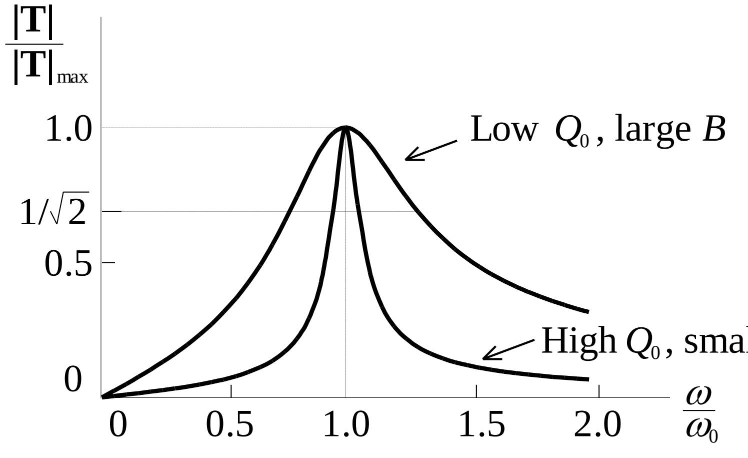
$$\omega_1 = \omega_0\left[\sqrt{1+\frac{1}{4Q_0^2}} - \frac{1}{2Q_0}\right]$$
$$\omega_2 = \omega_0\left[\sqrt{1+\frac{1}{4Q_0^2}} + \frac{1}{2Q_0}\right] \qquad (20.30)$$

Although individually complicated, their difference provides a very simple formula for the bandwidth:

$$B = \omega_2 - \omega_1 = \frac{\omega_0}{Q_0} \quad \text{rads}^{-1} \qquad (20.31)$$

Bandwidth defined for a parallel resonant circuit

This equation tells us that Q_0 and B are inversely related, as shown below:



The inverse relationship between B and Q_0

Figure 20.6

Circuits possessing a higher Q_0 have a narrower bandwidth – they have greater *frequency selectivity* or "higher quality". Such circuits were used extensively in receivers of the old analog broadcast systems, such as AM and FM radio and TV, to "tune into a station" whilst rejecting all others.

If we multiply the two half-power frequencies together, we can show that:

$$\omega_1\omega_2 = \omega_0^2 \tag{20.32}$$

and therefore:

$$\omega_0 = \sqrt{\omega_1\omega_2} \tag{20.33}$$

That is, the resonance frequency is the *geometric* mean of the two half-power frequencies.

For high-Q circuits ($Q_0 \geq 5$), we can show that:

$$\omega_{1,2} = \omega_0\left[\sqrt{1+\frac{1}{4Q_0^2}} \mp \frac{1}{2Q_0}\right] \approx \omega_0 \mp \frac{B}{2} \tag{20.34}$$

and:

$$\omega_0 = \sqrt{\omega_1\omega_2} \approx \frac{\omega_1+\omega_2}{2} \tag{20.35}$$

That is, for high-Q circuits, each half-power frequency is located approximately one-half bandwidth from the resonance frequency – the resonance frequency is approximately the *arithmetic* mean of the half-power frequencies.

20.6 Series Resonance

The series resonant circuit finds less use than the parallel circuit. Consider the series *RLC* circuit below:

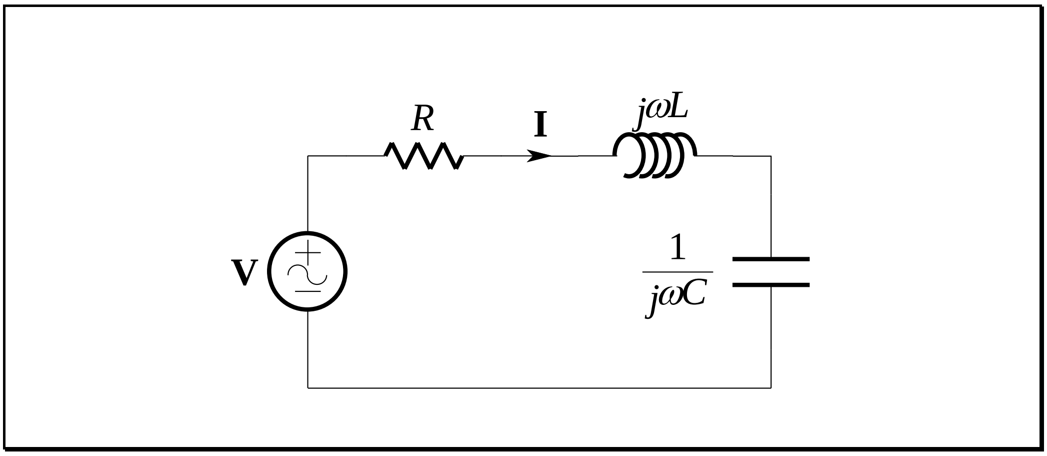


Figure 20.7

We can derive the important equations for the series resonant circuit by using a "dual language" on the parallel circuit. We find that resonance occurs when:

$$\omega_r = \frac{1}{\sqrt{LC}} = \omega_0 \tag{20.36}$$

which is the same as for the parallel *RLC* circuit. However, the quality factor at resonance for the series *RLC* circuit is different:

$$Q_0 = \frac{\omega_0 L}{R} = \frac{X_{L0}}{R} = \frac{X_{C0}}{R} = \frac{1}{R}\sqrt{\frac{L}{C}} \tag{20.37}$$

20.18

If the response is taken across the resistor so that it is proportional to the current, then we achieve a similar response to that of the parallel resonant circuit (a bandpass response). In this case the equations for the half-power frequencies, the bandwidth, and the resonance frequency are the same as before:

$$\omega_{1,2} = \omega_0 \left[\sqrt{1 + \frac{1}{4Q_0^2}} \mp \frac{1}{2Q_0} \right] \tag{20.38}$$

$$B = \omega_2 - \omega_1 = \frac{\omega_0}{Q_0} \quad \text{rads}^{-1} \tag{20.39}$$

$$\omega_0 = \sqrt{\omega_1 \omega_2} \tag{20.40}$$

The series resonant circuit is characterized by a low impedance at resonance. The series resonant circuit provides inductor and capacitor voltages which are greater than the source voltage by the factor Q_0. The series circuit thus provides voltage amplification at resonance.

20.7 Other Resonant Forms

The parallel and series *RLC* circuits of the previous two sections represent *idealized* resonant circuits – they are useful approximations to real physical circuits where the resistance of the wire making up the inductor and the losses in the capacitor's dielectric are small. The network shown below is a reasonably accurate model for the parallel combination of a physical inductor, capacitor and resistor. The resistor R_L represents the ohmic losses, core losses, and radiation losses of the physical coil. The resistor R represents the losses in the dielectric within the physical capacitor as well as the resistance of the physical resistor that is placed in parallel with the inductor and capacitor.

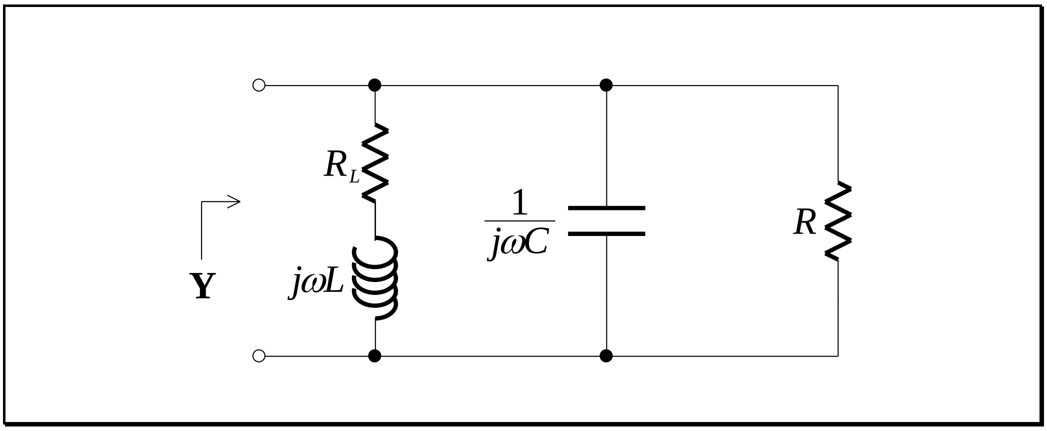


Figure 20.8

In this circuit, there is no way to combine elements and produce a simpler model. We have to resort to first principles to determine its resonant condition. The definition of resonance is unchanged, and we determine the resonance frequency by setting the imaginary part of the admittance to zero:

$$\operatorname{Im}\left[\mathbf{Y}(j\omega_r)\right] = 0$$

$$\operatorname{Im}\left(\frac{1}{R} + j\omega_r C + \frac{1}{R_L + j\omega_r L}\right) = 0 \tag{20.41}$$

Realizing the denominator of the inductor branch, we get:

$$\operatorname{Im}\left(\frac{1}{R} + j\omega_r C + \frac{R_L - j\omega_r L}{R_L^2 + \omega_r^2 L^2} \right) = 0 \tag{20.42}$$

Thus:

$$C = \frac{L}{R_L^2 + \omega_r^2 L^2} \tag{20.43}$$

and:

$$\omega_r = \sqrt{\frac{1}{LC} - \left(\frac{R_L}{L} \right)^2} \tag{20.44}$$

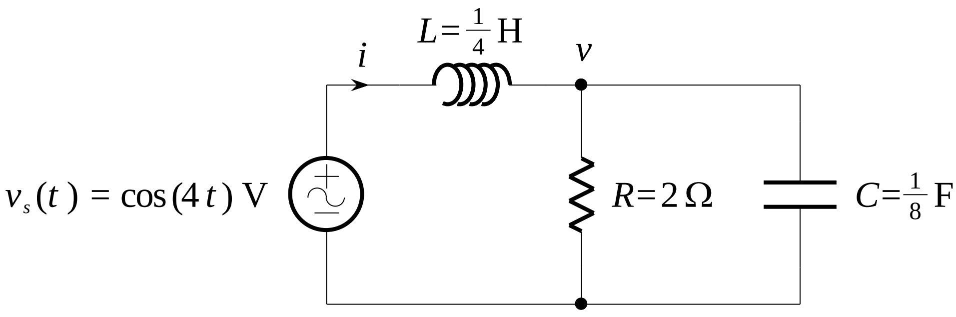
We note that ω_r is less than $1/\sqrt{LC}$, but sufficiently small values of the ratio R_L/L may result in a negligible difference between ω_r and $1/\sqrt{LC}$.

The maximum magnitude of the input impedance is *not* R, and it does *not* occur at ω_r (or at $\omega = 1/\sqrt{LC}$). The proof is algebraically cumbersome, but the theory is straightforward (set the derivative of the impedance magnitude to zero to find relative maxima and minima, etc.).

It should also be pointed out that if $R_L/L > 1/\sqrt{LC}$ then resonance will *never* occur, since Eq. (20.44) results in an imaginary quantity. Thus, we must be careful in any analysis we undertake to check the conditions under which a circuit may exhibit the phenomenon of resonance.

EXAMPLE 20.1 Resonance in a Circuit

Consider the simple *RLC* circuit shown below:



By KVL, for the mesh on the left:

$$L\frac{di}{dt} + v - v_s = 0$$

while, by KCL, at node *v*:

$$i = C\frac{dv}{dt} + \frac{v}{R}$$

Substituting the second expression into the first, we get:

$$LC\frac{d^2v}{dt^2} + \frac{L}{R}\frac{dv}{dt} + v - v_s = 0$$

from which:

$$\frac{d^2v}{dt^2} + \frac{1}{RC}\frac{dv}{dt} + \frac{1}{LC}v = \frac{1}{LC}v_s$$

Thus we see that the undamped natural frequency is:

$$\omega_0 = \frac{1}{\sqrt{LC}} = \frac{1}{\sqrt{\left(\frac{1}{4}\right)\left(\frac{1}{8}\right)}} = \sqrt{32} = 4\sqrt{2}\ \text{rads}^{-1}$$

However, the impedance seen by the source is:

$$\begin{aligned}\mathbf{Z} &= j\omega L + \frac{R(1/j\omega C)}{R + 1/j\omega C} \\ &= j\omega L + \frac{R}{1 + j\omega RC} \\ &= j\omega L + \frac{R(1 - j\omega RC)}{1 + \omega^2 R^2 C^2} \\ &= \frac{R}{1 + \omega^2 R^2 C^2} + j\left(\omega L - \frac{\omega R^2 C}{1 + \omega^2 R^2 C^2}\right)\end{aligned}$$

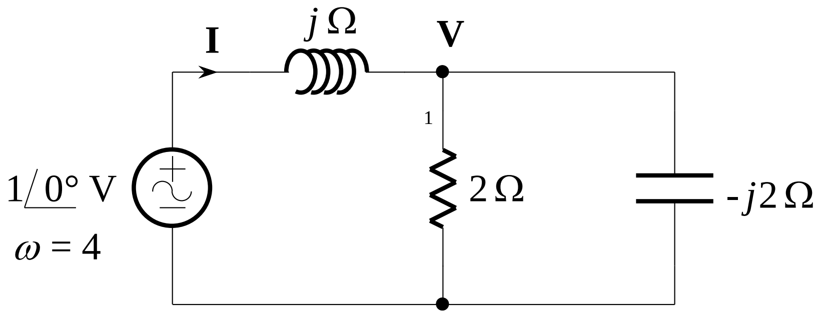
The imaginary part of **Z** vanishes when:

$$\begin{aligned}&\omega_r L - \frac{\omega_r R^2 C}{1 + \omega_r^2 R^2 C^2} = 0 \\ &\omega_r \left[L\left(1 + \omega_r^2 R^2 C^2\right) - R^2 C\right] = 0 \\ &\omega_r = 0 \quad \text{and} \quad \omega_r = \sqrt{\frac{R^2 C - L}{R^2 L C^2}} = \sqrt{\frac{1}{LC} - \frac{1}{R^2 C^2}}\end{aligned}$$

We can see that neither of these resonance frequencies is equal to the undamped natural frequency, i.e. $\omega_r \neq \omega_0$. The two resonance frequencies are:

$$\omega_r = 0 \quad \text{and} \quad \omega_r = \sqrt{\frac{1}{\left(\frac{1}{4}\right)\left(\frac{1}{8}\right)} - \frac{1}{2^2\left(\frac{1}{8}\right)^2}} = 4\ \text{rads}^{-1}$$

Since the frequency of the source is $\omega = 4\,\text{rads}^{-1}$, the circuit is in resonance. The circuit is shown below in the frequency-domain:



By nodal analysis:

$$\frac{1-\mathbf{V}}{j} = \frac{\mathbf{V}}{2} + \frac{\mathbf{V}}{-j2}$$

from which:

$$\mathbf{V} = \frac{2}{1+j} = \frac{2}{\sqrt{2}\angle 45^\circ} = \sqrt{2}\angle -45^\circ$$

and:

$$\mathbf{I} = \frac{1-\mathbf{V}}{j} = \frac{1-2/(1+j)}{j} = \frac{(-1+j)/(1+j)}{j} = \frac{-1+j}{-1+j} = 1\angle 0^\circ$$

So:

$$v(t) = \sqrt{2}\cos(4t - 45^\circ)$$

and:

$$i(t) = \cos(4t)$$

20.24

The energy stored in the inductor is:

$$w_L(t) = \frac{1}{2}Li^2(t) = \frac{1}{8}\cos^2(4t)$$

and the energy stored in the capacitor is:

$$w_C(t) = \frac{1}{2}Cv^2(t) = \frac{1}{8}\cos^2(4t - 45°)$$

Using the trigonometric identity:

$$\cos^2 x = \frac{1}{2}(1 + \cos 2x)$$

the total stored energy is:

$$\begin{aligned} w_L(t) + w_C(t) &= \frac{1}{16}[1 + \cos 8t] + \frac{1}{16}[1 + \cos(8t - 90°)] \\ &= \frac{1}{16}[2 + \cos 8t + \sin 8t] \\ &= \frac{1}{16}\left[2 + \sqrt{2}\cos(8t - 45°)\right] \end{aligned}$$

and this has a maximum value of:

$$[w_L(t) + w_C(t)]_{\max} = \frac{1}{16}\left[2 + \sqrt{2}\right] \text{ J}$$

The power dissipated by the resistor is:

$$P_R = \frac{1}{2}\frac{|\mathbf{V}|^2}{R} = \frac{1}{2}\left(\frac{2}{2}\right) = \frac{1}{2} \text{ W}$$

and the energy lost in a period is:

$$P_R T_0 = P_R\left(\frac{2\pi}{\omega_r}\right) = \frac{1}{2}\left(\frac{2\pi}{4}\right) = \frac{\pi}{4} \text{ J}$$

Thus the Q_0 of the circuit is:

$$Q_0 = 2\pi\frac{(1/16)\left(2 + \sqrt{2}\right)}{\pi/4} = 1 + \frac{1}{\sqrt{2}} \approx 1.707$$

20.8 The Second-Order Lowpass Frequency Response

Consider the series *RLC* circuit again, but this time the response is taken as the voltage across the capacitor:

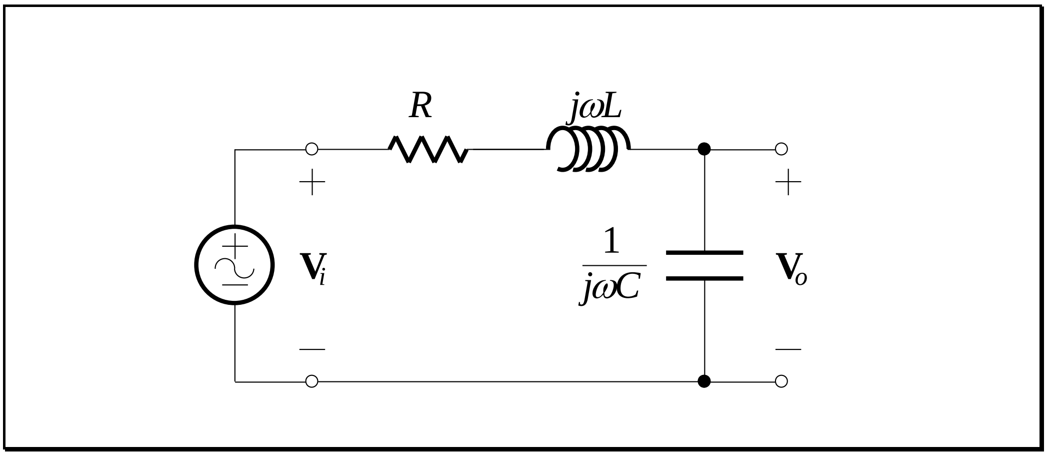


Figure 20.9

In this case the frequency response is given by the voltage divider rule:

$$\begin{aligned}\mathbf{T}(j\omega) &= \frac{\mathbf{V}_o}{\mathbf{V}_i} \\ &= \frac{1/j\omega C}{R + j\omega L + 1/j\omega C} \\ &= \frac{1/LC}{1/LC - \omega^2 + j\omega R/L}\end{aligned} \tag{20.45}$$

Noting that $\omega_0 = 1/\sqrt{LC}$ and $Q_0 = \omega_0 L/R$ for the series *RLC* circuit, this can be written as:

$$\mathbf{T}(j\omega) = \frac{\omega_0^2}{\omega_0^2 - \omega^2 + j\omega(\omega_0/Q_0)} \tag{20.46}$$

This has the form of a second-order *lowpass* frequency response – it passes low frequencies but attenuates high frequencies.

The magnitude response is:

The magnitude response of a lowpass second-order frequency response

$$\left|\mathbf{T}(j\omega)\right| = \frac{\omega_0^2}{\sqrt{\left(\omega_0^2 - \omega^2\right)^2 + \left(\omega\omega_0/Q_0\right)^2}} \tag{20.47}$$

and the phase is:

The phase response of a lowpass second-order frequency response

$$\angle\mathbf{T}(j\omega) = -\tan^{-1}\left(\frac{\omega\omega_0/Q_0}{\omega_0^2 - \omega^2}\right) \tag{20.48}$$

The magnitude and phase functions are plotted below for $Q_0 = 1.25$:

Typical magnitude and phase responses of a lowpass second-order frequency response

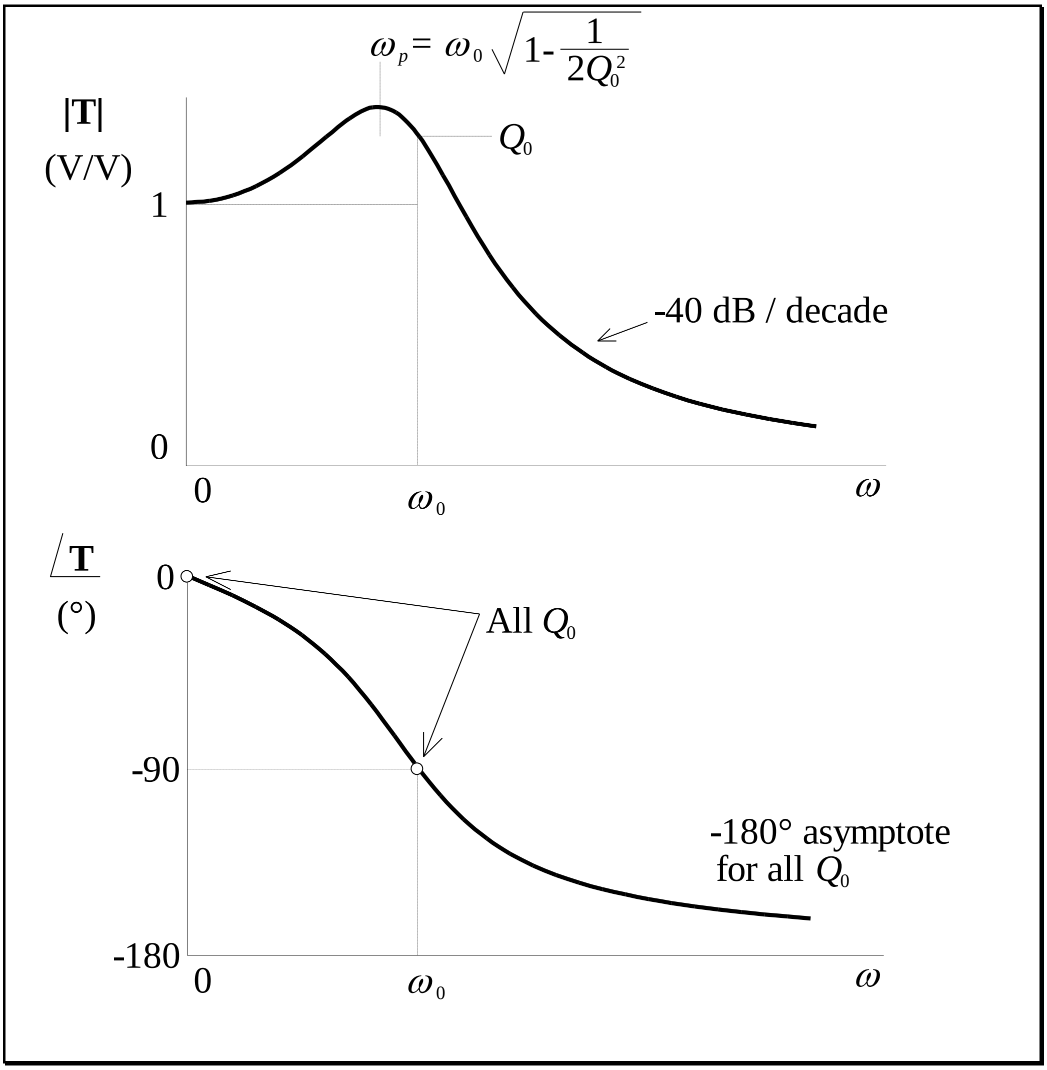


Figure 20.10

20.8.1 Peak Frequency

The peak of the frequency response **does not** correspond to the resonance frequency ω_0 (nor does it have any relation with ω_d, which arises in the description of the *time-domain* natural response). To obtain the peak of the frequency response, we find the relative maximum in the usual way.

To simplify the algebra, we let:

$$u = (\omega/\omega_0)^2 \tag{20.49}$$

Then the magnitude response can be written as:

$$|\mathbf{T}(j\omega)| = \frac{1}{\sqrt{(1-u)^2 + u/Q_0^2}} = \frac{1}{\sqrt{G(u)}} \tag{20.50}$$

where:

$$\begin{aligned} G(u) &= (1-u)^2 + u/Q_0^2 \\ &= u^2 + \left(1/Q_0^2 - 2\right)u + 1 \end{aligned} \tag{20.51}$$

We want to find ω_p so that $|\mathbf{T}(j\omega_p)| \to \max$, or equivalently, $G(u_p) \to \min$. To find u that minimizes $G(u)$, we let:

$$\frac{d}{du}G(u) = 2u + \left(1/Q_0^2 - 2\right) = 0 \tag{20.52}$$

Solving this for u we get:

$$u_p = 1 - \frac{1}{2Q_0^2} \tag{20.53}$$

Thus, the frequency at which the magnitude response reaches a peak is:

$$\begin{aligned} \omega_p &= \omega_0 \sqrt{1 - \frac{1}{2Q_0^2}}, & Q_0 &\geq 1/\sqrt{2} \\ \omega_p &= 0, & Q_0 &< 1/\sqrt{2} \end{aligned} \qquad (20.54)$$

Notice that the peak response always occurs *before* the resonance frequency for the lowpass response, and we approach $\omega_p \approx \omega_0$ for high Q_0 (say $Q_0 \geq 5$).

We can also see that a *relative* peak will **not** occur in the magnitude response if $Q_0 < 1/\sqrt{2}$ (for then ω_p is an imaginary quantity!). In this case, the absolute peak occurs at 0 Hz, or DC.

The special value of $Q_0 = 1/\sqrt{2}$ causes the relative and absolute "peak" to coincide at DC, and the magnitude response in this special case is known as *maximally flat* (all derivatives of the magnitude response at DC are zero).

At the peak frequency, the magnitude of the frequency response is:

$$\begin{aligned} \left|\mathbf{T}(j\omega_p)\right| &= \frac{Q_0}{\sqrt{1 - \frac{1}{4Q_0^2}}}, & Q_0 &\geq 1/\sqrt{2} \\ \left|\mathbf{T}(j\omega_p)\right| &= 1, & Q_0 &< 1/\sqrt{2} \end{aligned} \qquad (20.55)$$

For high Q_0 (say $Q_0 \geq 5$), the magnitude response is $\left|\mathbf{T}(j\omega_p)\right| \approx Q_0$.

20.8.2 Bandwidth

The bandwidth of the lowpass *RLC* circuit is the difference between the two half-power frequencies on each side of the peak frequency:

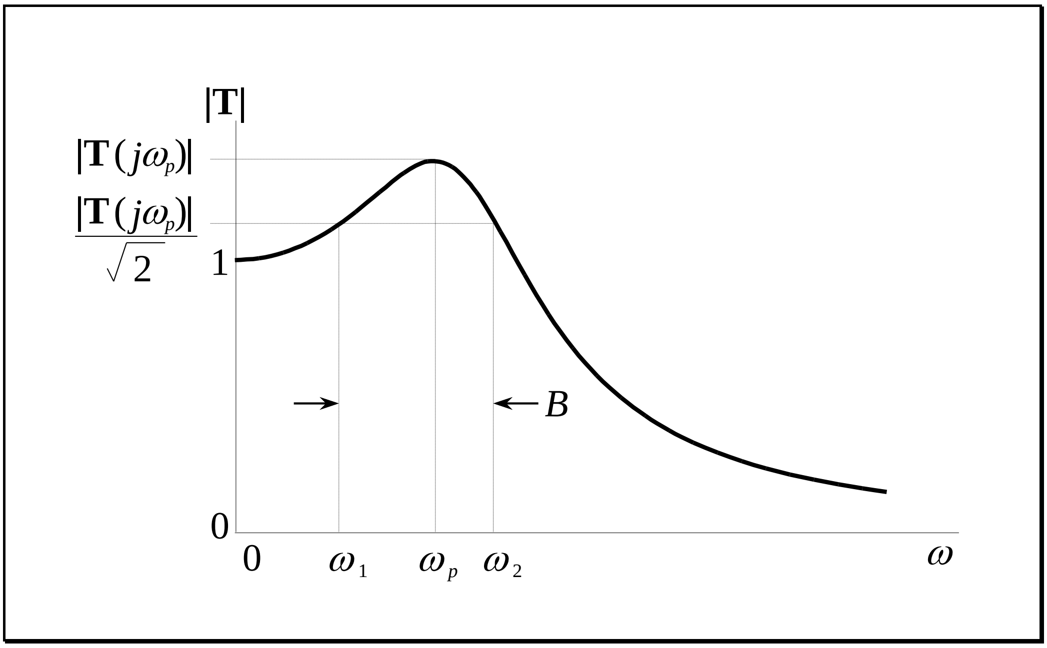


Figure 20.11

That is, the usual definition of bandwidth applies:

$$B = \omega_2 - \omega_1 \text{ rads}^{-1} \tag{20.56}$$

By definition, the two half-power frequencies must satisfy:

$$\frac{\left|\mathbf{T}(j\omega_{1,2})\right|^2}{\left|\mathbf{T}(j\omega_p)\right|^2} = \frac{1}{2} \tag{20.57}$$

However, as will be seen, the relative peak in the magnitude response only occurs when $Q_0 > 1/\sqrt{2} \approx 0.7071$, and when this is the case the peak is only greater than $\sqrt{2}$ when $Q_0 > \sqrt{1+1/\sqrt{2}} \approx 1.307$. Therefore, we need to consider three separate cases to determine the bandwidth of the lowpass frequency response.

Case I – Relative Peak > $\sqrt{2}$ **(**$Q_0 > \sqrt{1+1/\sqrt{2}}$ **)**

As shown before:

$$\left|\mathbf{T}(j\omega)\right|^2 = \frac{1}{G(u)} = \frac{1}{u^2 + \left(1/Q_0^2 - 2\right)u + 1} \tag{20.58}$$

and for the case of a relative peak:

$$\left|\mathbf{T}(j\omega_p)\right|^2 = \frac{Q_0^2}{1 - 1/\left(4Q_0^2\right)} \tag{20.59}$$

so we have:

$$\frac{\left|\mathbf{T}(j\omega_{1,2})\right|^2}{\left|\mathbf{T}(j\omega_p)\right|^2} = \frac{\left[1 - 1/\left(4Q_0^2\right)\right]/Q_0^2}{u^2 + \left(1/Q_0^2 - 2\right)u + 1} = \frac{1}{2} \tag{20.60}$$

This can be rewritten as:

$$u^2 + \left(1/Q_0^2 - 2\right)u + 1 - 2\left[1 - 1/\left(4Q_0^2\right)\right]/Q_0^2 = 0 \tag{20.61}$$

which can be solved for u to get:

$$u_{1,2} = 1 - \frac{1}{2Q_0^2} \mp \frac{1}{Q_0}\sqrt{1 - \frac{1}{4Q_0^2}} \tag{20.62}$$

The two half-power frequencies are therefore:

$$\begin{aligned} \omega_1 &= \omega_0\sqrt{u_1} \\ \omega_2 &= \omega_0\sqrt{u_2} \end{aligned} \tag{20.63}$$

and the bandwidth is thus:

$$B = \omega_2 - \omega_1 = \omega_0\left(\sqrt{u_2} - \sqrt{u_1}\right) \tag{20.64}$$

For high Q_0 (say $Q_0 \geq 5$), we can simplify this result in the following way. Firstly:

$$u_{1,2} \approx 1 \mp \frac{1}{Q_0} \tag{20.65}$$

then, using the binomial series:

$$(1+x)^{\frac{1}{2}} = 1 + \frac{1}{2}x - \frac{1}{8}x^2 + \cdots \tag{20.66}$$

we can approximate the $\sqrt{u}$ terms with the first two terms from the series:

$$\begin{aligned} B &= \omega_0\left(\sqrt{u_2} - \sqrt{u_1}\right) \\ &\approx \omega_0\left(\sqrt{1+\frac{1}{Q_0}} - \sqrt{1-\frac{1}{Q_0}}\right) \\ &\approx \omega_0\left(1+\frac{1}{2Q_0} - \left(1+\frac{1}{2Q_0}\right)\right) \\ &\approx \frac{\omega_0}{Q_0} \end{aligned} \tag{20.67}$$

The lowpass circuit in this case exhibits bandpass behaviour, and it is debatable whether we should still call it a lowpass filter. However, since the circuit still passes low frequencies down to DC (but at levels which are below half-power), the circuit is still classified as a lowpass filter. Perhaps the best name would be a "lowpass filter with band enhancement".

Case II – Relative Peak < $\sqrt{2}$ ($1/\sqrt{2} < Q_0 < \sqrt{1+1/\sqrt{2}}$)

From Eq. (20.62), it can be shown that when $Q_0 < \sqrt{1+1/\sqrt{2}} \approx 1.307$, then u_1 will be negative and the lower half-power frequency will cease to exist. In this case, the bandwidth is:

$$\begin{aligned} B &= \omega_2 = \omega_0\sqrt{u_2} \\ &= \omega_0\sqrt{1-\frac{1}{2Q_0^2}+\frac{1}{Q_0}\sqrt{1-\frac{1}{4Q_0^2}}} \end{aligned} \tag{20.68}$$

This is because the peak response falls below $\sqrt{2}$, and therefore $\left|\mathbf{T}(j\omega_p)\right|/\sqrt{2} < 1$. This case is illustrated below:

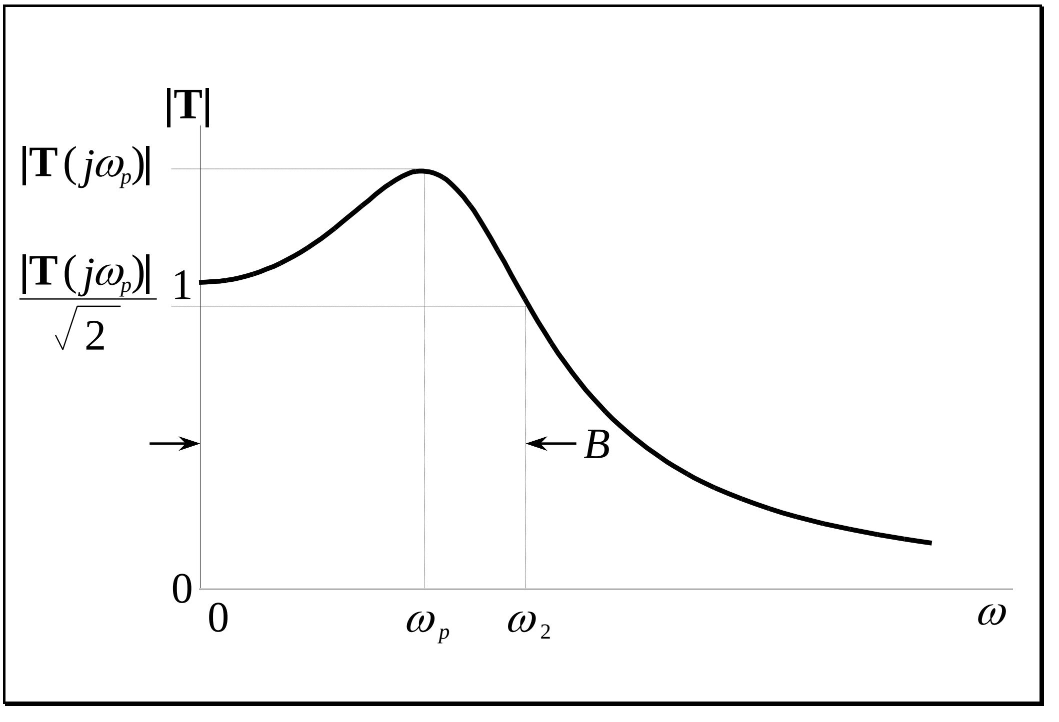


Figure 20.12

Case III – No Relative Peak ($Q_0 < 1/\sqrt{2}$)

If there is no relative peak in the magnitude response (for $Q_0 < 1/\sqrt{2} \approx 0.7071$), then the peak response occurs at DC with a magnitude of 1. We then have only one half-power frequency.

In a manner similar to the derivation for Case I, we can show that the bandwidth is given by:

$$B = \omega_0 \sqrt{1 - \frac{1}{2Q_0^2} + \sqrt{1 + \left(1 - \frac{1}{2Q_0^2}\right)^2}} \tag{20.69}$$

This case is illustrated below:

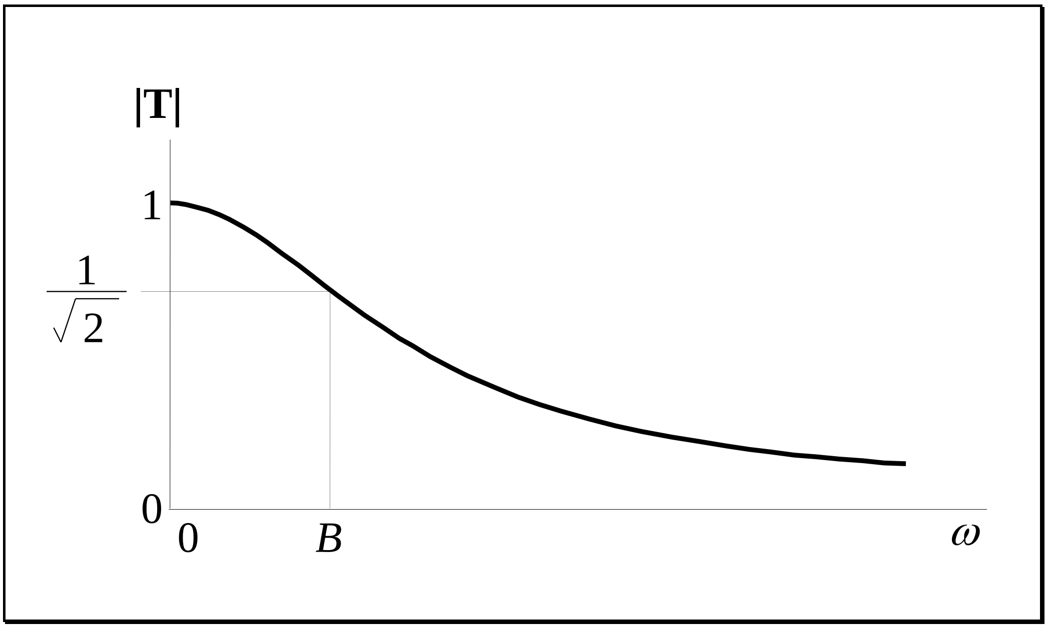


Figure 20.13

20.8.3 Bode Plots

The magnitude and phase Bode plots for a range of values of Q_0 are shown below for a normalised resonance frequency of $\omega_0 = 1$:

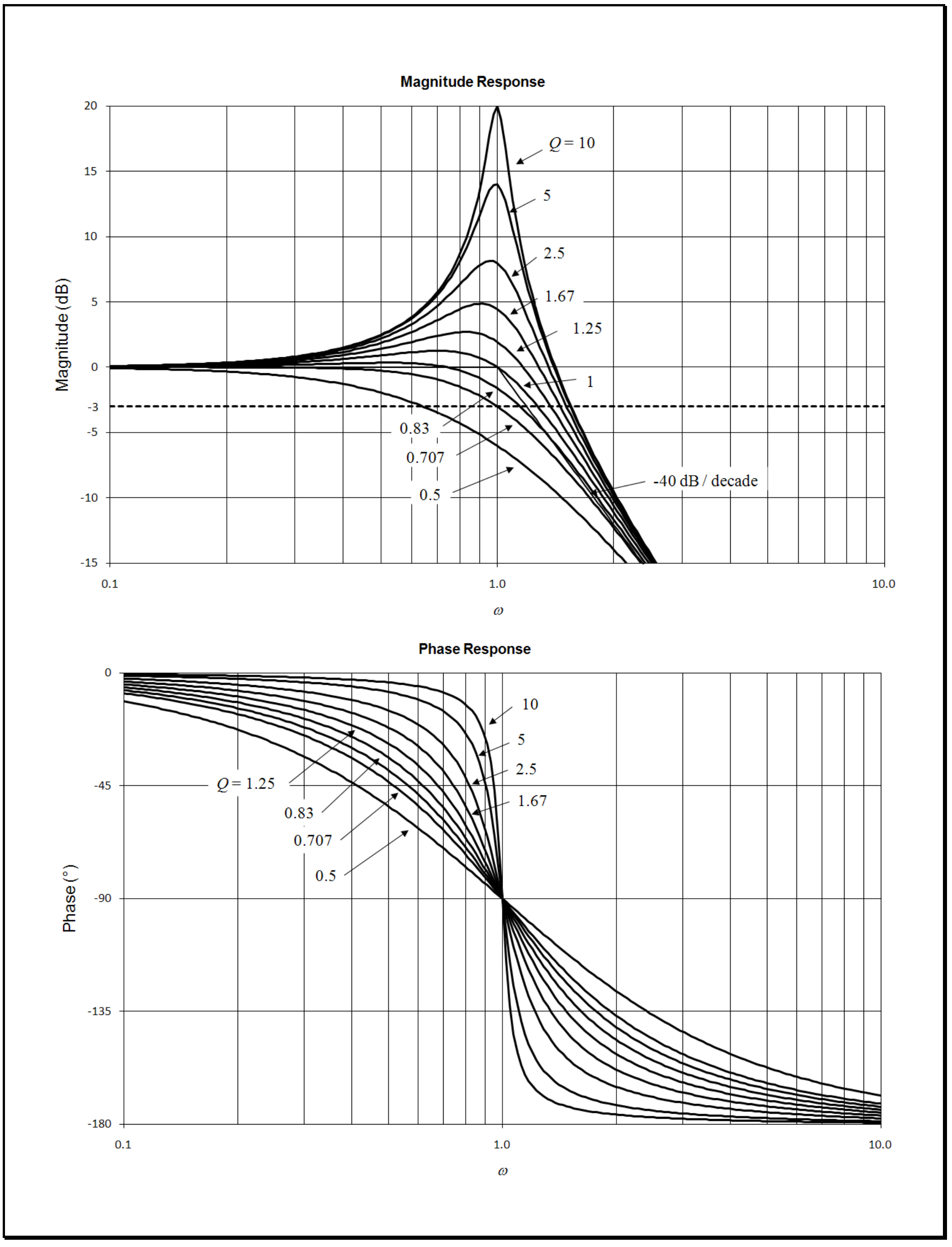


Figure 20.14

The asymptotic Bode magnitude plot decreases at the rate of -40 dB / decade, and this is sometimes described as *two-pole rolloff*. Note the symmetry in the phase response around -90°. *Do you know why?*

20.9 The Second-Order Highpass Frequency Response

Consider the series *RLC* circuit again, but this time the response is taken as the voltage across the inductor:

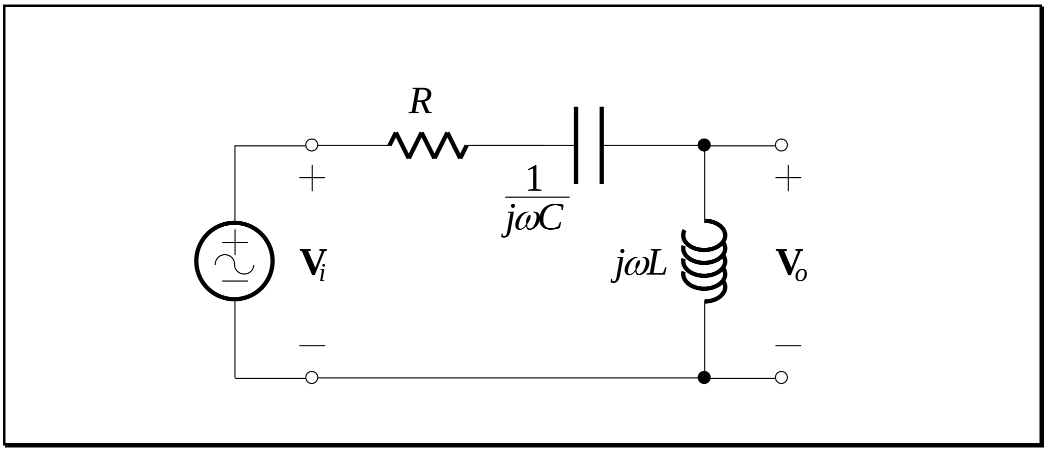


Figure 20.15

The frequency response is given by the voltage divider rule:

$$\mathbf{T}(j\omega) = \frac{\mathbf{V}_o}{\mathbf{V}_i}$$

$$= \frac{j\omega L}{R + j\omega L + 1/j\omega C}$$

$$= \frac{-\omega^2}{1/LC - \omega^2 + j\omega R/L} \quad (20.70)$$

Noting that $\omega_0 = 1/\sqrt{LC}$ and $Q_0 = \omega_0 L/R$ for the series *RLC* circuit, this can be written as:

$$\mathbf{T}(j\omega) = \frac{-\omega^2}{\omega_0^2 - \omega^2 + j\omega(\omega_0/Q_0)} \quad (20.71)$$

This has the form of a second-order *highpass* frequency response – it attenuates low frequencies but passes high frequencies.

The magnitude response is:

The magnitude response of a highpass second-order frequency response

$$\left|\mathbf{T}(j\omega)\right| = \frac{\omega^2}{\sqrt{\left(\omega_0^2 - \omega^2\right)^2 + \left(\omega\omega_0/Q_0\right)^2}} \tag{20.72}$$

and the phase is:

The phase response of a highpass second-order frequency response

$$\angle\mathbf{T}(j\omega) = 180° - \tan^{-1}\left(\frac{\omega\omega_0/Q_0}{\omega_0^2 - \omega^2}\right) \tag{20.73}$$

The magnitude and phase functions are plotted below for $Q_0 = 1.25$:

Typical magnitude and phase responses of a highpass second-order frequency response

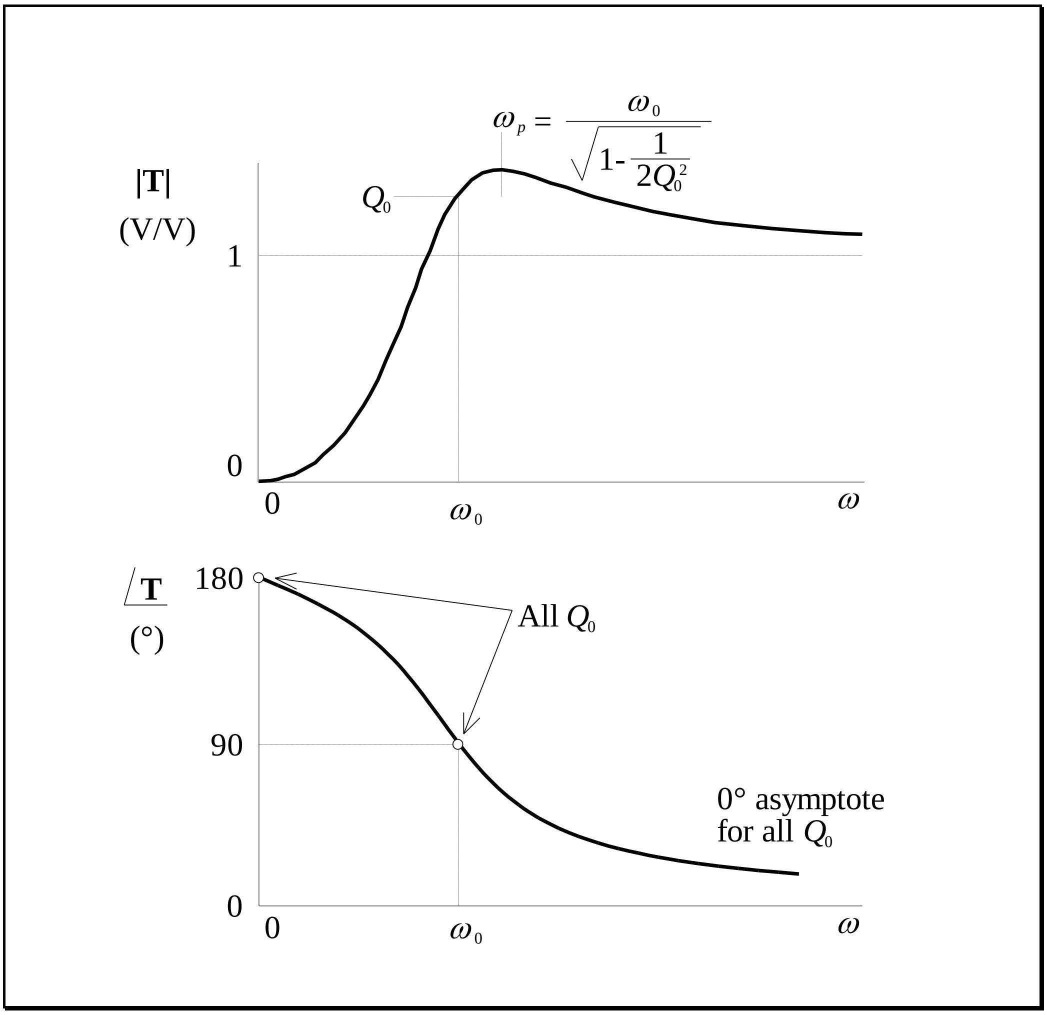


Figure 20.16

20.9.1 Peak Frequency

In a similar manner to the lowpass response, the peak of the frequency response for the highpass response **does not** correspond to the resonance frequency ω_0. If we proceed in a similar manner to that shown for the lowpass response, we find that the frequency at which the magnitude response reaches a peak is:

$$\begin{aligned} \omega_p &= \frac{\omega_0}{\sqrt{1-\dfrac{1}{2Q_0^2}}}, && Q_0 \geq 1/\sqrt{2} \\ \omega_p &= \infty, && Q_0 < 1/\sqrt{2} \end{aligned} \tag{20.74}$$

Notice that the peak response always occurs *after* the resonance frequency for the highpass response, and we approach $\omega_p \approx \omega_0$ for high Q_0 (say $Q_0 \geq 5$).

At the peak frequency, the magnitude of the frequency response is:

$$\begin{aligned} \left|\mathbf{T}(j\omega_p)\right| &= \frac{Q_0}{\sqrt{1-\dfrac{1}{4Q_0^2}}}, && Q_0 \geq 1/\sqrt{2} \\ \left|\mathbf{T}(j\omega_p)\right| &= 1, && Q_0 < 1/\sqrt{2} \end{aligned} \tag{20.75}$$

For high Q_0 (say $Q_0 \geq 5$), the magnitude response is $\left|\mathbf{T}(j\omega_p)\right| \approx Q_0$.

20.9.2 Bandwidth

If there is a relative peak in the magnitude response, then the two half-power frequencies are the same as for the lowpass case (the circuit exhibits bandpass behaviour) and the bandwidth is thus:

$$B = \omega_2 - \omega_1 = \omega_0\left(\sqrt{u_2} - \sqrt{u_1}\right) \tag{20.76}$$

If $Q_0 < \sqrt{1+1/\sqrt{2}} \approx 1.307$ then the peak is below $\sqrt{2}$ and the bandwidth is infinite.

20.10 Standard Forms of Second-Order Frequency Responses

The table below shows the possibilities and names associated with the second-order frequency response.

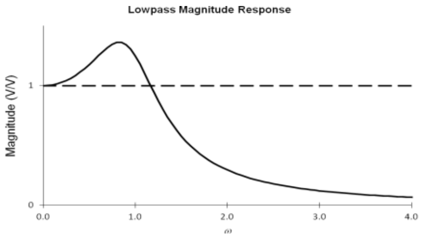
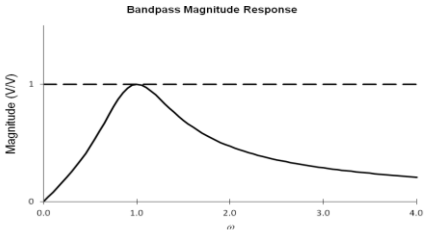
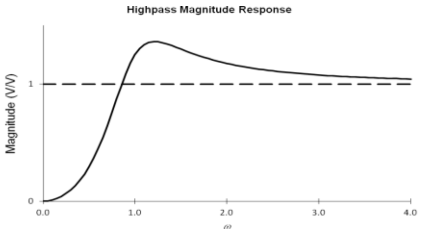
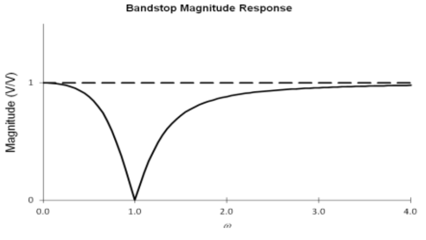
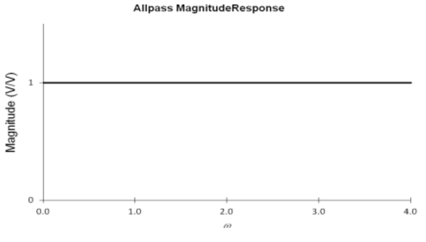
Frequency Response	Magnitude Response	Name
$\mathbf{T}_{LP} = \dfrac{\omega_0^2}{\omega_0^2 - \omega^2 + j\omega(\omega_0/Q_0)}$		Lowpass
$\mathbf{T}_{BP} = \dfrac{j\omega(\omega_0/Q_0)}{\omega_0^2 - \omega^2 + j\omega(\omega_0/Q_0)}$		Bandpass
$\mathbf{T}_{HP} = \dfrac{-\omega^2}{\omega_0^2 - \omega^2 + j\omega(\omega_0/Q_0)}$		Highpass
$\mathbf{T}_{BS} = \dfrac{\omega_0^2 - \omega^2}{\omega_0^2 - \omega^2 + j\omega(\omega_0/Q_0)}$		Bandstop “notch”
$\mathbf{T}_{AP} = \dfrac{\omega_0^2 - \omega^2 - j\omega(\omega_0/Q_0)}{\omega_0^2 - \omega^2 + j\omega(\omega_0/Q_0)}$		Allpass

Table 20.1 – Standard Forms of Second-Order Frequency Responses

20.11 Summary

- Resonance is a phenomenon that only occurs in 2nd-order or higher circuits, and even then, only under certain conditions. It occurs when the forcing function drives the circuit near one of its natural frequencies of oscillation.

- For the *RLC* circuit, we define the resonance frequency:

$$\omega_r = \frac{1}{\sqrt{LC}} = \omega_0$$

- For the *RLC* circuit, we define the quality factor at resonance:

parallel $$Q_0 = \omega_0 CR = \frac{R}{X_{C0}} = \frac{R}{X_{L0}} = R\sqrt{\frac{C}{L}}$$

series $$Q_0 = \frac{\omega_0 L}{R} = \frac{X_{L0}}{R} = \frac{X_{C0}}{R} = \frac{1}{R}\sqrt{\frac{L}{C}}$$

- The *RLC* circuit can be used to create a lowpass, bandpass or highpass filter.

- The two *half-power frequencies* ω_1 and ω_2 are those frequencies at which the magnitude of the response is $1/\sqrt{2}$ times its maximum value (or -3 dB down from the peak).

- The *bandwidth* of an *RLC* circuit is defined as the difference of the two half-power frequencies (if they exist):

$$B = \omega_2 - \omega_1 \quad \text{rads}^{-1} \qquad \text{or} \qquad B = f_2 - f_1 \quad \text{Hz}$$

- For a lowpass *RLC* circuit, if there is only one half-power frequency then the bandwidth is equal to it (the formulae above apply with $f_1 = \omega_1 = 0$).

20.12 References

Hayt, W. & Kemmerly, J.: *Engineering Circuit Analysis*, 3rd *Ed.*, McGraw-Hill, 1984.

Exercises

1.

Find ω_0 and Q_0 for a parallel resonant circuit in which:

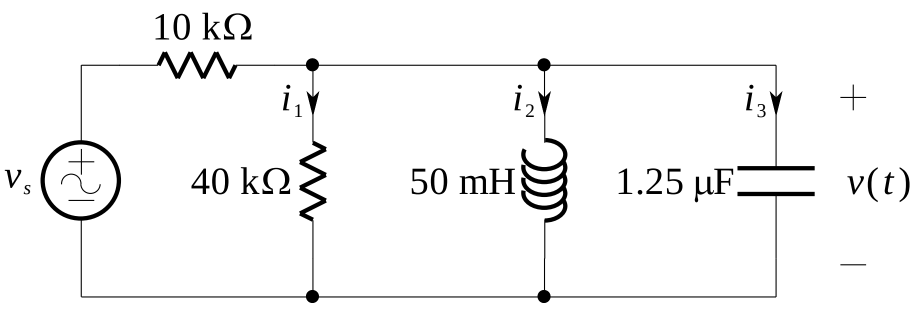
(a) $C = 1/4\,\mu\text{F}$, L = 4 H, $R = 20\,\text{k}\Omega$.

(b) $\alpha = 1000\,\text{s}^{-1}$, $C = 5\,\text{nF}$, $L = 1/72\,\text{H}$.

(c) $\alpha = 50\,\text{s}^{-1}$, $\omega_d = 600\ \text{rads}^{-1}$

2.

Let $v_s(t) = 100\cos(\omega_0 t)\ \text{V}$ in the circuit shown below:



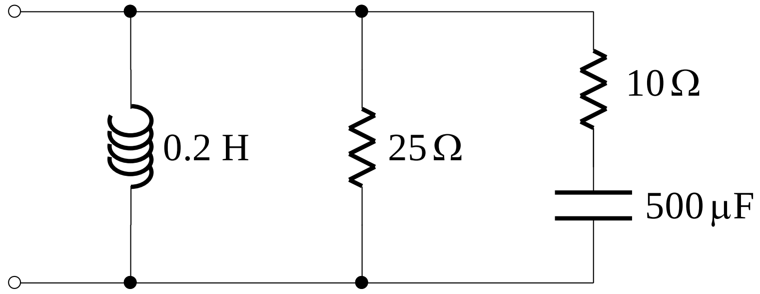
(a) Find the equivalent parallel *RLC* circuit and then determine ω_0, Q_0 and $v(t)$.

(b) Find $i_1(t)$, $i_2(t)$ and $i_3(t)$.

(c) Calculate the average power loss in the $10\,\text{k}\Omega$ resistor and the maximum energy stored in the inductor.

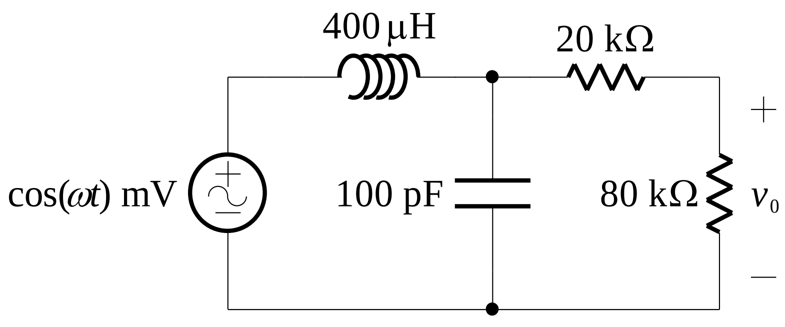
3.

Find the resonance frequency of the circuit shown:



4.

Consider the circuit shown below:



Find:

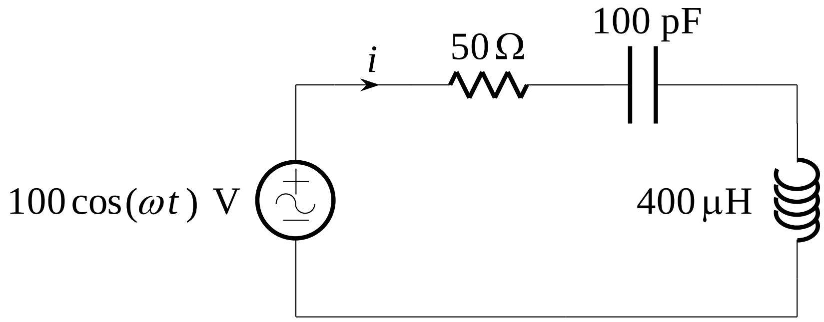
(a) ω_r (b) Q_0 (c) B (d) ω_1 (e) ω_2 (f) $\mathbf{V}_0$ at ω_r

5.

A parallel *RLC* circuit used in a radio frequency (RF) amplifier is intended to have an impedance magnitude of $5\,\mathrm{k}\Omega$ at resonance, $\omega_0 = 10^7\ \mathrm{rads}^{-1}$, and $3\,\mathrm{k}\Omega$ at a frequency 5 kHz below resonance. Specify *R*, *L* and *C*.

6.

In the series *RLC* circuit shown below:

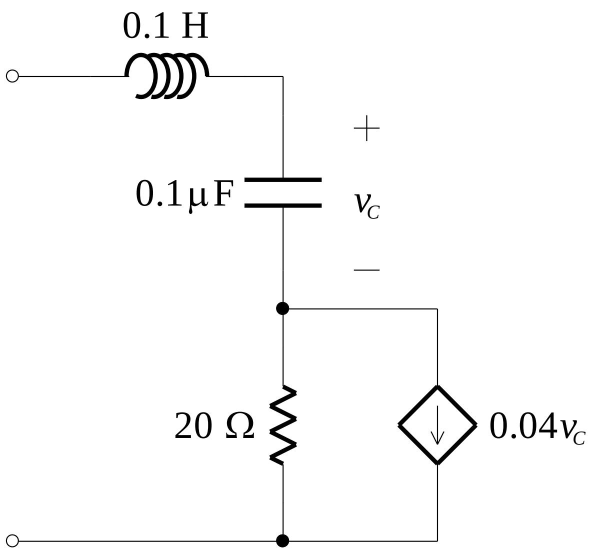


(a) At what value of ω is the amplitude of *i* a maximum?

(b) By how many rads^{-1} would ω have to be increased to reduce $|\mathbf{I}|$ by 5%?

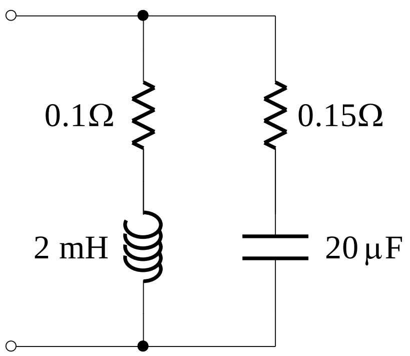
7.

Find the effective values of ω_0 and Q_0 for the network shown below:

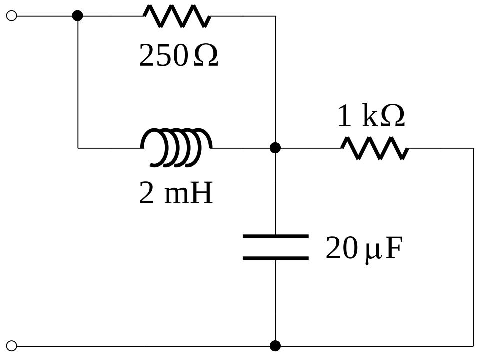


8.

Determine reasonably accurate values of ω_r and Q_0 for the resonant circuits shown below:



(a)



(b)

9.

A series *RLC* circuit has an impedance of $10 + j40\,\Omega$ at $\omega = 100\,\text{rads}^{-1}$. After it is scaled in magnitude and frequency by the same factor (i.e. $k_m = k_f$), it is found to have an impedance of $30 - j180\,\Omega$ at $\omega = 50\,\text{rads}^{-1}$. Determine the elements in the original network.

James Clerk Maxwell (1831-1879)

Maxwell produced a most spectacular work of individual genius – he unified electricity and magnetism. Maxwell was able to summarize all observed phenomena of electrodynamics in a handful of partial differential equations known as *Maxwell's equations*[3]:

$$\nabla \cdot \mathbf{E} = \frac{\rho}{\varepsilon}$$
$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}$$
$$\nabla \cdot \mathbf{B} = 0$$
$$\nabla \times \mathbf{B} = \mu \mathbf{J} + \mu\varepsilon \frac{\partial \mathbf{E}}{\partial t}$$

From these he was able to predict that there should exist electromagnetic waves which could be transmitted through free space at the speed of light. The revolution in human affairs wrought by these equations and their experimental verification by Heinrich Hertz in 1888 is well known: wireless communication, control and measurement - so spectacularly demonstrated by television and radio transmissions across the globe, to the moon, and even to the edge of the solar system!

James Maxwell was born in Edinburgh, Scotland. His mother died when he was 8, but his childhood was something of a model for a future scientist. He was endowed with an exceptional memory, and had a fascination with mechanical toys which he retained all his life. At 14 he presented a paper to the Royal Society of Edinburgh on ovals. At 16 he attended the University of Edinburgh where the library still holds records of the books he borrowed while still an undergraduate – they include works by Cauchy on differential equations, Fourier on the theory of heat, Newton on optics, Poisson on

---

[3] It was Oliver Heaviside, who in 1884-1885, cast the long list of equations that Maxwell had given into the compact and symmetrical set of four vector equations shown here and now universally known as "Maxwell's equations". It was in this new form ("Maxwell redressed," as Heaviside called it) that the theory eventually passed into general circulation in the 1890s.

mechanics and Taylor's scientific memoirs. In 1850 he moved to Trinity College, Cambridge, where he graduated with a degree in mathematics in 1854. Maxwell was edged out of first place in their final examinations by his classmate Edward Routh, who was also an excellent mathematician.

Maxwell stayed at Trinity where, in 1855, he formulated a "theory of three primary colour-perceptions" for the human perception of colour. In 1855 and 1856 he read papers to the Cambridge Philosophical Society "On Faraday's Lines of Force" in which he showed how a few relatively simple mathematical equations could express the behaviour of electric and magnetic fields.

In 1856 he became Professor of Natural Philosophy at Aberdeen, Scotland, and started to study the rings of Saturn. In 1857 he showed that stability could be achieved only if the rings consisted of numerous small solid particles, an explanation now confirmed by the Voyager spacecraft.

In 1860 Maxwell moved to King's College in London. In 1861 he created the first colour photograph – of a Scottish tartan ribbon – and was elected to the Royal Society. In 1862 he calculated that the speed of propagation of an electromagnetic wave is approximately that of the speed of light:

> *We can scarcely avoid the conclusion that light consists in the transverse undulations of the same medium which is the cause of electric and magnetic phenomena.*

Maxwell's famous account, "A Dynamical Theory of the Electromagnetic Field" was read before a largely perplexed Royal Society in 1864. Here he brought forth, for the first time, the equations which comprise the basic laws of electromagnetism.

All the mathematical sciences are founded on relations between physical laws and laws of numbers, so that the aim of exact science is to reduce the problems of nature to the determination of quantities by operations with numbers. – James Clerk Maxwell

Maxwell also continued work he had begun at Aberdeen, on the kinetic theory of gases (he had first considered the problem while studying the rings of Saturn). In 1866 he formulated, independently of Ludwig Boltzmann, the kinetic theory of gases, which showed that temperature and heat involved only molecular motion.

Maxwell was the first to publish an analysis of the effect of a capacitor in a circuit containing inductance, resistance and a sinusoidal voltage source, and to show the conditions for resonance. The way in which he came to solve this problem makes an interesting story:

> *Maxwell was spending an evening with Sir William Grove who was then engaged in experiments on vacuum tube discharges. He used an induction coil for this purpose, and found the if he put a capacitor in series with the primary coil he could get much larger sparks. He could not see why. Grove knew that Maxwell was a splendid mathematician, and that he also had mastered the science of electricity, especially the theoretical art of it, and so he thought he would ask this young man [Maxwell was 37] for an explanation. Maxwell, who had not had very much experience in experimental electricity at that time, was at a loss. But he spent that night in working over his problem, and the next morning he wrote a letter to Sir William Grove explaining the whole theory of the capacitor in series connection with a coil. It is wonderful what a genius can do in one night!*

Maxwell's letter, which began with the sentence, "Since our conversation yesterday on your experiment on magneto-electric induction, I have considered it mathematically, and now send you the result," was dated March 27, 1868. Preliminary to the mathematical treatment, Maxwell gave in this letter an unusually clear exposition of the analogy existing between certain electrical and mechanical effects. In the postscript, or appendix, he gave the mathematical theory of the experiment. Using different, but equivalent symbols, he derived and solved the now familiar expression for the current i in such a circuit:

$$L\frac{di}{dt} + Ri + \frac{1}{C}\int i\,dt = V\sin\omega t$$

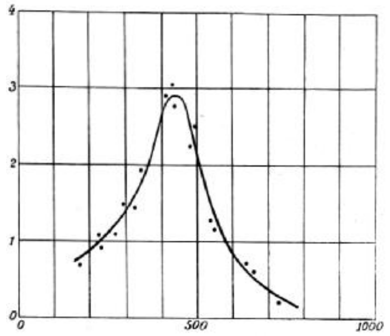
The solution for the current amplitude of the resulting sinusoid, in the steady-state is:

$$I = \frac{V}{\sqrt{R^2 + \left(\omega L - \dfrac{1}{\omega C}\right)^2}}$$

from which Maxwell pointed out that the current would be a maximum when:

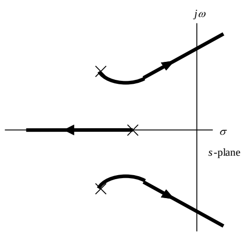
$$\omega L = \frac{1}{\omega C}$$

Following Maxwell, Heinrich Hertz later showed a thorough acquaintance with electrical resonance and made good use of it in his experimental apparatus that proved the existence of electromagnetic waves, as predicted by Maxwell's equations. In the first of his series of papers describing his experiment, "On Very Rapid Electric Oscillations", published in 1887, he devotes one section to a discussion of "Resonance Phenomena" and published the first electrical resonance curve:



The first electrical resonance curve published, by Hertz, 1887

When creating his standard for electrical resistance, Maxwell wanted to design a governor to keep a coil spinning at a constant rate. He made the system stable by using the idea of negative feedback. It was known for some time that the governor was essentially a centrifugal pendulum, which sometimes exhibited "hunting" about a set point – that is, the governor would oscillate about an equilibrium position until limited in amplitude by the throttle valve or the travel allowed to the bobs. This problem was solved by Airy in 1840 by fitting a damping disc to the governor. It was then possible to minimize speed fluctuations by adjusting the "controller gain". But as the gain was increased, the governors would burst into oscillation again. In 1868, Maxwell published his paper "On Governors" in which he derived the equations of motion of engines fitted with governors of various types, damped in several ways, and explained in mathematical terms the source of the oscillation. He was also able to set bounds on the parameters of the system that would ensure stable operation. He posed the problem for more complicated control systems, but thought that a general solution was insoluble. It was left to Routh some years later to solve the general problem of linear system stability: "It has recently come to my attention that my good friend James Clerk Maxwell has had difficulty with a rather trivial problem…".



In 1870 Maxwell published his textbook *Theory of Heat*. The following year he returned to Cambridge to be the first Cavendish Professor of Physics – he designed the Cavendish laboratory and helped set it up.

The four partial differential equations describing electromagnetism, now known as Maxwell's equations, first appeared in fully developed form in his *Treatise on Electricity and Magnetism* in 1873. The significance of the work was not immediately grasped, mainly because an understanding of the atomic nature of electromagnetism was not yet at hand.

The Cavendish laboratory was opened in 1874, and Maxwell spent the next 5 years editing Henry Cavendish's papers.

Maxwell died of abdominal cancer, in 1879, at the age of forty-eight. At his death, Maxwell's reputation was uncertain. He was recognised to have been an exceptional scientist, but his theory of electromagnetism remained to be convincingly demonstrated. About 1880 Hermann von Helmholtz, an admirer of Maxwell, discussed the possibility of confirming his equations with a student, Heinrich Hertz. In 1888 Hertz performed a series of experiments which produced and measured electromagnetic waves and showed how they behaved like light. Thereafter, Maxwell's reputation continued to grow, and he may be said to have prepared the way for twentieth-century physics.

References

Blanchard, J.: *The History of Electrical Resonance*, Bell System Technical Journal, Vol. 20 (4), p. 415, 1941.

21 Second-Order Op-Amp Filters

Contents

21.2

Introduction

With the advent of op-amps and circuit miniaturization, engineers developed what is known as a *universal filter*. It's frequency response takes the form of a biquadratic equation, and so it is also known as a *biquad*. Depending on the connections made and the point at which the output is taken, the universal filter can deliver lowpass, highpass, bandpass, bandstop (notch) and allpass responses. It is one of the most useful circuits to the electrical engineer and is widely available.

21.1 Filter Design Parameters

Jargon fills a special need for the engineer. It is shorthand that permits the expression of ideas quickly and compactly. "Design a lowpass filter with an ω_0 of 10,000 and a Q_0 of 5." We will explore how this can be carried out.

We begin with the *RLC* circuit shown below, which has the now familiar form of a voltage-divider circuit.

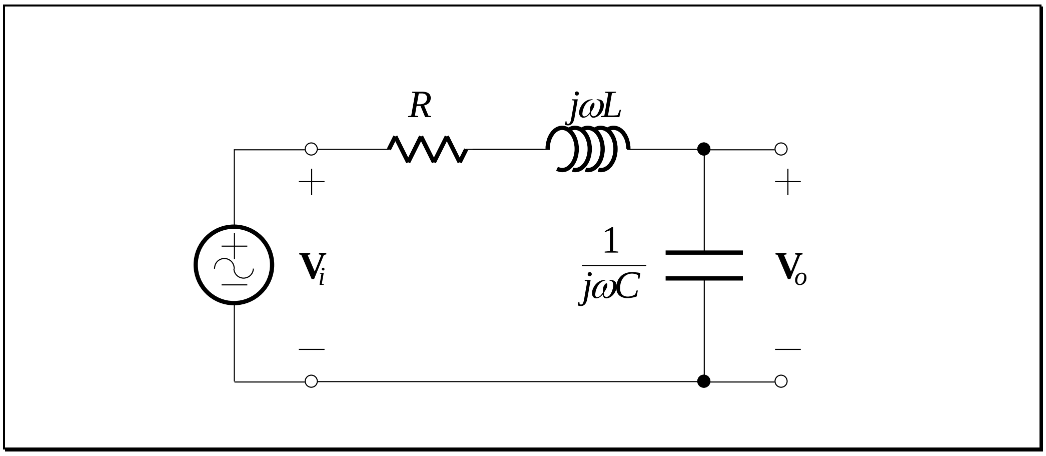


Figure 21.1

The frequency response for this circuit is given by:

$$\mathbf{T}(j\omega) = \frac{\mathbf{V}_o}{\mathbf{V}_i} = \frac{1/LC}{1/LC - \omega^2 + j\omega R/L} \tag{21.1}$$

This result can be put into a *standard form* by noting that $\omega_0 = 1/\sqrt{LC}$ and $Q_0 = \omega_0 L/R$ for the series *RLC* circuit. We can then write:

$$\mathbf{T}(j\omega) = \frac{\omega_0^2}{\omega_0^2 - \omega^2 + j\omega(\omega_0/Q_0)} \tag{21.2}$$

The two parameters ω_0 and Q_0 uniquely specify the standard form of the second-order frequency response.

ω_0 and Q_0 uniquely specify a second-order frequency response

The historical identification of Q_0 with *RLC* circuits is no longer appropriate, since we can identify many kinds of circuits with the parameter Q_0. We can now make the association of ω_0 and Q_0 with any second-order circuit, as suggested by the figure below:

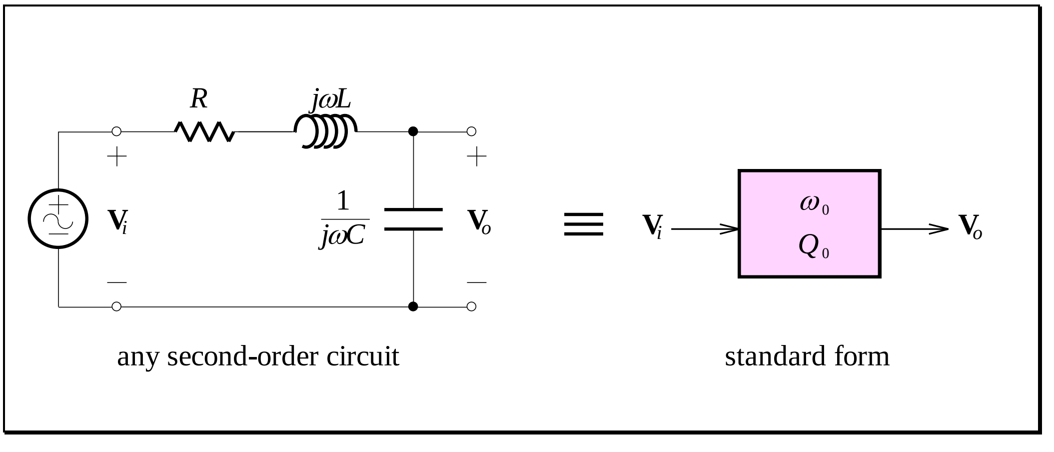


Figure 21.2

21.2 The Lowpass Biquad Circuit

The standard form of a lowpass second-order frequency response, as in Eq. (21.2), does not recognise the availability of gain that is possible with active circuits. Also, an active circuit may be inverting or non-inverting. A more general form for $\mathbf{T}(j\omega)$ is therefore:

Standard form of a lowpass second-order frequency response with gain

$$\mathbf{T}(j\omega) = \frac{\pm H\omega_0^2}{\omega_0^2 - \omega^2 + j\omega(\omega_0/Q_0)} \tag{21.3}$$

We seek a circuit that will implement this second-order frequency response, as well as any other "biquadratic" frequency response. (A biquadratic, or biquad frequency response is similar to the way a bilinear frequency response was defined – a biquadratic function is a ratio of second-order polynomials).

Normalising so that $\omega_0 = 1$, and anticipating an inverting realisation for the frequency response, we have:

Standard form of a normalised lowpass second-order frequency response with gain

$$\mathbf{T}(j\omega) = \frac{-H}{1 - \omega^2 + j\omega(1/Q_0)} = \frac{\mathbf{V}_o}{\mathbf{V}_i} \tag{21.4}$$

We can manipulate this equation so that it has a form that can be identified with simple circuits we have seen before. We first rewrite Eq. (21.4) as:

$$[1 + j\omega(j\omega + 1/Q_0)]\mathbf{V}_o = -H\mathbf{V}_i \tag{21.5}$$

Dividing by $j\omega(j\omega + 1/Q_0)$, it becomes:

$$\left[1 + \frac{1}{j\omega(j\omega + 1/Q_0)}\right]\mathbf{V}_o = \frac{-H}{j\omega(j\omega + 1/Q_0)}\mathbf{V}_i \tag{21.6}$$

We can manipulate further to form:

$$\mathbf{V}_o = \left[\frac{-1}{j\omega}\mathbf{V}_o + \frac{-H}{j\omega}\mathbf{V}_i \right] \cdot (-1) \cdot \left(\frac{-1}{j\omega + 1/Q_0} \right) \quad (21.7)$$

Second-order frequency response made from first-order parts

The (–1) term can be realised by an inverting circuit of gain 1. The factor $1/(j\omega + 1/Q_0)$ is realised by a "lossy" inverting integrator. Two operations are indicated by the remaining factor. The circuit realisation must produce a sum of voltages, and it must have a frequency response of the form $(-1/j\omega)$.

The three circuits that provide for these three operations are shown below:

The three first-order circuits that make a second-order circuit

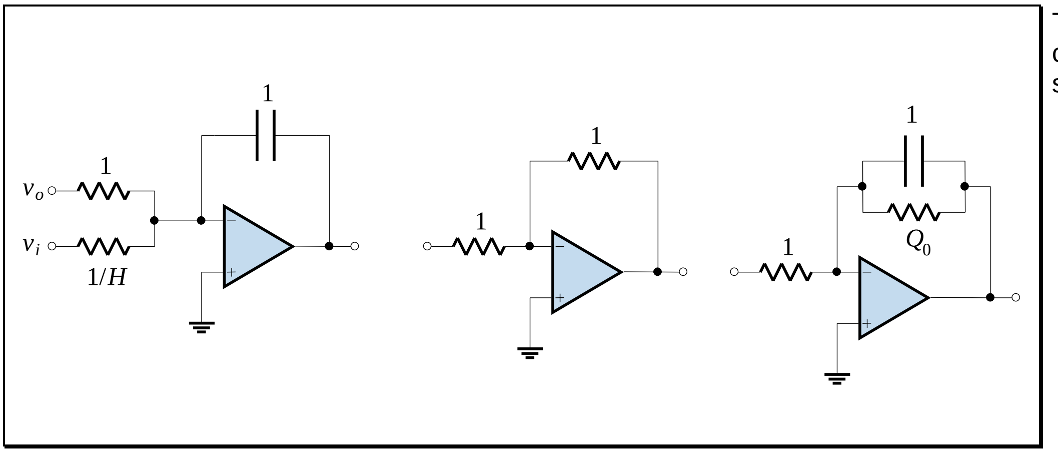


Figure 21.3

If we connect the three circuits together, including a *feedback* connection of the output v_o to the input, the result is a scaled version of the *Tow-Thomas biquad circuit*:

The normalised Tow-Thomas biquad circuit

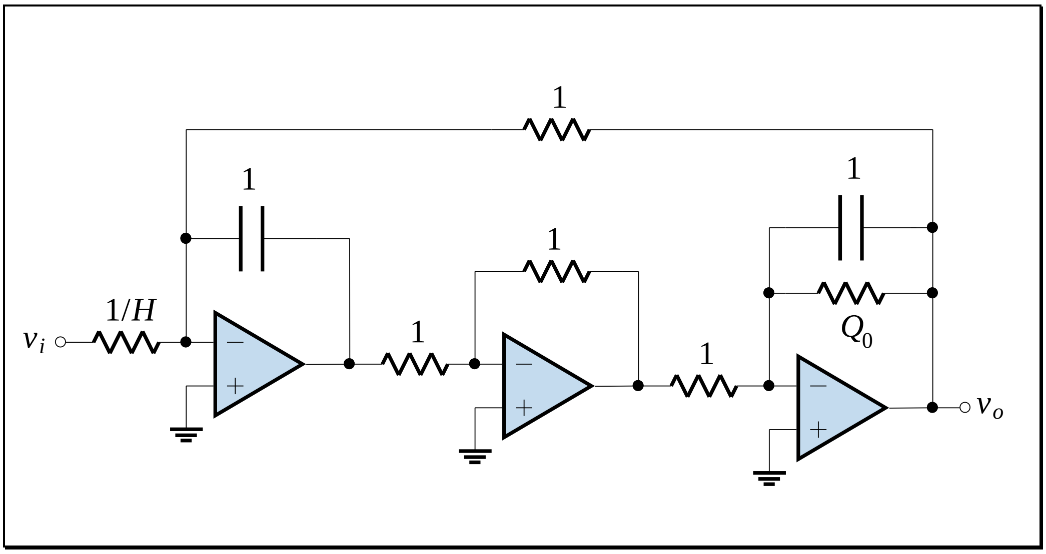


Figure 21.4

There are many circuits that implement biquadratic frequency responses. The Tow-Thomas circuit is one of them, the Kerwin-Huelsman-Newcomb (KHN) circuit is another. For brevity, we will simply refer to the Tow-Thomas biquad circuit as "the biquad".

With the elements identified by *R*'s and *C*'s, we have:

The Tow-Thomas biquad circuit

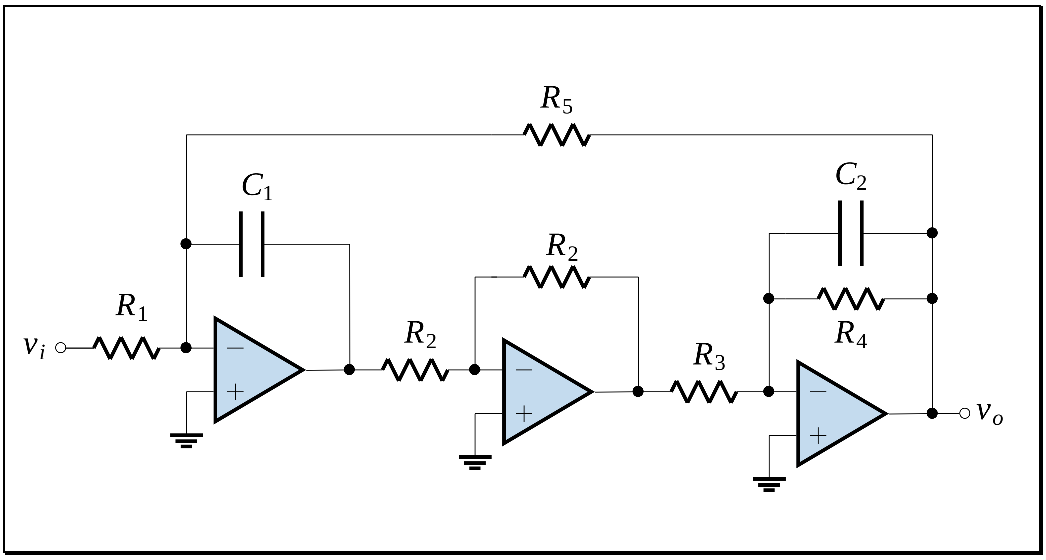


Figure 21.5

Show that the frequency response is:

$$\mathbf{T}(j\omega) = \frac{-1/R_1R_3C_1C_2}{1/R_3R_5C_1C_2 - \omega^2 + j\omega(1/R_4C_2)} \tag{21.8}$$

The biquad's frequency response

Comparing this with Eq. (21.3), we have:

$$\omega_0 = \frac{1}{\sqrt{R_3R_5C_1C_2}} \tag{21.9}$$

$$Q_0 = \sqrt{\frac{R_4^2C_2}{R_3R_5C_1}} \tag{21.10}$$

$$H = \frac{R_5}{R_1} \tag{21.11}$$

The biquad's design equations

An important property of the biquad is that it can be *orthogonally* tuned. Using the above equations, we can devise a *tuning algorithm*:

The biquad can be orthogonally tuned

1. R_3 can be adjusted to the specified value of ω_0.

2. R_4 can then be adjusted to give the specified value of Q_0 without changing ω_0.

3. R_1 can then be adjusted to give the specified value of H without affecting either ω_0 or Q_0.

The biquad's tuning algorithm

Other advantages of the circuit are:

- the input impedance is purely resistive

- there is effectively "pre-amplification" built-in to the topology via the gain setting resistor R_1 (the incoming signal amplitude is amplified and then filtered, which eliminates more "noise" than filtering and *then* amplifying).

EXAMPLE 21.1 Design of a Lowpass Biquad Circuit

We require a circuit that will provide an $\omega_0 = 1000\ \text{rads}^{-1}$, a $Q_0 = 0.866$ and a DC gain of $H = 2$. We set $\omega_0 = 1$ and use the biquad circuit of Figure 21.4 with the values of Q_0 and H specified above.

We then perform frequency scaling to meet the specifications, by setting $k_f = 1000$. The biquad circuit then becomes:

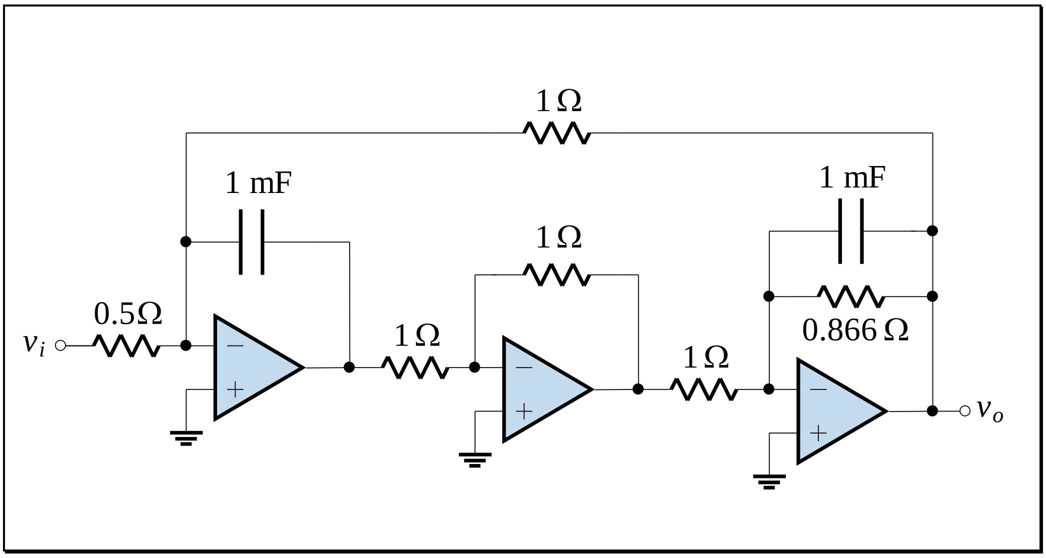


Figure 21.6

We then select $k_m = 10\ 000$ to give convenient element values. A realistic circuit that meets the specifications is then:

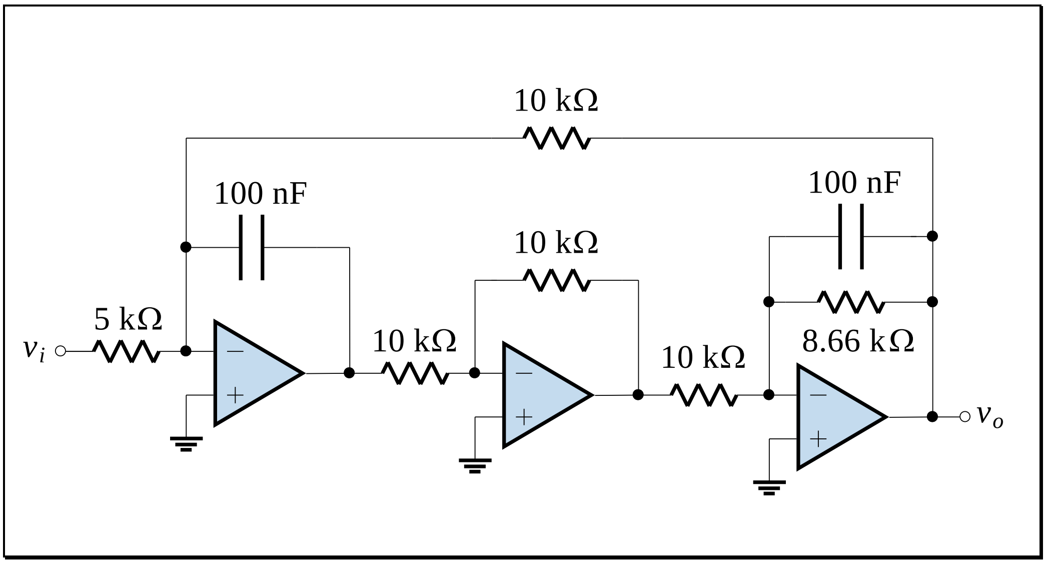
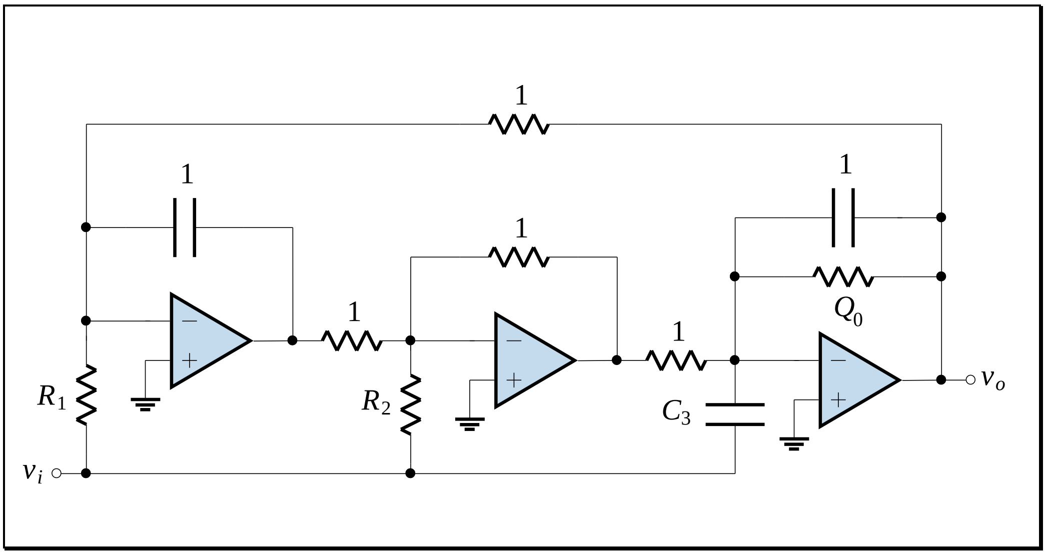


Figure 21.7

21.3 The Universal Biquad Circuit

By applying a feedforward scheme to the lowpass Tow-Thomas biquad circuit, a "universal" filter can be implemented. A universal filter is one that can be made either a lowpass, highpass, bandpass, notch or allpass filter by appropriate selection of component values.

A universal filter can implement any biquadratic frequency response



The normalised Tow-Thomas universal filter

Figure 21.8

In terms of the quantities in Figure 21.8, *show that:*

$$\frac{\mathbf{V}_o}{\mathbf{V}_i} = -\frac{1/R_1 + C_3\omega^2 - j\omega(1/R_2)}{1-\omega^2 + j\omega(1/Q_0)} \tag{21.12}$$

If we choose $C_3 = 1$ and $R_1 = R_2 = \infty$, then the first and third terms in the numerator vanish, leaving only the $-\omega^2$ term. Writing this result we have:

$$\mathbf{T}(j\omega) = \frac{-\omega^2}{1-\omega^2 + j\omega(1/Q_0)} \tag{21.13}$$

The universal biquad circuit can implement a highpass second-order frequency response

which means we have now created a highpass filter.

The universal biquad can implement many second-order frequency responses

Starting with the universal biquad circuit, it is possible to realise a lowpass, highpass, bandpass, bandstop or allpass filter by making simple changes such as the removal of a resistor. The normalised design values for the various responses are given in the table below, where H is the passband gain.

Table of design values for a universal filter

Table 21.1 – Design Values for the Tow-Thomas Universal Filter

Filter Type	Design Values		
	R_1	R_2	C_3
Lowpass	$1/H$	∞	0
Bandpass	∞	Q_0/H	0
Highpass	∞	∞	H
Notch	$(\omega_0/\omega_n)^2/H$	∞	H
Allpass	$1/H$	Q_0/H	H

An example of a commercially available universal filter is the UAF42 from Texas Instruments. This uses the "state-variable" biquad topology rather than the Tow-Thomas. It requires just four external resistors to set the filter parameters, and makes available lowpass, highpass and bandpass outputs:

The UAF42 universal filter from Texas Instruments

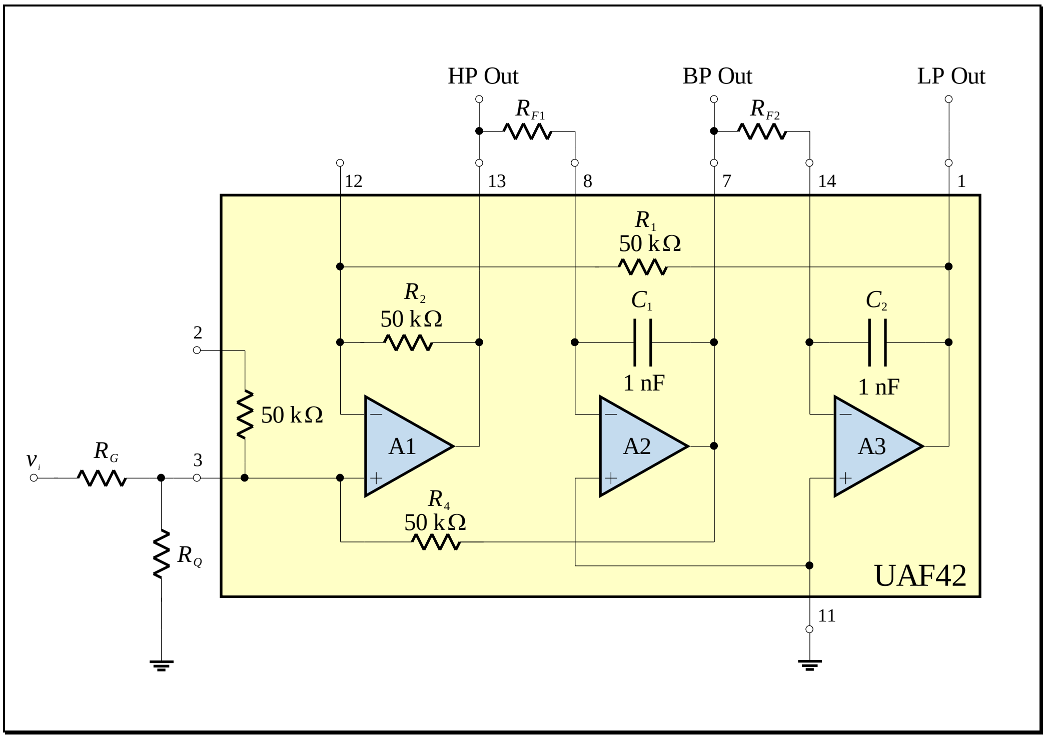


Figure 21.9

21.4 Approximating the Ideal Lowpass Filter

The ideal lowpass filter characteristic is the brick wall filter. Although we cannot achieve the ideal, it provides a basis on which we can rate an approximation. We want $|\mathbf{T}|$ to be as constant as possible in the passband, so that different low frequency signals are passed through the filter without a change in amplitude. In the stopband we require n-pole rolloff, where n is large, in contrast to the $n = 2$ rolloff for the biquad circuit. We want the transition from passband to stopband to be as abrupt as possible.

The features we want when approximating the ideal lowpass filter

This is summarised in the figure below:

Approximating the ideal lowpass filter

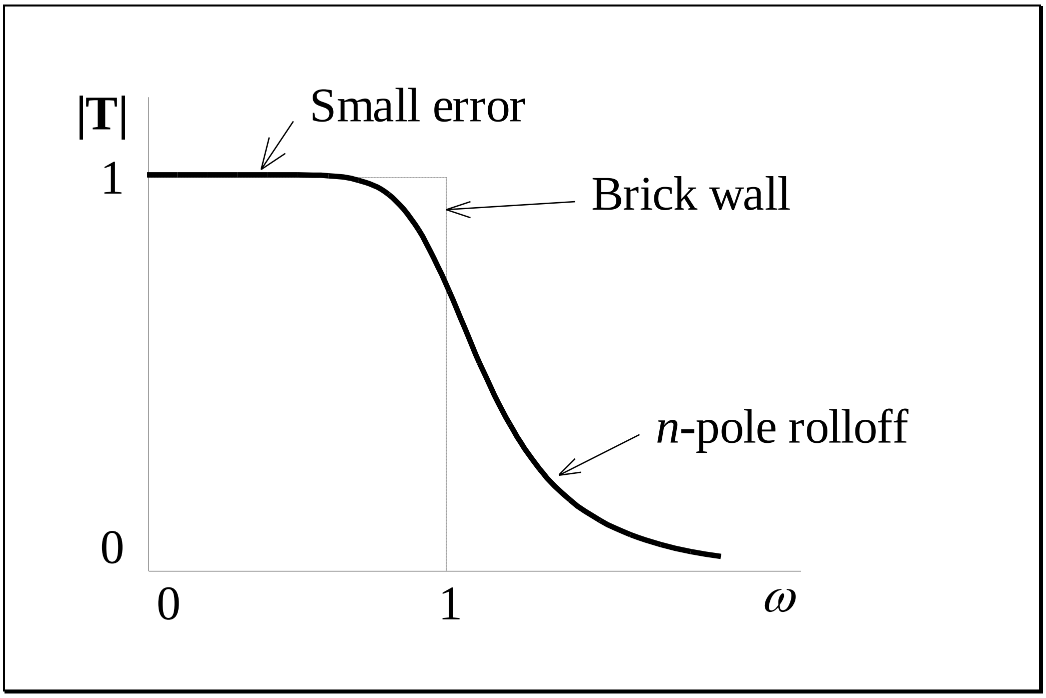


Figure 21.10

The method we will use in the approach to this problem is illustrated below:

We achieve the approximation to the ideal lowpass filter by cascading

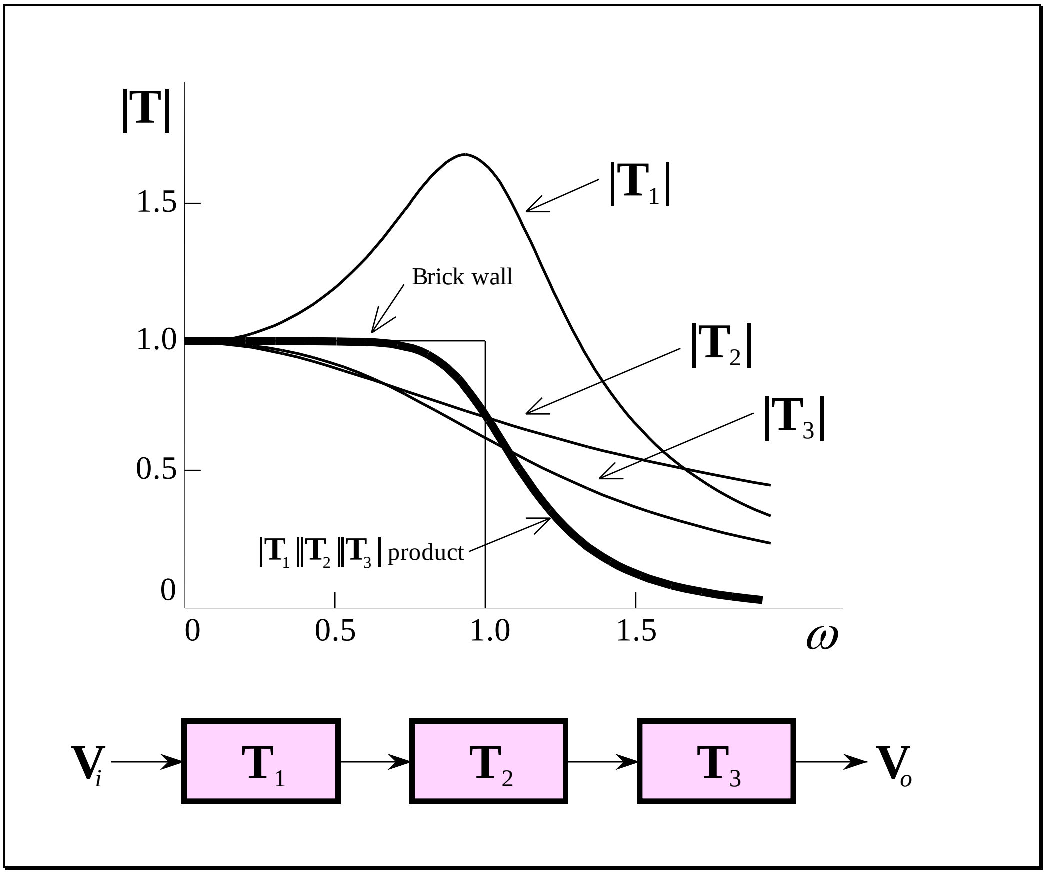


Figure 21.11

The cascaded circuits have the same ω_0 but different Q_0

We will connect modules in cascade such that the overall frequency response is of the form given in Figure 21.10. For the example in Figure 21.11, large values of $|\mathbf{T}_1|$ are just overcome by the small values of $|\mathbf{T}_2|$ and $|\mathbf{T}_3|$ to achieve the approximation to the brick wall. The frequency responses have the same value of ω_0, but different values of Q_0. Determining the required values of Q_0 is a part of *filter design*.

21.5 The Butterworth Lowpass Filter

The Butterworth[1] response is the name given to the following magnitude function:

$$\left|\mathbf{T}_n(j\omega)\right| = \frac{1}{\sqrt{1+(\omega/\omega_0)^{2n}}} \tag{21.14}$$

The Butterworth magnitude response defined

Normalising such that $\omega_0 = 1$ gives:

$$\left|\mathbf{T}_n(j\omega)\right| = \frac{1}{\sqrt{1+\omega^{2n}}} \tag{21.15}$$

The normalised Butterworth magnitude response

From this equation we can observe some interesting properties of the Butterworth response:

1. $\left|\mathbf{T}_n(j0)\right| = 1$ for all n.

1. $\left|\mathbf{T}_n(j1)\right| = 1/\sqrt{2} \approx 0.707$ for all n.

2. For large ω, $\left|\mathbf{T}_n(j\omega)\right|$ exhibits n-pole rolloff.

3. $\left|\mathbf{T}_n(j\omega)\right|$ has all derivatives but one equal to zero near $\omega = 0$. The response is known as *maximally flat*.

Properties of the normalised Butterworth magnitude response

[1] Stephen Butterworth was a British engineer who described this type of response in connection with electronic amplifiers in his paper "On the Theory of Filter Amplifiers", *Wireless Eng.*, vol. 7, 1930, pp. 536-541.

These properties are shown below:

Butterworth lowpass magnitude responses

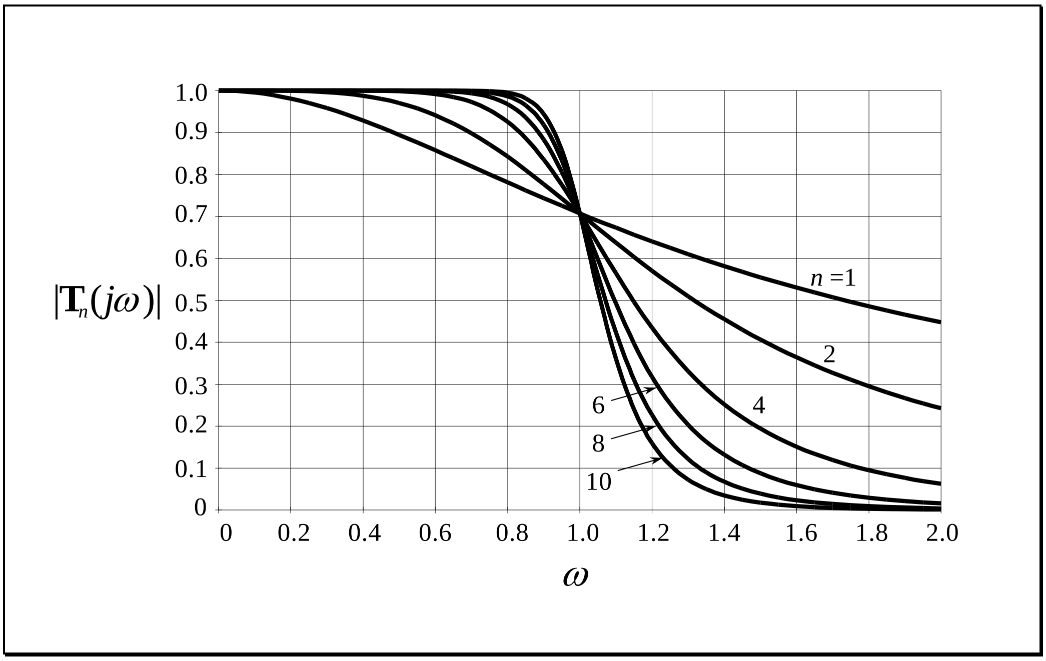


Figure 21.12

The Butterworth response is well-known and has been studied extensively. Since all normalised Butterworth responses have $\omega_0 = 1$, we can draw up a table of Q_0 values for each value of *n*:

Table 21.2 – Q_0 for Butterworth Filters

Table of Q_0 for Butterworth filters

n								
2	3	4	5	6	7	8	9	10
0.707	1.000	0.541 1.307	0.618 1.618	0.518 0.707 1.932	0.555 0.802 2.247	0.510 0.601 0.900 2.563	0.532 0.653 1.000 2.879	0.506 0.561 0.707 1.101 3.196

Note: For *n* odd, there is an additional first-order factor with a cutoff frequency at $\omega_0 = 1$.

EXAMPLE 21.2 Butterworth Lowpass Filter Design

We are required to implement a 5th-order Butterworth lowpass filter with an $\omega_0 = 1248\ \text{rads}^{-1}$.

From Table 20.2, for $n = 5$ the required values of Q_0 are 0.618 and 1.618. There is also a first-order term with a cutoff frequency at ω_0.

The realisation to meet these specifications can be shown in block diagram form:

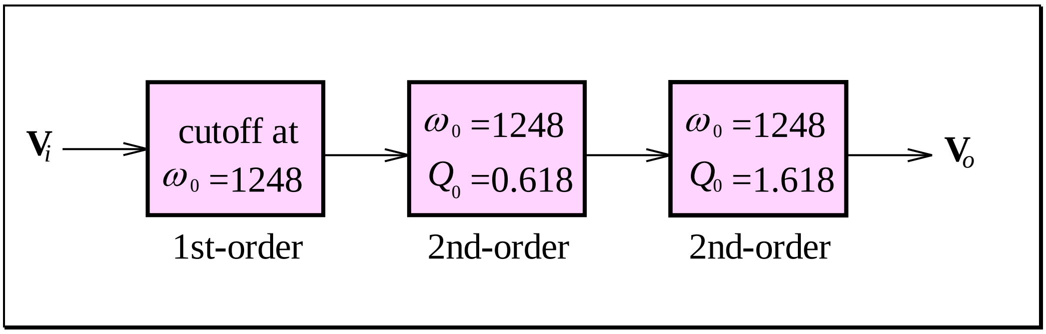


Figure 21.13

Each second-order block could be realised using the biquad circuit. The first-order circuit can be realised with a simple buffered *RC* circuit. If we use the circuit of Figure 21.4 for the biquad, then we must frequency scale using $k_f = 1248$. We then perform magnitude scaling to achieve practical element values.

The Butterworth response is not the only response that can be used to approximate the ideal brickwall filter. Other common responses are the Chebyshev response, inverse Chebyshev response and Cauer (or elliptic) response. Each has their advantages and disadvantages.

21.6 Summary

- The universal biquad circuit is a ready-made module which provides a variety of second-order frequency responses – lowpass, highpass, bandpass, notch or allpass.

- The design of a biquad circuit is specified by just three parameters – the resonance frequency ω_0, the quality factor at resonance Q_0, and the passband gain *H*.

- Butterworth filters are easily designed using a table of Q_0 values and can be implemented as a cascade of first-order and second-order circuits.

21.7 References

Huelsman, L. P.: *Active and Passive Analog Filter Design*, McGraw-Hill, Singapore, 1993.

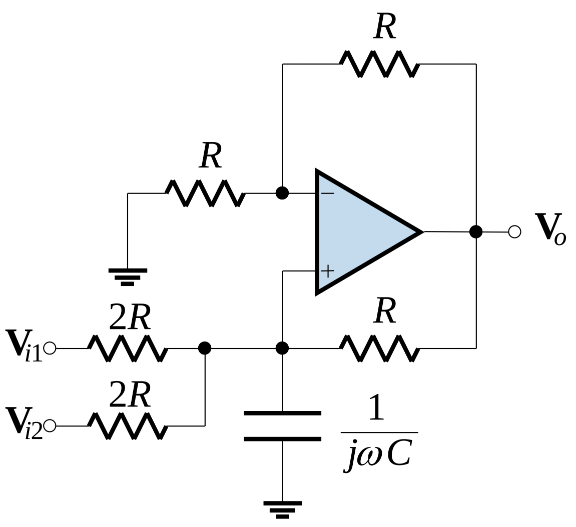
Sedra, A. & Smith, K.: *Microelectronic Circuits*, Saunders College Publishing, Sydney, 1991.

Van Valkenburg, M. E.: *Analog Filter Design*, Holt-Saunders, Tokyo, 1982.

Exercises

1.

Consider the circuit shown below:



(a) Show that:

$$\mathbf{V}_o = \frac{1}{j\omega CR}\mathbf{V}_{i1} + \frac{1}{j\omega CR}\mathbf{V}_{i2}$$

(b) Show that the use of this circuit in the lowpass biquad circuit permits us to reduce the number of op-amps by one.

2.

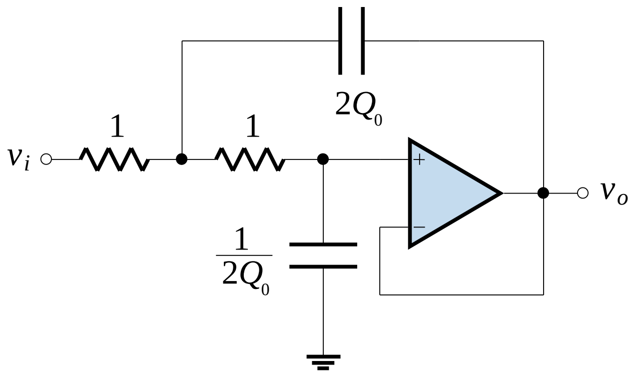
(a) Design a lowpass second-order filter with $\omega_0 = 10\ 000$, $Q_0 = 5$ and $H = 2$ using 1 nF capacitors.

(b) Design a highpass second-order filter with $\omega_0 = 5\ 000$, $Q_0 = 8$ and $H = 1$ using 10 nF capacitors.

21.18

3.

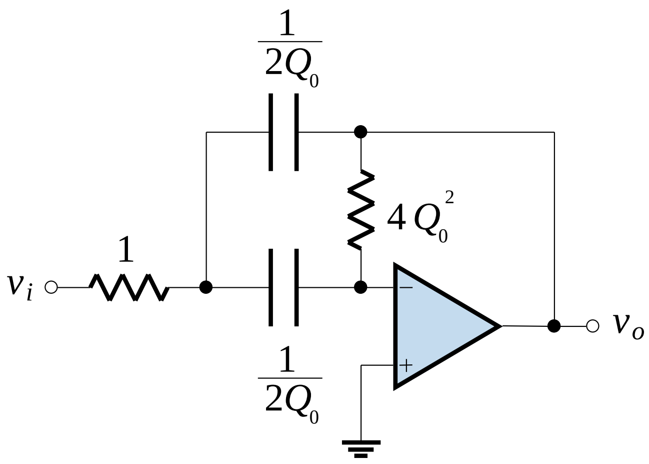
Derive the frequency response of the so-called Sallen-Key circuit:



This can be used instead of the biquad as a single op-amp lowpass second-order circuit.

4.

Derive the frequency response of the so-called Friend circuit:



This can be used instead of the biquad as a single op-amp bandpass second-order circuit.

22 Complex Frequency

Contents

Introduction

We have analysed circuits with DC sources and sinusoidal sources – DC can be considered a special case of a sinusoidal source (one that has zero frequency).

We have also seen that sinusoids are made up of exponentials with imaginary exponents, and that the exponential function plays a special and key role in the determination of the natural response of circuits.

It is only natural, then, to try and extend our study of circuits to a more general class of exponential functions – specifically to complex exponential functions, which then include DC, exponential, and sinusoidal functions as special cases.

The introduction of a complex exponential forcing function leads to a natural definition of *complex frequency*. Analysis of circuits using the concept of complex frequency leads to special insight into circuit behaviour. We will see that complex frequency unifies concepts and results in many advantages – from a deeper insight into understanding circuit behaviour, to analytical techniques associated with establishing the natural, forced and complete response of a circuit, and eventually to “intuition” and circuit design.

22.1 Complex Frequency

Consider an *exponentially damped sinusoidal function*, such as the voltage:

$$v(t) = V_m e^{\sigma t} \cos(\omega t + \theta) \tag{22.1}$$

(This is termed a *damped* function because σ is generally negative).

There are several special cases of this function.

When both σ and ω are zero, we have DC:

$$v(t) = V_m \cos\theta = V_0 \tag{22.2}$$

When only σ is zero, we have a sinusoid:

$$v(t) = V_m \cos(\omega t + \theta) \tag{22.3}$$

If only ω is zero, we have an exponential voltage:

$$v(t) = V_m e^{\sigma t} \cos\theta = V_0 e^{\sigma t} \tag{22.4}$$

The form of this exponential voltage reminds us of the complex exponential representation of a sinusoid:

$$\mathbf{V}_0 e^{j\omega t} \tag{22.5}$$

The only difference is that the exponent is real in one case, and imaginary in the other. Since σt must be dimensionless, we define σ to be a "frequency".

Perhaps we should generalise our complex exponential representation of a sinusoid by including this new frequency, σ? Thus, we define *complex frequency*:

$$\boxed{\mathbf{s} = \sigma + j\omega} \tag{22.6}$$

Now, suppose we declare a function of the form:

$$f(t) = \mathbf{K}e^{\mathbf{s}t} \tag{22.7}$$

where $\mathbf{K}$ and $\mathbf{s}$ are complex constants (independent of time). We say that the function is characterised by the complex frequency $\mathbf{s}$. For example, a constant voltage V_0 may be declared in the required format as:

$$v(t) = V_0 e^{\mathbf{0}t} \tag{22.8}$$

The complex frequency of a DC voltage or current is thus $\mathbf{s} = \mathbf{0}$.

The exponential function:

$$v(t) = V_0 e^{\sigma t} \tag{22.9}$$

is already in the required format, and we identify $\mathbf{s} = \sigma + j0$.

A sinusoidal function *cannot* be expressed in the required format. Given:

$$v(t) = V_m \cos(\omega t + \theta) \tag{22.10}$$

we use Euler's identity to express it as:

$$\begin{aligned} v(t) &= \tfrac{1}{2}V_m\left(e^{j(\omega t+\theta)} + e^{-j(\omega t+\theta)}\right) \\ &= \tfrac{1}{2}V_m e^{j\theta}e^{j\omega t} + \tfrac{1}{2}V_m e^{-j\theta}e^{-j\omega t} \\ &= \mathbf{K}_1 e^{\mathbf{s}_1 t} + \mathbf{K}_2 e^{\mathbf{s}_2 t} \end{aligned} \tag{22.11}$$

Therefore, we have the sum of *two* complex exponentials – a pair of conjugate, counter-rotating phasors. Therefore, *two* complex frequencies are present, one for each term. The complex frequency of the first term is $\mathbf{s} = \mathbf{s}_1 = j\omega$ and that of the second term is $\mathbf{s} = \mathbf{s}_2 = -j\omega$. Thus, $\mathbf{s}_1 = \mathbf{s}_2^*$ and the two values of $\mathbf{K}$ are also conjugate, $\mathbf{K}_1 = \tfrac{1}{2}V_m e^{j\theta}$ and $\mathbf{K}_2 = \mathbf{K}_1^* = \tfrac{1}{2}V_m e^{-j\theta}$.

Lastly, let us determine the complex frequency associated with the exponentially damped sinusoidal function. We again use Euler's identity:

$$\begin{aligned} v(t) &= V_m e^{\sigma t}\cos(\omega t+\theta) \\ &= \tfrac{1}{2}V_m e^{\sigma t}\left(e^{j(\omega t+\theta)}+e^{-j(\omega t+\theta)}\right) \\ &= \tfrac{1}{2}V_m e^{j\theta}e^{(\sigma+j\omega)t}+\tfrac{1}{2}V_m e^{-j\theta}e^{(\sigma-j\omega)t} \end{aligned} \tag{22.12}$$

We therefore find once again that a conjugate complex pair of frequencies is required to describe the exponentially damped sinusoid, $\mathbf{s}_1=\sigma+j\omega$ and $\mathbf{s}_2=\mathbf{s}_1^*=\sigma-j\omega$. In general, neither σ nor ω is zero, and we see that the exponentially varying sinusoidal waveform is the general case – the constant, exponential and sinusoidal waveforms are special cases.

As numerical illustrations, we should now recognize by inspection the complex frequencies associated with these voltages:

$$\begin{aligned} &v(t)=100 && \mathbf{s}=\mathbf{0} \\ &v(t)=5e^{-2t} && \mathbf{s}=-2+j0 \\ &v(t)=2\sin(500t) && \begin{cases}\mathbf{s}_1=j500 \\ \mathbf{s}_2=\mathbf{s}_1^*=-j500\end{cases} \\ &v(t)=4e^{-3t}\sin(6t+10°) && \begin{cases}\mathbf{s}_1=-3+j6 \\ \mathbf{s}_2=\mathbf{s}_1^*=-3-j6\end{cases} \end{aligned} \tag{22.13}$$

The reverse should also be true – given a complex frequency or a pair of conjugate complex frequencies, we should be able to identify the mathematical form of the function with which they are associated. Thus, the most special case $\mathbf{s} = \mathbf{0}$ defines a constant:

$$\mathbf{s} = \mathbf{0} \quad \Rightarrow \quad f(t) = \mathbf{K}e^{\mathbf{0}t} = \text{constant} \tag{22.14}$$

where $\mathbf{K}$ must be real if the function is to be real.

A negative real value, such as $\mathbf{s} = -5 + j0$, defines a decaying exponential:

$$\mathbf{s} = -5 + j0 \quad \Rightarrow \quad f(t) = \mathbf{K}e^{-5t} \tag{22.15}$$

where again $\mathbf{K}$ must be real if the function is to be real.

A purely imaginary value of $\mathbf{s}$, such as $\mathbf{s} = j10$, can never be associated with a real quantity. In order to construct a real function, it is necessary to consider conjugate values of $\mathbf{s}$, such as $\mathbf{s}_{1,2} = \pm j10$, with which must be associated conjugate values of $\mathbf{K}$. Loosely speaking, however, we may identify either of the complex frequencies $\mathbf{s} = j10$ or $\mathbf{s} = -j10$ with a sinusoidal voltage at the radian frequency 10 rad/s. The presence of the conjugate complex frequency is understood. Thus:

$$\left.\begin{aligned} \mathbf{s}_1 &= j10 \\ \mathbf{s}_2 &= -j10 \end{aligned}\right\} \quad \Rightarrow \quad f(t) = A\cos(10t + \theta) \tag{22.16}$$

In a similar manner, a general value for $\mathbf{s}$, such as $\mathbf{s} = -3 + j5$, can be associated with a real quantity only if it is accompanied by its conjugate:

$$\left.\begin{aligned} \mathbf{s}_1 &= -3 + j5 \\ \mathbf{s}_2 &= -3 - j5 \end{aligned}\right\} \quad \Rightarrow \quad f(t) = Ae^{-3t}\cos(5t + \theta) \tag{22.17}$$

In general, the complex frequency $\mathbf{s}$ describes an exponentially varying sinusoid. The real part of $\mathbf{s}$ is associated with the exponential variation – if it is negative, the function decays to zero as t increases, if positive, the function increases, and if it is zero, the sinusoidal amplitude is constant. The imaginary part of $\mathbf{s}$ describes the sinusoidal variation – it is specifically the radian frequency.

22.2 The Damped Sinusoidal Forcing Function

The general exponentially varying sinusoidal voltage:

$$v(t) = V_m e^{\sigma t} \cos(\omega t + \theta) \tag{22.18}$$

is expressible in terms of the complex frequency $\mathbf{s}$ by making use of Euler's identity:

$$v(t) = \text{Re}\left(V_m e^{\sigma t} e^{j(\omega t + \theta)}\right) \tag{22.19}$$

Collecting factors:

$$v(t) = \text{Re}\left(V_m e^{j\theta} e^{(\sigma + j\omega)t}\right) \tag{22.20}$$

we now substitute $\mathbf{s} = \sigma + j\omega$, and obtain:

$$v(t) = \text{Re}\left(V_m e^{j\theta} e^{\mathbf{s}t}\right) \tag{22.21}$$

We note the resemblance to the corresponding representation of the *undamped* sinusoid:

$$v(t) = \text{Re}\left(V_m e^{j\theta} e^{j\omega t}\right) \tag{22.22}$$

The only difference is that we now have $\mathbf{s}$ where we previously had $j\omega$. Thus, our approach will be to develop a *frequency-domain* description of the exponentially varying sinusoid in exactly the same way as we did for the undamped sinusoid: omit the Re notation and suppress $e^{\mathbf{s}t}$.

Thus, if we apply the forcing function:

$$\begin{aligned} & v(t) = V_m e^{\sigma t}\cos(\omega t + \theta) \\ \text{or} \quad & v(t) = \operatorname{Re}\left(V_m e^{j\theta} e^{\mathbf{s}t}\right) \end{aligned} \tag{22.23}$$

then we expect the forced response, say a current in some branch of the circuit, to be:

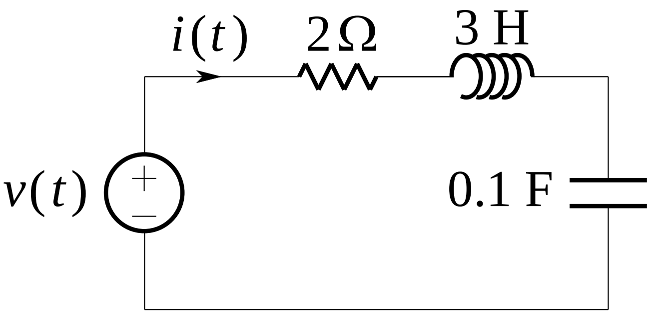
$$\begin{aligned} & i(t) = I_m e^{\sigma t}\cos(\omega t + \phi) \\ \text{or} \quad & i(t) = \operatorname{Re}\left(I_m e^{j\phi} e^{\mathbf{s}t}\right) \end{aligned} \tag{22.24}$$

where the complex frequency of the source and the response must be identical.

We recall that the application of a complex source results in a complex response, and that the real part of the source resulted in the real part of the response. Thus, given the real forcing function $v(t) = \operatorname{Re}\left(V_m e^{j\theta} e^{\mathbf{s}t}\right)$ we apply the complex forcing function $V_m e^{j\theta} e^{\mathbf{s}t}$. The resultant forced response, $I_m e^{j\phi} e^{\mathbf{s}t}$ is complex and it must have as its real part the desired time-domain forced response $i(t) = \operatorname{Re}\left(I_m e^{j\phi} e^{\mathbf{s}t}\right)$.

EXAMPLE 22.1 The Damped Sinusoidal Forcing Function

Consider the circuit:



We shall apply the forcing function:

$$v(t) = 60e^{-2t}\cos(4t + 10°)$$

and we desire the forced response:

$$i(t) = I_m e^{-2t}\cos(4t + \phi)$$

We first express the forcing function as the real part of a complex function:

$$\begin{aligned} v(t) &= 60e^{-2t}\cos(4t + 10°) \\ &= \mathrm{Re}\left(60e^{-2t}e^{j(4t+10°)}\right) \\ &= \mathrm{Re}\left(60e^{j10°}e^{(-2+j4)t}\right) \end{aligned}$$

or:

$$v(t) = \mathrm{Re}\left(\mathbf{V}e^{st}\right)$$

where:

$$\mathbf{V} = 60\angle 10° \quad \text{and} \quad \mathbf{s} = -2 + j4$$

After dropping the Re, we are left with the complex forcing function $\mathbf{V}e^{st}$. In a similar manner, we represent the unknown response by the complex quantity $\mathbf{I}e^{st}$ where $\mathbf{I} = I_m\angle\phi$.

22.10

From KVL we obtain:

$$Ri + L\frac{di}{dt} + \frac{1}{C}\int i\,dt = v(t)$$
$$2i + 3\frac{di}{dt} + 10\int i\,dt = v(t)$$

and we substitute the given complex forcing function and the assumed complex forced response:

$$2\mathbf{I}e^{st} + 3\mathbf{s}\mathbf{I}e^{st} + \frac{10}{\mathbf{s}}\mathbf{I}e^{st} = 60\angle 10°e^{st}$$

The common factor e^{st} is "suppressed" by dividing both sides by e^{st}:

$$2\mathbf{I} + 3\mathbf{s}\mathbf{I} + \frac{10}{\mathbf{s}}\mathbf{I} = 60\angle 10°$$

and:

$$\mathbf{I} = \frac{60\angle 10°}{2 + 3\mathbf{s} + 10/\mathbf{s}}$$

The left side of this equation is a current. On the right, the numerator is a voltage, and so the denominator must be interpreted as an impedance.

We now let $\mathbf{s} = -2 + j4$ and solve for the complex current $\mathbf{I}$:

$$\mathbf{I} = \frac{60\angle 10°}{2 + 3(-2 + j4) + 10/(-2 + j4)}$$

After manipulating the complex numbers, we find:

$$\mathbf{I} = 5.37\angle -106.6°$$

Thus, the forced response is:

$$i(t) = 5.37e^{-2t}\cos(4t - 106.6°)$$

22.3 Generalized Impedance and Admittance

In order to apply Kirchhoff's Laws directly to a circuit with a complex forcing function, we need to determine the relationship between the complex voltage across an element and the complex current through it, i.e. the impedance or admittance of an element.

We will consider the inductor, and then state the relationships for the other elements, since the derivations are similar.

The defining time-domain equation for the inductor is:

$$v(t) = L\frac{di(t)}{dt} \tag{22.25}$$

After applying the complex voltage and current equations, we obtain:

$$\mathbf{V}e^{\mathbf{s}t} = L\frac{d}{dt}\mathbf{I}e^{\mathbf{s}t} \tag{22.26}$$

Taking the indicated derivative:

$$\mathbf{V}e^{\mathbf{s}t} = \mathbf{s}L\mathbf{I}e^{\mathbf{s}t} \tag{22.27}$$

Suppressing $e^{\mathbf{s}t}$, we find the desired voltage-current relationship:

$$\mathbf{V} = \mathbf{s}L\mathbf{I} \tag{22.28}$$

Thus, the impedance of the inductor is:

$$\mathbf{Z}(\mathbf{s}) = \frac{\mathbf{V}}{\mathbf{I}} = \mathbf{s}L \tag{22.29}$$

Generalized impedance of an inductor

and the admittance is:

$$\mathbf{Y}(\mathbf{s}) = \frac{\mathbf{I}}{\mathbf{V}} = \frac{1}{\mathbf{s}L} \tag{22.30}$$

We still call **V** and **I** *phasors*. These complex quantities have a magnitude and angle which, along with a specific complex frequency value, enable us to characterize the exponentially varying sinusoidal waveform completely. The phasor is still a frequency-domain description, but we have now expanded its definition to include complex frequency.

The frequency-domain equivalent of an inductor is shown below:

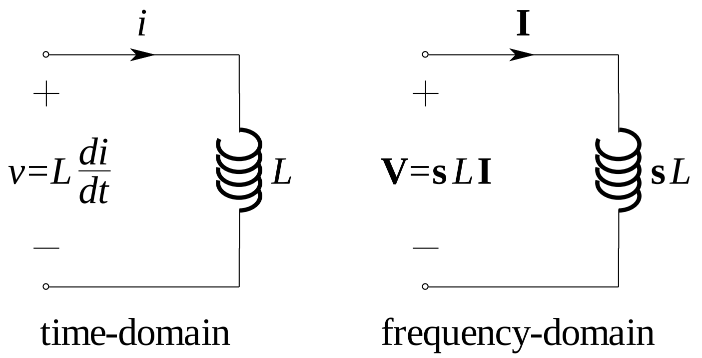


Figure 22.1

In the table below, we show how we can transform a resistor, inductor or capacitor from the time-domain into its frequency-domain impedance:

Generalized impedances of the three passive elements

Time-domain	*Frequency-domain*
R	R
L	$\mathbf{s}L$
C	$\frac{1}{\mathbf{s}C}$

Table 22.1 – Generalized impedances

EXAMPLE 22.2 Generalised Impedance

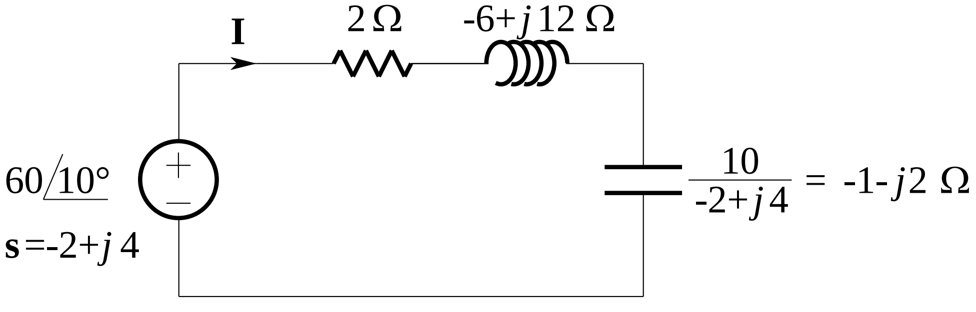
If we now reconsider the series *RLC* example in the frequency domain, the source voltage:

$$v(t) = 60e^{-2t}\cos(4t + 10°)$$

is transformed to the phasor voltage:

$$\mathbf{V} = 60\angle 10°$$

A phasor current $\mathbf{I}$ is assumed, and the impedance of each element at the complex frequency $\mathbf{s} = -2 + j4$ is determined and placed on a frequency-domain circuit diagram:



The phasor current is now easily found by dividing the phasor voltage by the sum of the three impedances:

$$\mathbf{I} = \frac{60\angle 10°}{2 + (-6 + j12) + (-1 - j2)} = \frac{60\angle 10°}{-5 + j10} = 5.37\angle -106.6° \text{ A}$$

The forced response is then:

$$i(t) = 5.37e^{-2t}\cos(4t - 106.6°)$$

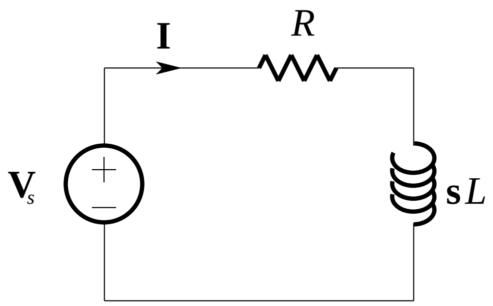
Thus, the previous result is obtained, but much more easily and rapidly.

22.4 Frequency Response as a Function of ω

Suppose we have a circuit which is excited by a single source, $\mathbf{V}_s = V_s\angle 0°$, and that the desired response is a current $\mathbf{I}$. This phasor response is a complex number, and it can be represented by two numbers – a magnitude and a phase. If the source varies as a function of *only* radian frequency ω, then we can plot the magnitude response as a function of ω and the phase response as a function of ω.

EXAMPLE 22.3 Frequency Response as a Function of ω

Suppose we have a series *RL* circuit:



The phasor voltage $\mathbf{V}_s = V_s\angle 0°$ is applied, and the response is the phasor current $\mathbf{I}$. The forced response is:

$$\mathbf{I} = \frac{\mathbf{V}_s}{R + \mathbf{s}L}$$

If $\mathbf{V}_s$ represents a sinusoidal forcing function, then $\mathbf{s} = j\omega$ and we have:

$$\mathbf{I} = \frac{\mathbf{V}_s}{R + j\omega L}$$

The magnitude of the response is:

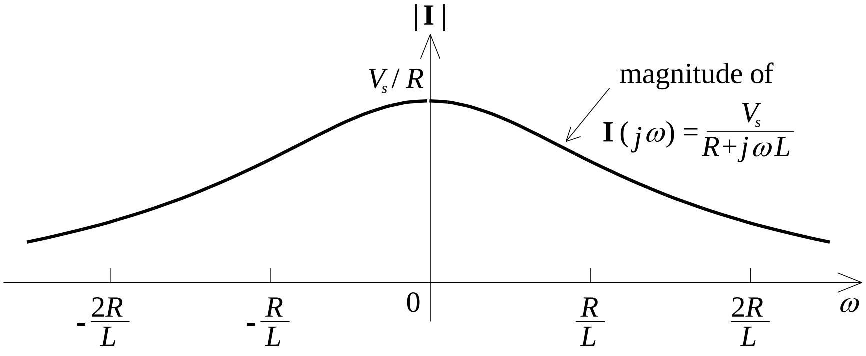
$$|\mathbf{I}| = \frac{V_s}{\sqrt{R^2 + \omega^2 L^2}}$$

and the phase of the response is:

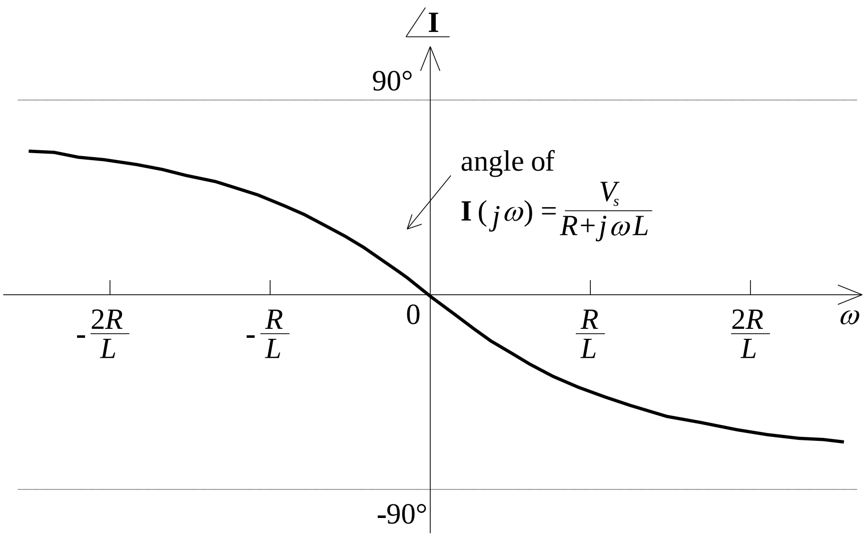
$$\angle \mathbf{I} = -\tan^{-1} \frac{\omega L}{R}$$

These are the analytical expressions for the magnitude and phase angle of the response as a function of ω – we can now present the same information graphically.

For the magnitude curve, we note that we are plotting the absolute value of some quantity and so the entire curve lies *above* the ω axis. The response is V_s/R at zero frequency, and the response approaches zero as frequency approaches infinity. We graph both positive and negative values of frequency:

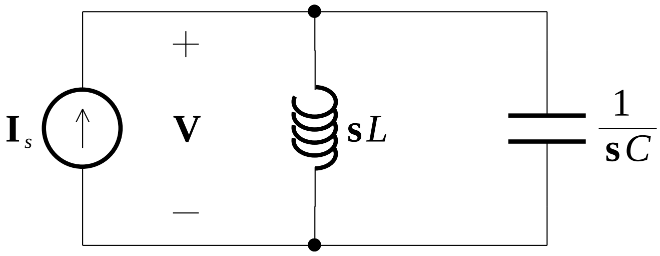


The second part of the response, the phase angle of $\mathbf{I}$ versus ω, is an inverse tangent function:



EXAMPLE 22.4 Frequency Response as a Function of ω

Suppose we have a parallel *LC* circuit:



The voltage response is easily obtained:

$$\mathbf{V} = \mathbf{I}_s \frac{(\mathbf{s}L)(1/\mathbf{s}C)}{\mathbf{s}L + 1/\mathbf{s}C} = \mathbf{I}_s \frac{\mathbf{s}L}{\mathbf{s}^2 LC + 1}$$

For a sinusoidal forcing function, then $\mathbf{s} = j\omega$ and the circuit presents an impedance to the current source of:

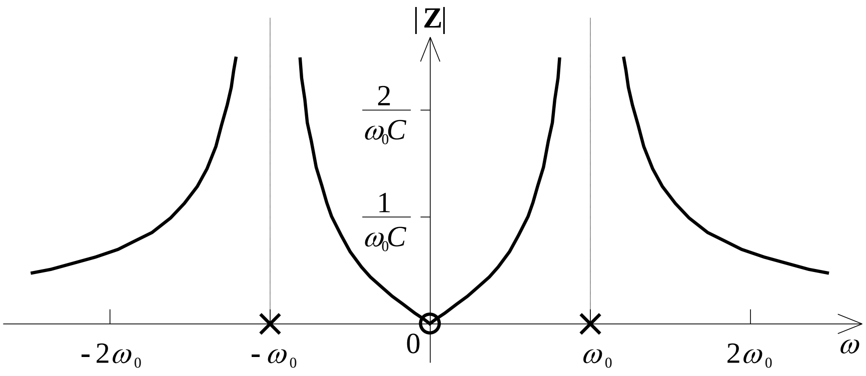
$$\mathbf{Z} = \frac{\mathbf{V}}{\mathbf{I}_s} = \frac{j\omega L}{-\omega^2 LC + 1} = -j\frac{1}{C}\frac{\omega}{\omega^2 - 1/LC}$$

By letting $\omega_0 = 1/\sqrt{LC}$, we can express the magnitude of the impedance in a form which enables those frequencies to be identified at which the response is zero or infinite:

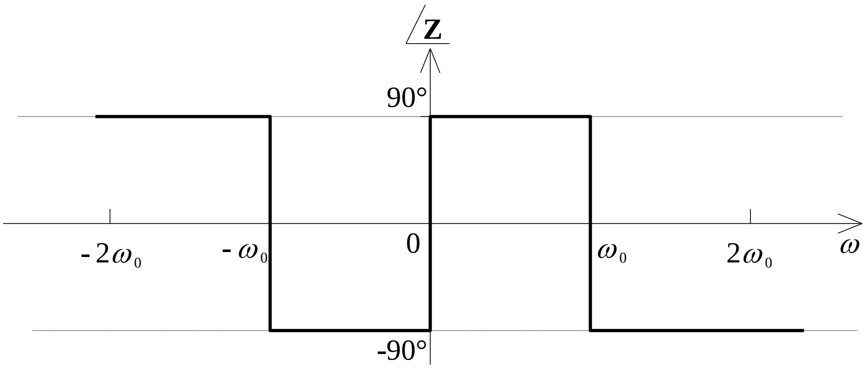
$$|\mathbf{Z}| = \frac{1}{C}\frac{|\omega|}{|(\omega - \omega_0)(\omega + \omega_0)|}$$

Such frequencies are called *critical frequencies*, and their early identification simplifies the construction of the response curve. We note that the response has zero amplitude when $\omega = 0$. When this happens we say that the response has a *zero* at $\omega = 0$, and we also describe the frequency at which it occurs as a zero. A response of infinite amplitude is noted at $\omega = \omega_0$ and $\omega = -\omega_0$. These frequencies are called *poles*, and the response is said to have a pole at each of these frequencies. Finally, we note that the response approaches zero as $\omega \to \infty$, and thus $\omega = \pm\infty$ is also a zero (it is customary to consider plus infinity and minus infinity as being the same point).

The locations of the critical frequencies should be marked on the ω axis, by using small circles for the zeros and crosses for the poles. The drawing of the graph is made easier by adding broken vertical lines as asymptotes at each pole location. The completed graph of magnitude versus ω is shown below:



The corresponding phase response is, from an inspection of the original equation:

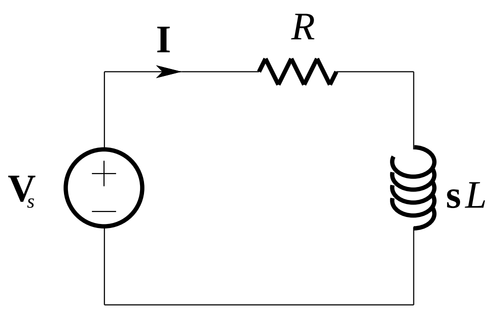


22.5 Frequency Response as a Function of σ

By analogy with the response due to a varying ω, we now consider a response due to a varying σ. Suppose we have a circuit which is excited by a single source, $\mathbf{V}_s = V_s\angle 0°$, and that the desired response is a current $\mathbf{I}$. This phasor response is a complex number, and it can be represented by two numbers – a magnitude and a phase. If the source varies as a function of *only* σ, then we can plot the magnitude response as a function of σ and the phase response as a function of σ.

EXAMPLE 22.5 Frequency Response as a Function of σ

Suppose we have a series *RL* circuit:



The phasor voltage $\mathbf{V}_s = V_s\angle 0°$ is applied, and the response is the phasor current $\mathbf{I}$. The forced response is:

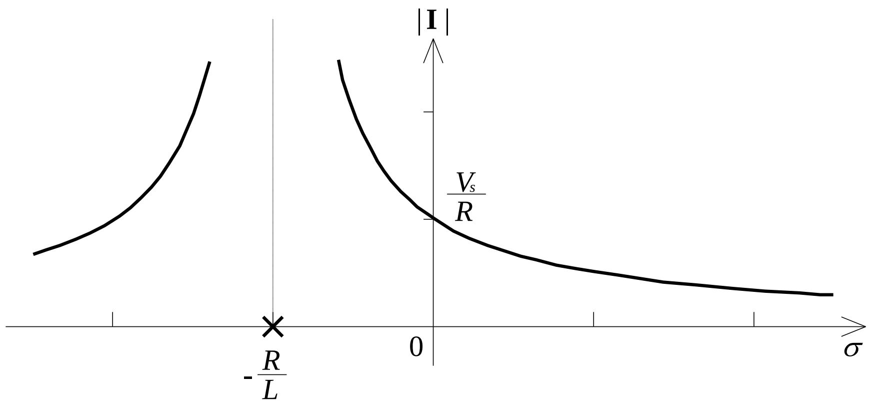
$$\mathbf{I} = \frac{\mathbf{V}_s}{R + \mathbf{s}L}$$

We now set $\omega = 0$, $\mathbf{s} = \sigma + j0$, thus restricting ourselves to time-domain sources of the form $v_s = V_s e^{\sigma t}$. Thus:

$$\mathbf{I} = \frac{\mathbf{V}_s}{R + \sigma L}$$

As σ varies, we can obtain a qualitative description of the circuit's behaviour. When σ is large and negative, the current response is negative and relatively small in amplitude. As σ increases, becoming a smaller negative number, the magnitude of the response increases, until we encounter $\sigma = -R/L$, a pole of the forced response, at which point the response is infinite. As σ continues to increase, the next noteworthy point is $\sigma = 0$, where we have the case of a DC source, $v_s = V_s$, and the forced response is obviously $\mathbf{I} = V_s / R$. Positive values of σ must all provide positive amplitude responses, and the response must decrease as σ get larger. Finally, as $\sigma \to \infty$, we have a zero-amplitude response, and therefore a zero. The only critical frequencies are the pole at $\sigma = -R/L$ and the zero at $\sigma = \pm\infty$.

This information is plotted graphically below:



The phase response, not shown, is either $0°$ or $180°$.

In the example above, a finite critical frequency occurred at $\sigma = -R/L$. That is, when the circuit is excited at this frequency by the voltage:

$$v_s = V_s e^{-Rt/L} \tag{22.31}$$

we get an infinite response. Why? The forcing function has a familiar form – it is the same form as the natural response of the circuit. We know that if we have a source-free circuit, then the response due to an initial current would be:

$$i = I_0 e^{-Rt/L} \tag{22.32}$$

Thus, a zero-amplitude input produces a finite response. Since the circuit is linear, we conclude that a non-zero-amplitude forcing function will produce an infinite forced response.

This result is quite general:

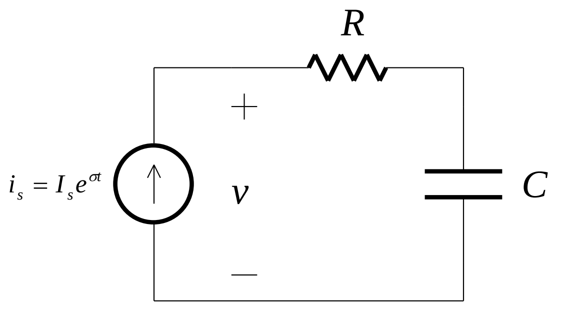
When any circuit is excited at a frequency which is a pole of the response, then an infinite response must result.	(22.33)

We also have:

The frequencies of the poles are directly related to the natural response of the circuit.	(22.34)

EXAMPLE 22.6 Frequency Response as a Function of σ

Suppose we have a series *RC* circuit, excited by an exponential current source $i_s = I_s e^{\sigma t}$:

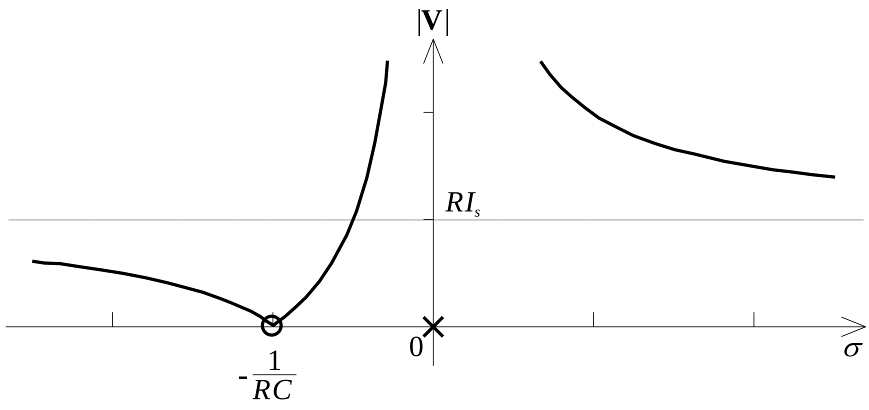


The voltage across the source is:

$$\mathbf{V} = \mathbf{ZI} = \left(R + \frac{1}{\sigma C} \right) I_s = RI_s \frac{\sigma + 1/RC}{\sigma}$$

We can now identify the poles and zeros of the response from the form of the equation – it indicates a pole at $\sigma = 0$ and a zero at $\sigma = -1/RC$.

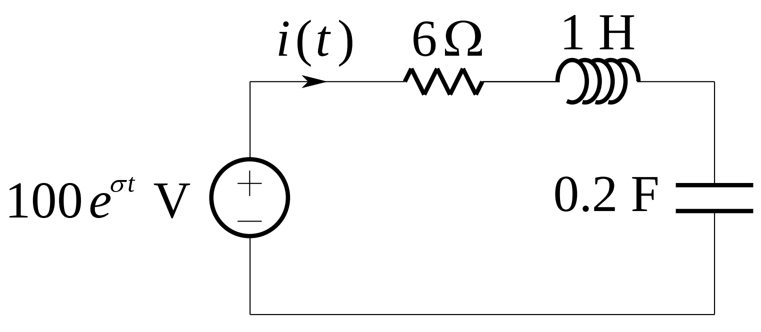
The response magnitude is plotted as a function of σ below:



The reason for the pole at zero frequency may again be explained on physical grounds. If $I_s = 0$, then the current source is an open-circuit and the response is the initial capacitor voltage. A non-zero I_s must therefore produce an infinite forced voltage response. In other words, if a constant current has been applied to the circuit forever, then the capacitor voltage would be infinite.

EXAMPLE 22.7 Frequency Response as a Function of σ

Suppose we have a series *RLC* circuit, excited by an exponential voltage source:

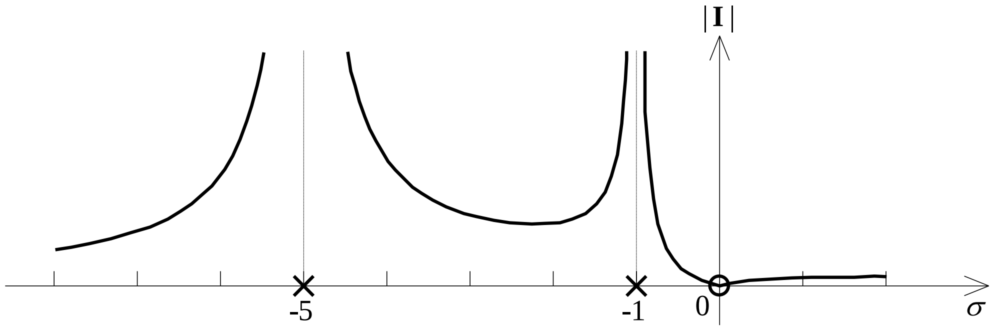


The current is easily found:

$$\mathbf{I} = \frac{100}{6+\sigma+5/\sigma} = 100\frac{\sigma}{(\sigma+1)(\sigma+5)}$$

22.22

The response curve is most easily found by first indicating the locations of all poles and zeros on the σ axis, and placing vertical asymptotes at the poles. We can then seek out relative minima and maxima, and make a sketch of the response:



The two poles may again be used to construct the natural response:

$$i_n(t) = A_1 e^{-t} + A_2 e^{-5t}$$

Compare this with the case of an overdamped series *RLC* circuit:

$$\begin{aligned}\mathbf{s}_{1,2} &= -\frac{R}{2L} \pm \sqrt{\left(\frac{R}{2L}\right)^2 - \frac{1}{LC}} \\ &= -3 \pm \sqrt{3^2 - 5} \\ &= -1 \quad \text{and} \quad -5\end{aligned}$$

Once again, at either of these two frequencies, a zero-amplitude forcing function may be associated with a finite-amplitude response which can be interpreted as the natural response of the circuit.

22.6 The Complex-Frequency Plane

Now that we have considered the forced response of a circuit as ω varies (with $\sigma = 0$) and as σ varies (with $\omega = 0$), we are prepared to develop a more general graphical representation by graphing quantities as functions of $\mathbf{s}$.

After we develop the concept of the complex-frequency plane, or $\mathbf{s}$-plane, we shall see how quickly the behaviour of a circuit can be obtained from a graphical representation of its critical frequencies.

The response of a circuit to a complex frequency $\mathbf{s}$ is a complex function of $\mathbf{s}$. Thus, we need a method of representing both the *magnitude* and *phase* of the response as $\mathbf{s}$ varies. Therefore, we represent the magnitude and phase individually, using a *three*-dimensional model built over the $\mathbf{s}$-plane, where the height at any point is equal to the value of the quantity we are representing – either the magnitude or the phase.

The $\mathbf{s}$-plane consists of a σ axis and a $j\omega$ axis, perpendicular to each other:

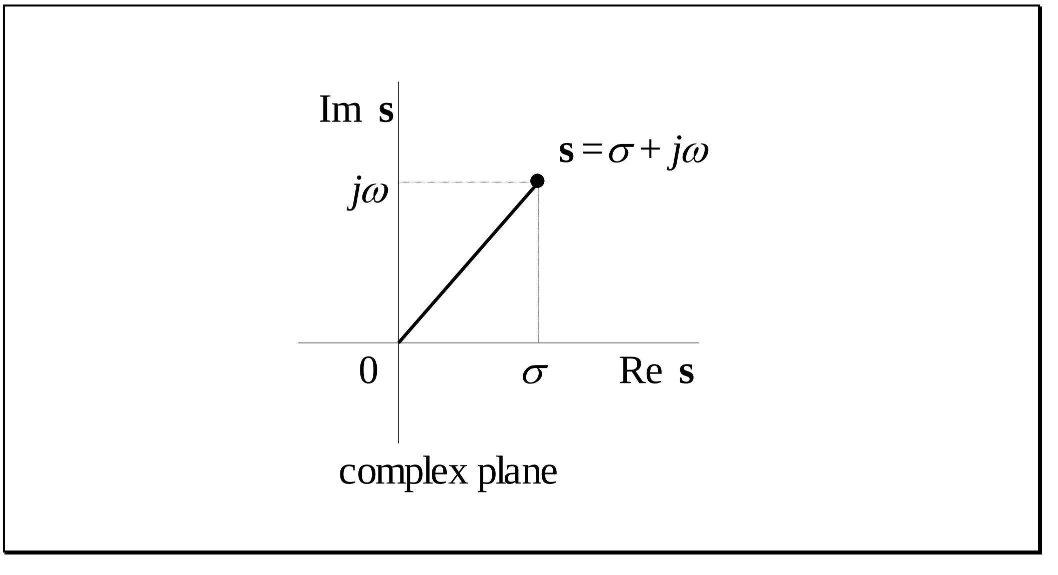


Figure 22.2

We are familiar with the type of time-domain function associated with a particular value of the complex frequency $\mathbf{s}$. It is now possible to associate the functional form of a forcing function or forced response with the various regions in the $\mathbf{s}$-plane.

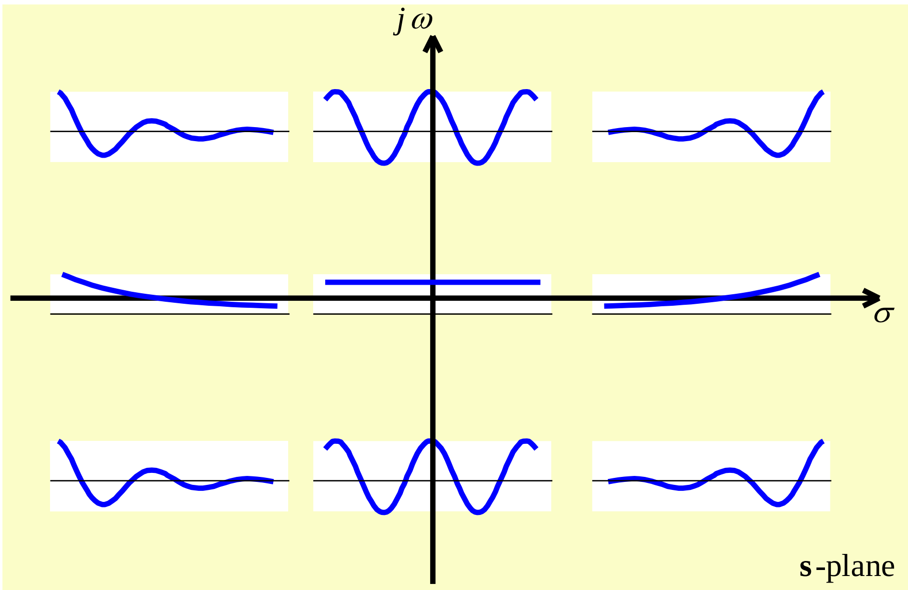


Figure 22.3

The origin, for example, must represent a DC quantity. Points lying on the σ axis must represent exponential functions, decaying for $\sigma < 0$, increasing for $\sigma > 0$. Pure sinusoids are associated with points on the positive or negative $j\omega$ axis. Points in the left half-plane (LHP) describe the frequencies of exponentially decreasing (or damped) sinusoids. The RHP contains points describing exponentially increasing sinusoids.

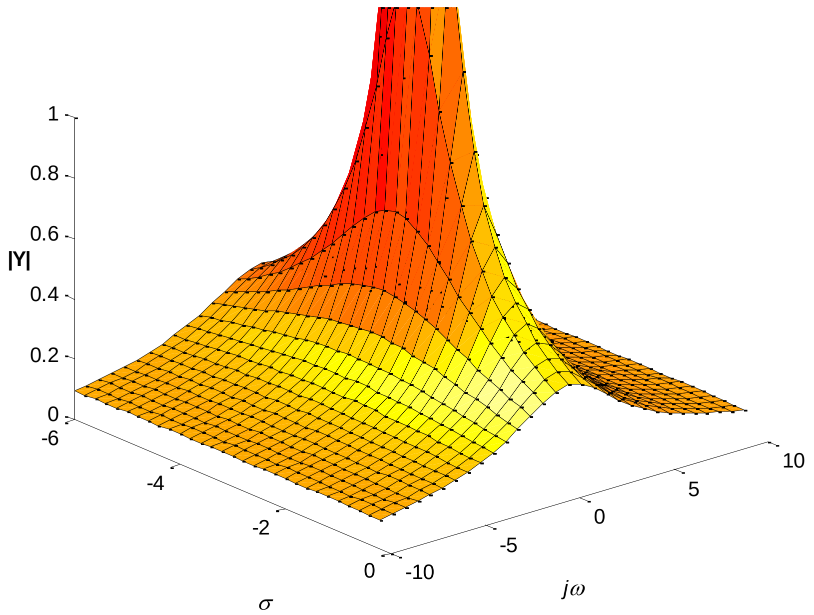
We may represent the magnitude of a function of $\mathbf{s}$ as a *surface* lying above (or just touching) the $\mathbf{s}$-plane.

EXAMPLE 22.8 Admittance Magnitude on the Complex Plane

Suppose we want to graph the magnitude of the admittance of the series combination of a 1 H inductor and a 3Ω resistor:

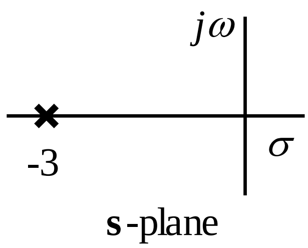
$$\mathbf{Y}(\mathbf{s}) = \frac{1}{\mathbf{s}+3}$$

We identify the pole immediately as $\mathbf{s} = -3 + j0$ – thus our 3D model of the magnitude of $\mathbf{Y}(\mathbf{s})$ must have infinite height over this point. When $\mathbf{s}$ is infinite, the magnitude of $\mathbf{Y}(\mathbf{s})$ must be zero. The model must have zero height at points infinitely far away from the origin. A cut-away view of the magnitude of $\mathbf{Y}(\mathbf{s})$ is shown below:



The graph of the magnitude of $\mathbf{Y}(\mathbf{s})$ over the **s**-plane forms a surface with poles and zeros

We can completely specify $\mathbf{Y}(\mathbf{s})$, apart from a constant gain factor, by drawing a so-called *pole-zero plot*:



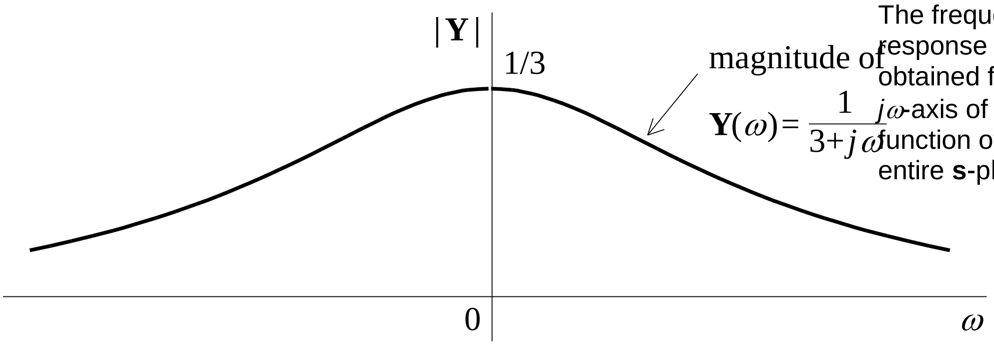
A pole-zero plot is a shorthand way of representing a complex function of **s**

22.26

A pole-zero plot locates all the critical points in the **s**-plane that completely specify the function $\mathbf{Y}(\mathbf{s})$ (to within an arbitrary constant), and it is a useful analytic and design tool.

One cut of the surface has been fortuitously placed along the imaginary axis. If we graph the height of the surface along this cut against ω, we get a picture of the magnitude of the frequency response versus ω:

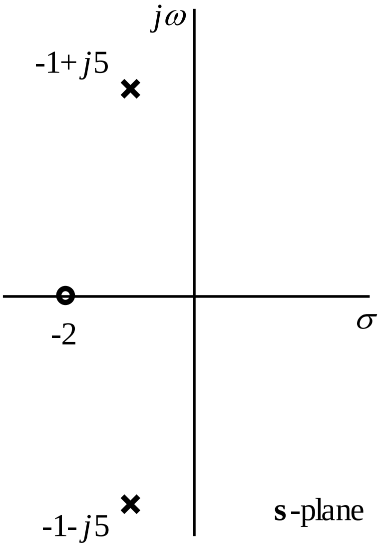
The frequency response is obtained from the $j\omega$-axis of a plot of a function over the entire **s**-plane



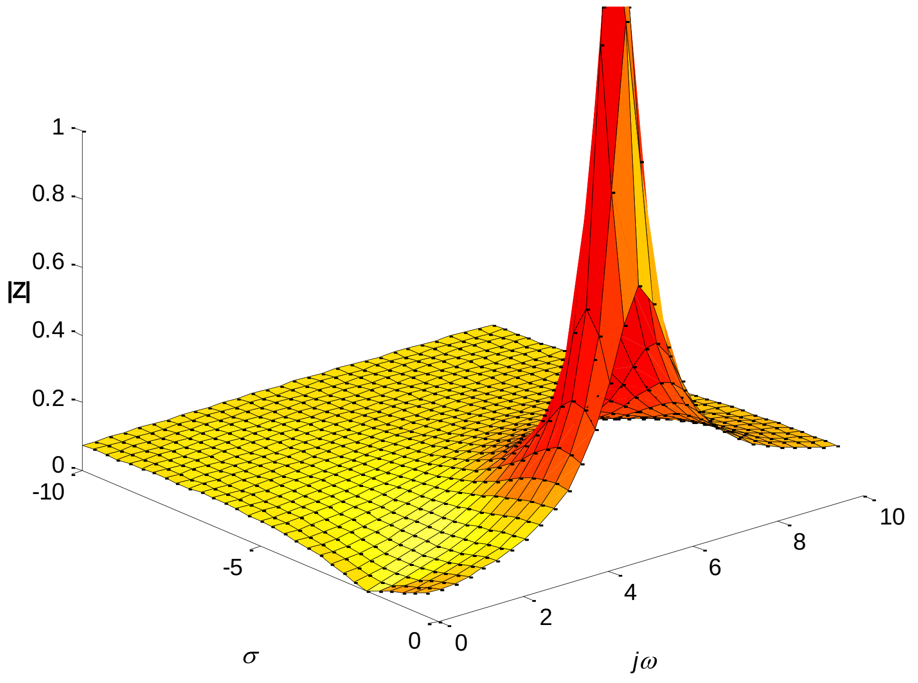
We can obtain a mental image of the surface that represents the magnitude of a function over the **s**-plane quite quickly if we imagine a rubber sheet model. At zeros, we secure the rubber sheet to the plane (tent pegs). At poles, we prop the rubber sheet up with infinitely high and thin poles.

EXAMPLE 22.9 Impedance Magnitude on the Complex Plane

We are given a pole-zero plot of some impedance $\mathbf{Z}(\mathbf{s})$.



If we visualize a rubber-sheet model, tacked down at $\mathbf{s} = -2 + j0$ and propped up at $\mathbf{s} = -1 + j5$ and $\mathbf{s} = -1 - j5$, we should see a terrain with two mountains and one valley. The portion of the model for the upper LHP is shown below:



22.28

We can build up the expression for $\mathbf{Z}(\mathbf{s})$ which leads to this pole-zero configuration. The zero requires a factor of $(\mathbf{s}+2)$ in the numerator, and the two poles require the factors $(\mathbf{s}+1-j5)$ and $(\mathbf{s}+1+j5)$ in the denominator. Except for a multiplying constant *k*, we now know the form of $\mathbf{Z}(\mathbf{s})$:

$$\begin{aligned}\mathbf{Z}(\mathbf{s}) &= k\frac{\mathbf{s}+2}{(\mathbf{s}+1-j5)(\mathbf{s}+1+j5)} \\ &= k\frac{\mathbf{s}+2}{\mathbf{s}^2+2\mathbf{s}+26}\end{aligned}$$

The plots $|\mathbf{Z}(\sigma)|$ versus σ and $|\mathbf{Z}(\omega)|$ versus ω may be obtained exactly from this expression, but the general form of the function is apparent from the pole-zero configuration and the rubber-sheet analogy. Portions of these two curves appear at the sides of the 3D model above.

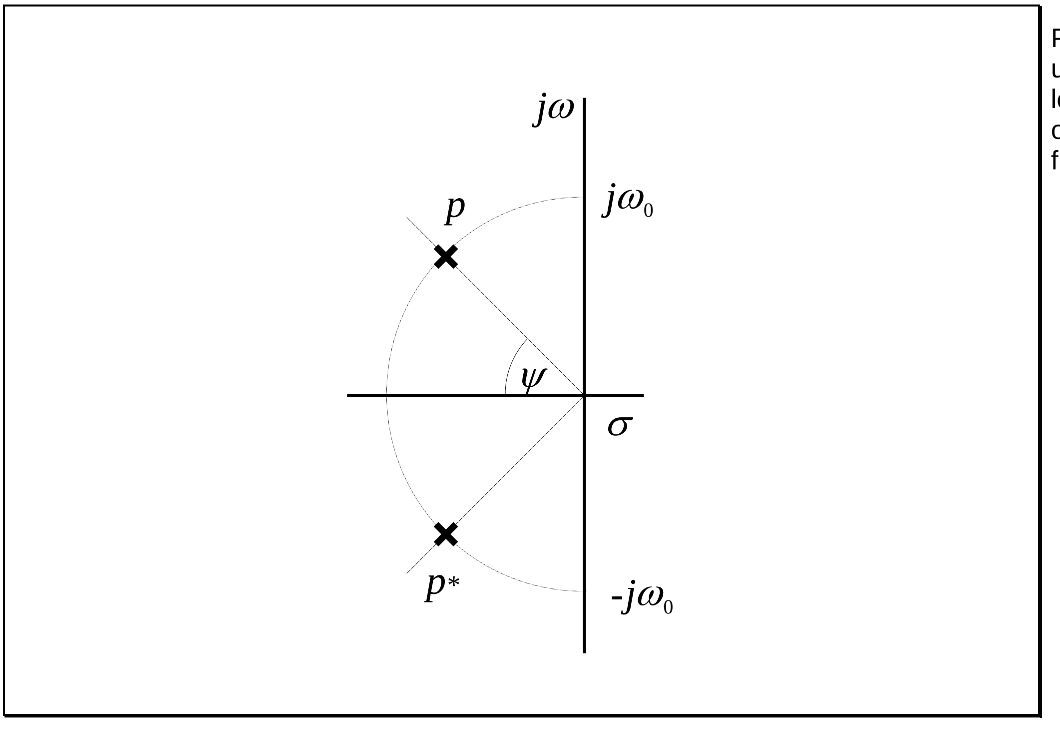
22.7 Visualization of the Frequency Response from a Pole-Zero Plot

The frequency response can be visualised in terms of the pole locations of the response function. For example, for an underdamped second-order lowpass system:

$$\mathbf{V}_o = \frac{\omega_0^2}{\mathbf{s}^2 + (\omega_0/Q_0)\mathbf{s} + \omega_0^2}\mathbf{V}_i \quad (22.35)$$

Standard form for a lowpass second-order response function

the poles are located on a circle of radius ω_0 and at an angle with respect to the negative real axis of $\psi = \cos^{-1}(1/2Q_0)$. These complex conjugate pole locations are shown below:



Pole locations for an underdamped lowpass second-order response function

Figure 22.4

In terms of the poles shown in Figure 22.4, the response function is:

$$\mathbf{V}_o = \frac{\omega_0^2}{(s-p)(s-p^*)}\mathbf{V}_i \quad (22.36)$$

Underdamped lowpass second-order response function using pole factors

When $\mathbf{s} = j\omega$ the two pole factors in this equation become:

Polar representation of the pole factors

$$j\omega - p = m_1 \angle \phi_1 \quad \text{and} \quad j\omega - p^* = m_2 \angle \phi_2 \tag{22.37}$$

In terms of these quantities, the magnitude response is:

Magnitude function written using the polar representation of the pole factors

$$\left|\mathbf{T}(j\omega)\right| = \left|\frac{\mathbf{V}_o}{\mathbf{V}_i}\right| = \frac{\omega_0^2}{m_1 m_2} \tag{22.38}$$

and the phase response is:

Phase function written using the polar representation of the pole factors

$$\angle \mathbf{T}(j\omega) = -(\phi_1 + \phi_2) \tag{22.39}$$

Vectors representing Eq. (22.37) are shown below:

Determining the magnitude and phase response from the **s**-plane

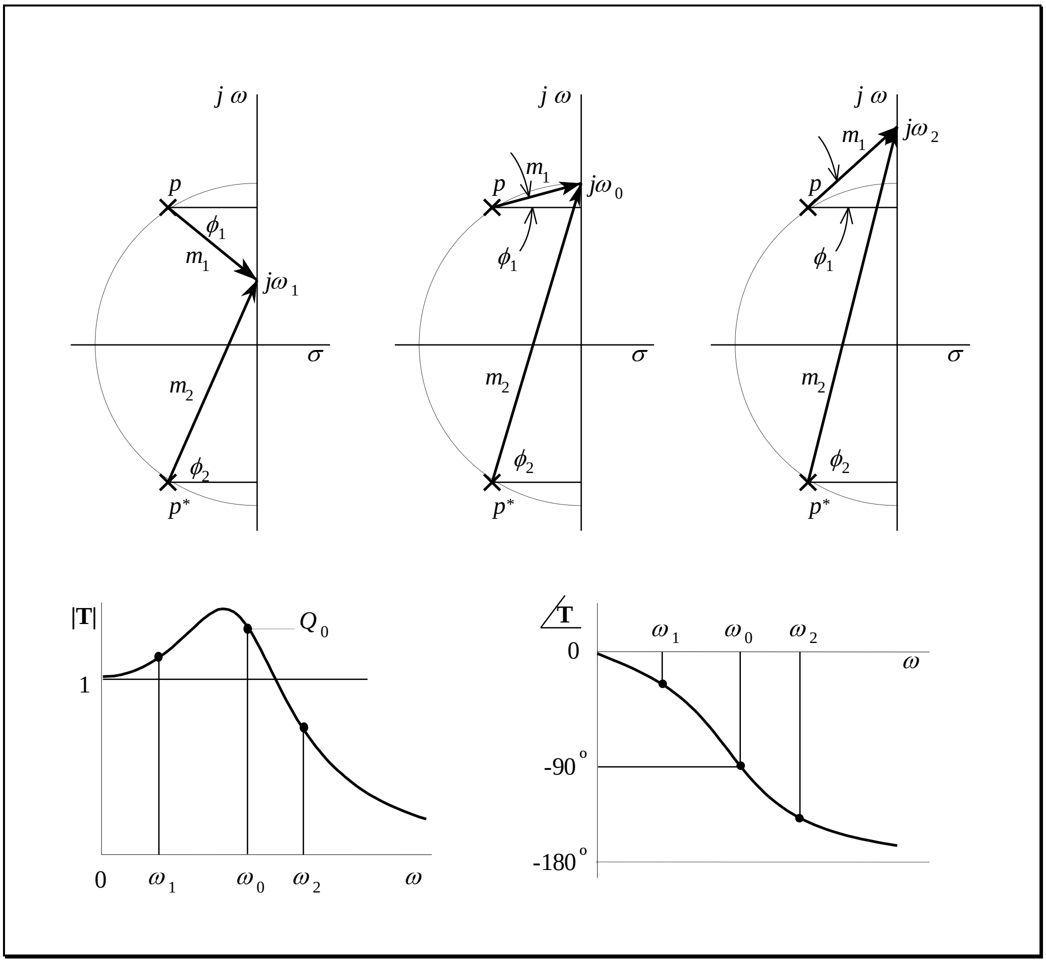


Figure 22.5

Figure 22.5 shows three different frequencies – one below ω_0, one at ω_0, and one above ω_0. From this construction we can see that the short length of m_1 near the frequency ω_0 is the reason why the magnitude function reaches a peak near ω_0. These plots are useful in visualising the frequency response of the circuit.

22.8 Summary

- Complex frequency is defined as $\mathbf{s} = \sigma + j\omega$.

- A whole class of forcing functions and responses (exponentially damped sinusoids) can be expressed using $f(t) = \mathbf{K}e^{\mathbf{s}t}$, where $\mathbf{K}$ is a complex constant. If $\mathbf{s}$ is complex, then a real function must be composed of a conjugate pair: $f(t) = \mathbf{K}e^{\mathbf{s}t} + \mathbf{K}^* e^{\mathbf{s}^* t}$.

- A complex forcing function for a voltage can be expressed as $\mathbf{V}e^{\mathbf{s}t}$, where $\mathbf{V} = V_m \angle \theta$ is a phasor with the complex frequency $\mathbf{s}$.

- The generalised impedances of the three passive elements are:

$$\mathbf{Z}_R = R, \qquad \mathbf{Z}_L = \mathbf{s}L, \qquad \mathbf{Z}_C = \frac{1}{\mathbf{s}C}$$

- Critical frequencies of a forced response occur at zeros and poles. A zero is a frequency at which the response is zero. A pole is a frequency at which the response is infinite.

- The poles of a circuit determine the form of its natural response.

- The frequency response of a circuit can be evaluated at each point in the **s**-plane. The response along the $j\omega$ axis represents the sinusoidal steady-state frequency response.

- A pole-zero plot encapsulates all of the information about a circuit's response, apart from a multiplicative constant. We can also use a pole-zero plot to visualize the frequency response.

22.9 References

Hayt, W. & Kemmerly, J.: *Engineering Circuit Analysis*, 3rd *Ed.*, McGraw-Hill, 1984.

Exercises

1.

Find the complex frequencies associated with the natural response of a source-free series *RLC* circuit in which $L = 10\ \text{H}$, $C = 250\ \mu\text{F}$, and $R =$:

(a) $500\ \Omega$ (b) $320\ \Omega$

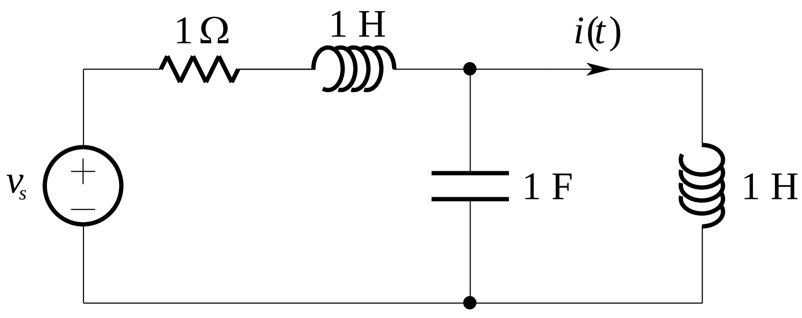
2.

(a) If $i_1(t) = 2e^{-100t}\cos(1000t - 35°)$ mA and $i_2(t) = 4e^{-100t}\sin(1000t + 20°)$ mA, find the phasor representing the sum of these two currents.

(b) Find the phasor corresponding to $i(t) = 2e^{-5t}(3\cos 50t + 2\sin 50t)$ mA.

3.

Consider the circuit shown below:



Find $i(t)$ at $t = 0.5\ \text{s}$ if:

(a) $v_s = 10$ V

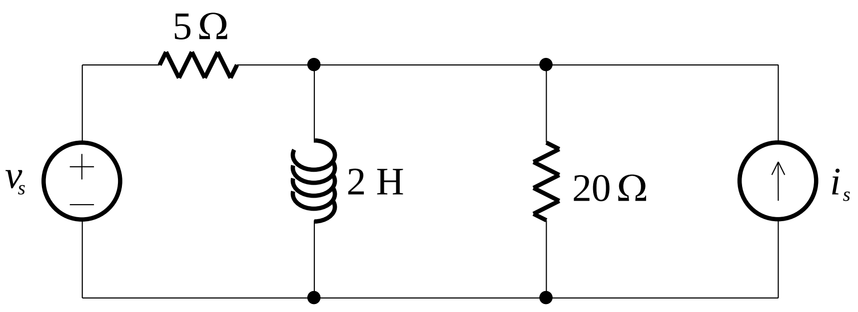
(b) $v_s = 10\cos(2t)$ V

(c) $v_s = 10e^{-3t}$ V

(d) $v_s = 10e^{-3t}\cos(2t)$ V

4.

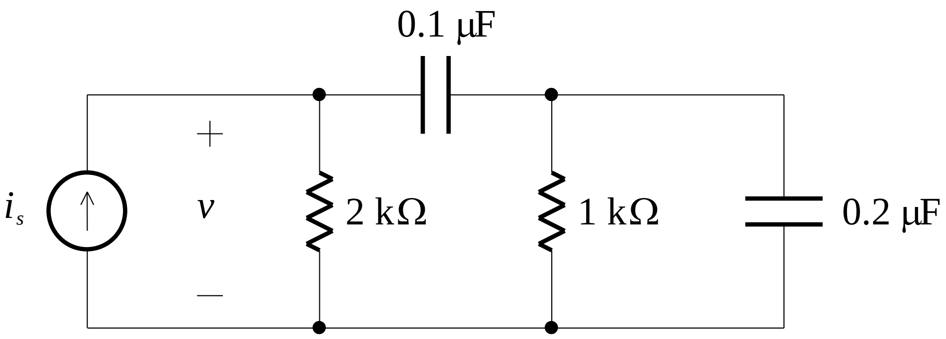
Consider the circuit shown below:



If $v_s = 20e^{-4t}\cos 3t$ V, while $i_s = 3e^{-3t}\cos 4t$ A, find the energy stored in the inductor at $t = 0$.

5.

Consider the circuit shown below:



Find all the critical frequencies of the ratio $\mathbf{V}/\mathbf{I}_s$, and graph the magnitude of the ratio as a function of σ (use a spreadsheet).

6.

One of the critical frequencies of the impedance of a series *RLC* circuit occurs at $\mathbf{s} = -4 + j10\ \text{s}^{-1}$. If the minimum impedance magnitude is 100 Ω, determine:

(a) R (b) L (c) C

7.

The pole-zero plot of an input impedance displays a zero at $\mathbf{s} = -50$ and poles at $\mathbf{s} = -20 \pm j30$.

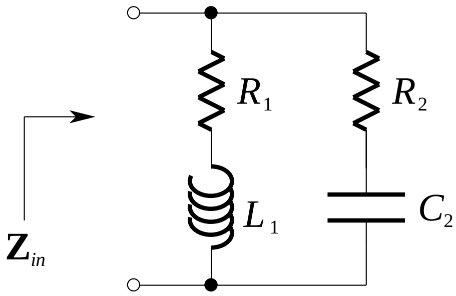
(a) Determine the magnitude and angle of each vector from a critical frequency to the point $\mathbf{s} = j20$.

(b) Calculate the (complex) ratio of $\mathbf{Z}_{\mathrm{in}}(j20)$ to $\mathbf{Z}_{\mathrm{in}}(j40)$.

8.

In 1945, Dr Hendrik W. Bode published *Network Analysis and Feedback Amplifier Design*, codifying in one classic book the filter and feedback-amplifier theory upon which much of the electronics industry still relies.

One of the many results his seminal work revealed was the constant-resistance network:



The figure shows only one of many forms of this circuit.

Show that, as long as you let $R_1 = R_2$ and scale the components such that the time constant L_1/R_1 equals the time constant R_2C_2, then the impedance of the whole circuit remains constant at *all* frequencies, and is equal to R_1.

The circuit occasionally sees application in digital systems as a terminating network, where C_2 represents the unavoidable capacitance of a logic gate input.

23 Specialty Amplifiers

Contents

23.2

Introduction

There are several popular types of *specialty amplifiers*, or amplifiers that are based in some way on op-amp techniques. In an overall application sense, they are not generally used as universally as op-amps. Examples of specialty amplifiers include difference amplifiers, instrumentation amplifiers, programmable gain amplifiers (PGAs) and isolation amplifiers. These will be looked at briefly because they are used in the important areas of data acquisition and distribution systems and embedded systems.

In addition, there are many other types of amplifiers such as audio and video amplifiers, cable drivers, high-speed variable gain amplifiers and various communication-related amplifiers.

23.1 Differential and Common-Mode Signals

A difference amplifier is one that responds to the difference between the two signals applied at its input and ideally rejects signals that are common to the two inputs. The representation of signals in terms of their differential and common-mode components is given below:

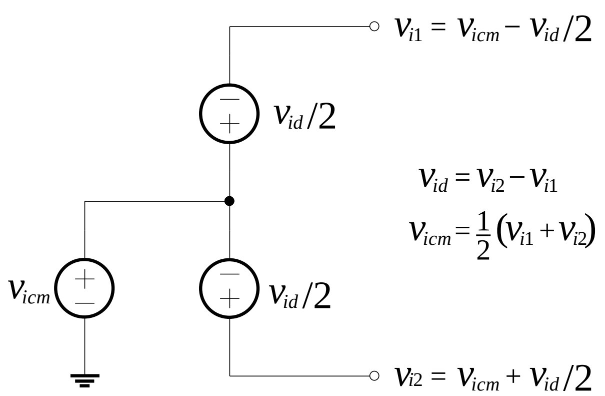


Figure 23.1 – Common-Mode and Differential Signals

Although the ideal difference amplifier will amplify only the differential input signal v_{id} and reject completely the common-mode signal v_{icm}, practical circuits will have an output voltage v_o given by:

$$v_o = A_d v_{id} + A_{cm} v_{icm} \tag{23.1}$$

where A_d denotes the amplifier differential gain and A_{cm} denotes its common-mode gain (ideally zero). One measure of a differential amplifier's performance is the degree of its rejection of common-mode signals in preference to differential signals. This is usually quantified by a measure known as the *common-mode rejection ratio* (CMRR), defined as:

$$\text{CMRR} = 20\log_{10}\frac{|A_d|}{|A_{cm}|}\ \text{dB} \tag{23.2}$$

23.2 Difference Amplifiers

A simple difference amplifier can be constructed with four resistors and an op-amp, as shown below:

Difference amplifier

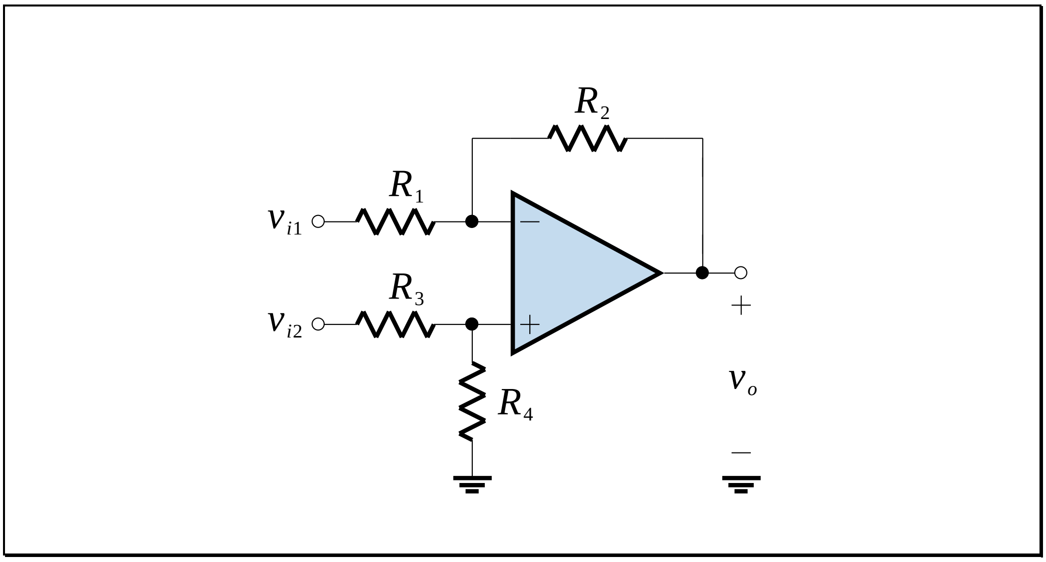


Figure 23.2 – A Difference Amplifier

It is often used in applications where a simple differential to single-ended conversion is required. Because of its popularity, this circuit will be examined in more detail, in order to understand its fundamental limitations before discussing instrumentation amplifiers.

We can use superposition to determine the output:

Analyzing a difference amplifier using superposition

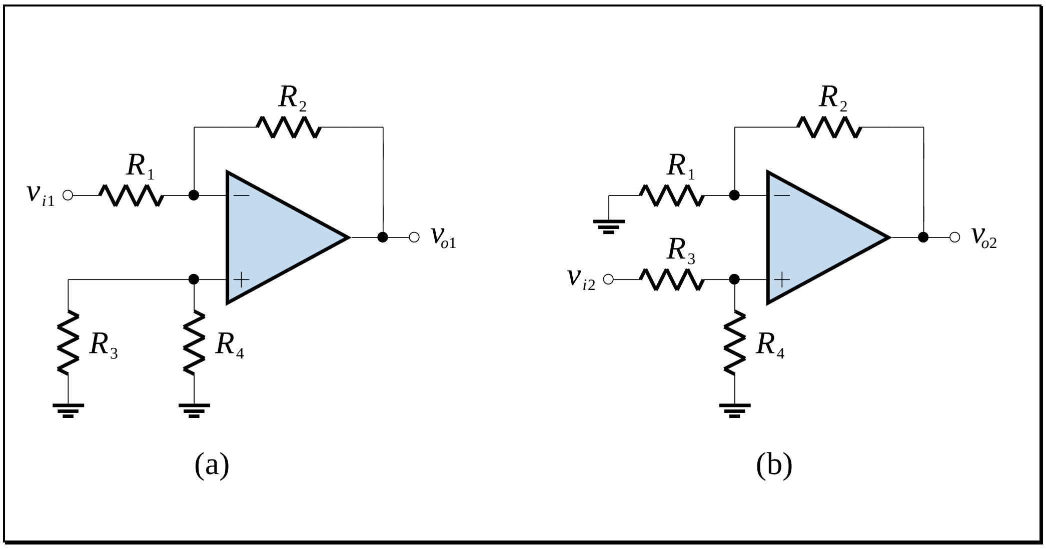


Figure 23.3 – Analyzing a Difference Amplifier

We have:

$$v_{o1} = -\frac{R_2}{R_1} v_{i1} \tag{23.3}$$

and, since we normally select $R_4/R_3 = R_2/R_1$, we also have:

$$v_{o2} = \left(1 + \frac{R_2}{R_1}\right) \frac{R_4}{R_3 + R_4} v_{i2} = \frac{R_2}{R_1} v_{i2} \tag{23.4}$$

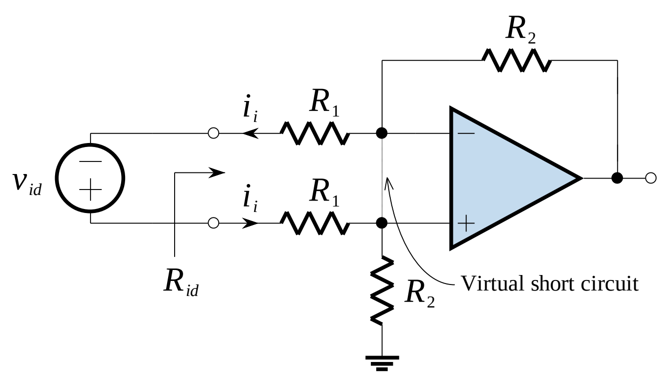
Thus:

$$v_o = \frac{R_2}{R_1}\left(v_{i2} - v_{i1}\right) = \frac{R_2}{R_1} v_{id} = A_d v_{id} \tag{23.5}$$

The differential gain of a difference amplifier

Thus, for exact resistor ratios $R_4/R_3 = R_2/R_1$, the common-mode gain is zero. Any mismatch, however, will lead to finite CMRR.

The differential resistance of the amplifier is calculated as shown below:



Determining the input resistance of a difference amplifier

Figure 23.4 – Determining the Input Resistance of a Difference Amp

Thus, the input resistance is:

$$R_{id} = 2R_1 \tag{23.6}$$

The input resistance of a difference amplifier

23.2.1 Difference Amplifier Deficiencies

If the amplifier is required to have a large differential gain R_2/R_1, then R_1 will be relatively small and the input resistance will be correspondingly low, a drawback of this circuit.

Another drawback is that it is not easy to vary the differential gain, since we need to maintain the resistors in the ratio $R_4/R_3 = R_2/R_1$.

It is also extremely sensitive to source impedance imbalance (the two inputs see different input impedances, so a balanced source is not necessarily a good thing).

A major difference amplifier deficiency is due to inevitable unequal resistor ratios

A net matching tolerance of 0.1% in the resistor ratios yields a worst case DC CMRR of 66 dB, which is quite low (compared to the op-amp which typically has a CMRR of 100 dB).

23.2.2 Difference Amplifier ICs

To overcome the matching tolerance problem on the resistors, it is usual to seek out an integrated circuit difference amplifier which includes on-chip laser trimmed precision thin film resistor networks. An example is the INA146 from Texas Instruments:

The Texas Instruments INA146, a difference amplifier with laser-trimmed resistors – a very flexible IC!

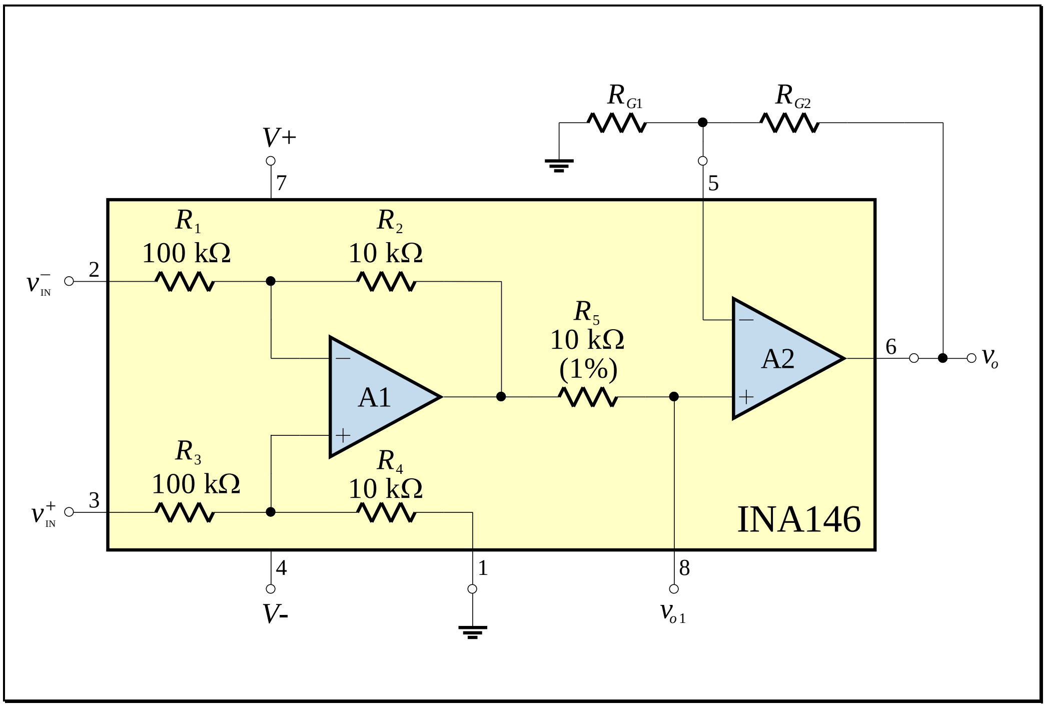
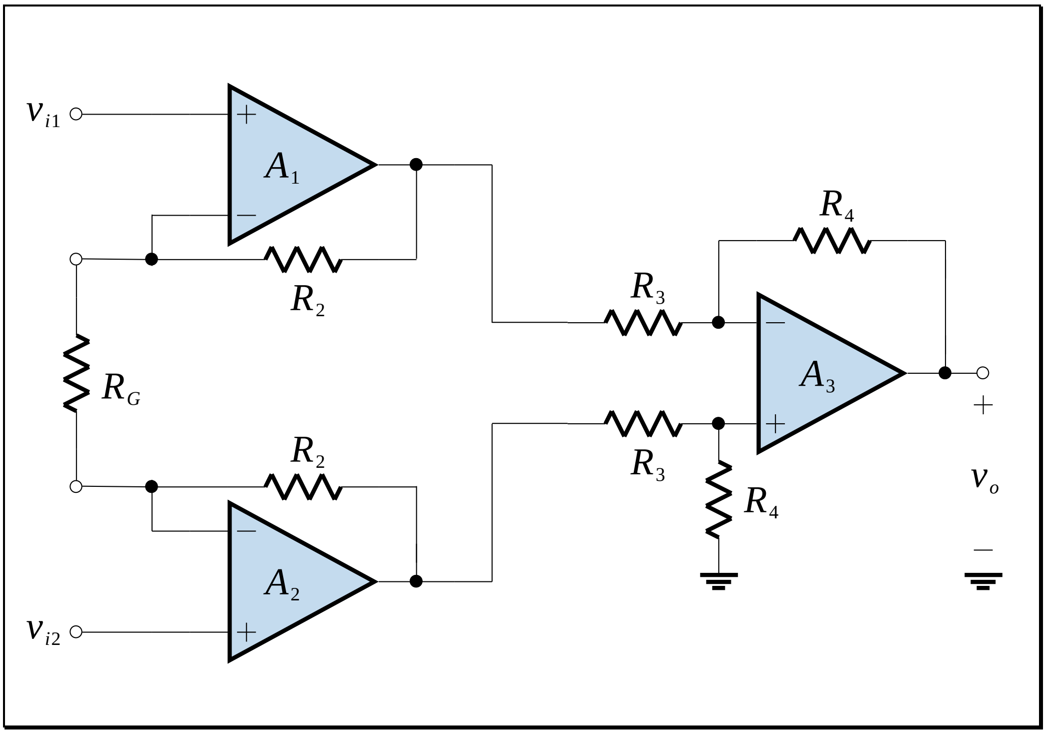


Figure 23.5 – The Texas Instruments INA146

23.3 Instrumentation Amplifiers

The most popular among all of the specialty amplifiers is the *instrumentation amplifier* (or simply *in-amp*). The in-amp is widely used in many industrial and measurement applications where DC precision and gain accuracy must be maintained within a noisy environment, and where large common-mode signals (usually at the AC power line frequency) are present.

The internal configuration of an in-amp is shown below:



A 3 op-amp instrumentation amplifier (in-amp)

Figure 23.6 – A 3 Op-Amp Instrumentation Amplifier

It has a pair of differential input terminals, and a single-ended output that works with respect to a reference or common terminal. The input impedances are balanced and high in value, typically greater than $1\,\mathrm{G}\Omega$. The in-amp uses an internal feedback resistor network, plus one gain set resistance, R_G, which is external to the package.

An analysis of the in-amp is shown below:

Analysis of the in-amp to determine the differential gain

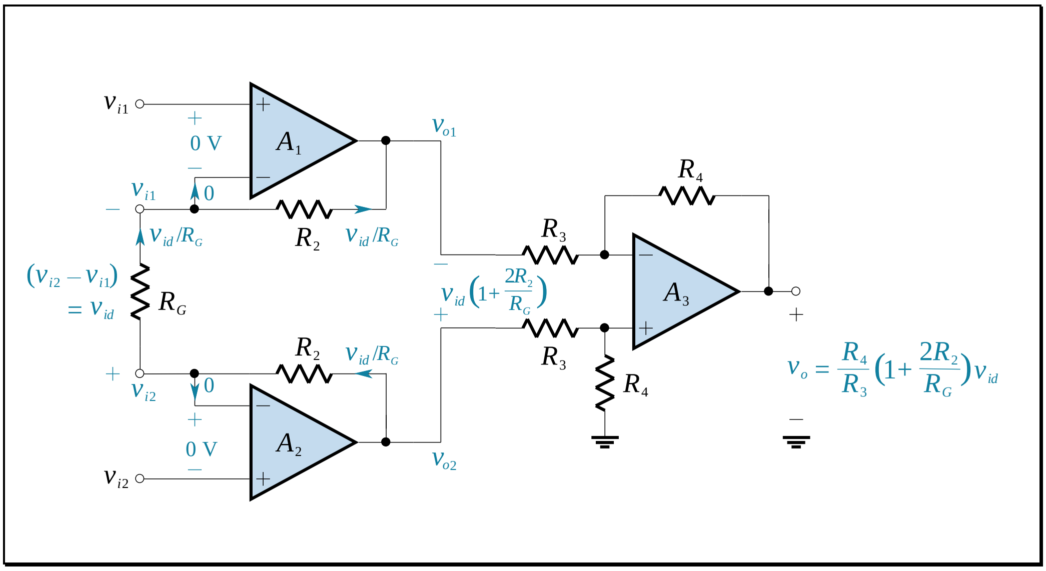


Figure 23.7 – Analysis of the In-Amp

The analysis shows that the gain of the in-amp is given by:

The differential gain of a difference amplifier

$$v_o = \frac{R_4}{R_3}\left(1 + \frac{2R_2}{R_G}\right)v_{id} = A_d v_{id} \tag{23.7}$$

Note that this formula is only valid for perfect op-amps and matched resistors. In this case, there is no output component due to the common-mode voltage, and so the common-mode gain is $A_{cm} = 0$. Real in-amps have CMRRs ranging from 90 to 130 dB, i.e. the common-mode gain is around a million times smaller than the differential gain!

23.3.1 In-Amp Advantages

The in-amp gain can be set by the user by selecting the external R_G. In-amp gain can also be preset via an internal R_G by pin selection (again isolated from the signal inputs). Typical in-amp gains range from 1 to 1000. The internal resistors are trimmed so that standard 1% or 0.1% resistors can be used to set the gain to popular values, such as 2, 5, 10, …, 200, 500, 1000.

Furthermore, common-mode signals are only amplified in the first stage by a factor of 1 regardless of gain. When the two input terminals are connected together to a common-mode input voltage v_{icm}, it is easy to see that an equal voltage appears at the negative input terminals of A_1 and A_2, causing the current through R_G to be zero. Thus there will be no current in the R_2 resistors, and the voltages at the output terminals of A_1 and A_2 will be equal to the input (i.e. v_{icm}). Thus the first stage does not amplify v_{icm}, it simply propagates v_{icm} to its two output terminals where they are subtracted to produce a zero common-mode output by A_3.

The in-amp has a very high CMRR

Another important feature of the in-amp is that the internal resistance network and R_G are *isolated* from the signal input terminals.

23.3.2 In-Amp Disadvantages

From Figure 23.7, we can see that for the in-amp to "work", the sum of the common-mode voltage and the signal voltage at the outputs of A_1 and A_2 must fall within the amplifier output voltage range. Otherwise, one of these op-amps will saturate and the output of the difference amplifier will not be proportional to the differential voltage.

Unlike the difference amplifier, there is no attenuation of the input voltage before it "meets" the op-amp terminals. Therefore, the input voltages to the in-amp must lie within the supply voltage range for the in-amp to "work".

The CMRR is a function of frequency – it falls off at around 10-100 Hz. In addition, the bandwidth (which is dependent on the gain) falls within the range 1-1000 kHz. This makes the in-amp only suitable for very low frequency applications, such as weigh scales, ECG and medical instrumentation and industrial process controls where the frequency of the sensor signals is naturally low.

23.3.3 In-Amp Application

An example of a commercially available in-amp is the AD620. An application of the in-amp is shown below, where it is used in a data acquisition system to amplify the output from a pressure transducer bridge (which is normally balanced) from a single 5 V supply:

Typical application of an in-amp (the AD620 from Analog Devices) as an amplifier for a pressure transducer

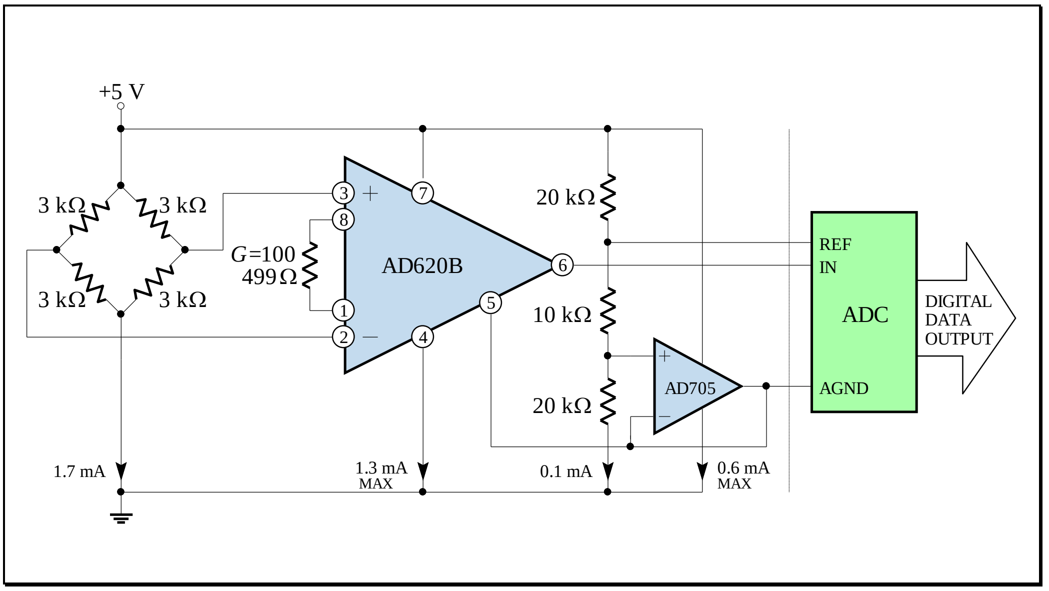


Figure 23.8 – A Pressure Monitor Circuit

The in-amp is ideally suited for this application because the bridge output is fundamentally balanced, and the in-amp presents it with a truly balanced high impedance load.

Full scale output voltages from a typical bridge circuit can range from approximately 10 mV to several hundred mV. Typical in-amp gains in the order of 100 to 1000 are therefore ideally suited for amplifying these small voltages to levels compatible with popular analog-to-digital (ADC) input voltage ranges (usually 1 V to 10 V full scale).

In addition, the in-amp's high CMRR at power line frequencies allows common-mode noise to be rejected, when the bridge must be located remotely from the in-amp.

23.4 Programmable Gain Amplifiers

Most data acquisition systems with wide dynamic range need some method of adjusting the input signal level to the analog-to-digital converter (ADC). Typical ADC full scale input voltage ranges lie between 1 V and 10 V. To achieve the rated precision of the converter, the maximum input signal should be fairly near its full scale voltage.

A PGA is used to increase the dynamic range of a system

Transducers however, have a very wide range of output voltages. High gain is needed for a small sensor voltage, but with a large output, a high gain will cause the amplifier or ADC to saturate. So, some type of predictably controllable gain device is needed.

Such a device has a gain that is controlled by a digital input. This device is known as a *programmable gain amplifier*, or PGA. Typical PGAs may be configured either for selectable *decade gains* such as 10, 100, 1000, etc., or they might also be configured for *binary gains* such as 1, 2, 4, 8, etc.

A PGA is usually located between a sensor and its ADC, as shown below:

A PGA's location in a system

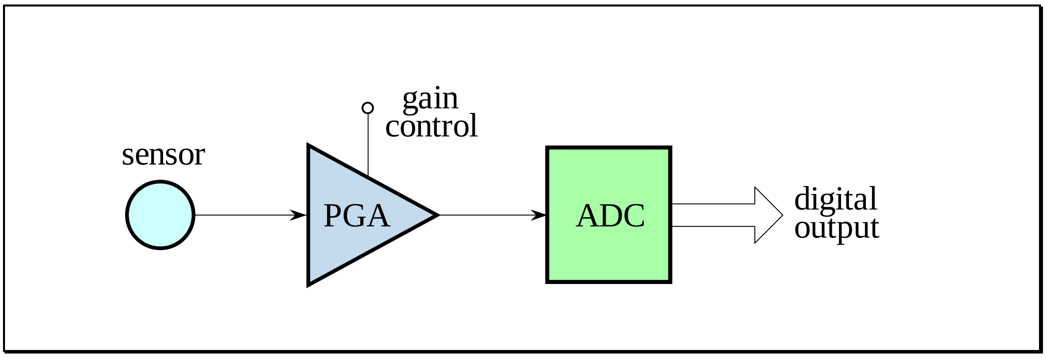


Figure 23.9 – Use of a Programmable Gain Amplifier

Thus, a digital processor can combine PGA gain information with the digital output of the ADC to increase its resolution. Some ADCs have on-chip PGAs.

23.4.1 PGA Design Issues

- How to switch the gain
- Effects of the switch on-resistance (R_{ON})
- Gain accuracy
- Gain linearity
- Bandwidth
- DC offset
- Gain and offset drift over temperature
- Settling time after switching gain

23.4.2 PGA Example

An example of a PGA with programmable gains of 1, 10, 100 and 1000 is shown below:

A PGA with a decade scale that is not influenced by switch resistance

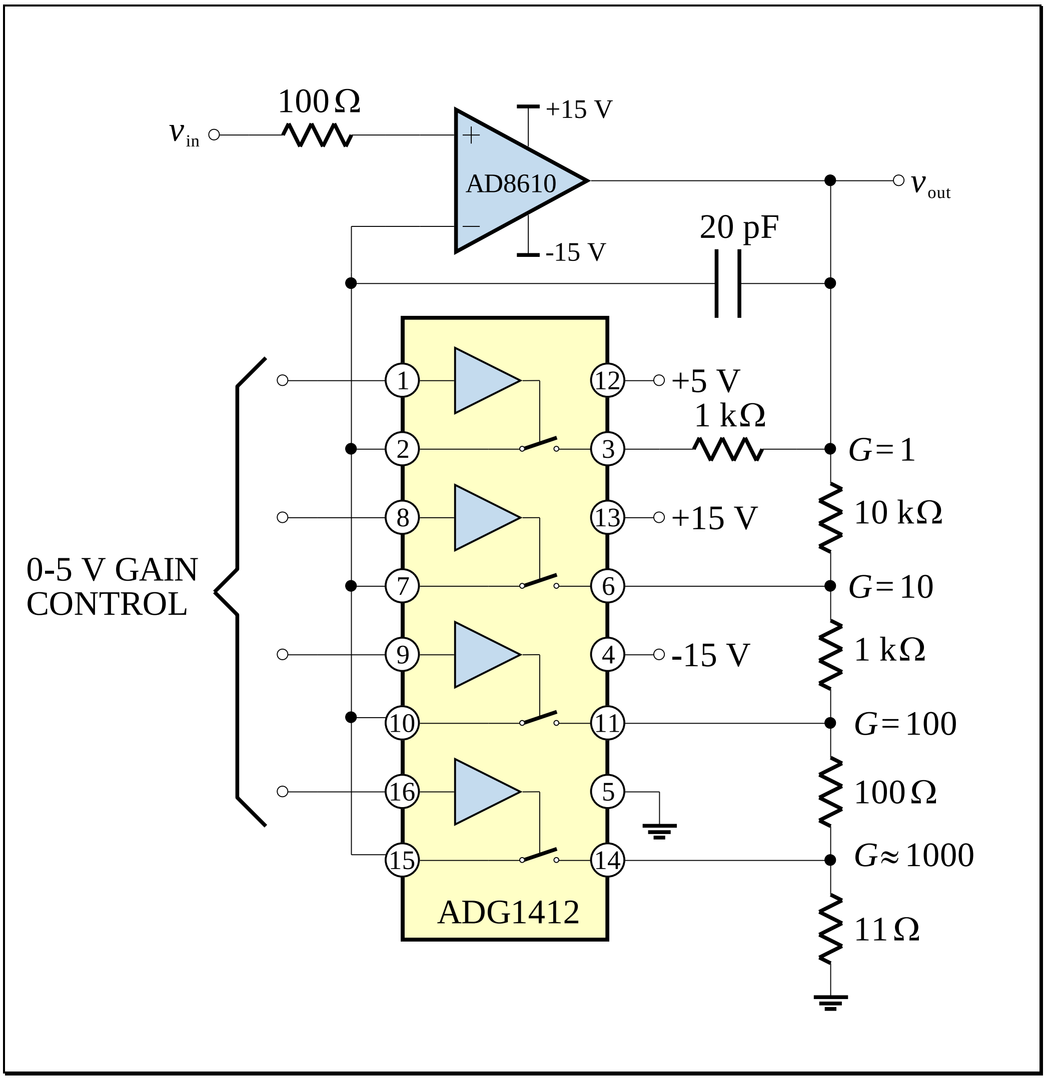


Figure 23.10 – A PGA with Decade Gains

23.5 Isolation Amplifiers

There are many applications where is it desirable, or even essential, for a sensor to have no direct ("galvanic") electrical connection with the system to which it is supplying data. This might be in order to avoid the possibility of dangerous voltages or currents from one half of the system doing damage in the other. Such a system is said to be *isolated*, and the arrangement that passes a signal without galvanic connections is known as an *isolation barrier*.

Electrical isolation is necessary in a wide variety of applications

Examples include the need to prevent the ignition of explosive gases by sparks at sensors and the protection from electric shock of patients whose ECG or EEG is being monitored. In the ECG case, protection may be required in *both* directions: the patient must be protected from accidental electric shock, but if the patient's heart should stop, the ECG machine must be protected from the very high voltages (> 7.5 kV) applied to the patient by the defibrillator which will be used to attempt to restart it.

A summary of applications for isolation amplifiers (both analog and digital) is given below:

- Sensor is at a high voltage relative to other circuitry (or may become so under fault conditions)
- Sensor may not carry dangerous voltages, irrespective of faults in other circuitry (e.g. patient monitoring and intrinsically safe equipment for use with explosive gases)
- To break ground loops (two "grounds" in the same system at different voltages, causing current between them)

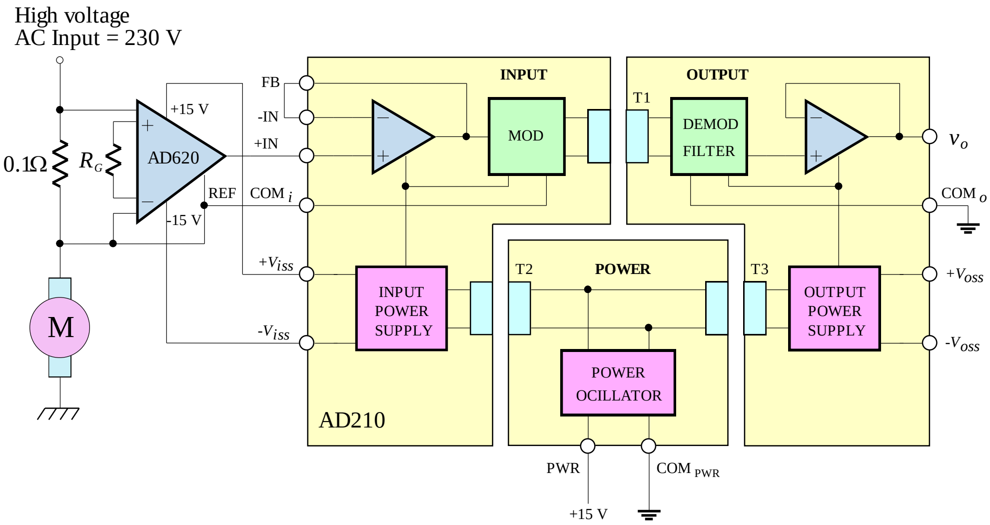
The most common isolation amplifiers use *transformers*, which exploit magnetic fields, and another common type uses small high voltage capacitors, exploiting electric fields. *Optoisolators*, which consist of an LED and a photocell, provide isolation by light. Different isolators have differing performance: some are sufficiently linear to pass high accuracy analog signals across an isolation barrier. With others, the signal may need to be converted to digital form before transmission for accuracy to be maintained.

EXAMPLE 23.1 Isolation Amplifier

An example of a 3-port isolation transformer is the Analog Devices AD210. It allows the user to select gains from 1 to 100, using external resistors, with the input section op-amp. It uses transformers to achieve 2500 V RMS (3500 V peak) isolation.

A typical application using the AD210 is shown below:

Motor control current sensing using an analog 3-port isolator – providing power, input and output isolation

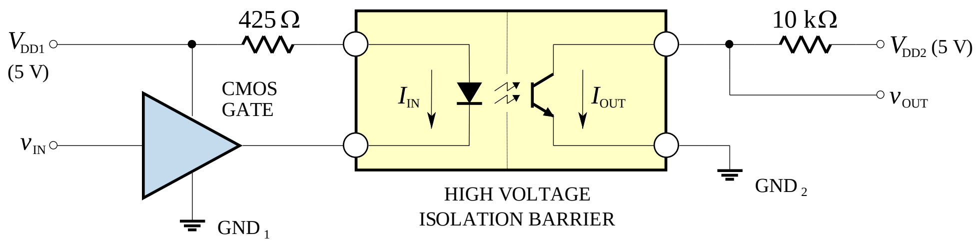


The AD210 is used with an AD620 in-amp in a high-side current-sensing system for motor control. Current is sensed on the high (230 V) side of the motor (as opposed to using a resistor on the ground side – so-called low-side sensing) so that any fault current from the motor to the chassis (earth) is also detected.

The input of the AD210, being isolated, can be directly connected to a 230 V power line without protection being necessary. The input section's isolated ±15 V powers the AD620, which senses the voltage drop in a small value current sensing resistor. The AD210 input stage op-amp is simply connected as a unity-gain follower. The 230 V RMS common-mode voltage is ignored by this isolated system.

EXAMPLE 23.2 Digital Isolator

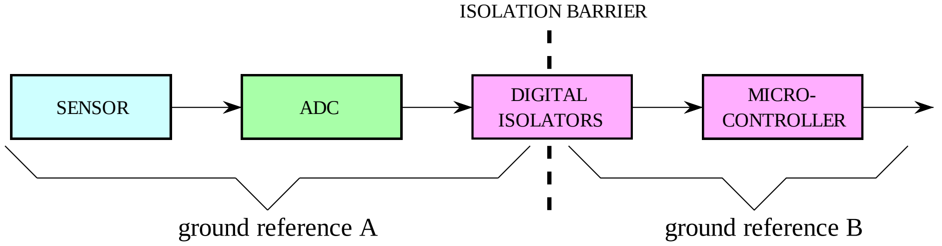
An example of a digital isolator is an LED / phototransistor optocoupler:



A digital isolator using a LED and phototransistor

A current of approximately 10 mA drives an LED transmitter, with light output received by a phototransistor. The light produced by the LED saturates the phototransistor (drives the collector-emitter voltage low). Input / output isolation of 5000–7000 V RMS is common. Although fine for digital signals, most optocouplers are too nonlinear for most analog applications. Also, since the phototransistor is being saturated, response times can be around 10-20 μs for slower (cheap) devices, limiting high speed applications.

The availability of low cost digital isolators solves most system isolation problems in data acquisition systems as shown below:



Practical application of digital isolation in a data acquisition system

In this system, digitizing the signal first using an ADC with *serial* output, then using digital isolation, eliminates the problem of analog isolation amplifiers (which are expensive).

23.6 Summary

- The two signals that appear at the inputs of difference amplifiers can be split into a *common-mode signal* and a *differential signal*. This aids in calculating the output of a difference amplifier, and gives rise to a measure of difference amplifier performance, known as the common-mode rejection ratio, or CMRR. It is usually expressed in dB.

- *Difference amplifiers* are used to amplify the difference between two signals, and to reject the common-mode. However, they are very sensitive to the values (i.e. tolerance, drift due to temperature, etc.) of the resistors used in the circuit. Therefore, only integrated circuit packages with laser-trimmed resistors in close proximity are used in practice.

- *Instrumentation amplifiers* are an arrangement of op-amps that overcomes some of the deficiencies of the standard difference amplifier. They have a very large balanced input impedance, and can provide large differential gain whilst rejecting the common-mode signal. They are readily available in integrated circuit packages, and are commonly used in bridge circuits.

- *Programmable gain amplifiers* are used in data acquisition systems where the transducer exhibits a very large range of output voltages. They can be controlled digitally.

- *Isolation amplifiers* provide electrical isolation between sensors and electronic circuitry, and are used in environments where a direct electrical connection would be hazardous, such as in mining and biomedical applications.

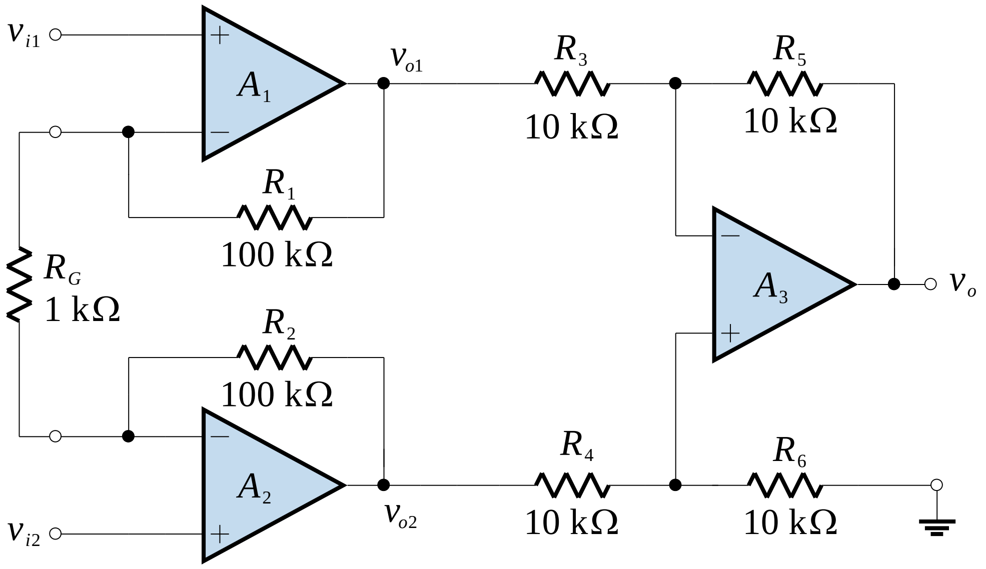
23.7 References

Jung, W: *Op-Amp Applications*, Analog Devices, 2002.

Exercises

1.

Consider the instrumentation amplifier implementation shown below:



(a) Determine expressions for the output voltages of A_1 and A_2 in terms of the input voltages v_{i1} and v_{i2}.

(b) Find the overall differential voltage gain A_d of the in-amp.

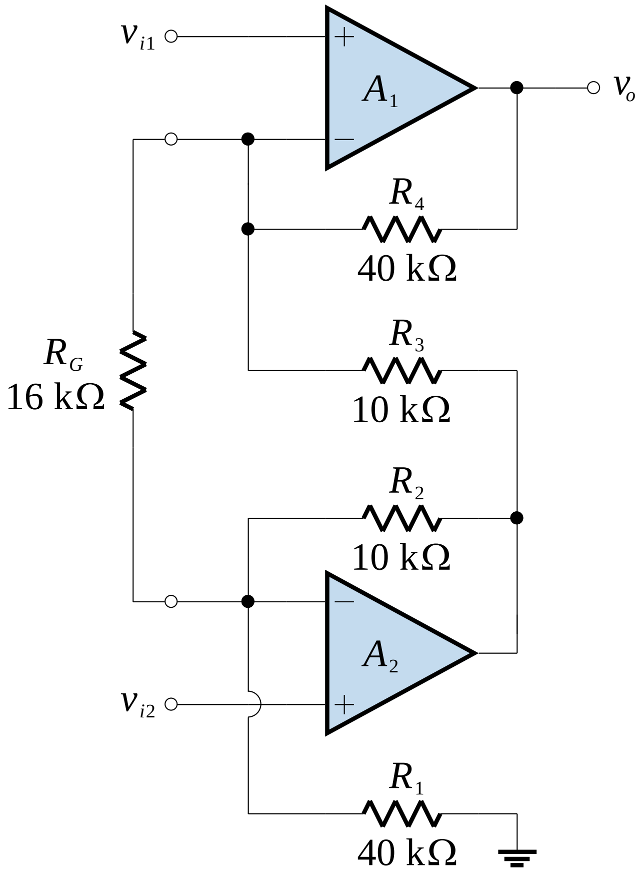
(c) The following voltages are applied to the in-amp: $v_{i1} = 230\ \text{mV}$, $v_{i2} = 235\ \text{mV}$. Determine the final output voltage.

(d) What value of R_G must be used to change the gain of the in-amp to 1000?

23.18

2.

Consider the two-op-amp instrumentation amplifier shown below:



Determine the value of the overall differential voltage gain A_d of the in-amp.

24 Transfer Functions

Contents

Introduction

In any particular circuit, the phasor ratio of the desired forced response to the forcing function, written in terms of the complex frequency **s**, is called the *transfer function*. Thus, a transfer function is an input-output description of the behaviour of a circuit, and it does not include any information concerning the internal structure of the circuit and its behaviour (we have already seen that the *RLC* circuit can be replaced by a biquad circuit utilising op-amps – both have the same input-output behaviour within certain bounds).

The transfer function is intimately related to the characteristic equation – it *completely characterises* a circuit. We can thus dispense with circuit schematics, and start to think in terms of cascaded and interconnected "blocks" that are described by transfer functions – so-called *block diagrams*.

The transfer function will also be seen to hold information about the *form* of the circuit's natural response. Thus, given a transfer function of a circuit, we can write down an expression for the natural response by inspection. If we are given a forcing function and the initial conditions, we can then determine the *complete* response.

We shall also see that circuits are not special – we can model any system described by linear differential equations (e.g. mechanical, hydraulic, electrical, thermal, fluid) with block diagrams.

24.1 Transfer Functions

Let us analyse the following simple circuit:

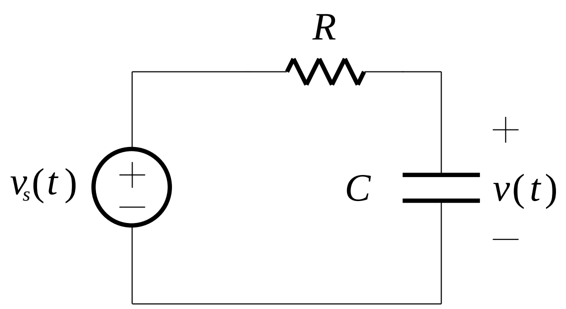


Figure 24.1

We can find the forced response of this circuit by working in the frequency-domain (if the source has a functional form that we can transform to the frequency-domain, such as an exponentially damped sinusoid). Then we have:

$$\mathbf{V} = \frac{1/\mathbf{s}C}{R+1/\mathbf{s}C}\mathbf{V}_s = \frac{1/RC}{\mathbf{s}+1/RC}\mathbf{V}_s \tag{24.1}$$

A *transfer function* is the ratio of the desired forced response to the forcing function, using phasor notation and the complex frequency **s**. It is usually designated $\mathbf{T}(\mathbf{s})$:

$$\mathbf{T}(\mathbf{s}) = \frac{\mathbf{V}}{\mathbf{V}_s} \tag{24.2}$$

With this notation, we can see that the transfer function is a *complex function* of the *complex variable* $\mathbf{s}$. Thus, in this case, we have:

$$\mathbf{T}(\mathbf{s}) = \frac{1/RC}{\mathbf{s}+1/RC} \tag{24.3}$$

Several important conclusions can be drawn from the *form* of the transfer function.

24.1.1 Characteristic Equation

Setting the denominator of the transfer function to zero, we get the *characteristic equation* of the circuit, $\mathbf{s}+1/RC=0$. This corresponds exactly to that obtained from the original source-free homogeneous differential equation for the circuit, $dv/dt+v/RC=0$. Thus we see that the *denominator* of the transfer function seems to come directly from the source-free circuit itself, and is thus dependent only on the topology and types of passive circuit components. We will see later that the form of the natural (i.e. source-free) response is directly related to the characteristic equation.

24.1.2 Pole-Zero Plot

Roots of the characteristic equation give us the poles of the transfer function. In this case there is one pole at $\mathbf{s}=-1/RC+j0$. Zeros are obtained by finding those frequencies for which the transfer function is zero. In this case, there is a zero at $\mathbf{s}=\pm\infty$. When plotting poles and zeros, we normally don't show poles and zeros at infinity, and so the pole-zero plot corresponding to this particular transfer function is:

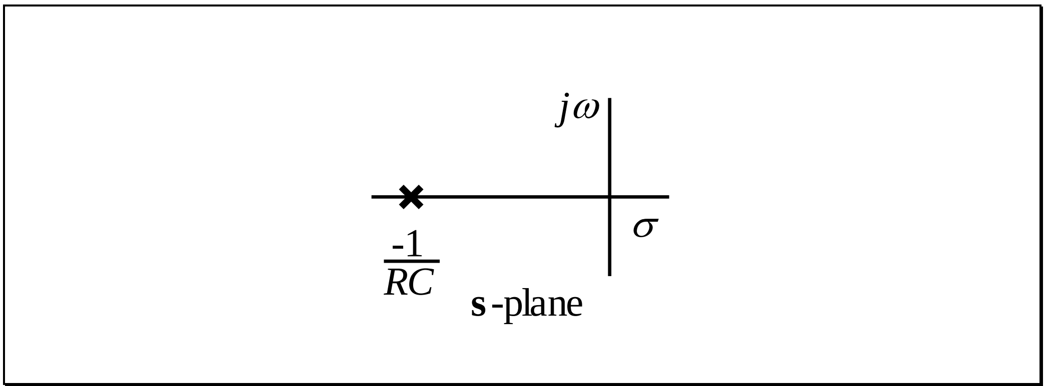


Figure 24.2

This pole-zero plot conveys exactly the same information as the transfer function, apart from the multiplicative factor $1/RC$. For example, from the plot above we know the transfer function has the form $\dfrac{K}{\mathbf{s}+1/RC}$. In many instances, just the form of the transfer function tells us a lot about the circuit's behaviour, so engineers make a lot of use of pole-zero plots.

24.1.3 Transfer Function Form

The form of a transfer function is such that it can always be written as a ratio of two polynomials. The numerator polynomial can be factored into *m* "zero" terms, the denominator can be factored into *n* "pole" terms.

$$\mathbf{T}(\mathbf{s}) = K \frac{(\mathbf{s}-\mathbf{z}_1)(\mathbf{s}-\mathbf{z}_2)\cdots(\mathbf{s}-\mathbf{z}_m)}{(\mathbf{s}-\mathbf{p}_1)(\mathbf{s}-\mathbf{p}_2)\cdots(\mathbf{s}-\mathbf{p}_n)} \tag{24.4}$$

The number of pole terms determines the order of the circuit. For our simple example *RC* circuit, we have a first-order circuit, since there is one pole. There are no zeros (except for the implicit one at $\mathbf{s} = \pm\infty$).

Some of the poles in the transfer function may occur as complex conjugate pairs. For example, if the first two pole terms are complex conjugates, then the transfer function can be written as:

$$\begin{aligned}\mathbf{T}(\mathbf{s}) &= K \frac{(\mathbf{s}-\mathbf{z}_1)(\mathbf{s}-\mathbf{z}_2)\cdots(\mathbf{s}-\mathbf{z}_m)}{(\mathbf{s}-\mathbf{p}_1)(\mathbf{s}-\mathbf{p}_1^*)\cdots(\mathbf{s}-\mathbf{p}_n)} \\ &= K \frac{(\mathbf{s}-\mathbf{z}_1)(\mathbf{s}-\mathbf{z}_2)\cdots(\mathbf{s}-\mathbf{z}_m)}{\left(\mathbf{s}^2-(\mathbf{p}_1+\mathbf{p}_1^*)\mathbf{s}+\mathbf{p}_1\mathbf{p}_1^*\right)\cdots(\mathbf{s}-\mathbf{p}_n)}\end{aligned} \tag{24.5}$$

If we let $\mathbf{p}_1 = x + jy$ then $\mathbf{p}_1^* = x - jy$, then $\mathbf{p}_1 + \mathbf{p}_1^* = 2x$ and $\mathbf{p}_1\mathbf{p}_1^* = x^2 + y^2$, both of which are real quantities. Thus, if we define the real numbers $a = -2x$ and $b = x^2 + y^2$, then the transfer function can be written as:

$$\mathbf{T}(\mathbf{s}) = K \frac{(\mathbf{s}-\mathbf{z}_1)(\mathbf{s}-\mathbf{z}_2)\cdots(\mathbf{s}-\mathbf{z}_m)}{(\mathbf{s}^2+a\mathbf{s}+b)\cdots(\mathbf{s}-\mathbf{p}_n)} \tag{24.6}$$

This is an extremely important observation – it tells us that a transfer function can be written with real coefficients as a product of first-order and second-order factors! Thus, if we become familiar with the properties of first-order and second-order circuits, we can handle circuits of *any order*!

24.1.4 Relationship to Differential Equation

Returning to the original circuit, we can do KVL around the loop and write the describing differential equation in the form:

$$\frac{dv}{dt} + \frac{v}{RC} = \frac{v_s}{RC} \tag{24.7}$$

or, using the D operator:

$$\left(D + \frac{1}{RC}\right)v = \frac{v_s}{RC} \tag{24.8}$$

Now, putting aside all mathematical formality, we substitute **s** for D, and transform into the frequency-domain:

$$\left(\mathbf{s} + \frac{1}{RC}\right)\mathbf{V} = \frac{1}{RC}\mathbf{V}_s \tag{24.9}$$

We now get:

$$\frac{\mathbf{V}}{\mathbf{V}_s} = \mathbf{T}(\mathbf{s}) = \frac{1/RC}{\mathbf{s} + 1/RC} \tag{24.10}$$

Thus, we can directly convert a differential equation into the frequency-domain by setting all derivatives to **s**. It also turns out that integrals in the time-domain turn into $1/\mathbf{s}$ factors when converting to the frequency-domain.

A more mathematically formal way to transform from the differential equation (the time-domain) to the transfer function (the frequency-domain) will come later with a study of the Laplace Transform.

24.1.5 Circuit Abstraction

The transfer function is a *relationship between the input and output of a circuit only*, expressed in the frequency-domain using the notion of complex frequency and phasors. It tells us nothing about internal voltages, currents, consumed power, topology or even components. Thus, we could represent the *RC* circuit as a simple "black box" that converts one phasor into another:

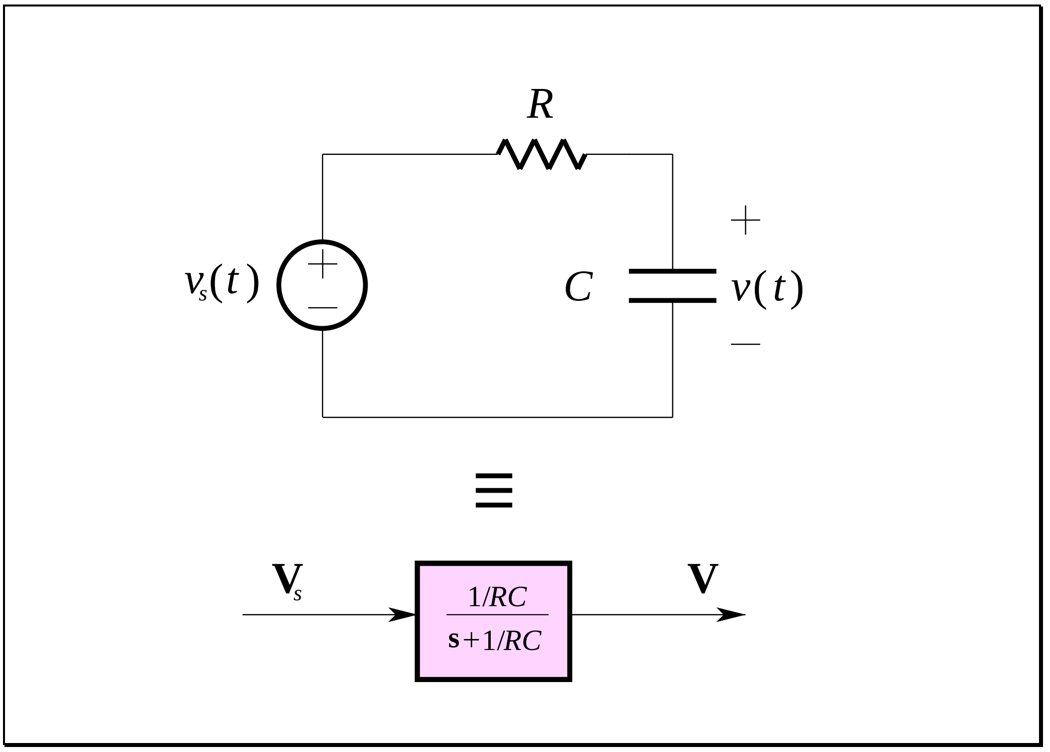


Figure 24.3

The representation of a circuit in terms of its input signal, transfer function and output signal is known as a *block diagram*. The block diagram is a concept that is used often across all disciplines of engineering, and can be applied to any linear system, not just electrical circuits. It is a way to characterize a system without resorting to writing differential equations – instead we represent a system by its transfer function and work with input and output phasors. The only drawback to this approach is that we must work with algebraic equations involving complex numbers – but most of us would prefer this to solving differential equations!

24.2 Forced Response

For now, we can only consider cases where the forced response can be expressed as an exponentially damped sinusoid (in the most general case). This includes DC, exponential and sinusoidal waveforms as special cases.

To determine the forced response, the transfer function is evaluated at the complex frequency of the forcing function, say $\mathbf{s} = \mathbf{s}_f$. Then, to determine the phasor representation of the desired response, we simply multiply the phasor representation of the forcing function by the transfer function (which is now just a complex number).

That is, the output phasor representing the forced response is given by:

$$\mathbf{V}_f = \mathbf{T}(\mathbf{s})\big|_{\mathbf{s}=\mathbf{s}_f} \mathbf{V}_s \tag{24.11}$$

For example, suppose the forcing function is $v_s(t) = V_0$, a DC voltage. Then we know that this forcing function has a complex frequency $\mathbf{s} = \mathbf{0}$, so evaluation of the transfer function at $\mathbf{s} = \mathbf{0}$ gives:

$$\mathbf{T}(\mathbf{s})\big|_{\mathbf{s}=\mathbf{0}} = \left.\frac{1/RC}{\mathbf{s} + 1/RC}\right|_{\mathbf{s}=\mathbf{0}} = 1 \tag{24.12}$$

Thus, the desired forced response voltage phasor is:

$$\mathbf{V}_f = \mathbf{T}(\mathbf{s})\big|_{\mathbf{s}=\mathbf{0}} \mathbf{V}_s = 1 \cdot V_0 = V_0 \tag{24.13}$$

Converting back to the time-domain, we find that the forced response is V_0, which we know is true from other circuit analysis methods.

If we had a sinusoidal forcing function, $v_s(t) = V_m \cos(\omega t + \theta)$, then we evaluate the transfer function at $\mathbf{s} = j\omega$:

$$\begin{aligned}
\mathbf{T}(\mathbf{s})\big|_{\mathbf{s}=j\omega} &= \left.\frac{1/RC}{\mathbf{s} + 1/RC}\right|_{\mathbf{s}=j\omega} \\
&= \frac{1/RC}{j\omega + 1/RC} \\
&= \frac{1}{1 + j\,\omega/\omega_0}, \quad \omega_0 = 1/RC
\end{aligned} \tag{24.14}$$

Thus, the desired forced response voltage phasor is:

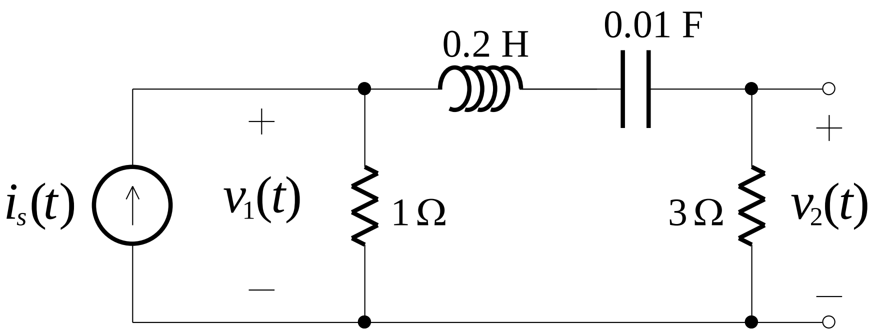
$$\begin{aligned}
\mathbf{V}_f &= \mathbf{T}(\mathbf{s})\big|_{\mathbf{s}=j\omega} \mathbf{V}_s \\
&= \frac{1}{1 + j\,\omega/\omega_0} \cdot V_m \angle\theta \\
&= \frac{V_m \angle\theta}{1 + j\,\omega/\omega_0}
\end{aligned} \tag{24.15}$$

In more advanced circuit analysis, when the Laplace Transform is introduced, we will be able to handle almost any type of forcing function, rather than just exponentially damped sinusoids (or its special cases), by replacing the forcing function phasor (which is a constant) by a forcing function which is a mathematical expression containing **s**.

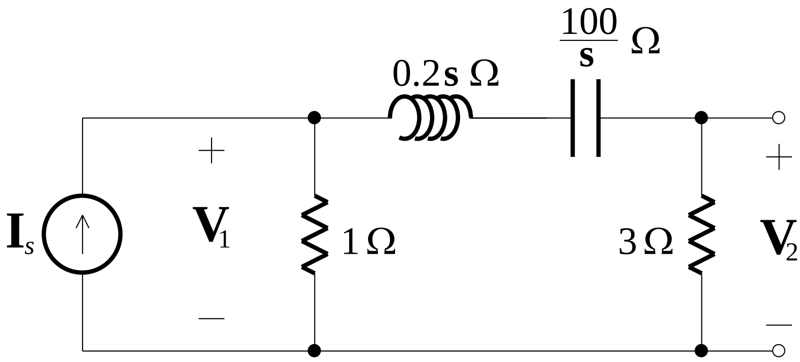
24.10

EXAMPLE 24.1 Forced Response

Consider the circuit:



We transform to the frequency-domain:



and, using the current-divider rule, we write down the transfer functions:

$$\frac{\mathbf{V}_1}{\mathbf{I}_s} = \mathbf{T}_1(\mathbf{s}) = \frac{3 + 0.2\mathbf{s} + 100/\mathbf{s}}{4 + 0.2\mathbf{s} + 100/\mathbf{s}} = \frac{\mathbf{s}^2 + 15\mathbf{s} + 500}{\mathbf{s}^2 + 20\mathbf{s} + 500}$$

$$\frac{\mathbf{V}_2}{\mathbf{I}_s} = \mathbf{T}_2(\mathbf{s}) = 3\frac{1}{4 + 0.2\mathbf{s} + 100/\mathbf{s}} = \frac{15\mathbf{s}}{\mathbf{s}^2 + 20\mathbf{s} + 500}$$

Note that the denominators are the same!

Suppose we want to find $\mathbf{V}_2$ for $\mathbf{I}_s = 0.2\angle 0° \text{ A}$ and $\mathbf{s} = \mathbf{0}$. Then:

$$\mathbf{V}_2 = \mathbf{T}_2(\mathbf{0})\mathbf{I}_s = \frac{15 \cdot \mathbf{0}}{\mathbf{0}^2 + 20 \cdot \mathbf{0} + 500} 0.2\angle 0° = \mathbf{0}$$

Thus, the output voltage for a DC current source is zero, as can be seen by inspection of the circuit – the capacitor acts as an open circuit.

Suppose now that $\mathbf{s} = \infty$. Then:

$$\mathbf{V}_2 = \mathbf{T}_2(\infty)\mathbf{I}_s = \frac{15/\mathbf{s}}{1 + 20/\mathbf{s} + 500/\mathbf{s}^2} 0.2\angle 0° = \frac{15/\infty}{1 + 20/\infty + 500/\infty^2} 0.2\angle 0° = \mathbf{0}$$

Thus, the output voltage for an infinitely high frequency source (say a sinusoidal source) is zero, as can be seen by inspection of the circuit – the inductor acts as an open circuit.

Suppose now that $\mathbf{s} = -10$. Then:

$$\mathbf{V}_2 = \mathbf{T}_2(-10)\mathbf{I}_s = \frac{15 \cdot (-10)}{(-10)^2 + 20 \cdot (-10) + 500} 0.2\angle 0° = -0.075$$

Thus $v_2(t) = -0.075 e^{-10t}$. Note that this is a forced response due to the forcing function $v_2(t) = 0.2 e^{-10t}$, and **not** the natural response.

Now apply a sinusoidal forcing function, with $\mathbf{s} = j10$. Then:

$$\begin{aligned}
\mathbf{V}_2 &= \mathbf{T}_2(j10)\mathbf{I}_s \\
&= \frac{15 \cdot (j10)}{(j10)^2 + 20 \cdot (j10) + 500} 0.2\angle 0° \\
&= \frac{j3}{8 + j4} 0.2\angle 0° \\
&= 0.06708 \angle 63.4°
\end{aligned}$$

Thus $v_2(t) = 67.08 \cos(10t + 63.4°) \text{ mV}$.

Lastly, let $\mathbf{s} = -10 + j20$. Then since we are exciting the system at the same frequency as a pole, the response must be infinite.

24.3 Frequency Response

Frequency response, by definition, is the sinusoidal steady-state response, i.e. the forced response to a sinusoid. As shown above, we can derive the frequency response directly from the transfer function. All we have to do is evaluate the transfer function at an arbitrary sinusoidal frequency $\mathbf{s} = j\omega$:

$$\mathbf{T}(\mathbf{s})\big|_{\mathbf{s}=j\omega} = \mathbf{T}(j\omega) = \text{frequency response} \tag{24.16}$$

For example, for the simple *RC* circuit, we have:

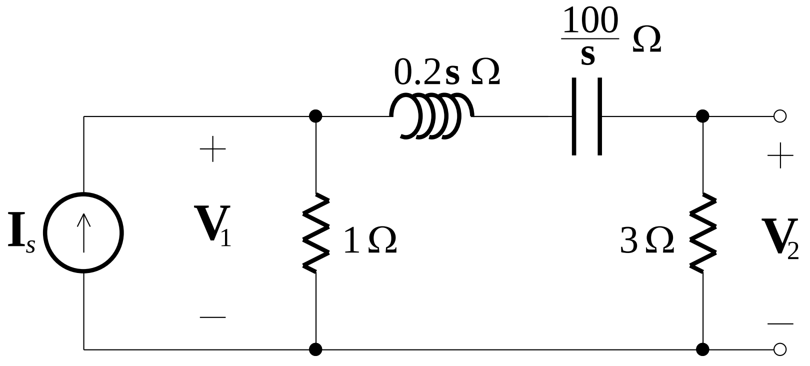
$$\mathbf{T}(\mathbf{s})\big|_{\mathbf{s}=j\omega} = \frac{1}{1 + j\,\omega/\omega_0}, \quad \omega_0 = 1/RC \tag{24.17}$$

Thus, it is a simple matter to derive expressions for the frequency response from the transfer function. We can now determine the magnitude response and phase response for the simple *RC* circuit:

$$\begin{aligned} \left|\mathbf{T}(j\omega)\right| &= \frac{1}{\sqrt{1 + (\omega/\omega_0)^2}} \\ \angle\mathbf{T}(j\omega) &= -\tan^{-1}(\omega/\omega_0) \end{aligned} \tag{24.18}$$

EXAMPLE 24.2 Frequency Response from Transfer Function

For the circuit seen previously, in the frequency-domain:



we can use the transfer function to establish the frequency response quite quickly. For example, we know:

$$\mathbf{T}_2(\mathbf{s}) = \frac{15\mathbf{s}}{\mathbf{s}^2 + 20\mathbf{s} + 500}$$

Then the frequency response is given by:

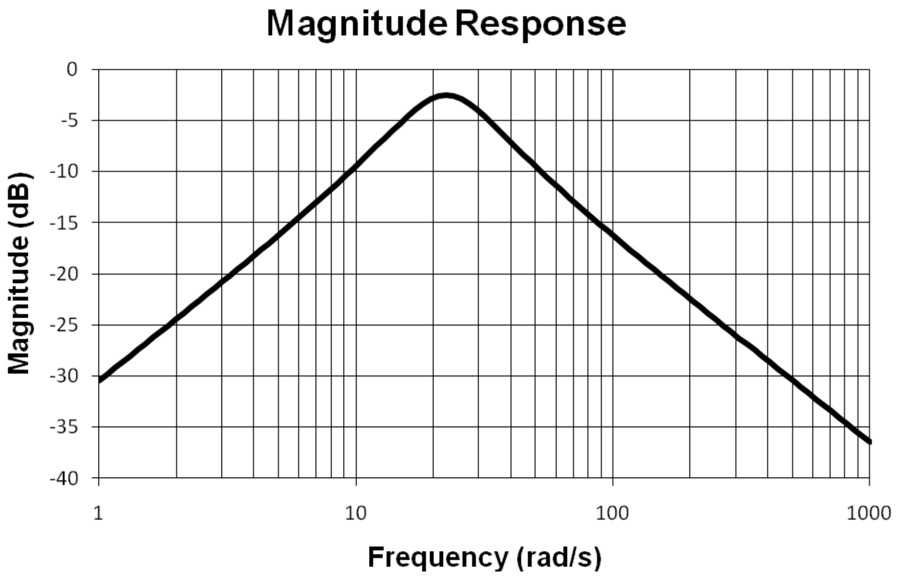
$$\begin{aligned}\mathbf{T}_2(j\omega) &= \frac{15 \cdot j\omega}{(j\omega)^2 + 20 \cdot j\omega + 500} \\ &= \frac{j15\omega}{500 - \omega^2 + j20\omega}\end{aligned}$$

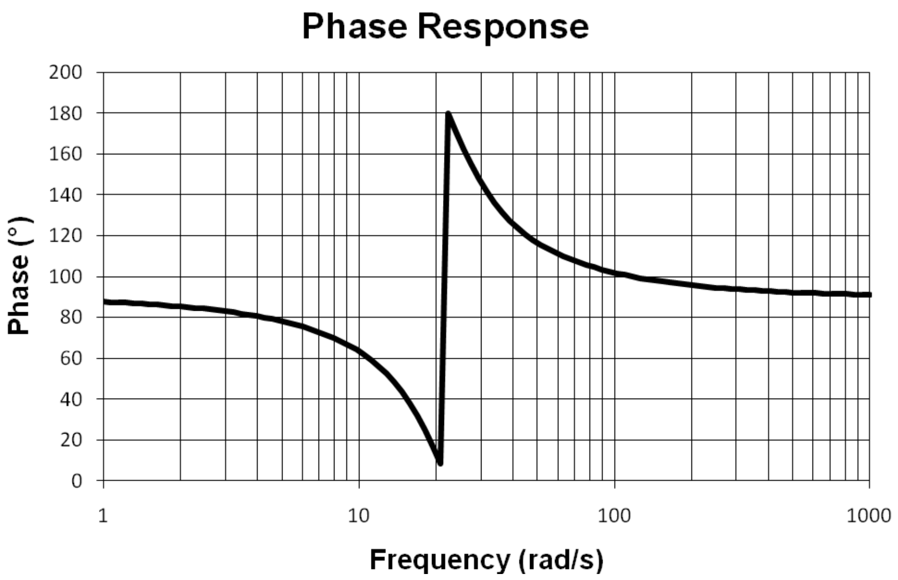
The magnitude and phase responses are, respectively:

$$\left|\mathbf{T}_2(j\omega)\right| = \frac{15\omega}{\sqrt{\left(500 - \omega^2\right)^2 + 400\omega^2}}$$

$$\angle \mathbf{T}_2(j\omega) = 90° - \tan^{-1}\left(\frac{20\omega}{500 - \omega^2}\right)$$

Bode plots of these responses are shown below:





The circuit, with the output taken as $v_2(t)$, is thus a bandpass filter.

24.4 Natural Response

The natural response is, by definition, of a form independent of the forcing function. Thus, we should be able to find the natural response by setting the forcing function to zero. In the example *RC* circuit, we work in the frequency-domain and set $\mathbf{V}_s = \mathbf{0}$. Thus, we have:

$$\begin{aligned}\mathbf{V} &= \frac{1/RC}{\mathbf{s}+1/RC}\mathbf{V}_s \\ &= \frac{1/RC}{\mathbf{s}+1/RC}\mathbf{0}\end{aligned} \tag{24.19}$$

At first glance, it appears as though the response phasor must be zero. This is true for most frequencies, but what happens when the forcing function happens to be at a complex frequency $\mathbf{s} = -1/RC + j0$?

We know that if we excite the circuit at $\mathbf{s} = -1/RC + j0$ with a finite voltage, we get an infinite voltage (the definition of a pole). Mathematically, we have:

$$\mathbf{V} = \frac{\mathbf{V}_s}{\mathbf{0}} = \infty \tag{24.20}$$

So we conclude that if we "excite" the circuit at the same frequency with a zero voltage, that we get a finite response. Mathematically, we have:

$$\begin{aligned}\mathbf{V} &= \frac{\mathbf{0}}{\mathbf{0}} \\ \mathbf{V}\cdot\mathbf{0} &= \mathbf{0} \\ \text{finite}\cdot\mathbf{0} &= \mathbf{0}\end{aligned} \tag{24.21}$$

Thus, the natural response must occur at frequencies corresponding to the poles of the transfer function.

24.16

For the example *RC* circuit, the pole of the transfer function occurs and infinite voltage results when the operating frequency is $\mathbf{s} = -1/RC + j0$. A finite voltage at this frequency thus represents the natural response:

$$\mathbf{V}_n = A \quad \text{at} \quad \mathbf{s} = -1/RC + j0 \tag{24.22}$$

Transforming this natural response to the time-domain:

$$v_n(t) = Ae^{-t/RC} \tag{24.23}$$

For an arbitrary transfer function:

$$\mathbf{T}(\mathbf{s}) = K\frac{(\mathbf{s}-\mathbf{z}_1)(\mathbf{s}-\mathbf{z}_2)\cdots(\mathbf{s}-\mathbf{z}_m)}{(\mathbf{s}-\mathbf{p}_1)(\mathbf{s}-\mathbf{p}_2)\cdots(\mathbf{s}-\mathbf{p}_n)} \tag{24.24}$$

the poles of $\mathbf{T}(\mathbf{s})$ occur at $\mathbf{s} = \mathbf{p}_1, \mathbf{p}_2, \ldots \mathbf{p}_n$, and if the input forcing function is operating at any of these frequencies, an infinite forced response will result. Thus, a finite response at each of these frequencies is possible for the natural response. We thus have the form of the natural response by inspection of the poles of the transfer function. The most general expression is:

$$\begin{aligned} f_n(t) = {} & A_1 e^{p_1 t} + \cdots \\ & + B_1 e^{bt} + B_2 t e^{bt} + B_2 t^2 e^{bt} + \cdots \\ & + C_1 e^{\alpha_1 t}\cos(\omega_{d1} t) + D_2 e^{\alpha_1 t}\sin(\omega_{d1} t) + \cdots \\ & + (E_1 + E_2 t + \cdots)e^{\alpha_2 t}\cos(\omega_{d2} t) + (F_1 + F_2 t + \cdots)e^{\alpha_2 t}\sin(\omega_{d2} t) + \cdots \end{aligned} \tag{24.25}$$

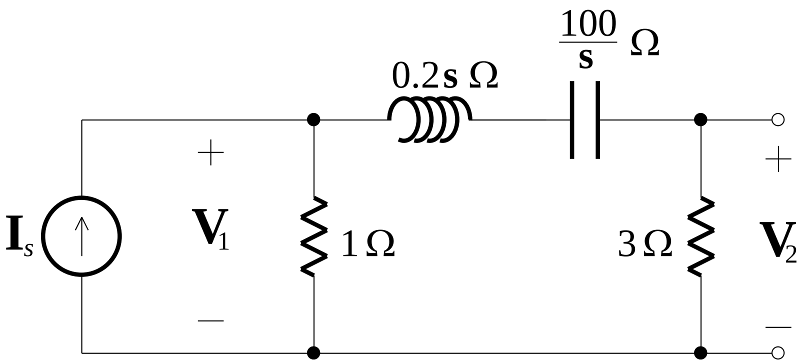
or, in words:

$$\begin{aligned} \text{natural response} = {} & \text{reponse due to real and distinct poles} \\ & + \text{response due to real and repeated poles} \\ & + \text{reponse due to complex and distinct poles} \\ & + \text{response due to complex and repeated poles} \end{aligned} \tag{24.26}$$

The coefficients of the natural response must be evaluated using a knowledge of the forcing function and the initial conditions.

EXAMPLE 24.3 Frequency Response from Transfer Function

Consider the circuit seen previously in the frequency-domain:



The transfer functions were derived as:

$$\frac{\mathbf{V}_1}{\mathbf{I}_s} = \mathbf{T}_1(\mathbf{s}) = \frac{3+0.2\mathbf{s}+100/\mathbf{s}}{4+0.2\mathbf{s}+100/\mathbf{s}} = \frac{\mathbf{s}^2+15\mathbf{s}+500}{\mathbf{s}^2+20\mathbf{s}+500}$$

$$\frac{\mathbf{V}_2}{\mathbf{I}_s} = \mathbf{T}_2(\mathbf{s}) = 3\frac{1}{4+0.2\mathbf{s}+100/\mathbf{s}} = \frac{15\mathbf{s}}{\mathbf{s}^2+20\mathbf{s}+500}$$

Consideration of either transfer function will determine the form of the natural response. First, we find the poles by setting the denominator to zero, and thus obtain the roots by solving the *characteristic* equation:

$$\begin{aligned} \mathbf{s}^2+20\mathbf{s}+500 &= 0 \\ (\mathbf{s}+10)^2+400 &= 0 \\ \mathbf{s} &= -10 \pm j20 \end{aligned}$$

Thus, we have two complex and distinct poles. The *form* of the natural response for either $v_1(t)$ or $v_2(t)$ is thus:

$$v_n(t) = Ae^{-10t}\cos(20t) + Be^{-10t}\sin(20t)$$

To evaluate A and B, we need to know the initial conditions and the forcing function.

Note that the denominators of the two transfer functions in the preceding example are the same. This is not a coincidence. Provided that one portion of a circuit is not separated from the rest, each transfer function will have the same denominator regardless of which voltage or current is chosen as the output variable. This should not be surprising however, for the denominator (via the characteristic equation) determines those values of **s**, called *poles* or *natural frequencies*, that determine the natural response – and the form of the natural response is the same throughout a nonseparated circuit.

If the natural response is desired for a circuit that contains no forcing function, then we can insert any source we like into the circuit, evaluate the transfer function, and then determine the natural response by inspection of the poles.

If the circuit already contains a source, then we are allowed to set it to zero, apply a forcing function in a more convenient location, and then determine the poles from the resulting transfer function.

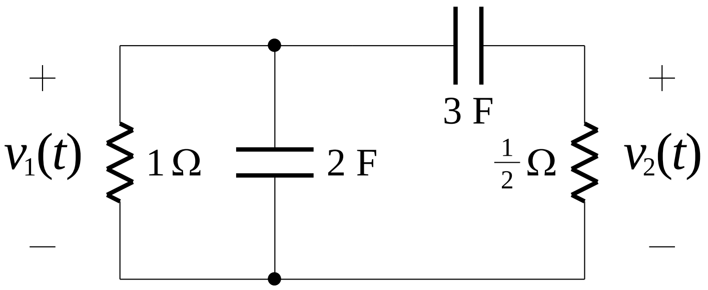
These methods will work since the poles of the transfer function are a characteristic of the circuit only, and not of the forcing function – we will obtain the *same poles* for *any* of the many source locations that are possible – apart from two special cases:

(a) if a forcing function is naively applied in a portion of the circuit that is separated from the response, then the transfer function $\mathbf{T}(\mathbf{s}) = \mathbf{0}$ will result. A case of this type can often be recognised by an inspection of the circuit before the source is installed.

(b) if a forcing function is naively applied in a portion of the circuit so that it *is* the response, then the transfer function $\mathbf{T}(\mathbf{s}) = 1$ will result. A case of this type will occur if a voltage source is placed in parallel with a desired voltage response or a current source is placed in series with a desired current response.

EXAMPLE 24.4 Natural Response from the Transfer Function

Consider the source-free circuit:



We seek expressions for $v_1(t)$ and $v_2(t)$ for $t > 0$, given the initial conditions, $v_1(0) = v_2(0) = 11\ \text{V}$.

Firstly, note that this is **not** a single time constant (STC) circuit, since we cannot reduce the circuit further. We thus need to investigate this circuit by examination of a transfer function.

Let us install a current source $\mathbf{I}_s$ in parallel with the 1Ω resistor, and find the transfer function $\mathbf{T}(\mathbf{s}) = \mathbf{V}_1/\mathbf{I}_s$, which also happens to be the input impedance seen by the current source.

We have:

$$\mathbf{V}_1 = \frac{\mathbf{I}_s}{1 + 2\mathbf{s} + 6\mathbf{s}/(3\mathbf{s}+2)} = \frac{(3\mathbf{s}+2)\mathbf{I}_s}{6\mathbf{s}^2 + 13\mathbf{s} + 2}$$

$$\mathbf{T}(\mathbf{s}) = \frac{\mathbf{V}_1}{\mathbf{I}_s} = \frac{\tfrac{1}{2}(\mathbf{s}+\tfrac{2}{3})}{(\mathbf{s}+2)(\mathbf{s}+\tfrac{1}{6})}$$

Thus, $v_1(t)$ must be of the form:

$$v_1(t) = Ae^{-2t} + Be^{-t/6}$$

24.20

The solution is completed by using the given initial conditions to establish the values of *A* and *B*. Since $v_1(0)$ is given as 11, then:

$$11 = A + B$$

The necessary additional equation is obtained by differentiating v_1 and applying KCL at the bottom node:

$$\left.\frac{dv_1}{dt}\right|_{t=0} = \frac{i_C}{C} = -\frac{11+22}{2} = -2A - \tfrac{1}{6}B$$

Thus, $A = 8$ and $B = 3$, and the desired solution is:

$$v_1(t) = 8e^{-2t} + 3e^{-t/6}$$

The natural frequencies comprising v_2 are the same as those of v_1, and a similar procedure to evaluate the arbitrary constants leads to:

$$v_2(t) = 12e^{-2t} - e^{-t/6}$$

24.5 Complete Response

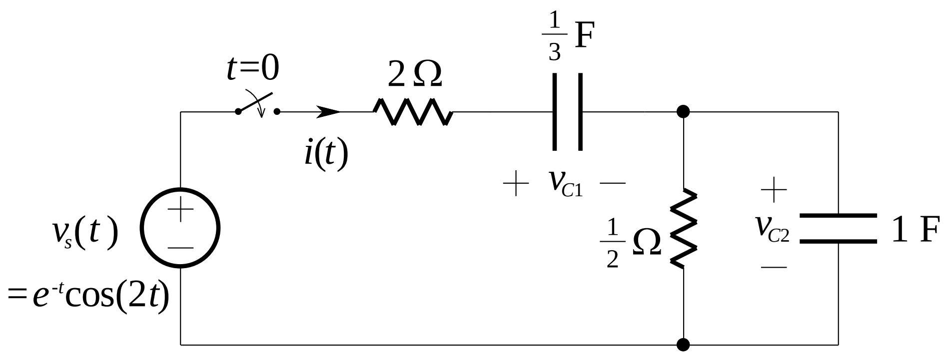
As we know, the complete response is the sum of the forced response plus natural response:

$$\begin{matrix}\text{complete}\\ \text{response}\end{matrix} = \begin{matrix}\text{forced}\\ \text{response}\end{matrix} + \begin{matrix}\text{natural}\\ \text{response}\end{matrix} \tag{24.27}$$

The transfer function can be used to give us both the forced response and natural response, and therefore the complete response!

EXAMPLE 24.5 Complete Response Using the Transfer Function

We wish to find the complete response $i(t)$ of the following circuit:



The switch is in an open position prior to $t=0$, and thus all currents and voltages to the right of the switch are assumed to be zero. At $t=0$ the switch is closed, and the current through the switch is to be found. This response is composed of both a forced response and a natural response:

$$i(t) = i_f(t) + i_n(t)$$

Each may be found through a knowledge of the transfer function, $\mathbf{T}(\mathbf{s}) = \mathbf{I}/\mathbf{V}_s$, which is also the input admittance of the circuit to the right of the switch.

We have:

$$\mathbf{I} = \frac{\mathbf{V}_s}{2 + 3/\mathbf{s} + 1/(\mathbf{s}+2)} = \frac{\mathbf{s}(\mathbf{s}+2)\mathbf{V}_s}{2\mathbf{s}(\mathbf{s}+2) + 3(\mathbf{s}+2) + \mathbf{s}}$$

24.22

After combining and factoring:

$$\mathbf{T}(\mathbf{s}) = \frac{\mathbf{I}}{\mathbf{V}_s} = \frac{\mathbf{s}(\mathbf{s}+2)}{2(\mathbf{s}+1)(\mathbf{s}+3)}$$

In order to find the forced response, the frequency-domain voltage source $\mathbf{V}_s = 1\angle 0°$ at $\mathbf{s} = -1 + j2$ may be multiplied by the transfer function, evaluated at $\mathbf{s} = -1 + j2$:

$$\mathbf{I}_f = \mathbf{T}(\mathbf{s})\big|_{\mathbf{s}=-1+j2} \mathbf{V}_s = \frac{(-1+j2)(1+j2)}{2(j2)(2+j2)} \cdot 1 = \frac{-5}{j8(1+j)}$$

and thus:

$$\mathbf{I}_f = \frac{5\sqrt{2}}{16}\angle 45°$$

Transforming to the time-domain, we have:

$$i_f(t) = \frac{5\sqrt{2}}{16} e^{-t}\cos(2t + 45°)$$

The form of the natural response can be written by inspection of the poles of the transfer function:

$$i_n(t) = Ae^{-t} + Be^{-3t}$$

The complete response is therefore:

$$i(t) = \frac{5\sqrt{2}}{16} e^{-t}\cos(2t + 45°) + Ae^{-t} + Be^{-3t}$$

The solution is completed by using the given initial conditions to establish the values of A and B. Since the voltage across both capacitors is initially zero, the initial source voltage of 1 V must appear across the $2\,\Omega$ resistor. Thus:

$$i(0) = \frac{1}{2} = \frac{5\sqrt{2}}{16}\frac{1}{\sqrt{2}} + A + B$$

Again, it is necessary to differentiate and then to obtain an initial condition for di/dt. From the expression for the complete response, we first find:

$$\left.\frac{di}{dt}\right|_{t=0} = \frac{5\sqrt{2}}{16}\left(-\frac{2}{\sqrt{2}}-\frac{1}{\sqrt{2}}\right) - A - 3B$$
$$= -\frac{15}{16} - A - 3B$$

The initial value of this rate of change is obtained by analysing the circuit. However, those rates of change which are most easily found are the derivatives of the capacitor voltages, since $i = C\,dv/dt$, and the initial values of the capacitor currents should not be too difficult to find. KVL around the circuit gives:

$$2i = v_s - v_{C1} - v_{C2}$$

Dividing by 2 and taking the derivative:

$$\frac{di}{dt} = \frac{1}{2}\frac{dv_s}{dt} - \frac{1}{2}\frac{dv_{C1}}{dt} - \frac{1}{2}\frac{dv_{C2}}{dt}$$

The first term on the right-hand side is obtained by differentiation of the source function and evaluation at $t = 0$, the result is $-1/2\ \text{As}^{-1}$.The second term is numerically equal to $-3/2$ of the initial current through the $1/3\ \text{F}$ capacitor, or $-3/4\ \text{As}^{-1}$. Similarly, the last term is $-1/4\ \text{As}^{-1}$. Thus:

$$\left.\frac{di}{dt}\right|_{t=0} = -\frac{1}{2} - \frac{3}{4} - \frac{1}{4} = -\frac{3}{2}$$

We may now use our two equations in A and B to determine the unknown coefficients of the natural response:

$$A = 0 \quad \text{and} \quad B = 3/16$$

The complete response is therefore:

$$i(t) = 5/16\,\sqrt{2}e^{-t}\cos(2t + 45°) + 3/16\,e^{-3t}$$

24.6 Summary

- A *transfer function* is the ratio of the desired forced response to the forcing function, using phasor notation and the complex frequency $\mathbf{s}$. It is usually designated $\mathbf{T}(\mathbf{s})$.

- A transfer function is always the ratio of two polynomials in $\mathbf{s}$. It can be written as:

$$\mathbf{T}(\mathbf{s}) = K\frac{(\mathbf{s}-\mathbf{z}_1)(\mathbf{s}-\mathbf{z}_2)\cdots(\mathbf{s}-\mathbf{z}_m)}{(\mathbf{s}-\mathbf{p}_1)(\mathbf{s}-\mathbf{p}_2)\cdots(\mathbf{s}-\mathbf{p}_n)}$$

  where $\mathbf{z}_i$ is termed a *zero* and $\mathbf{p}_j$ is termed a *pole*.

- A pole-zero plot provides an alternative description of a transfer function (to within an arbitrary constant *K*) as it specifies all zeros and poles.

- The transfer function can be derived from the differential equation describing a circuit by replacing the *D* operator with $\mathbf{s}$.

- The transfer function can be derived by working directly in the frequency-domain with the concept of generalized impedance (R, $\mathbf{s}L$ and $1/\mathbf{s}C$).

- The transfer function can be used to determine the forced response phasor – by evaluating it at the complex frequency of the forcing function and multiplying the resultant complex number by the input phasor.

- The transfer function can be used to determine the frequency response of a circuit by letting $\mathbf{s} = j\omega$:

$$\mathbf{T}(\mathbf{s})\big|_{\mathbf{s}=j\omega} = \mathbf{T}(j\omega) = \text{frequency response}$$

- The transfer function can be used to determine the natural response of a circuit, since the poles of a circuit determine the form of its natural response.

- The transfer function is a complete description of a circuit – from it we can derive the forced response, the natural response, the complete response, and the frequency response!

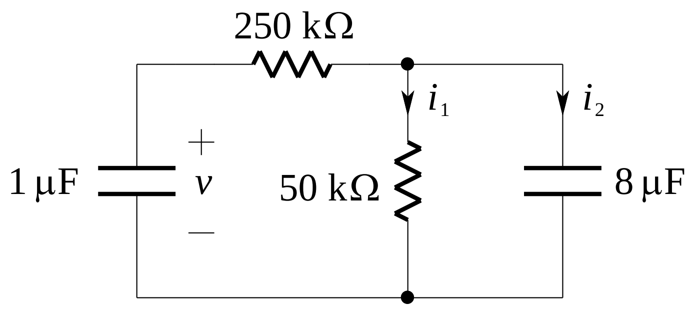
24.7 References

Hayt, W. & Kemmerly, J.: *Engineering Circuit Analysis*, 3rd *Ed*., McGraw-Hill, 1984.

Exercises

1.

The capacitors in the circuit shown below carry charge at $t = 0$.

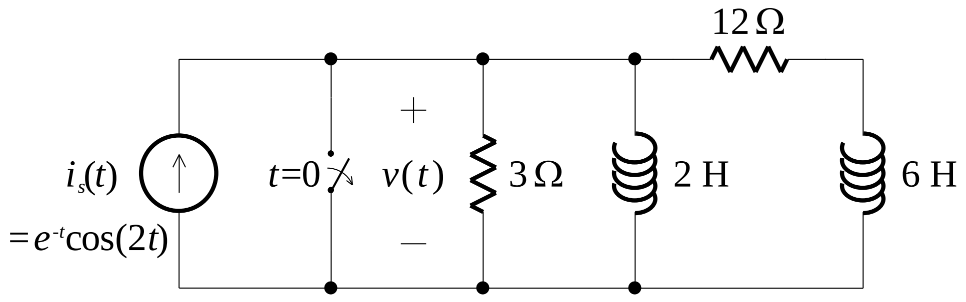


Find the frequencies present in the specified response for $t > 0$:

(a) $i_{1n}(t)$ (b) $i_{2n}(t)$ (c) $v_n(t)$

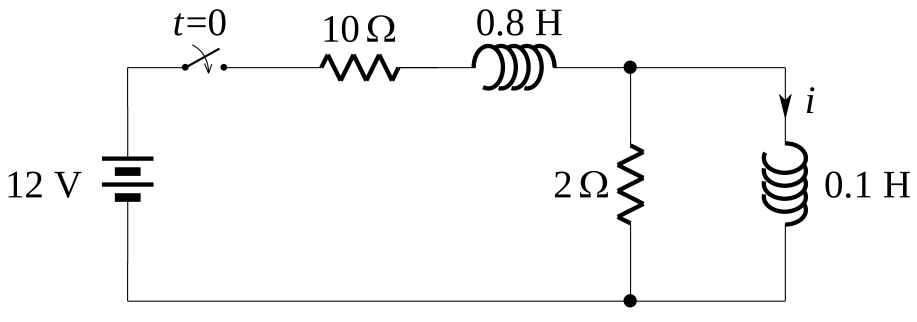
2.

Find $v(t)$ for all values of time in the circuit shown below:



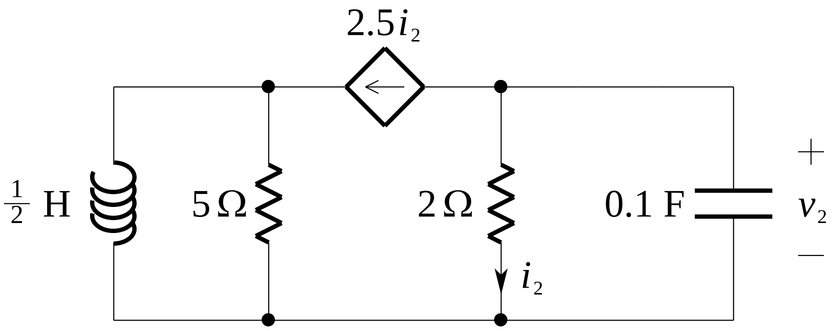
3.

Find $i(t)$ for all values of time in the circuit shown below:



4.

Consider the circuit shown below:



Determine:

(a) The transfer function $\mathbf{V}_2/\mathbf{I}_{s1}$, if $\mathbf{I}_{s1}$ is in parallel with the inductor with its arrow directed upward.

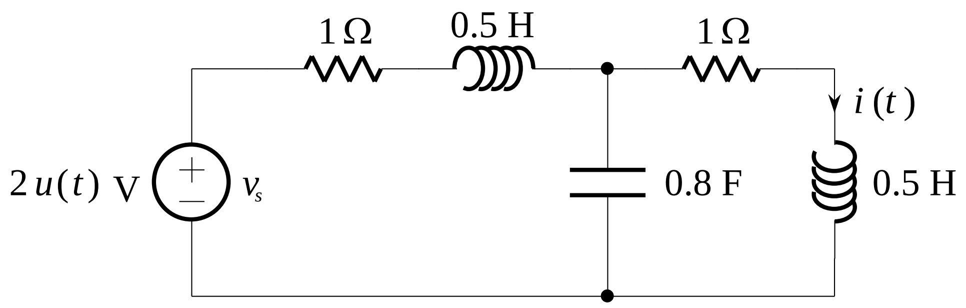
(b) The transfer function $\mathbf{V}_2/\mathbf{I}_{s2}$, if $\mathbf{I}_{s2}$ is in parallel with the capacitor with its arrow directed upward.

(c) The transfer function $\mathbf{V}_2/\mathbf{V}_{s1}$, if $\mathbf{V}_{s1}$ is in series with the inductor with its positive reference on top.

(d) Specify the form of the natural response $v_{2n}(t)$.

5.

Consider the following circuit:



(a) Write the transfer function for $\mathbf{I}/\mathbf{V}_s$ and find the three natural frequencies (possibly complex) associated with the response $i(t)$.

Hint: There is a pole at $\mathbf{s} = -2$.

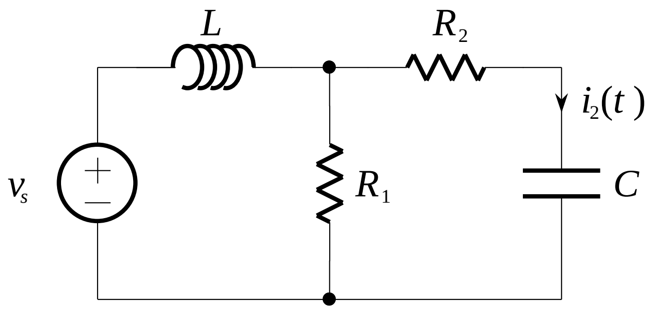
(b) If $i(t)$ is represented as a forced response plus a natural response, $i_f(t)+i_n(t)$, find $i_f(t)$ by working in the frequency-domain.

(c) Write down the form of the natural response, $i_n(t)$.

(d) Determine the complete response $i(t)$.

6.

There is no initial energy stored in the following circuit:



(a) Transform the circuit into the **s**-domain and formulate mesh-current equations.

(b) Solve the mesh equations to establish the transfer function $\mathbf{I}_2/\mathbf{V}_s$.

(c) Find $i_2(t)$ if $v_s(t)=12u(t)\,\text{V}$, $R_1=1\,\text{k}\Omega$, $R_2=2\,\text{k}\Omega$, $L=4\,\text{H}$ and $C=500\text{ nF}$.

Pierre Simon de Laplace (1749-1827)

The application of mathematics to problems in physics became a primary task in the century after Newton. Foremost among a host of brilliant mathematical thinkers was the Frenchman Laplace. He was a powerful and influential figure, contributing to the areas of celestial mechanics, cosmology and probability.

Laplace was the son of a farmer of moderate means, and while at the local military school, his uncle (a priest) recognised his exceptional mathematical talent. At sixteen, he began to study at the University of Caen. Two years later he travelled to Paris, where he gained the attention of the great mathematician and philosopher Jean Le Rond d'Alembert by sending him a paper on the principles of mechanics. His genius was immediately recognised, and Laplace became a professor of mathematics.

He began producing a steady stream of remarkable mathematical papers. Not only did he make major contributions to difference equations and differential equations but he examined applications to mathematical astronomy and to the theory of probability, two major topics which he would work on throughout his life. His work on mathematical astronomy before his election to the Académie des Sciences included work on the inclination of planetary orbits, a study of how planets were perturbed by their moons, and in a paper read to the Academy on 27 November 1771 he made a study of the motions of the planets which would be the first step towards his later masterpiece on the stability of the solar system.

In 1773, before the Academy of Sciences, Laplace proposed a model of the solar system which showed how perturbations in a planet's orbit would not change its distance from the sun. For the next decade, Laplace contributed a stream of papers on planetary motion, clearing up discrepancies in the orbit's of Jupiter and Saturn, he showed how the moon accelerates as a function of the Earth's orbit, he introduced a new calculus for discovering the motion of celestial bodies, and even a new means of computing planetary orbits which led to astronomical tables of improved accuracy.

The 1780s were the period in which Laplace produced the depth of results which have made him one of the most important and influential scientists that the world has seen. Laplace let it be known widely that he considered himself the best mathematician in France. The effect on his colleagues would have been only mildly eased by the fact that Laplace was right!

In 1784 Laplace was appointed as examiner at the Royal Artillery Corps, and in this role in 1785, he examined and passed the 16 year old Napoleon Bonaparte.

In 1785, he introduced a field equation in spherical harmonics, now known as Laplace's equation, which is found to be applicable to a great deal of phenomena, including gravitation, the propagation of sound, light, heat, water, electricity and magnetism.

"Your Highness, I have no need of this hypothesis. "
- Laplace, to Napoleon on why his works on celestial mechanics make no mention of God.

Laplace presented his famous nebular hypothesis in 1796 in *Exposition du systeme du monde*, which viewed the solar system as originating from the contracting and cooling of a large, flattened, and slowly rotating cloud of incandescent gas. The *Exposition* consisted of five books: the first was on the apparent motions of the celestial bodies, the motion of the sea, and also atmospheric refraction; the second was on the actual motion of the celestial bodies; the third was on force and momentum; the fourth was on the theory of universal gravitation and included an account of the motion of the sea and the shape of the Earth; the final book gave an historical account of astronomy and included his famous nebular hypothesis which even predicted black holes. Laplace stated his philosophy of science in the *Exposition*:

> *If man were restricted to collecting facts the sciences were only a sterile nomenclature and he would never have known the great laws of nature. It is in comparing the phenomena with each other, in seeking to grasp their relationships, that he is led to discover these laws...*

Exposition du systeme du monde was written as a non-mathematical introduction to Laplace's most important work. Laplace had already discovered the invariability of planetary mean motions. In 1786 he had proved that the eccentricities and inclinations of planetary orbits to each other always remain small, constant, and self-correcting. These and many of his earlier results formed the basis for his great work the *Traité du Mécanique Céleste* published in 5 volumes, the first two in 1799.

The first volume of the *Mécanique Céleste* is divided into two books, the first on general laws of equilibrium and motion of solids and also fluids, while the second book is on the law of universal gravitation and the motions of the centres of gravity of the bodies in the solar system. The main mathematical approach was the setting up of differential equations and solving them to describe the resulting motions. The second volume deals with mechanics applied to a study of the planets. In it Laplace included a study of the shape of the Earth which included a discussion of data obtained from several different expeditions, and Laplace applied his theory of errors to the results.

In 1812 he published the influential study of probability, *Théorie analytique des probabilités*. The work consists of two books. The first book studies generating functions and also approximations to various expressions occurring in probability theory. The second book contains Laplace's definition of probability, Bayes's rule (named by Poincaré many years later), and remarks on mathematical expectation. The book continues with methods of finding probabilities of compound events when the probabilities of their simple components are known, then a discussion of the method of least squares, and inverse probability. Applications to mortality, life expectancy, length of marriages and probability in legal matters are given.

After the publication of the fourth volume of the *Mécanique Céleste*, Laplace continued to apply his ideas of physics to other problems such as capillary action (1806-07), double refraction (1809), the velocity of sound (1816), the theory of heat, in particular the shape and rotation of the cooling Earth (1817-1820), and elastic fluids (1821).

Many original documents concerning his life have been lost, and gaps in his biography have been filled by myth. Some papers were lost in a fire that destroyed the chateau of a descendant, and others went up in flames when Allied forces bombarded Caen during WWII.

Laplace died on 5 March, 1827 at his home outside Paris.

25 Sensor Signal Conditioning

Contents

25.2

Introduction

Sensors translate a physical quantity to an electrical quantity

A *sensor* is a device that receives a signal or stimulus and responds with an electrical signal. Sensors and their associated circuits are used to measure various physical properties such as temperature, force, pressure, flow, position, light intensity, etc. These properties act as the stimulus to the sensor, and the sensor output is conditioned and processed to provide the corresponding measurement of the physical property.

Sensors which measure different properties may have the same type of electrical output. For example, a Resistance Temperature Detector (RTD) is a variable resistance, as is a resistive strain gauge. Both RTDs and strain gauges are often placed in bridge circuits, and the conditioning circuits are therefore quite similar. Therefore bridges and their conditioning circuits will be looked at in detail.

Active sensors require an external source of excitation. Examples include RTDs and strain gauges.

Passive or *self-generating* sensors do not require external power. Examples include thermocouples and photodiodes.

The full-scale outputs of most sensors are relatively small voltages, currents, or resistance changes, and therefore their outputs must be properly conditioned before further analog or digital processing can occur. Amplification, level translation, galvanic isolation, impedance transformation, linearization and filtering are fundamental signal-conditioning functions that may be required.

25.1 Sensors

There are many types of sensors – we will briefly look at those which lend themselves to measurement systems, data acquisition systems and process control systems.

Some typical sensors and their output formats are shown in the table below:

Some typical sensors and their output formats

Property	Sensor	Active / Passive	Output
Temperature	Thermocouple Silicon RTD Thermistor	Passive Active Active Active	Voltage Voltage / Current Resistance Resistance
Force / Pressure	Strain Gauge Piezoelectric	Active Passive	Resistance Voltage
Acceleration	Accelerometer	Active	Capacitance
Position	Linear Variable Differential Transformer (LVDT)	Active	AC Voltage
Light Intensity	Photodiode	Passive	Current

Table 25.1 – Typical Sensors

25.2 Process Control Systems

A common application of sensors is within process control systems. One example would be control of a physical property, such as temperature. A sample block diagram of how this might be implemented is given below:

A typical process control system, showing where the sensor fits into the overall system

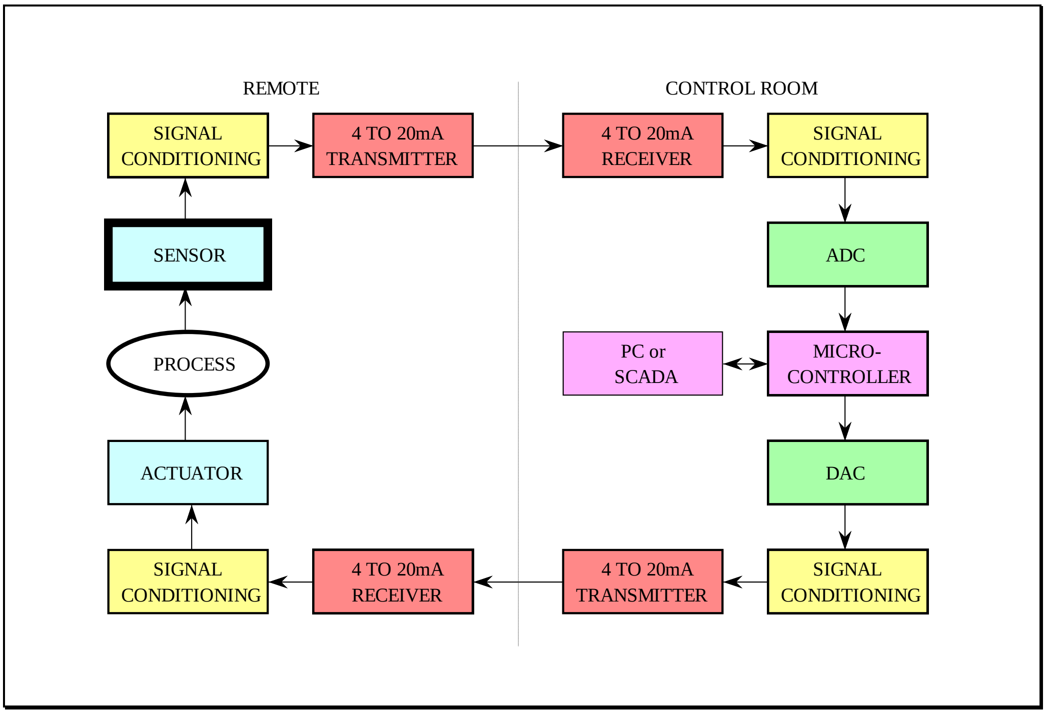


Figure 25.1 – A Typical Process Control System

In this system, an output from a temperature sensor is conditioned, transmitted over some distance, received, and then digitized by an analog-to-digital converter (ADC). The microcontroller or host computer determines if the temperature is above or below the desired value, and outputs a digital value to the digital-to-analog converter (DAC). The DAC output is conditioned and drives the remotely located *actuator*, in this case a heater. The interface between the control room and the remote process is via an "industry standard" 4-20 mA current loop.

25.3 Programmable Logic Controllers

Many industrial processes are controlled at the "remote" end, with the microcontroller taking the form of a Programmable Logic Controller, (PLC). A PLC is a digital computer used for automation of processes, such as control of machinery on factory assembly lines, chemical processes, etc. Unlike general-purpose computers, the PLC is designed for multiple input and output arrangements, extended temperature ranges, immunity to electrical noise, and resistance to vibration and impact. They are generally programmed with proprietary software from the PLC vendor.

A Programmable Logic Controller (PLC)

Figure 25.2 – A Programmable Logic Controller (PLC)

PLCs communicate with a Supervisory Control and Data Acquisition (SCADA) system using "industrial Ethernet" or a vendor proprietary protocol. The control room has a human interface to the SCADA system.

A process control scheme that uses PLCs and a SCADA system

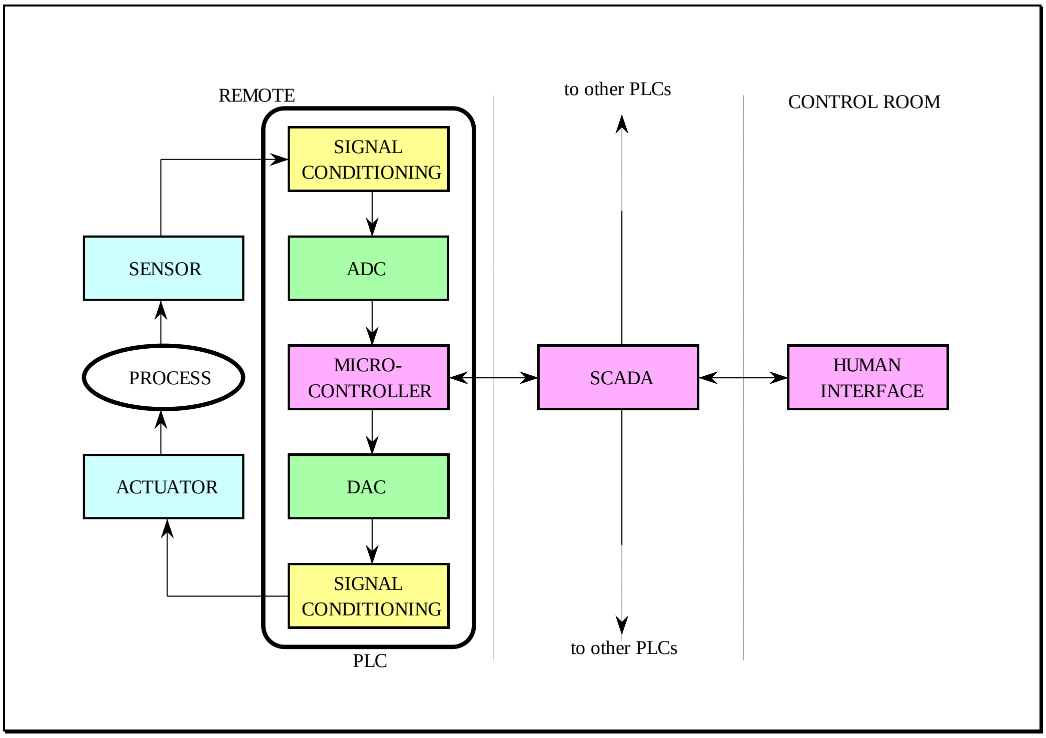


Figure 25.3 – A Process Control System with PLCs and SCADA

25.4 Smart Transducers

By including the microcontroller (with integrated ADC and DAC), the sensor, and the actuator into one device, a "smart transducer" can be implemented with self-contained calibration and linearization features, among others.

A smart transducer

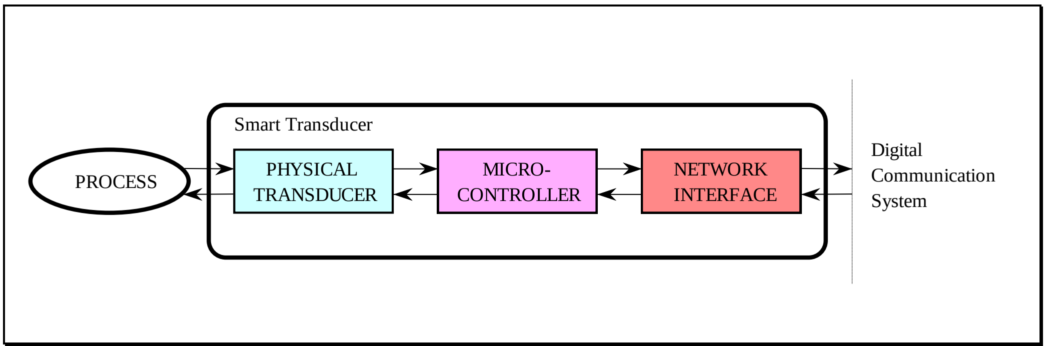
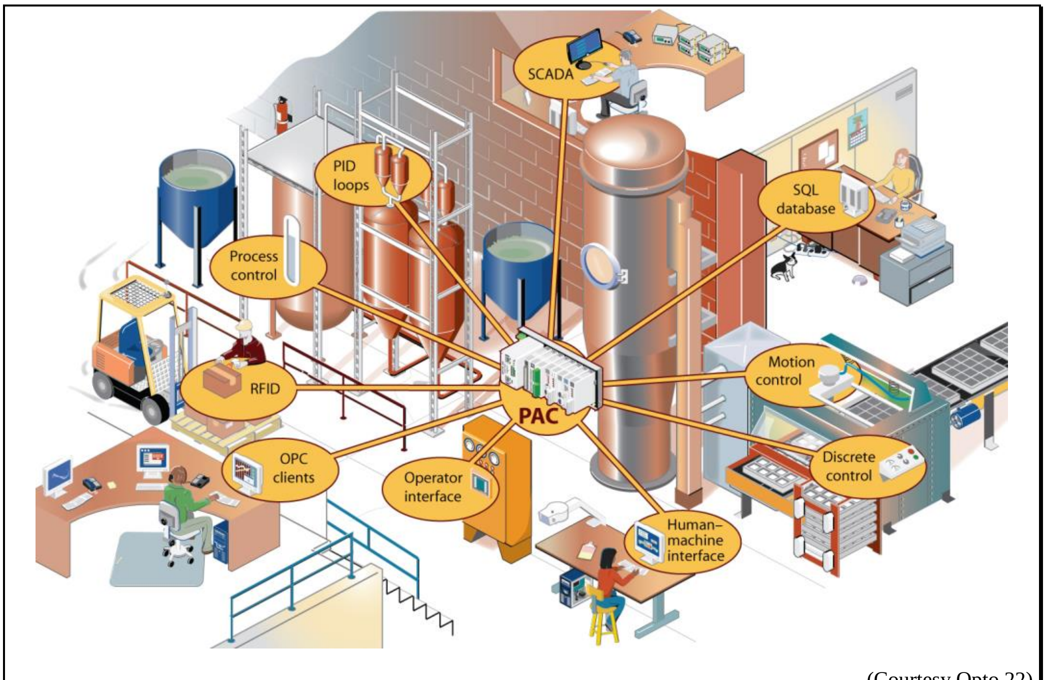


Figure 25.4 – A Smart Transducer

25.5 Programmable Automation Controllers

Modern industrial automation is starting to use a new device known as a Programmable Automation Controller (PAC). PACs are used to interface with simple sensors and actuators, just like a PLC, but they also have advanced control features, network connectivity, device interoperability and enterprise data integration capabilities. PACs are multifunctional, handling the digital, analog, and serial signal types common in all types of industrial applications. The same hardware can be used for data collection, remote monitoring, process control and discrete and hybrid manufacturing. PACs also use standard IT components and protocols, with Ethernet and TCP/IP being very common.

A PAC in a modern industrial application



(Courtesy Opto 22)

Figure 25.5 – A PAC in a Modern Industrial Application

In the example shown above, the PAC is operating in multiple domains to monitor and manage a production line, a chemical process, a test bench, and shipping activities. To do so, the PAC must simultaneously manage analog values such as temperatures and pressures; digital on/off states for valves, switches, and indicators; and serial data from inventory tracking and test equipment. At the same time, the PAC is exchanging data with an OLE for Process Control (OPC) server, an operator interface, and a Structured Query Language (SQL) database. Simultaneously handling these tasks without the need for additional processors or “middleware” is a hallmark of a PAC.

25.6 Bridge Circuits

Resistive elements are some of the most common sensors. They are inexpensive, and relatively easy to interface with signal-conditioning circuits. Resistive elements can be made sensitive to temperature, strain (by pressure or by flex) and light. Using these basic elements, many complex physical phenomena can be measured.

Resistive elements form the basis for many types of physical measurements

Sensor element resistance can range from less than 100 Ω to several hundred $k\Omega$, depending on the sensor design and the physical environment to be measured. The table below shows the wide range of sensor resistances used in bridge circuits.

Sensor	Resistance Range
Strain Gauges	120 Ω, 350 Ω, 3500 Ω
Weigh-Scale Load Cells	350 Ω - 3500 Ω
Pressure Sensors	350 Ω - 3500 Ω
Relative Humidity	100 $k\Omega$ - 10 $M\Omega$
Resistance Temperature Devices (RTDs)	100 Ω, 1000 Ω
Thermistors	100 Ω - 10 $M\Omega$

Table 25.2 – Sensor Resistance Ranges

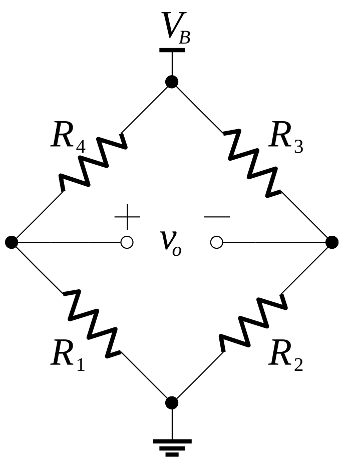
The change in a resistive element is small and therefore accurate measurement is essential

Resistive sensors produce relatively small percentage changes in resistance in response to a change in a physical variable. For example, RTDs and strain gauges present a significant measurement challenge because the typical change in resistance over the entire operating range may be less than 1% of the nominal resistance value. Accurately measuring small resistance changes is therefore critical when applying resistive sensors.

Rather than trying to measure the sensor output directly, a *bridge* circuit is often used. The ability to balance the bridge initially (and zero the output, v_o) is a significant advantage of the bridge, since it is much easier to measure small changes in voltage Δv_o from a null voltage than from an elevated voltage v_o, which may be as much as 1000 times greater than Δv_o.

A *resistance bridge*, or *Wheatstone bridge*, shown in the figure below, is used to measure small resistance changes accurately.

The Wheatstone bridge is the basis of many types of sensor



The Wheatstone bridge:

$$v_o = \left(\frac{R_1}{R_1 + R_4} - \frac{R_2}{R_2 + R_3} \right) V_B$$

At balance,

$$v_o = 0 \text{ if } \frac{R_1}{R_4} = \frac{R_2}{R_3}$$

Figure 25.6 – The Wheatstone Bridge

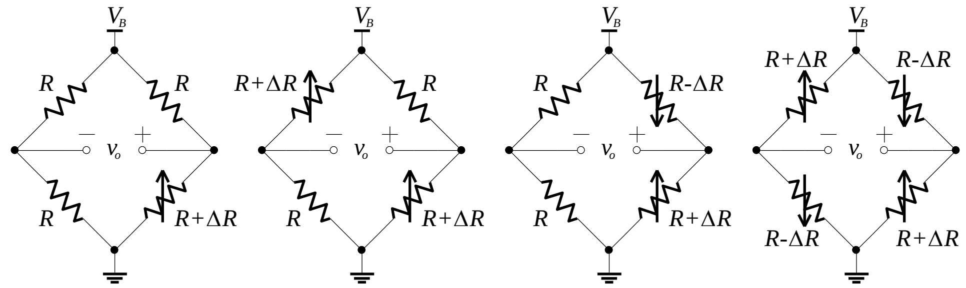
It consists of four resistors connected to form a quadrilateral, a source of excitation voltage V_B (or, alternately, a current) connected across one of the diagonals, and a voltage detector connected across the other diagonal. The detector measures the difference between the outputs of the two voltage dividers connected across V_B. The general formula for the output v_o is:

$$v_o = \left(\frac{R_1}{R_1 + R_4} - \frac{R_2}{R_2 + R_3} \right) V_B \quad (25.1)$$

For sensor applications, the deviation of one or more of the resistors in a bridge from an initial value is measured as an indication of the change in the measured variable. In this case, the output voltage change is an indication of the resistance change. Since very small resistance changes are common, the output voltage change may be as small as tens of millivolts, even with an excitation voltage of $V_B = 10\,\text{V}$ (a typical value).

In many bridge applications, there may not just be a single variable element, but two, or even four elements, all of which may vary. The figure below shows those bridges most commonly suited for sensor applications:

Bridges that are commonly suited for sensor applications



	(A)	(B)	(C)	(D)
v_o:	$\dfrac{V_B}{4}\left[\dfrac{\Delta R}{R+\frac{\Delta R}{2}}\right]$	$\dfrac{V_B}{2}\left[\dfrac{\Delta R}{R+\frac{\Delta R}{2}}\right]$	$\dfrac{V_B}{2}\left[\dfrac{\Delta R}{R}\right]$	$V_B\left[\dfrac{\Delta R}{R}\right]$
	(A) Single-Element Varying	(B) Two-Element Varying (1)	(C) Two-Element Varying (2)	(D) All-Element Varying

Figure 25.7 – Bridges for Sensors

Note that since the bridge output is always directly proportional to V_B, the measurement accuracy can be no better than that of the accuracy of the excitation voltage.

In each case, the value of the fixed bridge resistor R is chosen to be equal to the nominal value of the variable resistor(s). The deviation of the variable resistor(s) about the nominal value is assumed to be proportional to the quantity being measured, such as strain (in the case of a strain gauge), or temperature (in the case of an RTD).

The single-element varying bridge is most suited to temperature sensing

The *single-element varying* bridge (A) is most suited to temperature sensing using RTDs or thermistors. All the resistances are nominally equal, but one of them (the sensor) is variable by an amount ΔR. As the equation indicates, the relationship between the bridge output and ΔR is not linear. Also, in practice, most sensors themselves will exhibit a certain specified amount of nonlinearity, which must be taken into account. Software calibration is used to remove the linearity error in digital systems.

The *two-element varying* bridge (C) requires two identical elements that vary in *opposite* directions. For example, this could correspond to two identical strain gauges: one mounted on top of a flexing surface, and one on the bottom. This configuration is linear, and the terms $R+\Delta R$ and $R-\Delta R$ can be viewed as two sections of a linear potentiometer.

The *all-element varying* bridge (D) produces the most signal for a given resistance change, and is inherently linear. It is also an "industry-standard" configuration for load cells constructed from four identical strain gauges.

Load cells – bridge elements are "strain gauges" arranged in a particular orientation

Figure 25.8 – Load Cells

25.6.1 Bridge Design Issues

- Selecting the configuration (1, 2, 4-element varying)
- Selection of voltage or current excitation
- Stability of excitation or ratiometric operation
- Bridge sensitivity: Full-Scale Output / Excitation
  1 mV/V to 10 mV/V typical
- Full-scale bridge outputs
  10 mV to 100 mV typical
- Precision, low-noise amplification / conditioning techniques required
- Linearization techniques may be required
- Remote sensors present challenges

Regardless of the absolute level, the stability of the excitation voltage or current directly affects the accuracy of the bridge output. Therefore stable references and / or *ratiometric* drive techniques are required to maintain highest accuracy.

Ratiometric refers to the use of the bridge drive voltage of a voltage-driven bridge as the reference input to the ADC that digitizes the amplified bridge output voltage.

25.6.2 Amplifying and Linearizing Bridge Outputs

The output of a single-element varying bridge may be amplified with an instrumentation amplifier (in-amp):

The output of a single-element varying bridge is normally amplified with an in-amp

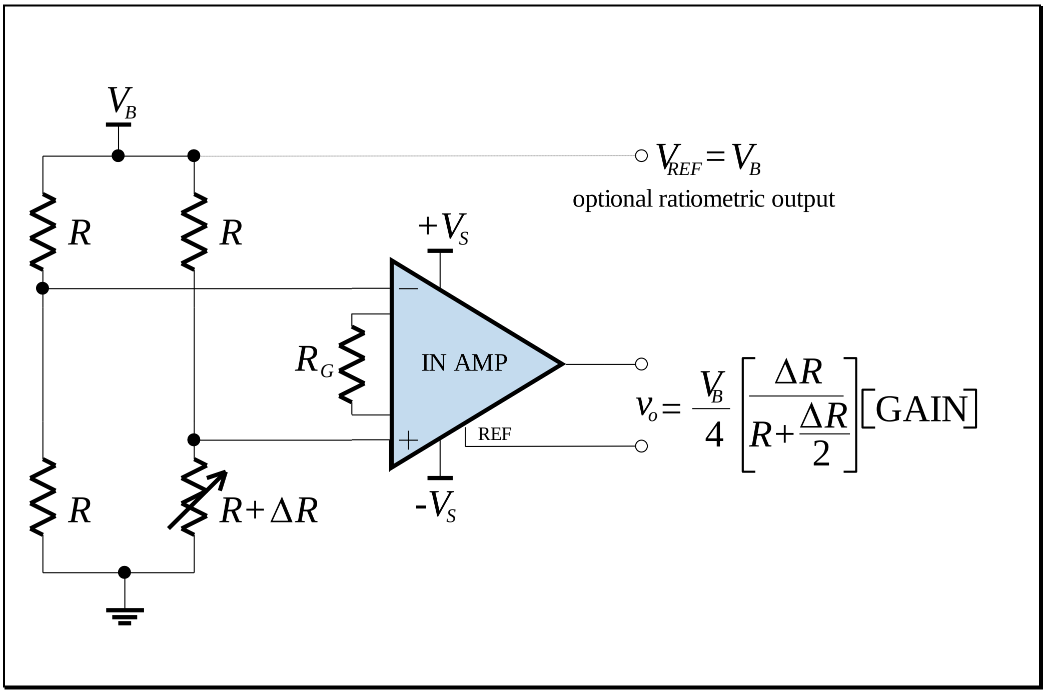


Figure 25.9 – Bridge Output Using an In-Amp

The in-amp provides a large and accurate gain that is set with a single resistor, R_G. The in-amp also provides dual, high-impedance loading to the bridge nodes – it does not unbalance or load the bridge. Using modern in-amps with gain ranging from 10-1000, excellent common-mode rejection and gain accuracy can be achieved with this circuit.

However, due to the intrinsic characteristics of the bridge, the output is still nonlinear. In a system where the output of the in-amp is digitized using an ADC and fed into a microcontroller, this nonlinearity can be corrected in software.

The bridge in this example is voltage driven, by the voltage V_B. This voltage can optionally be used for an ADC reference voltage, in which case it is an additional output of the circuit, V_{REF}.

An analog circuit for linearizing a single-element varying bridge

An analog circuit for linearizing a single-element bridge is shown below:

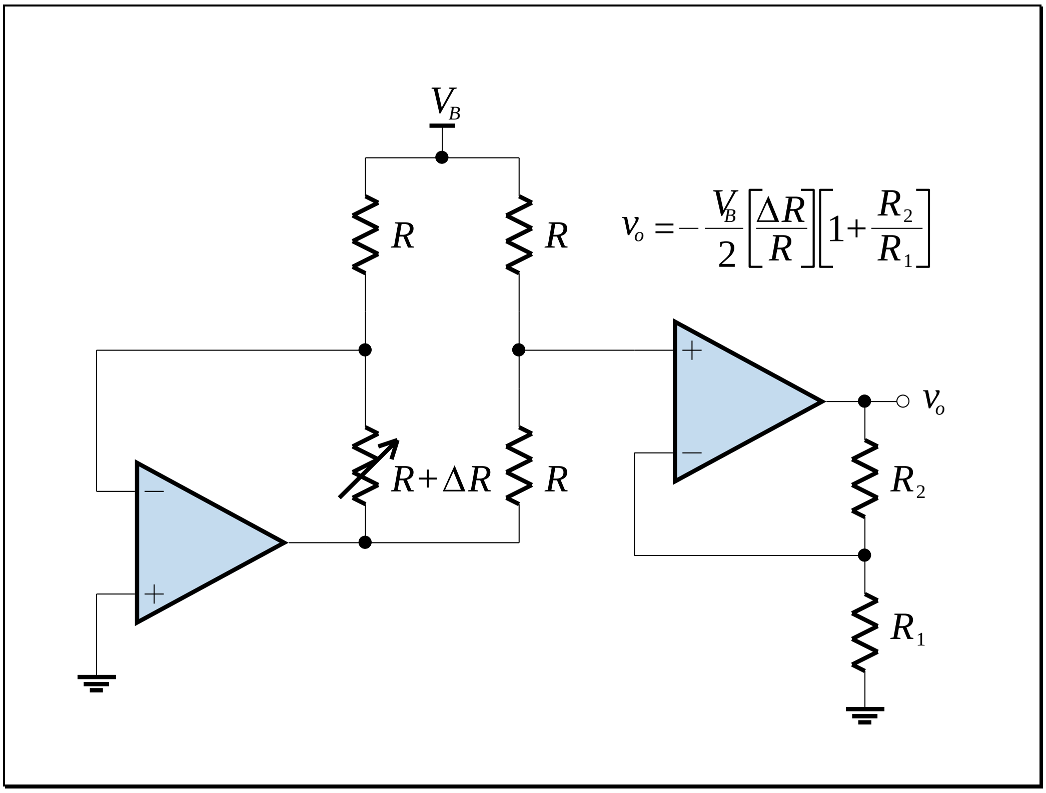


Figure 25.10 – Linearizing a One-Element Varying Bridge

The top node of the bridge is excited by the voltage V_B. The bottom of the bridge is driven in complementary fashion by the left op-amp, which maintains a constant current of V_B/R in the varying resistance element, $R+\Delta R$, which is the mechanism for linearity improvement. Also, the bridge left-side centre node is "ground-referenced" by the op-amp, making this configuration suppress common-mode voltages.

The output signal is taken from the right-hand leg of the bridge, and is amplified by a second op-amp connected as a noninverting gain stage.

The circuit requires two op-amps operating on dual supplies. In addition, paired resistors R_1 and R_2 must be ratio matched and stable for overall accurate and stable gain. The circuit is practical if a dual precision op-amp (with high gain, low offset / noise and high stability) is used.

A circuit for linearizing a two-element bridge is shown below:

An analog circuit for linearizing a two-element varying bridge

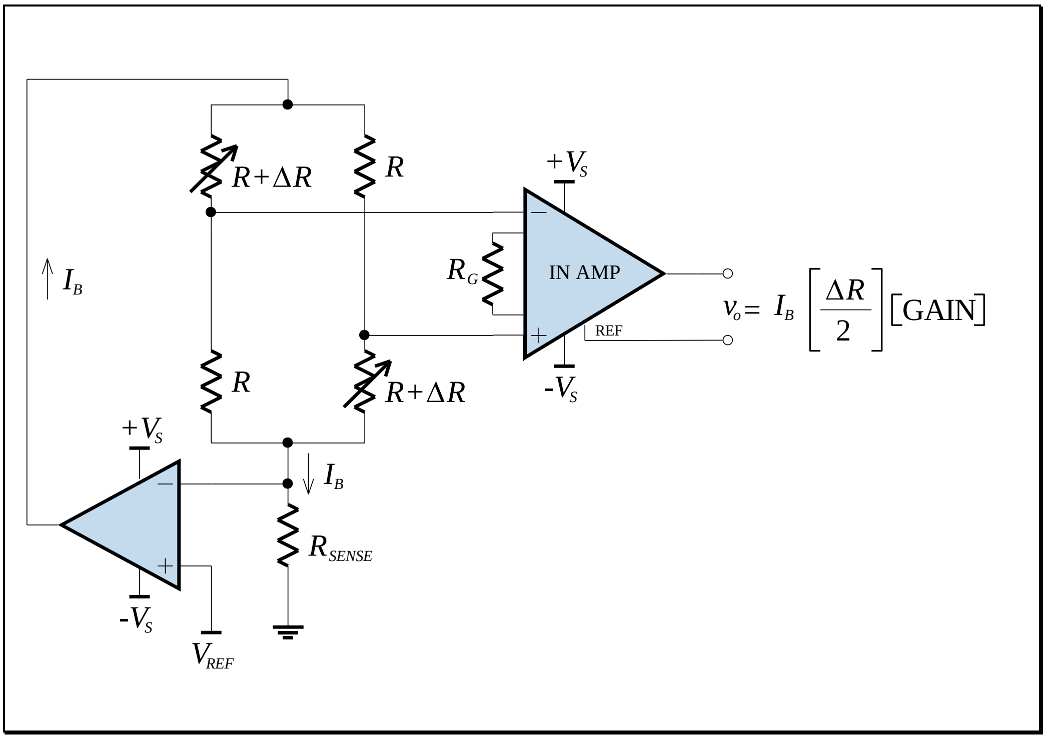


Figure 25.11 – Linearizing a Two-Element Varying Bridge

This circuit uses an op-amp, a sense resistor, and a voltage reference, set up in a feedback loop containing the sensing bridge. The net effect of the loop is to maintain a constant current through the bridge of $I_B = V_{REF}/R_{SENSE}$. The current through each leg of the bridge remains constant $(I_B/2)$ as the resistance changes. Therefore the output is a linear function of ΔR. An in-amp provides the additional gain. If ratiometric operation of an ADC is desired, the V_{REF} voltage can be used to drive the ADC.

25.6.3 Driving Remote Bridges

Wiring resistance is a big problem for remotely located bridges

Wiring resistance and noise pickup are the biggest problems associated with remotely located bridges. The figure below shows a $350\ \Omega$ strain gauge which is connected to the rest of the bridge circuit by 30 m of twisted pair copper wire. The temperature coefficient of the copper wire is 0.385 %/°C. The figure shows nominal resistor values at $25\ ^\circ\mathrm{C}$.

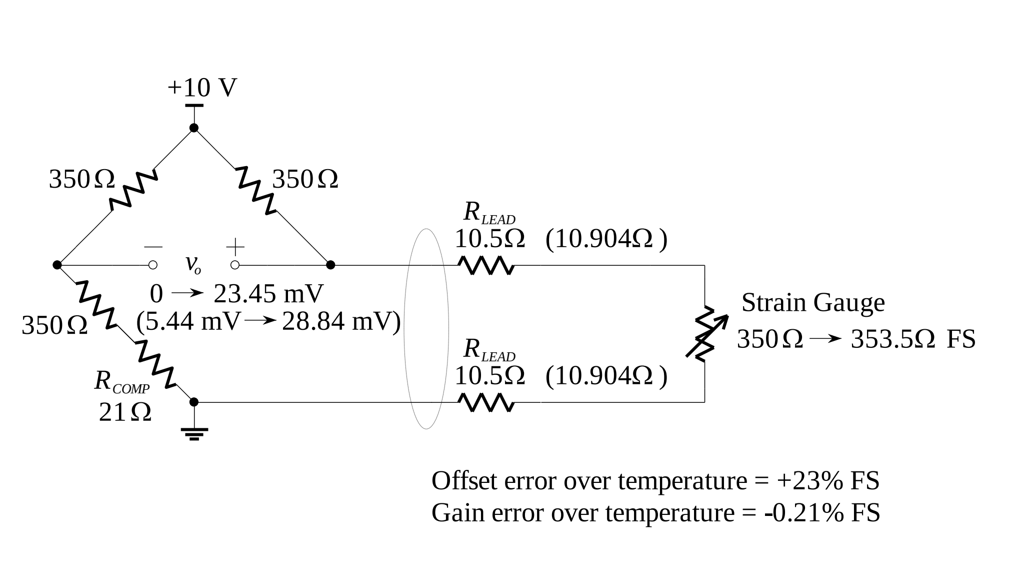
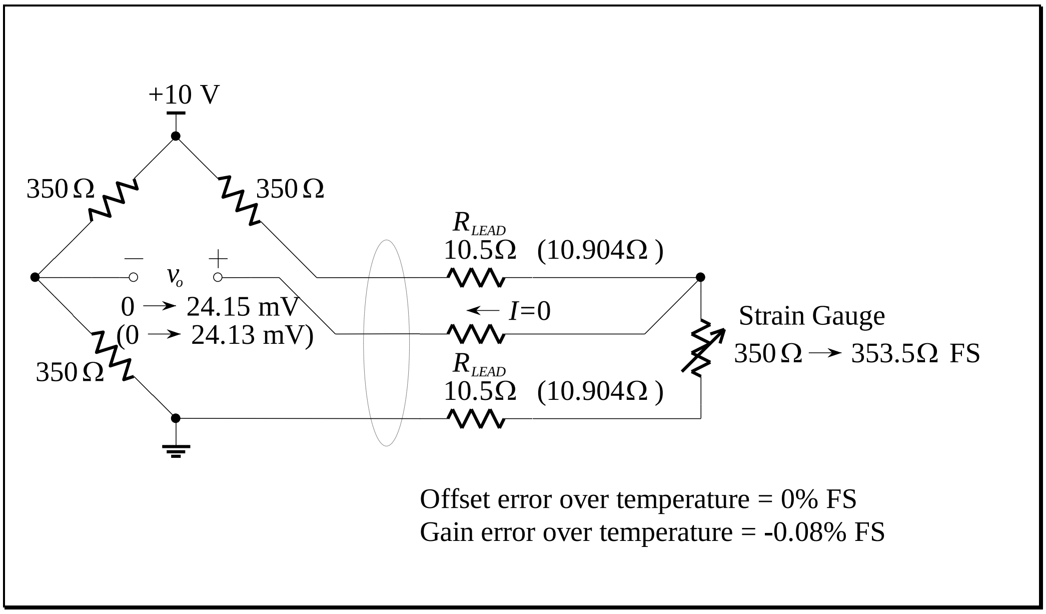


Figure 25.12 – A Bridge with Wiring Resistance

The full-scale variation of the strain gauge resistance above its nominal $350\ \Omega$ value is +1% ($+3.5\ \Omega$), corresponding to a full-scale strain gauge resistance of $353.5\ \Omega$ which causes a bridge output voltage of +23.45 mV. Notice that the addition of the 21Ω resistor, R_{COMP}, is used to compensate for the wiring resistance and balances the bridge when the strain gauge is $350\ \Omega$.

Assume that the cable temperature increase $+10\ ^\circ\mathrm{C}$ above a nominal room temperature of $25\ ^\circ\mathrm{C}$. The values in parentheses in the diagram indicate the values at $+35\ ^\circ\mathrm{C}$. With no strain, the additional lead resistance produces an offset of +5.44 mV in the bridge output. Full-scale strain produces a bridge output of +28.84 mV (a change of +23.4 mV from no strain). Thus, the increase in temperature produces an offset voltage error of +5.44 mV, or +23% full-scale, and a gain error of -0.05 mV (23.4 mV – 23.45 mV), or -0.21% full-scale.

The effects of wiring resistance on the bridge output can be minimized by the 3-wire connection shown below:



Wiring resistance for a remotely located bridge can be overcome using a 3-wire connection

Figure 25.13 – A 3-Wire Connection for a Bridge

We assume that the bridge output voltage is measured by a high impedance device, therefore there is no current in the sense lead. The sense lead measures the voltage output of a divider: the top half is the bridge resistor plus the lead resistance, and the bottom half is strain gauge resistance plus the lead resistance. The nominal sense voltage is therefore independent of the lead resistance. When the strain gauge resistance increases to full-scale ($353.5\,\Omega$), the bridge output increases to 24.15 mV.

Increasing the temperature to $+35\,°\text{C}$ increases the lead resistance in each half of the divider. The full-scale bridge output voltage decreases to +24.13 mV, but there is no offset error. The gain error due to the temperature increase of $+10\,°\text{C}$ is therefore only -0.02 mV, or -0.08% full-scale.

The three-wire method works well for remotely located resistive elements which make up one leg of a single-element varying bridge. However, four-element varying bridges are generally housed in a complete assembly, as in the case of a load cell. When these bridges are remotely located from the conditioning electronics, special techniques must be used to maintain accuracy.

Most four-element varying bridges (such as load cells) are six-lead assemblies: two leads for the bridge output, two leads for the bridge excitation, and two *sense* leads. To take full advantage of the additional accuracy that these two extra leads allow, a method called Kelvin or 4-wire sensing is employed, as shown below:

A 6-lead bridge uses Kelvin sensing to overcome lead resistance in both the sensing terminals and the excitation terminals

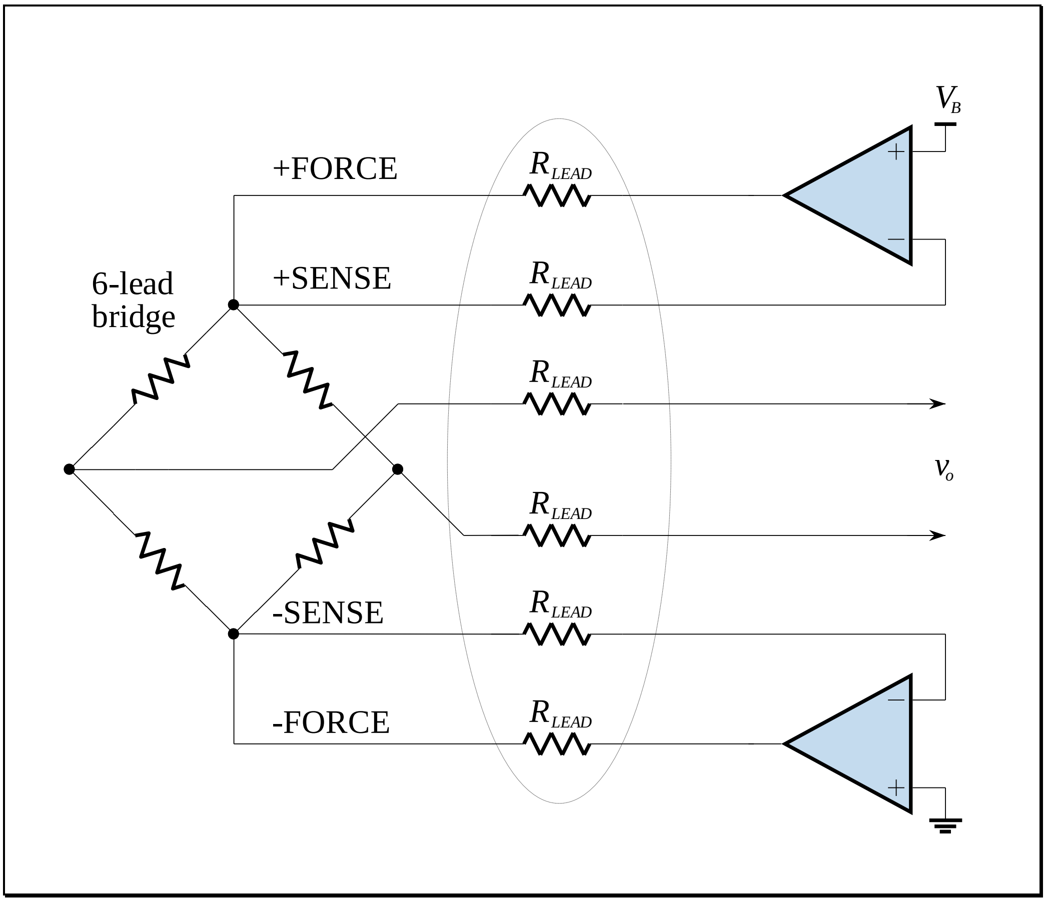


Figure 25.14 – A 6-Lead Bridge Using Kelvin Sensing

In this setup the drive voltage V_B is not applied directly to the bridge, but goes instead to the input of the upper precision op-amp, which is connected in a feedback loop around the bridge (+) terminal. Although there may be a substantial voltage drop in the +FORCE lead resistance of the remote cable, the op-amp will automatically correct for it, since it has a feedback path through the +SENSE lead. The net effect is that the upper node of the remote bridge is maintained at a precise level of V_B. A similar situation occurs with the bottom precision op-amp, which drives the bridge (-) terminal to ground level. Again, the voltage drop in the –FORCE lead is relatively immaterial, because of the sensing at the –SENSE terminal.

25.6.4 Integrated Bridge Transducers

A very powerful combination of bridge circuit techniques is shown below:

A bridge transducer ADC can be used to connect directly to a 6-lead bridge

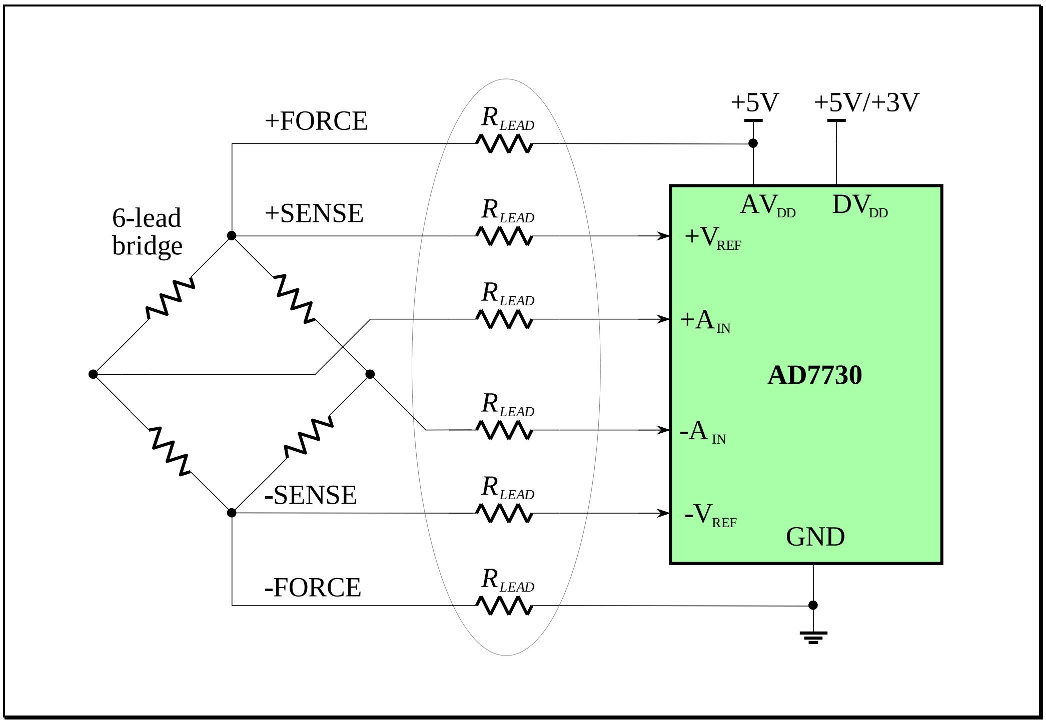


Figure 25.15 – An Integrated Bridge Transducer

This is an example of a basic DC operated bridge, utilising ratiometric conversion with a high performance ADC, combined with Kelvin sensing to minimize errors due to wiring resistance.

The Analog Devices AD7730 Bridge Transducer ADC can be driven from a single supply voltage of 5 V, which in this case is used to excite the bridge. Both the analog input and the reference input to the ADC are high impedance and fully differential. By using the +SENSE and –SENSE outputs from the bridge as the differential reference voltage to the ADC, there is no loss in measurement accuracy as the actual bridge excitation voltage varies.

The AD7730 is one of a family of “sigma-delta” ADCs with high resolution (24 bits) and internal programmable gain amplifiers (PGAs) and is marketed as a “Bridge Transducer ADC”. The chip has a self-calibration feature which allow offset and gain errors due to the ADC to be minimized. A system calibration feature allows offset and gain errors to be reduced to a few microvolts.

25.7 Strain, Force, Pressure and Flow Measurements

A strain gauge can be used to measure strain, force, pressure and flow

The most popular electrical elements used in force measurements include the resistance strain gauge, the semiconductor strain gauge, and piezoelectric transducers. The strain gauge measures force indirectly by measuring the deflection it produces in a calibrated carrier. Pressure can be converted into a force using an appropriate transducer, and strain gauge techniques can then be used to measure pressure. Flow rates can be measured using differential pressure measurements, which also make use of strain gauge technology.

Measurement	Sensor
Strain	Strain gauge, Piezoelectric transducer
Force	Load cell
Pressure	Diaphragm to force to strain gauge
Flow	Differential pressure techniques

Table 25.3 – Sensors used for Typical Measurements

The resistance-based strain gauge uses a resistive element which changes in length, hence resistance, as the force applied to the base on which it is mounted causes stretching or compression. It is the most well known transducer for converting force into an electrical variable.

A *bonded* strain gauge consists of a thin wire or conducting film arranged in a coplanar pattern and cemented to a base or carrier. The basic form of this type of gauge is shown below:

A bonded strain gauge

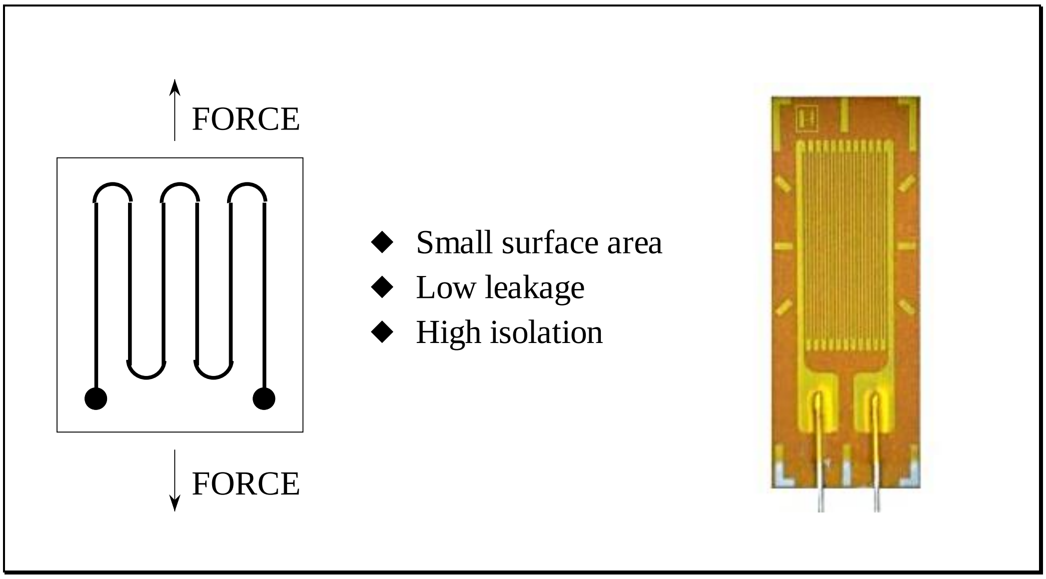


Figure 25.16 – A Bonded Strain Gauge

The strain gauge is normally mounted so that as much as possible of the length of the conductor is aligned in the direction of the stress that is being measured. Lead wires are attached to the base and brought out for interconnection.

Semiconductor strain gauges have a greater sensitivity and higher-level output than wire strain gauges. They can also be produced to have either positive or negative changes when strained. However, they are temperature sensitive, are difficult to compensate, and the change in resistance is nonlinear.

25.8 High Impedance Sensors

Many popular sensors have output impedances greater than several megohms, and thus the associated signal-conditioning circuitry must be carefully designed to meet the challenges of low bias current, low noise, and high gain. A few examples of high impedance sensors are:

High impedances sensors…

- Photodiode preamplifiers
- Piezoelectric sensors
- Humidity monitors
- pH monitors
- Chemical sensors
- Smoke detectors

Very high gain is usually required to convert the output signal of these sensors into a usable voltage. For example, a photodiode application typically needs to detect outputs down to 30 pA of current, and even a gain of 10^6 will only yield 30 mV. To accurately measure photodiode currents in this range, the bias current of the op-amp should be no more than a few picoamps. A high performance JFET-input op-amp is normally used to achieve this specification.

…require special interfacing circuits

Special circuit layout techniques are required for the signal conditioning circuitry. For example, circuit layouts on a printed circuit board (PCB) typically need very short connections to minimise leakage and parasitic elements. Inputs tend to be "guarded" with ground tracks to isolate sensitive amplifier inputs from voltages appearing across the PCB.

25.9 Temperature Sensors

Temperature measurement is critical in many electronic devices, especially expensive laptop computers and other portable devices – their densely packed circuitry dissipates considerable power in the form of heat. Knowledge of system temperature can also be used to control battery charging, as well as to prevent damage to expensive microprocessors.

Temperature is an extremely important physical property to measure

Accurate temperature measurements are required in many other measurement systems, for example within process control and instrumentation applications. Some popular types of temperature sensors and their characteristics are indicated in the table below:

Popular types of temperature sensor and their characteristics

Sensor	Range	Accuracy	Excitation	Feature
Thermocouple	-184°C to +2300°C	High accuracy and repeatability	Needs cold junction compensation	Low-voltage
RTD	-200°C to +850°C	Fair linearity	Requires excitation	Low cost
Thermistor	0°C to +100°C	Poor linearity	Requires excitation	High sensitivity
Semiconductor	-55°C to +150°C	Linearity: 1°C Accuracy: 1°C	Requires excitation	10 mV/K, 20 mV/K or 1μA/K typical output

Table 24.1 – Popular Types of Temperature Sensors

In most cases, because of low-level and/or nonlinear outputs, the sensor output must be properly conditioned and amplified before further processing can occur. Sensor outputs may be digitized directly by high resolution ADCs – linearization and calibration can then be performed in software, reducing cost and complexity.

Resistance Temperature Devices (RTDs) are accurate, but require excitation current and are generally used within bridge circuits.

Thermistors have the most sensitivity, but are also the most nonlinear. They are popular in portable applications for measurement of battery and other critical system temperatures.

Modern *semiconductor temperature sensors* offer both high accuracy and linearity over about a -55°C to +150°C operating range. Internal amplifiers can scale output to convenient values, such as 10 mV/°C.

25.10 Summary

- A *sensor* is a device that receives a signal or stimulus and responds with an electrical signal. The full-scale outputs of most sensors are relatively small voltages, currents, or resistance changes, and therefore their outputs must be properly conditioned before further analog or digital processing can occur.

- Amplification, level translation, galvanic isolation, impedance transformation, linearization and filtering are fundamental signal-conditioning functions that may be required with sensors.

- A *resistance bridge*, or *Wheatstone bridge*, is used to measure small resistance changes accurately. There are a variety of different bridge circuits, and a variety of amplifying and linearizing techniques to suit each type.

- There are a variety of methods for interfacing to remote bridges. Many integrated bridge transducers are available as “one-chip” solutions to bridge driving and measurement.

- There are many types of sensors – their use in a certain application requires an understanding of their physical construction and operation, as well as the required performance and cost demanded by the overall system.

25.11 References

Jung, W: *Op-Amp Applications*, Analog Devices, 2002.

Exercises

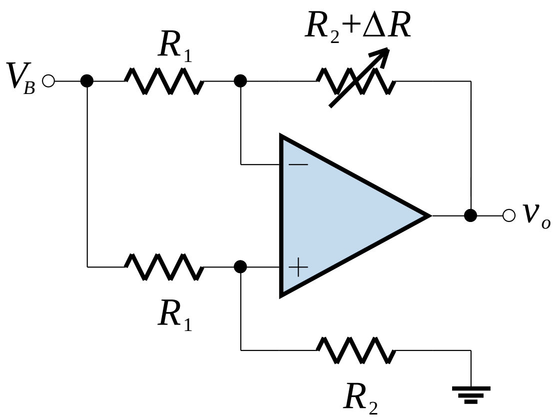
1.

For temperature measurements only one active transducer is used and so it is not possible to have a linear output if it is placed in a bridge.

(a) Show that the output from a single-element varying bridge is given by:

$$v_o = \frac{V_B}{4}\frac{\Delta R}{R + \frac{\Delta R}{2}}$$

(b) Since the active transducer resistance change can be rather large (up to 100% or more for RTDs), the nonlinearity of the bridge output characteristic (the formula above) can become quite significant. It is therefore desired to linearize the output of a temperature transducer using the following circuit:



Derive an equation for the output voltage.

26 System Modelling

Contents

26.2

Introduction

In order to understand, analyse and design complex systems, we must obtain quantitative *mathematical models* of these systems. Since most systems are dynamic in nature, the descriptive equations are usually *differential equations*. If the system stays “within bounds”, then the equations are usually treated as linear differential equations, and the method of transfer functions can be used to simplify the analysis.

In practice, the complexity of systems and the ignorance of all the relevant factors necessitate the introduction of *assumptions* concerning the system operation. Therefore, we find it useful to consider the physical system, delineate some necessary assumptions, and linearize the system. Then, by using the physical laws describing the linear equivalent system, we can obtain a set of linear differential equations. Finally, utilizing mathematical tools, such as the transfer function, we obtain a solution describing the operation of the system.

In summary, we:

1. Define the system and its components.
2. List the necessary assumptions and formulate the mathematical model.
3. Write the differential equations describing the model.
4. Solve the equations for the desired output variables.
5. Examine the solutions and the assumptions.
6. Reanalyse or design.

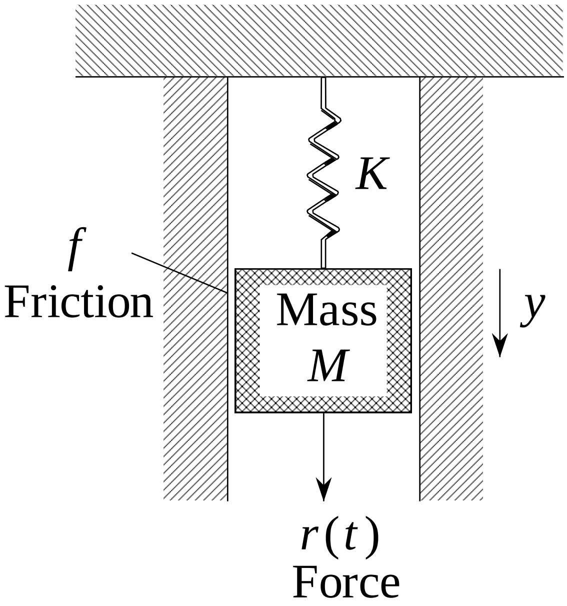
26.1 Differential Equations of Physical Systems

The differential equations describing the dynamic performance of a physical system are obtained by utilizing the physical laws – this approach applies equally well to electrical, mechanical, fluid and thermodynamic systems.

For mechanical systems, Newton's laws are applicable.

EXAMPLE 26.1 Spring-Mass-Damper Mechanical System

Consider the simple spring-mass-damper mechanical system shown below:



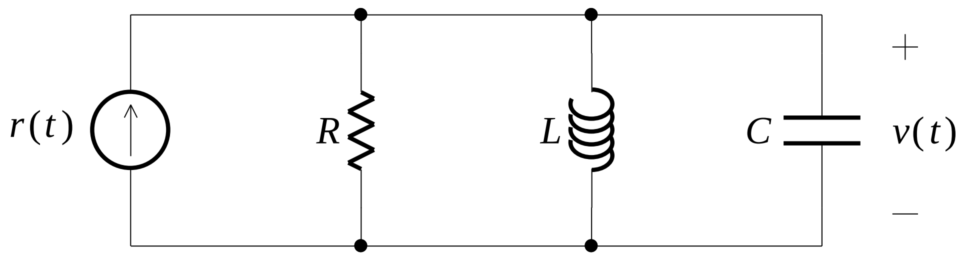
This is described by Newton's second law of motion (this system could represent, for example, a car's shock absorber). We therefore obtain:

$$M\frac{d^2y}{dt^2} + f\frac{dy}{dt} + Ky = r$$

where K is the spring constant of the ideal spring and f is the friction constant.

EXAMPLE 26.2 Parallel *RLC* Circuit

Consider the electrical *RLC* circuit below:



This is described by Kirchhoff's current law. We therefore obtain:

$$C\frac{dv}{dt}+\frac{v}{R}+\frac{1}{L}\int v dt = r$$

In order to reveal the close similarity between the differential equations for the mechanical and electrical systems, we can rewrite the mechanical equation in terms of velocity:

$$v=\frac{dy}{dt}$$

Then we have:

$$M\frac{dv}{dt}+fv+K\int v dt = r$$

The equivalence is immediately obvious where velocity $v(t)$ and voltage $v(t)$ are equivalent variables, usually called *analogous* variables, and the systems are analogous systems.

The concept of analogous systems is a very useful and powerful technique for system modelling. Analogous systems with similar solutions exist for electrical, mechanical, thermal and fluid systems. The existence of analogous systems and solutions allows us to extend the solution of one system to all analogous systems with the same describing differential equation.

26.2 Linear Approximations of Physical Systems

Many physical systems are linear within some range of variables. However, all systems ultimately become nonlinear as the variables are increased without limit. For example, the spring-mass-damper system is linear so long as the mass is subjected to small deflections $y(t)$. However, if $y(t)$ were continually increased, eventually the spring would be overextended and break. Therefore, the question of linearity and the range of applicability must be considered for each system.

A necessary condition for a system to be linear can be determined in terms of a forcing function $x(t)$ and a response $y(t)$. A system is linear *if and only if*:

$$ax_1(t)+bx_2(t)\Rightarrow ay_1(t)+by_2(t) \tag{26.1}$$

That is, linear systems obey the *principle of superposition*, [excitation by $x_1(t)+x_2(t)$ results in $y_1(t)+y_2(t)$] and they also satisfy the *homogeneity* property [excitation by $ax_1(t)$ results in $ay_1(t)$].

It may come as a surprise that a system obeying the relation $y=mx+b$ is not linear, since it does not satisfy the homogeneity property. However, the system may be considered linear about an *operating point* (x_0, y_0) for small changes Δx and Δy. When $x=x_0+\Delta x$ and $y=y_0+\Delta y$, we have:

$$\begin{aligned} y &= mx+b \\ y_0+\Delta y &= mx_0+m\Delta x+b \end{aligned} \tag{26.2}$$

and, since $y_0=mx_0+b$, then $\Delta y=m\Delta x$, which is linear.

In general, we can often linearize nonlinear elements by assuming *small-signal conditions*. This approach is the normal approach used to obtain linear equivalent circuits for electronic circuits and transistors.

26.6

Consider a general element of a system which can be described by a relationship between the excitation variable $x(t)$ and the response $y(t)$:

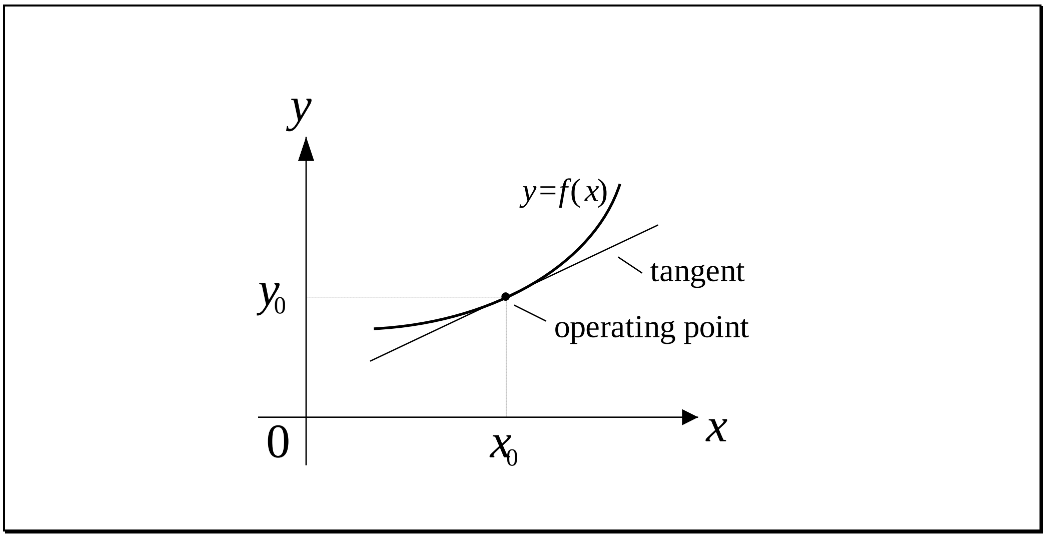


Figure 26.1

The normal operating point is designated by (x_0, y_0). A *Taylor series* expansion about the operating point gives:

$$y = f(x_0 + \Delta x) = f(x_0) + \Delta x \left.\frac{df}{dx}\right|_{x=x_0} + \cdots \tag{26.3}$$

The tangent to the curve at the operating point is a good approximation to the curve for small Δx. Thus, for a small region about the operating point, a reasonable *first-order* approximation of the element is:

$$y \approx f(x_0) + \Delta x \left.\frac{df}{dx}\right|_{x=x_0} = y_0 + m\Delta x \tag{26.4}$$

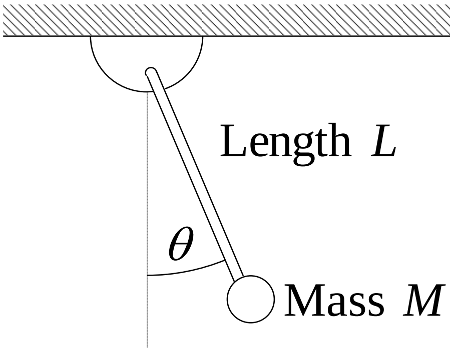
which can be written as the approximate linear equation:

$$\Delta y \approx m\Delta x \tag{26.5}$$

This *linear approximation* is only accurate for a range of *small signals* which depends on the actual nonlinear element's characteristic.

EXAMPLE 26.3 Pendulum Oscillator

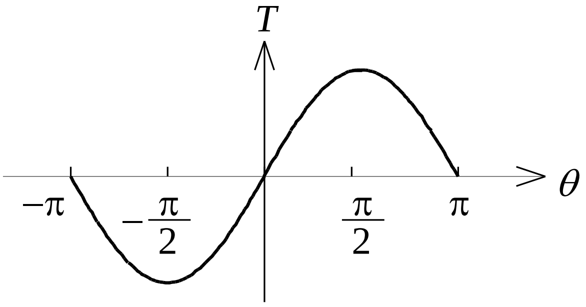
Consider the pendulum oscillator:



The magnitude of the torque on the mass is:

$$T = MgL\sin\theta$$

where g is the gravity constant. The equilibrium position for the mass is $\theta_0 = 0°$. The nonlinear relation between T and θ is shown graphically below:



The first derivative evaluated at equilibrium provides the linear approximation which is:

$$\begin{aligned}\Delta T &\approx MgL\frac{d}{d\theta}\sin\theta\Big|_{\theta=\theta_0}(\theta-\theta_0)\\ &= MgL(\cos 0°)(\theta - 0°)\\ &= MgL\theta\end{aligned}$$

This approximation is reasonably accurate for $-\pi/4 \le \theta \le \pi/4$. For example, the response of the linear model for the swing through $\pm 30°$ is within 2% of the actual nonlinear pendulum response.

26.3 The Transfer Function

We have already seen that the transfer function, which represents the input/output relationship of a system in terms of complex frequency, has a simple relationship to the differential equation describing the system. Differentiation in the time-domain turns into multiplication by **s** in the frequency-domain:

$$\frac{d}{dt} \Leftrightarrow \mathbf{s} \tag{26.6}$$

Similarly, integration in the time-domain turns into division by **s** in the frequency-domain.

$$\int_{0^+}^{t} d\tau \Leftrightarrow \frac{1}{\mathbf{s}} \tag{26.7}$$

EXAMPLE 26.4 Transfer Function from a Differential Equation

For the simple spring-mass-damper mechanical system, the describing differential equation was found to be:

$$M\frac{d^2y}{dt^2} + f\frac{dy}{dt} + Ky = r$$

Transforming to the frequency-domain, we get:

$$M\mathbf{s}^2\mathbf{Y} + f\mathbf{s}\mathbf{Y} + K\mathbf{Y} = \mathbf{R}$$

Thus, the transfer function is:

$$\frac{\mathbf{Y}}{\mathbf{R}} = \frac{1}{M\mathbf{s}^2 + f\mathbf{s} + K}$$

Alternatively, the transfer function of a system can be obtained by analysis performed entirely in the frequency-domain (as we do for the transfer function of circuits).

26.4 Block Diagrams

We can represent a transfer function graphically, with a *block diagram*. This shows the relationship between the forced response and the forcing function – an input/output relationship. For example, the simple *RC* circuit can be represented by:

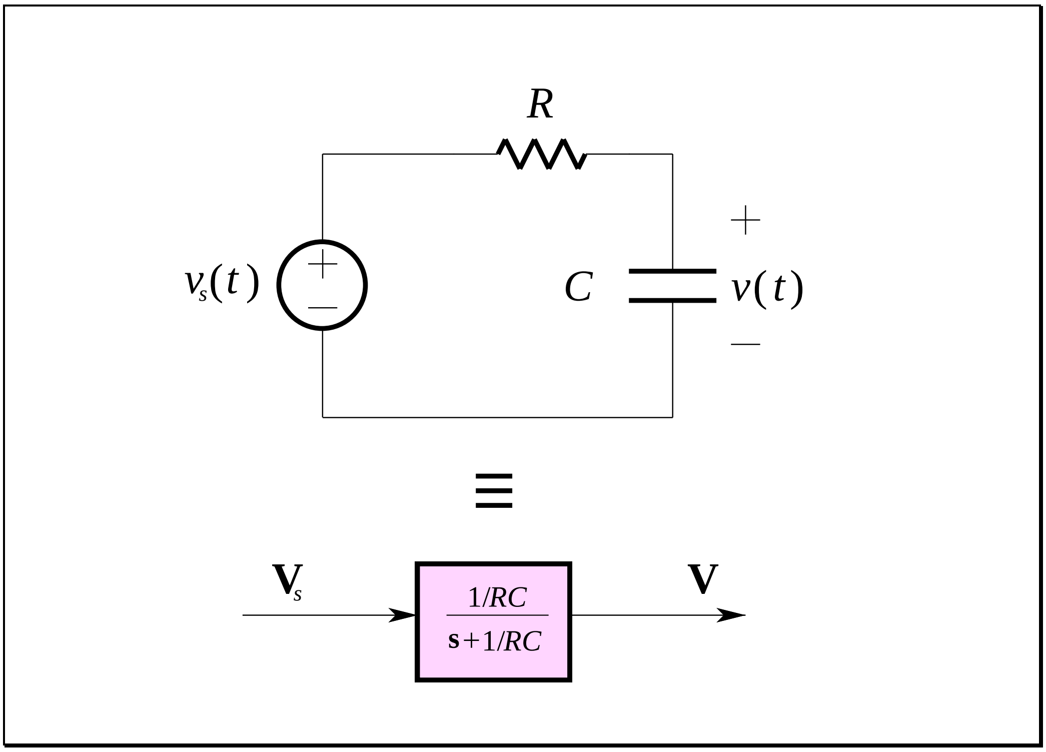
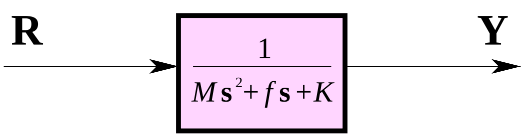


Figure 26.2

EXAMPLE 26.5 Block Diagram of a Spring-Mass-Damper System

The block diagram for the simple spring-mass-damper mechanical system is:



With this representation, we understand that we have to multiply the input phasor by the transfer function in the box to obtain the output phasor. In general, we use the following notation, where $\mathbf{G}(\mathbf{s})$ is the transfer function, $\mathbf{X}$ is the input, and $\mathbf{Y}$ is the output:

A block represents multiplication with a transfer function

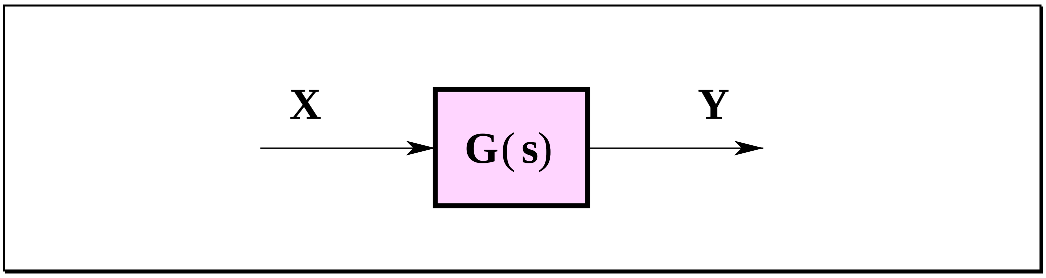


Figure 26.3

Blocks can be connected in cascade, but only if the outputs are "buffered", i.e. the connection does not cause the transfer function of each individual block to be different from the "unloaded" or open condition:

Cascading blocks implies multiplying the transfer functions

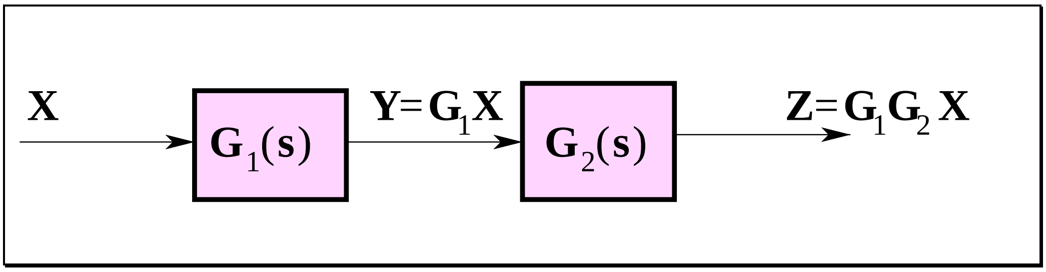


Figure 26.4

Obviously, op-amp circuits are ideal for cascading.

Most systems have several blocks interconnected by various forward and backward paths. Signals (e.g. voltages, currents) in block diagrams can not only be transformed by a transfer function, they can also be added and subtracted.

Addition and subtraction of signals in a block diagram

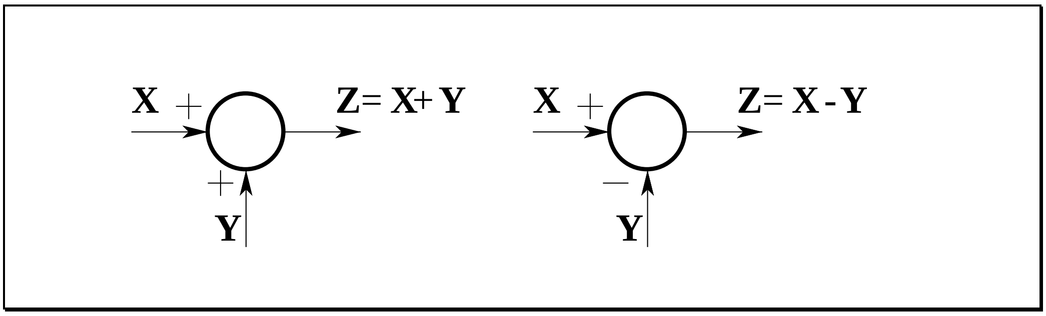
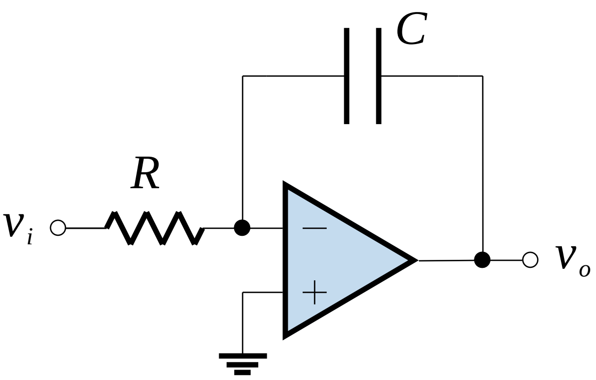


Figure 26.5

EXAMPLE 26.6 Block Diagram of an Ideal Integrator Op-Amp Circuit

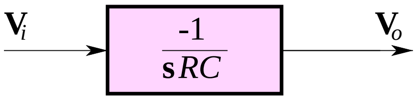
Consider the ideal integrator op-amp circuit:



We know that the gain can be expressed as:

$$\frac{\mathbf{V}_o}{\mathbf{V}_i} = -\frac{\mathbf{Z}_2}{\mathbf{Z}_1} = -\frac{1/\mathbf{s}C}{R}$$

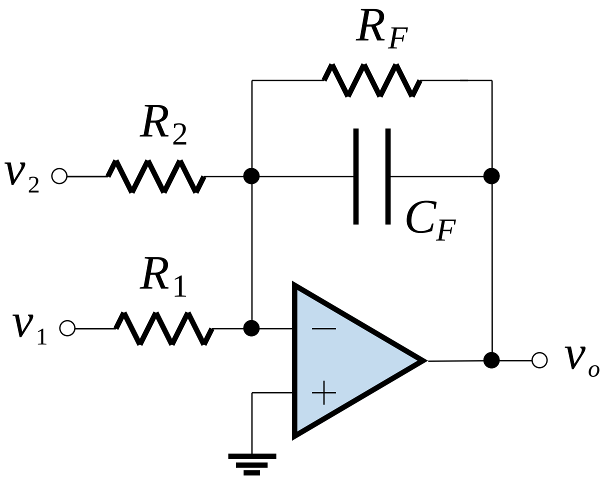
Therefore, the op-amp integrator can be represented by:



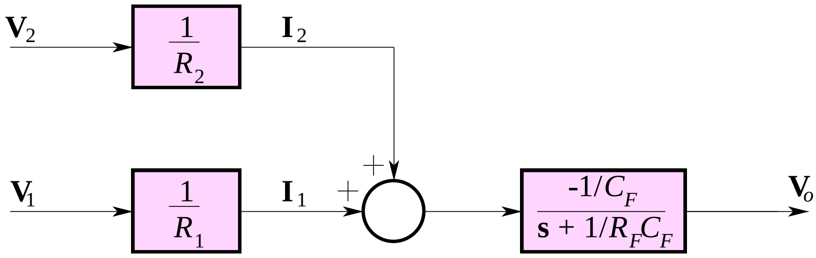
Thus, an integrator circuit has a $1/\mathbf{s}$ term in its transfer function. This makes intuitive sense, since multiplication by $\mathbf{s}$ represents differentiation, and so division by $\mathbf{s}$ must be representative of integration.

EXAMPLE 26.7 Block Diagram of a Summing Lossy Integrator

The op-amp circuit:

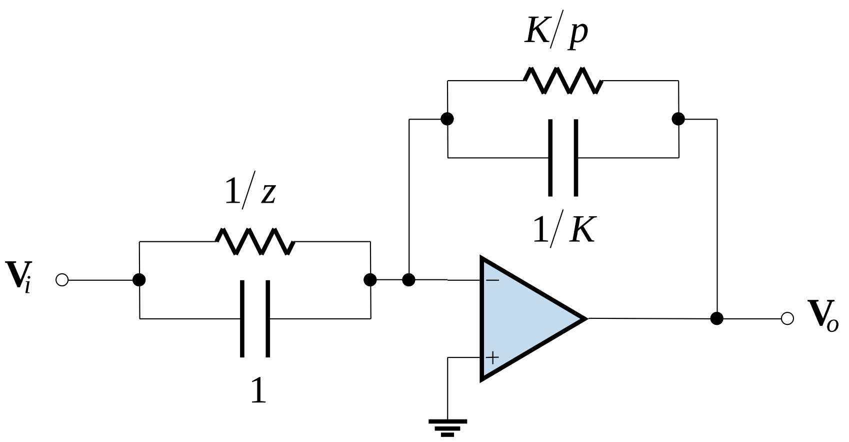


can be represented by:

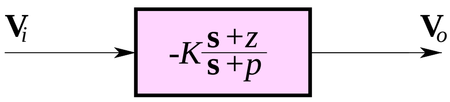


EXAMPLE 26.8 Block Diagram of a Bilinear Op-Amp Circuit

The bilinear op-amp circuit:



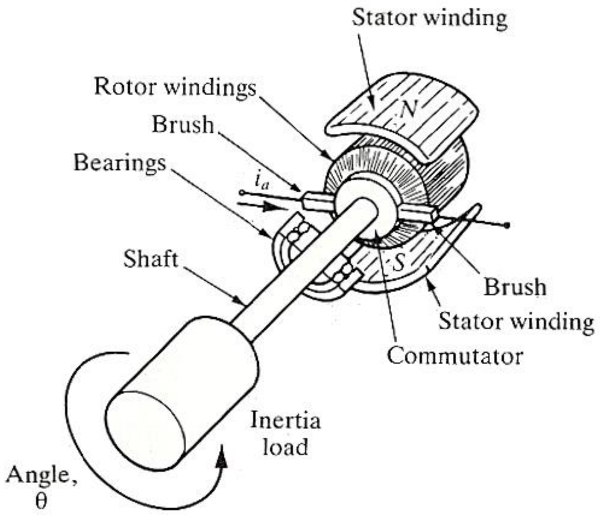
can be represented by:



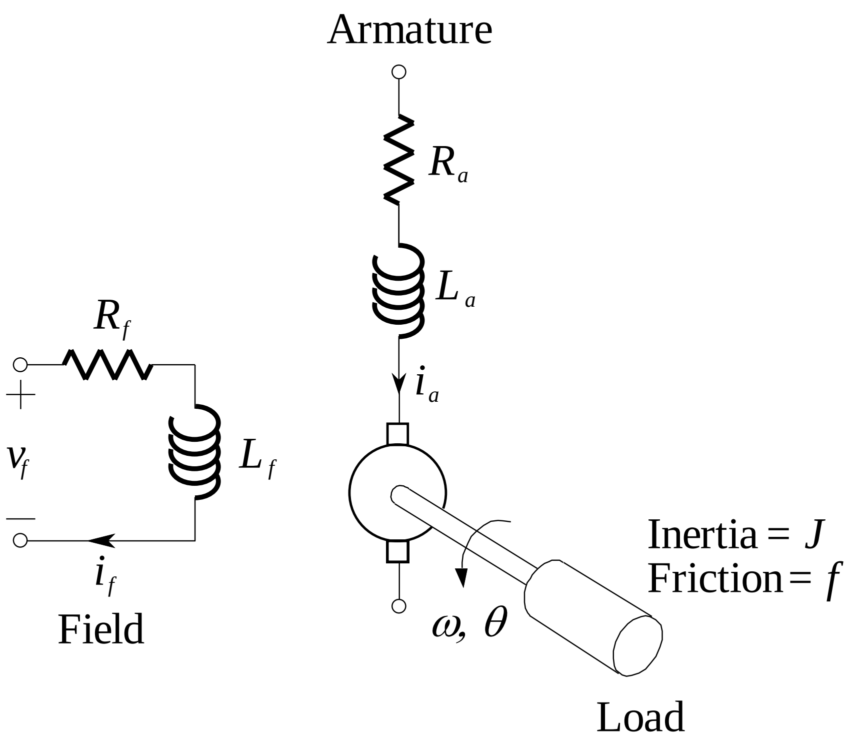
It should be obvious that we can now establish a "cookbook" approach to circuit design – if we have a transfer function that we want to implement, we simply find the appropriate op-amp circuit in the "cookbook".

EXAMPLE 26.9 Block Diagram of a Field-Controlled DC Motor

The DC motor is a power actuator device that delivers energy to a load as shown below:



A schematic diagram of the DC motor is shown below:



For a field-controlled DC motor, the armature current is assumed to be constant and the torque developed by the motor is assumed to be linearly related to the field current:

$$T_m = K_m i_f$$

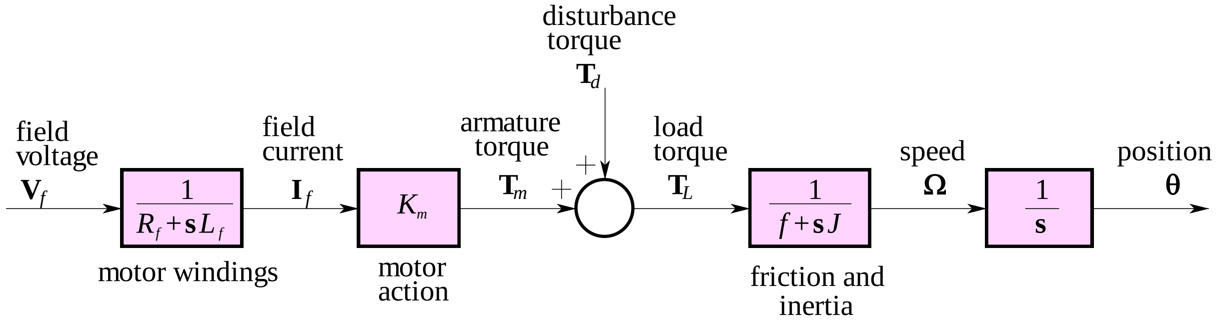
The motor torque is delivered to the load, which can also be subjected to external disturbances (e.g. wind-gust forces for a tracking antenna):

$$T_L = T_m + T_d$$

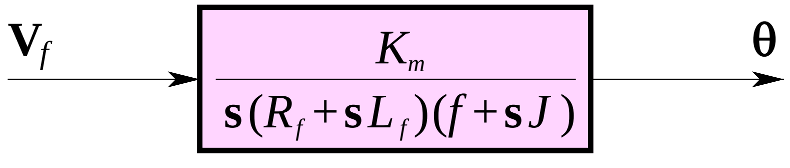
The load torque for a rotating system with inertia and friction comes from the rotational form of Newton's second law:

$$T_L = J\frac{d\omega}{dt} + f\omega, \qquad \omega = \frac{d\theta}{dt}$$

The blocks are derived from the differential equations governing the various parts of the system. The block diagram of a field-controlled DC motor is therefore:



If there is no disturbance torque, then the model of the field-controlled DC motor is:



26.5 Feedback

Perhaps the most important block diagram is that of a *feedback connection*, shown below:

Standard form for the feedback connection

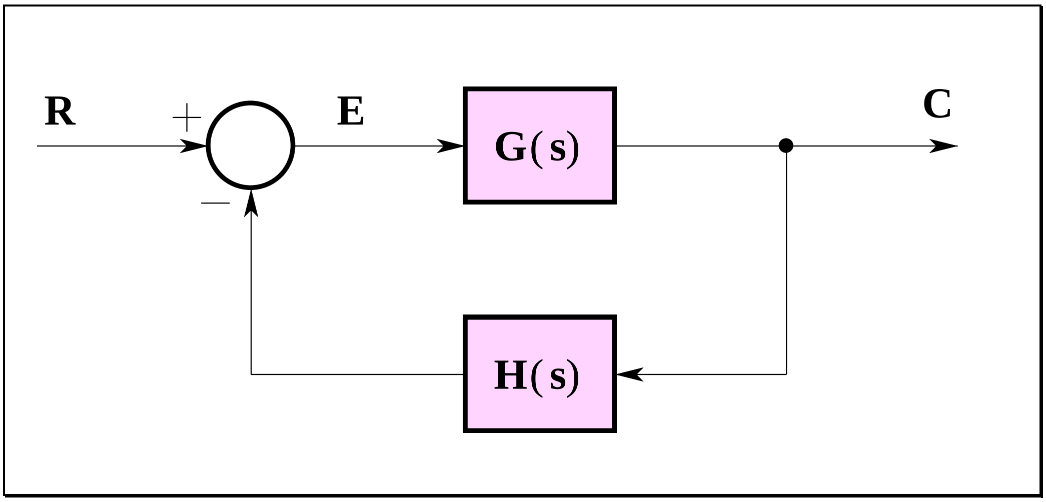


Figure 26.6

We have the following definitions:

$\mathbf{G}(\mathbf{s})$ = forward path transfer function

$\mathbf{H}(\mathbf{s})$ = feedback path transfer function

$r(t)$ = **r**eference, input, or desired output

$c(t)$ = **c**ontrolled variable, or output

$e(t)$ = actuating **e**rror signal

$r(t)-c(t)$ = system error

$\frac{\mathbf{C}}{\mathbf{R}}$ = closed-loop transfer function

$\mathbf{G}(\mathbf{s})\mathbf{H}(\mathbf{s})$ = loop gain

To find the *closed-loop* transfer function, we solve the following two equations which are self-evident from the block diagram:

$$\begin{aligned} \mathbf{C} &= \mathbf{G}(\mathbf{s})\mathbf{E} \\ \mathbf{E} &= \mathbf{R} - \mathbf{H}(\mathbf{s})\mathbf{C} \end{aligned} \tag{26.8}$$

Then the output $\mathbf{C}$ is given by:

$$\begin{aligned} \mathbf{C} &= \mathbf{G}(\mathbf{s})\mathbf{R} - \mathbf{G}(\mathbf{s})\mathbf{H}(\mathbf{s})\mathbf{C} \\ \mathbf{C}[1+\mathbf{G}(\mathbf{s})\mathbf{H}(\mathbf{s})] &= \mathbf{G}(\mathbf{s})\mathbf{R} \end{aligned} \tag{26.9}$$

and therefore:

$$\frac{\mathbf{C}}{\mathbf{R}} = \frac{\mathbf{G}(\mathbf{s})}{1+\mathbf{G}(\mathbf{s})\mathbf{H}(\mathbf{s})} \tag{26.10}$$

Transfer function for the standard negative feedback connection

Note that for negative feedback we get $1+\mathbf{G}(\mathbf{s})\mathbf{H}(\mathbf{s})$ and for positive feedback we get $1-\mathbf{G}(\mathbf{s})\mathbf{H}(\mathbf{s})$.

Negative and positive feedback

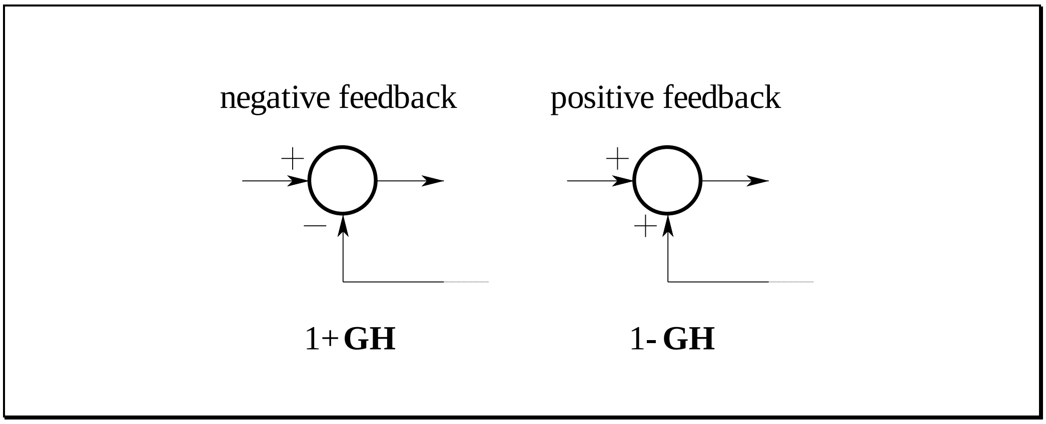
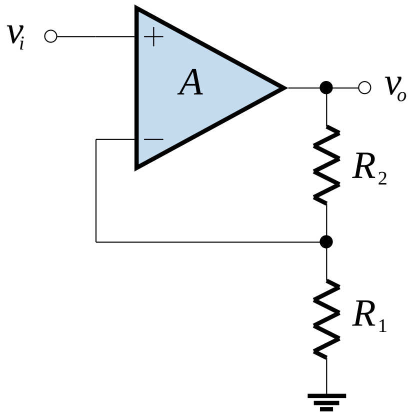


Figure 26.7

EXAMPLE 26.10 Block Diagram of a Noninverting Amplifier

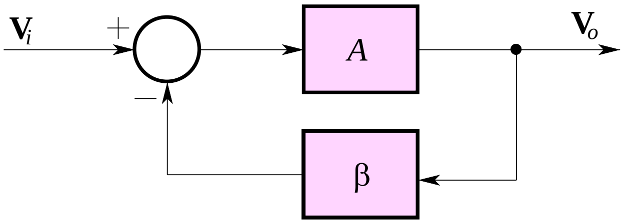
Consider a noninverting amplifier where the op-amp is modelled as having infinite input impedance, zero output impedance, and a large but finite gain A:



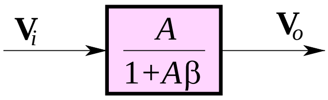
The voltage fed back to the inverting terminal is negative feedback. Let the proportion of the voltage fed back be given by:

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

Then a block diagram that models the non-inverting amplifier is:



The transfer function can be simplified to:



For large values of A such that $A\beta >> 1$, the transfer function reduces to:

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

26.6 Summary

- We model linear time-invariant systems by making necessary simplifying assumptions before applying the basic physical laws. The result is a linear differential equation describing the system.

- A linear differential equation can be turned into a transfer function by replacing the derivative operator with the complex frequency **s**. Alternatively, the transfer function of a system can be obtained by performing analysis directly in the frequency-domain.

- A transfer function can be represented graphically in the form of a block diagram. Block diagrams can be connected together in cascade, and signals can be added and subtracted.

- Negative feedback in a system is a very important concept. For a forward path transfer function **G** and a feedback path transfer function **H**, the closed-loop transfer function is:

$$\mathbf{T}(\mathbf{s}) = \frac{\mathbf{G}}{1 + \mathbf{GH}}$$

26.7 References

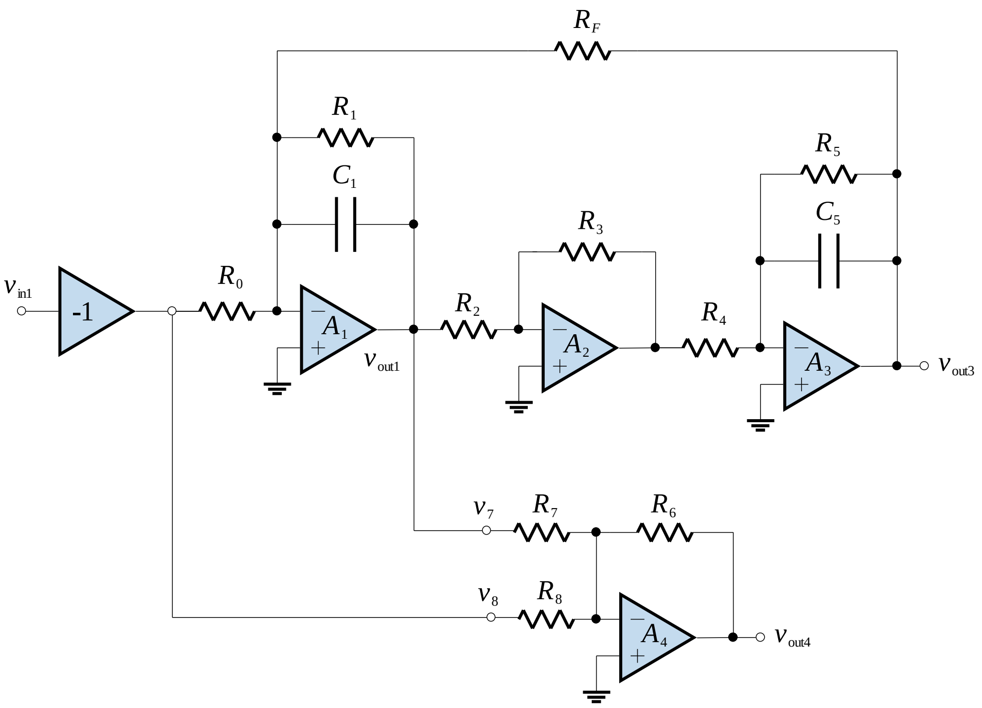
Dorf, R.: *Modern Control Systems*, 5th *Ed*., Addison-Wesley, 1989.

26.20

Exercises

1.

Consider the following circuit:



(a) Construct a block diagram of the circuit.

(b) Show that the transfer function $\mathbf{T}_1(\mathbf{s}) = \mathbf{V}_{\text{out3}}/\mathbf{V}_{\text{in1}}$ is:

$$\mathbf{T}_1(\mathbf{s}) = \frac{R_3/R_0R_2R_4C_1C_5}{\mathbf{s}^2 + \left(\frac{1}{R_1C_1} + \frac{1}{R_5C_5}\right)\mathbf{s} + \left(\frac{R_3}{R_2R_4R_FC_1C_5} + \frac{1}{R_1R_5C_1C_5}\right)}$$

(c) By comparing with the "standard form" of a second-order all-pole transfer function:

$$\mathbf{T}_1(\mathbf{s}) = \frac{K_1\omega_0^2}{\mathbf{s}^2 + 2\alpha\mathbf{s} + \omega_0^2}$$

write expressions for K_1, α and ω_0 in terms of R's and C's.

(d) For the *special* case of when $R_1 = R_5$, $C_1 = C_5$ and $R_F = R_4$, and defining $\omega_d = \sqrt{\omega_0^2 - \alpha^2}$, show that the poles of the transfer function are located at $p_{1,2} = -\alpha \pm j\omega_d$ where:

$$\alpha = \frac{1}{R_1 C_1} \quad \omega_d = \sqrt{\frac{R_3}{R_2} \frac{1}{R_4 C_1}}$$

(e) For the *special* case of when $R_1 = R_5$, $C_1 = C_5$ and $R_F = R_4$, show that the transfer function $\mathbf{T}_2(\mathbf{s}) = \mathbf{V}_{\text{out1}} / \mathbf{V}_{\text{in1}}$ can be put in the form:

$$\mathbf{T}_2(\mathbf{s}) = \frac{K_2(\mathbf{s} + \alpha)}{\mathbf{s}^2 + 2\alpha\mathbf{s} + \omega_0^2}$$

(f) For the *special* case of when $R_1 = R_5$, $C_1 = C_5$, $R_F = R_4$, $R_6 = R_8$ and $\frac{2R_0}{R_1} = \frac{R_8}{R_7}$, show that the transfer function $\mathbf{T}_3(\mathbf{s}) = \mathbf{V}_{\text{out4}} / \mathbf{V}_{\text{in1}}$ can be put in the form:

$$\mathbf{T}_3(\mathbf{s}) = \frac{K_3(\mathbf{s}^2 + \omega_0^2 - 2\alpha^2)}{\mathbf{s}^2 + 2\alpha\mathbf{s} + \omega_0^2}$$

(g) Draw pole-zero plots for each of the three transfer functions.

(h) Perform a simulation of the circuit and determine each transfer function’s frequency response. Hence, classify each of the transfer functions in terms of their frequency response (e.g. lowpass, highpass, notch, etc.).

27 Revision

Contents

27.1 Basic Laws

Know

- Current, voltage, power and circuits.
- Circuit elements and types of circuits.
- Ohm's Law.
- Kirchhoff's Current Law (KCL).
- Kirchhoff's Voltage Law (KVL).
- Resistance and source combination.
- Voltage and current dividers.

Know the passive and active sign conventions and have confidence in applying KCL and KVL to any circuit. Know and use Ohm's Law. Know circuit analysis short-cuts with series / parallel combinations, voltage and current dividers.

27.2 Amplifiers

Know

- Amplifiers.
- The ideal op-amp.
- Negative feedback.
- The noninverting amplifier.
- The inverting amplifier.

Know what parameters are used for an ideal op-amp, the concept and application of negative feedback and the concept of the "virtual short-circuit". You should know the gain formulas for both noninverting and inverting amplifiers, and be able to derive them if necessary.

27.3 Nodal and Mesh Analysis

Know

- Nodal analysis.
- Mesh analysis.

Know that nodal analysis applies KCL and finds nodal voltages, whilst mesh analysis applies KVL and finds mesh current. Know how to perform, by hand, analysis on a three-node or three-mesh circuit composed of resistors, ideal independent and ideal dependent sources.

27.4 Circuit Analysis Techniques

Know

- Source transformations.
- Linearity and superposition.
- Thévenin's and Norton's Theorems.

You should be able to convert between a real voltage source and a real current source (a special application of Thévenin's and Norton's Theorems). Know when linearity applies (and does not apply), and be able to use superposition appropriately (e.g. DC and AC sources, common-mode and differential mode). Know how to apply Thévenin's and Norton's Theorems to any linear circuit, even if it contains dependent sources.

27.5 Linear Op-Amp Applications

Know

- Summing amplifier.
- Difference amplifier.
- Integrator.
- Differentiator.

Know the basic functionality of the summing and difference amplifier, integrator and differentiator. You should be able to derive output equations from first principles if needed, using the basic circuit laws and the concept of the virtual short-circuit.

27.6 Reactive Components

Know

- Capacitor and inductor construction and *v*-*i* relationships.
- Stored energy relationships for inductors and capacitors.
- Series/parallel connection of inductors and capacitors.
- Circuit analysis with inductors and capacitors.

Familiar

- Physical characteristics of inductors; parasitic effects.
- Physical characteristics of capacitors; parasitic effects.

Aware

- Duality.

Know the *v*-*i* relationships and the energy stored for C and L, and how they can be combined in series and parallel. Know how to write nodal or mesh equations for any circuit. Know how to perform DC analysis for any circuit.

Familiar with the fact that real components exhibit other characteristics, and we initially study idealized forms of components. The most important nonideality of the passive components is the resistance of the inductor windings.

Aware of the concept of duality – be able to recognise circuits that are "duals".

27.7 Diodes and Basic Diode Circuits

Know

- Ideal diode.

Familiar

- Diode models.
- Rectifier circuits.
- Limiting and clamping circuits.
- LEDs.

Know the operation of an ideal diode.

Familiar with the fact that real signal diodes exhibit a 0.7 V voltage drop when conducting, the principles of rectification and why we rectify, application of diodes to limiting and clamping voltage signals, light emitting diodes have a forward voltage drop different to 0.7 V (depends on the colour).

27.8 Source-Free *RC* and *RL* Circuits

Know

- Differential operators.
- Properties of differential operators.
- The characteristic equation.
- The simple *RC* circuit.
- Properties of the exponential response.
- Single time constant *RC* circuits.
- The simple *RL* circuit.
- Single time constant *RL* circuits.

Know how to write a differential equation using the D operator, write a characteristic equation by inspection of a differential equation, analyse any first-order *RC* or *RL* circuit to obtain the natural response, expressed in terms of the appropriate time constant.

27.9 Nonlinear Op-Amp Applications

Familiar

- Op-amp comparators.
- The superdiode.
- Precision half-wave and full-wave rectifiers.
- Precision peak detector.
- Limiters.
- Clamps

Familiar with nonlinear op-amp circuits such as the comparator, superdiode, precision rectification, limiters and clamps.

27.10 The First-Order Step Response

Know

- The unit-step forcing function.
- The driven *RC* circuit.
- Forced and natural response.
- *RC* circuits.
- Analysis procedure for single time constant *RC* circuits.
- *RL* circuits.
- Analysis procedure for single time constant *RL* circuits.

Know the definition of a unit-step as applied to a voltage or current source, that a driven circuit has a forced response and a natural response, and that these correspond to the particular solution and complementary solution to the circuit's describing differential equation, how to analyse any first-order circuit with a step (DC) forcing function.

27.11 Op-Amp Imperfections

Know

- DC imperfections (offset voltage, bias and offset currents).
- Finite open-loop gain.

Familiar

- Finite bandwidth.
- Output voltage saturation.
- Output current limits.
- Slew rate.
- Full-power BW.

Know how to calculate the effect of DC imperfections on an op-amp circuit's output, and derive gain expressions if the op-amp has finite gain.

Familiar with other limitations of the op-amp, as observed in the laboratory.

27.12 The Phasor Concept

Know

- Sinusoidal signals.
- Sinusoidal steady-state response.
- Complex forcing function.
- Phasors.
- Phasor relationships for R, L and C.
- Impedance.
- Admittance.

Know the general expression for a sinusoid, the concepts of amplitude and phase, that a sinusoidal input yields a sinusoidal output in the steady-state ("sinusoid in = sinusoid out"), complex forcing functions can be created from real sinusoidal forcing functions, the concept of a phasor (visualize a rotating phasor, and how to obtain the real part), phasor relationships for R, L and C and the concepts of impedance and admittance.

27.13 Circuit Simulation

Familiar

- DC analysis.
- AC analysis.
- Transients.

Familiar with the schematic capture process, simulation in the time-domain (transient analysis) and frequency-domain (AC analysis).

27.14 The Sinusoidal Steady-State Response

Know

- Nodal analysis.
- Mesh analysis.
- Superposition.
- Source transformations.
- Thévenin's Theorem.
- Norton's theorem.
- Phasor diagrams.
- Power in AC circuits.

Know how to apply nodal and mesh analysis to circuits in the frequency-domain, apply superposition, source transformations, Thévenin's and Norton's Theorems and draw phasor diagrams. Know the different forms of power such as instantaneous, average, real, reactive and complex.

27.15 Amplifier Characteristics

Familiar

- Circuit models for amplifiers.
- Cascaded amplifiers.
- Efficiency.
- I/O impedances.
- Ideal amplifiers.
- Frequency response.
- Linear distortion.
- Transfer characteristic.
- Nonlinear distortion.

Familiar with the circuit models used for real amplifiers, cascading amplifiers, efficiency, input and output impedances, ideal amplifier characteristics, frequency response including the concepts of bandwidth and half-power, linear amplitude and phase distortion, nonlinear distortion.

27.16 Frequency Response

Know

- Decibels.
- Cascading two-ports.
- Logarithmic frequency scales.
- Bode plots.
- First-order lowpass filters.
- First-order highpass filters.

Aware

- Fourier analysis, cascade filters.

Know the definition of the decibel and its application to amplifier circuits, how to cascade circuits, logarithmic scales, Bode plots, how to derive the equations for the frequency response for first-order lowpass and highpass filters, and sketch their responses.

27.17 First-Order Op-Amp Filters

Know

- Bilinear frequency response.
- First-order lowpass filters.
- First-order highpass filters.

Know the form of the bilinear frequency response, and be able to derive a suitable op-amp implementation of it, how to implement lowpass and highpass filters using cascaded op-amp circuits.

27.18 The Second-Order Step Response

Know

- Solution of the homogeneous linear differential equation.
- Source-free parallel *RLC* circuit.
- Overdamped parallel *RLC* circuit.
- Critical damping.
- Underdamped parallel *RLC* circuit.
- Source-free series *RLC* circuit.
- Complete response of the *RLC* circuit.

Know the solution of the second-order homogeneous linear differential equation from inspection of the roots of the characteristic equation, the natural response of the parallel *RLC* circuit for the overdamped, critically damped and underdamped cases, the natural response of the series *RLC* circuit and the complete response of any *RLC* circuit.

27.19 Waveform Generation

Familiar

- Open-loop comparator.
- Comparator with hysteresis.
- Astable multivibrator.
- Waveform generator.

Familiar with the open-loop comparator, hysteresis, the astable multivibrator and waveform generation from laboratory experience.

27.20 Second-Order Frequency Response

Familiar

- Lowpass, bandpass and highpass responses.
- Resonance.
- Bandwidth.
- Quality factor.

Familiar with the concept of lowpass, bandpass and highpass frequency responses for second-order systems. The concept of resonance, bandwidth and quality factor for a second-order system and the relationships between them. The concept of frequency selectivity and its application to filtering signals in the frequency-domain.

27.21 Second-Order Op-Amp Filters

Familiar

- Filter design parameters.
- The lowpass biquad circuit.
- The universal biquad circuit.

Aware

- Approximating the ideal lowpass filter.
- The Butterworth lowpass filter.

Familiar with op-amp circuits that implement second-order systems, such as the universal biquad filter, and their advantages and disadvantages. Aware of the concept of cascade filter design, the ideal filter and various practical response types such as the Butterworth response.

27.22 Complex Frequency

Know

- Complex frequency.
- Damped sinusoidal forcing function.
- Generalized impedance and admittance.
- Frequency response.
- The complex-frequency plane.

Aware

- Visualization of the frequency response from a pole-zero plot.

Know the concept of complex frequency, the damped sinusoidal forcing function, generalized impedance and admittance, frequency response for both ω and σ. Aware of visualization of the frequency response from a pole-zero plot.

27.23 Specialty Amplifiers

Know

- Differential and common-mode signals.
- Difference amplifiers.

Aware

- Instrumentation amplifiers.
- Programmable gain amplifiers.
- Isolation amplifiers.

Know the concept of the differential and common-mode signals, and why we split signals up in this way. Know the difference amplifier and its limitations.

Aware of the existence of other types of amplifier, which are based on the op-amp, such as the instrumentation amplifier which is suited to specific applications, such as those found in data acquisition and distribution systems.

27.24 Transfer Functions

Know

- Transfer functions.
- Forced response.
- Frequency response.
- Natural response.
- Complete response.

Know the concept of the transfer function, how to find the forced response and natural response from it.

27.25 Sensor Signal Conditioning

Familiar

- Bridge circuits.

Aware

- Strain, force, pressure and flow measurements.
- High impedance sensors.
- Temperature sensors.

Familiar with bridge circuits and their linearization.

Aware of a few sensors and their signal-conditioning circuits, with an emphasis on bridge circuits which are found in process control systems and data acquisition systems.

27.26 System Modelling

Familiar

- Differential equations of physical systems.
- Linear approximations of physical systems.
- The transfer function.
- Block diagrams.
- Feedback

Familiar with the concept of system modelling and linear approximations, derivation of the transfer function from the describing differential equation, block diagrams, modelling of electromechanical systems and the principles of feedback.

Matrices - Quick Reference Guide

Definitions

Symbol	Description
a_{ij}	Element of a matrix. i is the row, j is the column.
$\mathbf{A}=\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}=[a_{ij}]$	$\mathbf{A}$ is the representation of the matrix with elements a_{ij}.
$\mathbf{x}=\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$	$\mathbf{x}$ is a column vector with elements x_i.
$\mathbf{0}=\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$	Null matrix, every element is zero.
$\mathbf{I}=\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$	Identity matrix, diagonal elements are one.
$\lambda\mathbf{I}=\begin{bmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{bmatrix}$	Scalar matrix.
$\mathbf{\Lambda}=\begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix}$	Diagonal matrix, $a_{ij}=0\ (i\neq j)$.

Multiplication

Multiplication	Description
$\mathbf{Z}=k\mathbf{Y}$	Multiplication by a scalar: $z_{ij}=ky_{ij}$
$\mathbf{z}=\mathbf{Ax}$	Multiplication by a vector: $z_i=\sum_{j=1}^{n}a_{ij}x_j$
$\mathbf{Z}=\mathbf{AB}$	Matrix multiplication: $z_{ij}=\sum_{k=1}^{n}a_{ik}b_{kj}$.
$\mathbf{AB}\neq\mathbf{BA}$	In general, matrix multiplication is not commutative.

M.2

Operations

Terminology	Description
$\mathbf{A}^t = \begin{bmatrix} a_{11} & a_{21} & a_{31} \\ a_{12} & a_{22} & a_{32} \\ a_{13} & a_{23} & a_{33} \end{bmatrix}$	Transpose of **A** (interchange rows and columns): $a_{ij}^t = a_{ji}$.
$\lvert\mathbf{A}\rvert = \det \mathbf{A} = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$	Determinant of **A**. If $\lvert\mathbf{A}\rvert = 0$, then **A** is *singular*. If $\lvert\mathbf{A}\rvert \neq 0$, then **A** is *non-singular*.
$\alpha_{ij} = \begin{vmatrix} a_{11} & \cdots & a_{1j} & \cdots & a_{1n} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{i1} & \cdots & a_{ij} & \cdots & a_{in} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{n1} & \cdots & a_{nj} & \cdots & a_{nn} \end{vmatrix}$	Minor of a_{ij}. Delete the row and column containing the element a_{ij} and obtain a new determinant.
$A_{ij} = (-1)^{i+j}\alpha_{ij}$	Cofactor of a_{ij}.
$\operatorname{adj}\mathbf{A} = \begin{bmatrix} A_{11} & A_{21} & A_{31} \\ A_{12} & A_{22} & A_{32} \\ A_{13} & A_{23} & A_{33} \end{bmatrix}$	Adjoint matrix of **A**. Replace every element a_{ij} by its cofactor in $\lvert\mathbf{A}\rvert$, and then transpose the resulting matrix.
$\mathbf{A}^{-1} = \dfrac{\operatorname{adj}\mathbf{A}}{\lvert\mathbf{A}\rvert}$	Reciprocal of **A**: $\mathbf{A}^{-1}\mathbf{A} = \mathbf{A}\mathbf{A}^{-1} = \mathbf{I}$. Only exists if **A** is square and non-singular. Formula is only used for 3x3 matrices or smaller.

Linear Equations

Terminology	Description
$a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = b_1$ $a_{21}x_1 + a_{22}x_2 + a_{23}x_3 = b_2$ $a_{31}x_1 + a_{32}x_2 + a_{33}x_3 = b_3$	Set of linear equations written explicitly.
$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$	Set of linear equations written using matrix elements.
$\mathbf{Ax} = \mathbf{b}$	Set of linear equations written using matrix notation.
$\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$	Solution to set of linear equations.

Eigenvalues

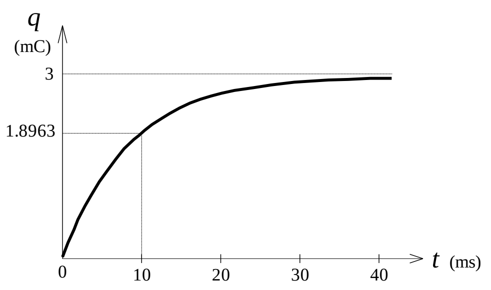
Equations	Description
$\mathbf{Ax} = \lambda\mathbf{x}$	λ are the eigenvalues. $\mathbf{x}$ are the column eigenvectors.
$\lvert\mathbf{A} - \lambda\mathbf{I}\rvert = 0$	Finding eigenvalues.

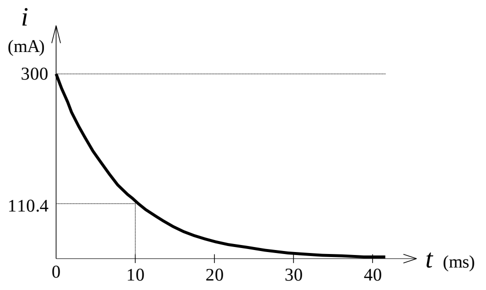
Answers

1.1

(a) To the left (b) $300e^{-100t}$ mA (c) 4.055 ms (d) 0.3820 A/mm^2

(e)





(f) 1.860×10^{16}

1.2

(a) 10 μC (b) 10 μC (c) 12.71 mA

1.3

(a) 975 C (b) 383 C

1.4

(a) 31.1 kC (b) 48 W (c) 373 kJ (d) 24.9 W

1.5

8.0 C

1.6

(a) 120.8 V (b) 8.453 kW (c) 754.0 W/mm^2

1.7

(a) 10 V (b) 5 A (c) 50 W

1.8

(a) -3 A (b) 3 V (c) 15 W

1.9

-30 W

1.10

$P_{v_x} = 68.48\text{ W}$, $P_{50\text{ V}} = -168.9\text{ W}$, $P_{0.2v_x} = -13.70\text{ W}$, $P_{10\,\Omega} = 114.1\text{ W}$

1.11

(a) -20 V, -20 mA, 50 mA (b) $26\tfrac{2}{3}\text{ V}$, $26\tfrac{2}{3}\text{ mA}$, $3\tfrac{1}{3}\text{ mA}$

1.12

(a) $6\tfrac{2}{3}\text{ V}$, $3\tfrac{1}{3}\text{ A}$, $66\tfrac{2}{3}\text{ W}$ (b) 20 V, 10 A, 200 W

1.13

(a) $8\,\Omega$ (b) $3.7\,\Omega$

1.14

(a) 2.5 A (b) 4 V

1.15

$$v_3 = \frac{R_1 R_3}{R_1 + R_2 + R_3} i_s \qquad i_1 = \frac{R_2 + R_3}{R_1 + R_2 + R_3} i_s$$

1.16

5.5 V, 3.975 A

1.17

(a) 30 W (b) -2 A

2.1

(a) 0 V, 1 mA, 1 mA, -10 V, -10 mA, -11 mA

(b) -10 V/V (c) -10 A/A (d) 100 W/W

2.2

(c) $-\frac{R_2}{R_1}$ (b) R_1 (c) same results, thanks to the VSC

2.3

R_1 can be any non-zero value, $R_2 = 200\,\Omega$

2.4

(a) $-R$ (b) $0\,\Omega$

2.5

(a) $-\left[R_1 + \left(1+\frac{R_1}{R_2}\right)R_3\right]$ (b) $0\,\Omega$ (c) $1+\frac{R_1}{R_2}$

2.6

(a) $\left(1+\frac{R_2}{R_1}\right)R_i$ (b) R_i (c) $\frac{R_i}{R_1}$

A.4

3.1

(a) -33 (b) 17, -34, -11

3.2

2 A

3.3

25.64 W

3.4

(a) 6 A (b) 3 A

3.5

-3.5 mA

3.6

20 mA, -80 mW

4.1

80 W

4.2

4 A

4.3

(a) 150 V (b) 110 V

4.4

(a) 15 A, $2\,\Omega$ (b) $2\,\Omega$ (c) 112.5 W

4.5

381 mW

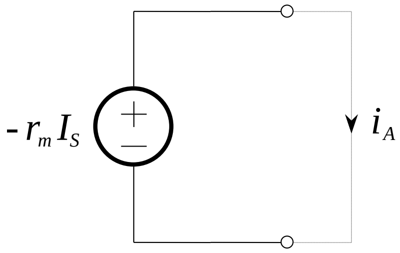
4.6

$8/9\,\Omega$

4.7

(a) $i_A = \dfrac{R_2 I_S}{r_m - R_2}$

(b) It is indeterminant – $i_A = \pm\infty\,\text{A}$. For this special case you can show that KVL is violated – the circuit becomes:



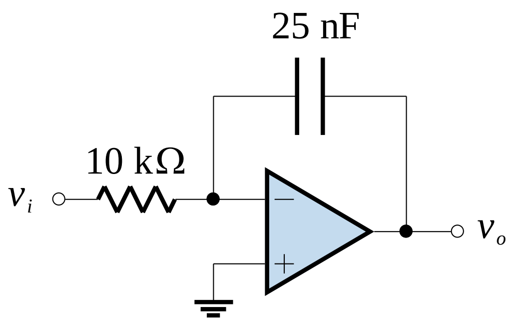
Thus, the "branch" containing i_A, which has 0 V across it, appears in parallel with a Thévenin equivalent ideal voltage source equal to $V_{Th} = -r_m I_S$, which is impossible.

Note that there is no Norton equivalent circuit!

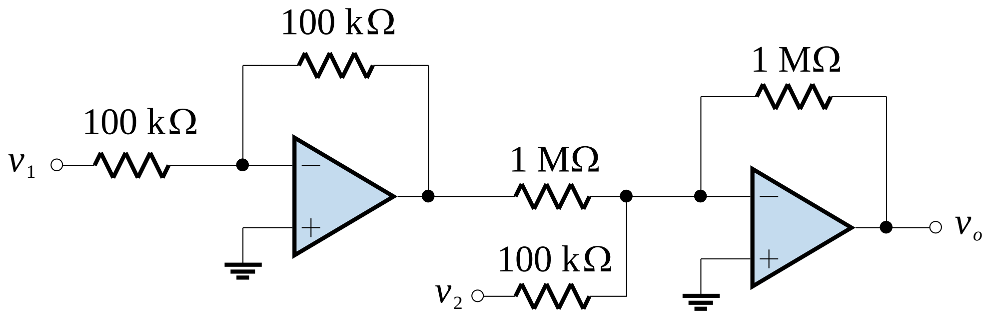
(c) In a practical setting, it is impossible to create ideal circuit elements, so the real circuit's behaviour would be perfectly explainable and measureable if modelled correctly (inclusion of finite wire resistance, linear range of power supplies, etc).

A large but finite current i_A would result – either briefly (before a fuse blows or a protection device trips), or continuously (limited by the power supply's output current and voltage capability).

5.1



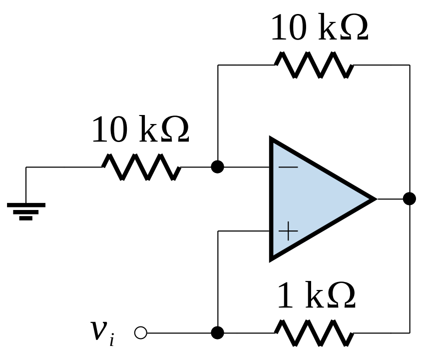
5.2



5.3

$$v_o = v_1 + v_2 - 2v_3$$

5.4



$v_i = 10$ mV when the source is attached

5.5

$$i_o = -\frac{v_i}{R}$$

6.1

(a) 20.6 ms (b) 177.7 ms

6.2

$2\sqrt{t-0.001}$ A, $0.01/\sqrt{t-0.001}$ V

6.3

(a) 9.6 V, 192 mW, 1.152 mJ (b) 16 V, 0 W, 3.20 mJ

6.4

(a) 2 nF (b) 2.4 nF

6.5

(a) $\frac{12}{7}$ μF (b) $\frac{12}{11}$ μF (c) 9 μF

6.6

(a) $60\cos 10t$ V (b) $5+2\sin 10t$ A

6.7

(a) $2\times 10^{-4}\frac{dv}{dt}+2\times 10^{-4}v+25\int_0^t v\,dt=i_s+i_L(0)$

(b) $0.04\frac{di}{dt}+5000i+5000\int_0^t i\,dt=v_s-v_C(0)$

6.8

Use v_1, v_2, v_C left to right.

$$v_1-v_C=3v_1$$

$$10\int_0^t (v_1-v_2)dt+0.5+0.8e^{-100t}+\frac{v_1}{20}+\frac{(v_C-v_2)}{50}+2\times 10^{-4}\frac{dv_C}{dt}=0$$

$$10\int_0^t (v_2-v_1)dt-0.5-0.8e^{-100t}+\frac{(v_2-v_C)}{50}=0$$

8.1

$100\ \Omega,\ \ 2\ \mu\text{F}$

8.2

(b) 50 mC (b) 38.9 mC

8.3

76.10 V

8.4

$3e^{t/0.003}$ mA

8.5

$t > 122.6\ \text{s}$ (over 2 minutes)

8.6

$-6.75e^{-t/40}\ \mu\text{A}$

8.7

153.7 ms

8.8

(a) 9.6 A (b) 2.4 A (c) $9.6e^{-2t}$ A (d) $2.4e^{-4t}$ A (e) $19.2\left(e^{-4t} - e^{-2t}\right)$ V

8.9

(a) 28.95 ms (b) 144.3 ms

8.10

(a) 800 mA (b) 280 mA

8.11

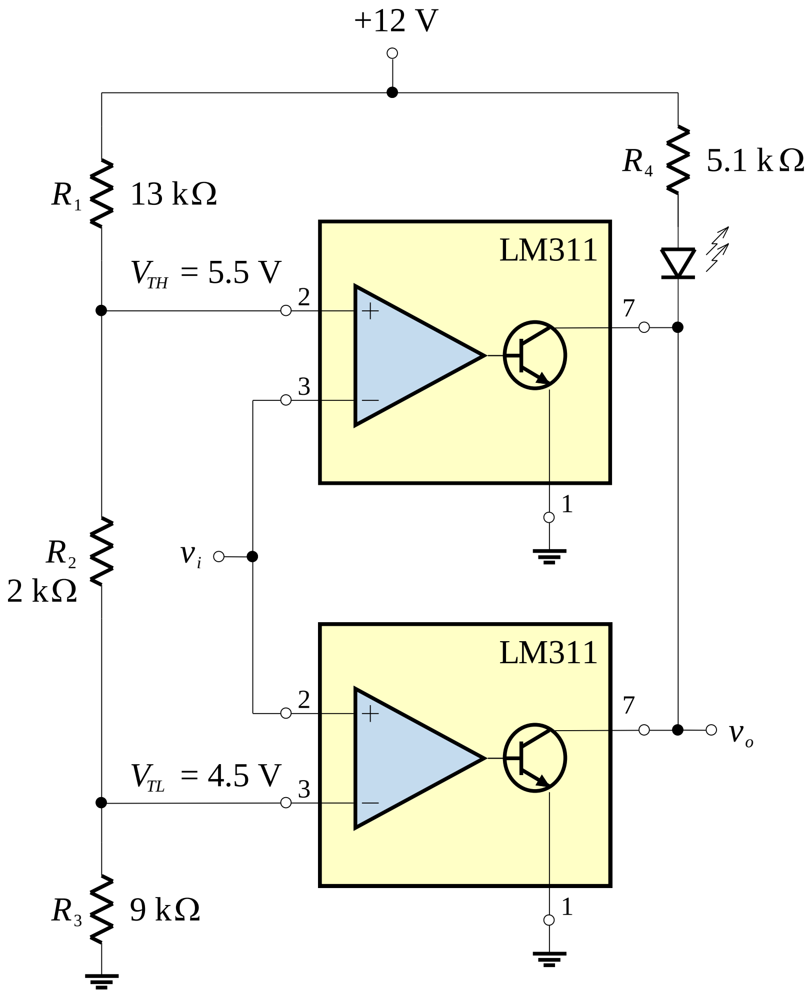
$10e^{-80t}$ V

8.12

$60 + 4e^{-250t} - 40e^{-200t}$ mA

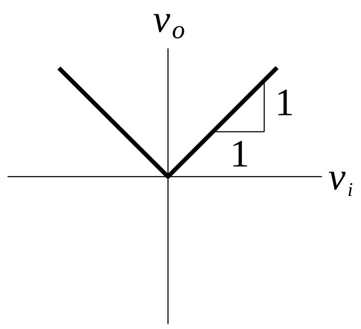
A.10

9.1



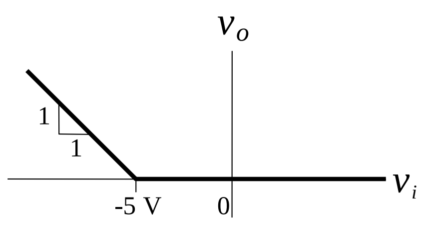
+12 V
R1 13 kΩ
R4 5.1 kΩ
LM311
VTH = 5.5 V
2
3
7
1
R2 2 kΩ
vi
LM311
2
3
7
1
VTL = 4.5 V
vo
R3 9 kΩ

9.2



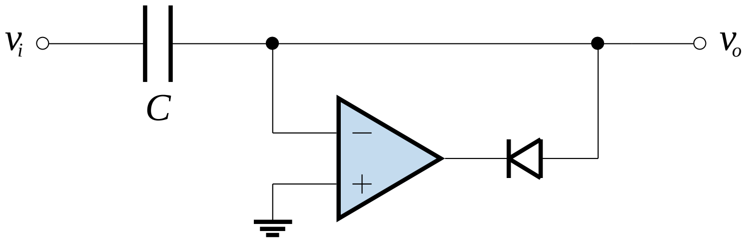
vo
1
1
vi

9.3

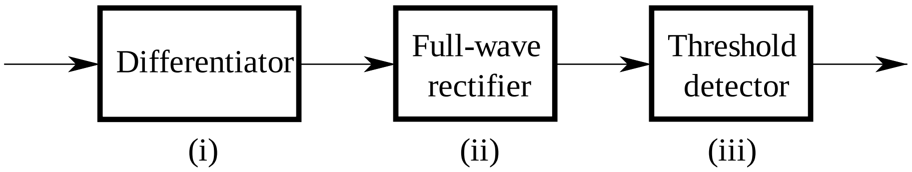


vo
1
1
vi
-5 V
0

9.4


v_i
C
v_o

9.5


Differentiator
Full-wave rectifier
Threshold detector
(i)
(ii)
(iii)

A.12

10.1

$$20\sum_{n=0}^{\infty}(-1)^n u(t-0.2n)\text{ V}$$

10.2

$$50\left(1-e^{-10t}\right)u(\text{t})\text{ V},\ \left(20-25e^{-10t}\right)u(\text{t})\text{ V}$$

10.3

$$20\left(1-e^{-10^6 t/3}\right)u(\text{t})\text{ V}$$

10.4

(a) $100u(-\text{t})+\left(40+60e^{-6250t}\right)u(\text{t})\text{ V}$

(b) 100 V

10.5

$76.28\text{ k}\Omega,\ 62.13\ \mu\text{F}$

10.6

$10+30e^{-40000t}\text{ V}$

10.7

22.31 mA, 9.812 mA

10.8

(a) $0.4\left(1-e^{-1250t}\right)$ A

(b) $10/23\left(e^{-100t}-e^{-1250t}\right)$ A

(c) $0.2\left[\sqrt{2}\cos(1250t-45°)-e^{-1250t}\right]$ A

10.9

$\left(-40-200e^{-15000t}\right)u(\text{t})$ V

10.10

(a) 0 W (b) 200 W (c) 131.7 W (d) 0 W

11.1

(a) $A_{OL} \geq 200799 \text{ V/V}$

(b) $V_{OS} \leq 497.5\,\mu\text{V}$

(c) $R_2 \leq 500\,\text{k}\Omega$

12.1

(a) 12.5 ms, 80 Hz, 502.7 rad/s

(b) $27.77\cos(160\pi t - 43.92°)$ V

(c) $66.08°$

12.2

(a) 8.00 and $-38.68°$ (b) 11.17 ms (c) 89.54 Hz (d) 562.6 rad/s

(e) 8.00 and $-128.7°$ (f) 0.8983

12.3

$412.3\cos(500t - 116.0°)$ V

12.4

$80\cos(2000t - 36.87°)$ mA

12.5

$\sqrt{2}\cos(5t - 45°) + 1.342\cos(10t - 63.44°)$ A

12.6

(a) $18.83\angle 133.5°$ (b) $5.584\angle -56.87°$

(c) $-0.07248 + j0.04702$ (d) $0.05243 + j0.1838$

12.7

$15.98\cos(1000t - 71.23°) + 4.598\cos(500t + 38.22°)$ V

12.8

(a) $95.79\cos(\omega t + 94.01°)$ mA (b) $25.14\angle -111.6°$ mA

12.9

(a) 143.5 W (b) -135.0 W

12.10

(a) 39.99 W (b) 9.512 W (c) -9.512 W

12.11

(a) $26.00\angle -47.38°$ mA (b) $0.7692\angle 37.38°$ A

12.12

(a) 384.2 and 65.92 Hz (b) 203.8 and 124.3 Hz

12.13

1.25 mH

12.14

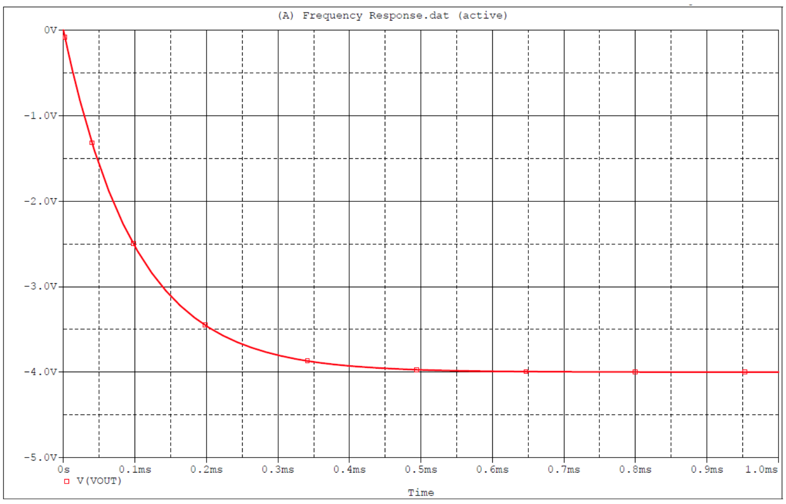
125 nF

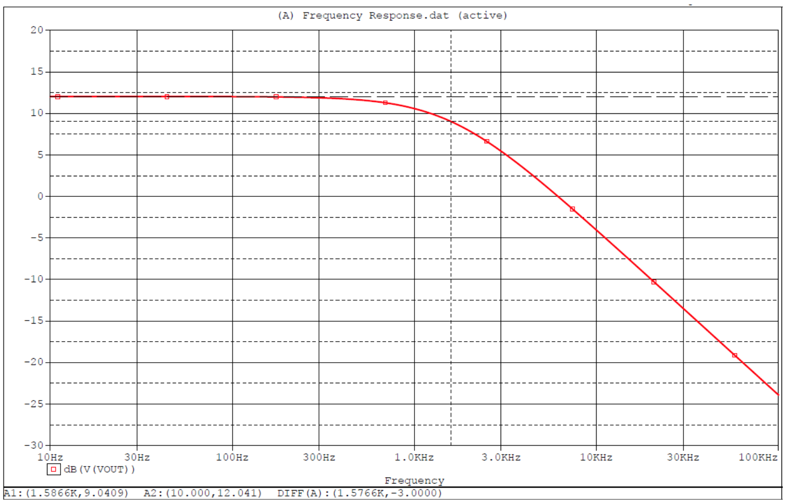
12.15

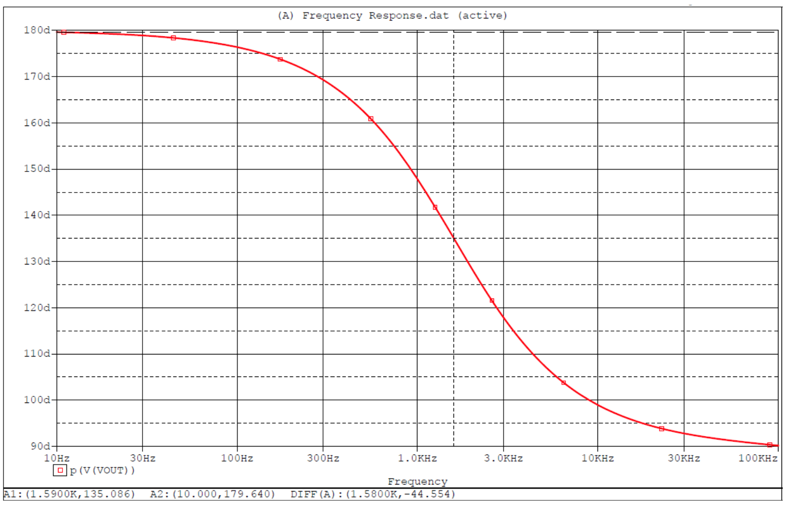
1 Ω, 1 H

A.16

13.1


(A) Frequency Response.dat (active)
0V
-1.0V
-2.0V
-3.0V
-4.0V
-5.0V
0s
0.1ms
0.2ms
0.3ms
0.4ms
0.5ms
0.6ms
0.7ms
0.8ms
0.9ms
1.0ms
V(VOUT)
Time


(A) Frequency Response.dat (active)
20
15
10
5
0
-5
-10
-15
-20
-25
-30
10Hz
30Hz
100Hz
300Hz
1.0KHz
3.0KHz
10KHz
30KHz
100KHz
dB(V(VOUT))
Frequency
A1:(1.5866K,9.0409) A2:(10.000,12.041) DIFF(A):(1.5766K,-3.0000)


(A) Frequency Response.dat (active)
180d
170d
160d
150d
140d
130d
120d
110d
100d
90d
10Hz
30Hz
100Hz
300Hz
1.0KHz
3.0KHz
10KHz
30KHz
100KHz
p(V(VOUT))
Frequency
A1:(1.5900K,135.086) A2:(10.000,179.640) DIFF(A):(1.5800K,-44.554)

14.1

(a) $7.809\angle -128.7^\circ$ V (b) $4\sqrt{6}\ \Omega$

14.2

(a) $6.468\cos\left(10^4 t + 44.04^\circ\right)$ V (b) $3.234\cos\left(10^4 t + 44.04^\circ\right)$ A

14.3

(a) $11.09\cos\left(10^4 t + 33.69^\circ\right)$ A (b) $3.288\cos\left(10^4 t + 170.5^\circ\right)$ A

14.4

(a) $-4 + \sqrt{2}\cos\left(10^6 t - 45^\circ\right)$ V (b) $\sqrt{2}\cos\left(10^6 t + 135^\circ\right)$ V

14.5

$6 + j17\ \Omega$

14.6

$1.581\angle -18.43^\circ$ A

14.7

65.05° and -60.72°, or -112.8° and 13.00°

14.8

$p_{R_1} = 4.8$ mW, $p_L = 110.9$ mW, $p_C = -110.9$ mW and $p_{R_2} = 19.2$ mW

14.9

(a) $P_{20} = 10$ kW, $P_{10} = 5$ kW (b) $P_{20} = 3.125$ kW, $P_{10} = 6.25$ kW

14.10

6.622 mA

14.11

(a) $\frac{300}{2\sqrt{2}} \approx 106.1\,\Omega$ (b) 61.24 V RMS

14.12

(a) 4.471 A RMS (b) 0.9150 lagging

14.13

(a) 46.86 kW (b) 33.32 kvar (c) $57.5\angle 35.41°$ kVA

(d) 57.5 kVA (e) $92\angle 35.41°\,\Omega$

14.14

There are two possible solutions:

$$\mathbf{S}_V = 98.04 - j35.57\text{ VA} \qquad \mathbf{S}_V = -118.0 - j124.4\text{ VA}$$
$$\mathbf{S}_I = 81.96 - j124.4\text{ VA} \quad \text{or} \quad \mathbf{S}_I = 298.0 - j35.57\text{ VA}$$

16.1

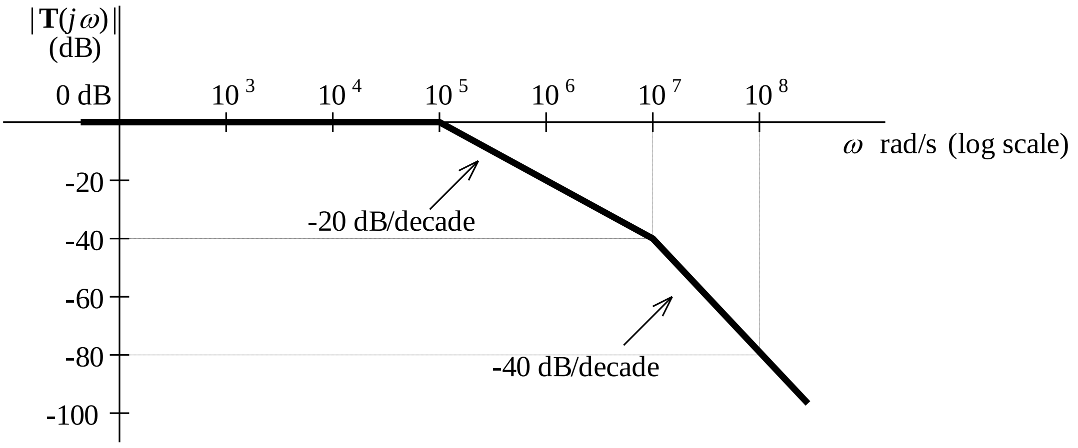
(a) 2 kHz (b) 2.5 Hz

16.2

(a) 0 dB (b) 32.04 dB (c) -6.021 dB

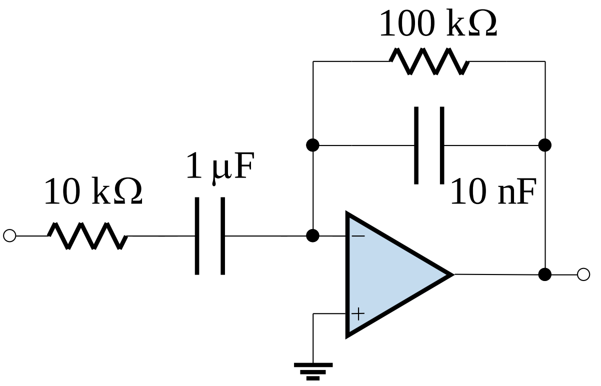
A.20

17.1



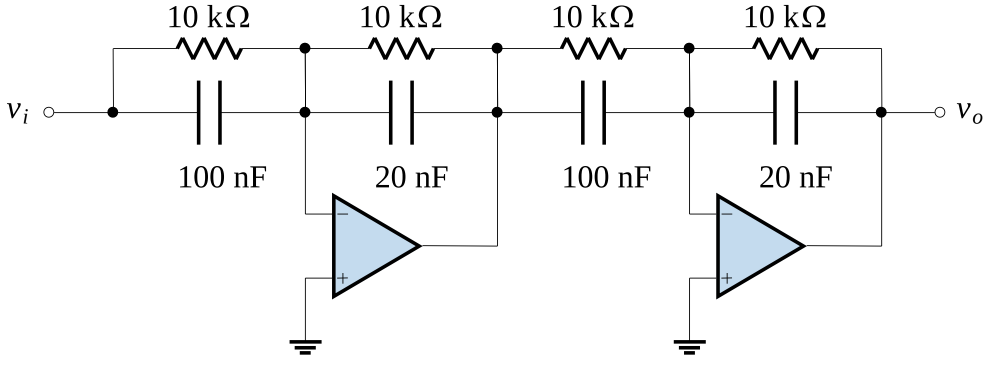
17.2

One possible solution is:



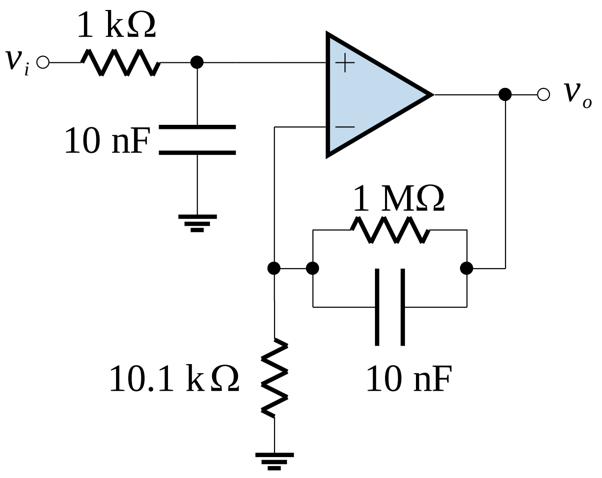
17.3

One possible solution is:



17.4

One possible solution is:



18.1

(a) $200e^{-2t} - 100e^{-8t}$ V, $t > 0$ (b) $-100e^{-2t} + 200e^{-8t}$ V, $t > 0$

18.2

(a) $-40e^{-500t} + 55e^{-2000t}$ V, $t > 0$ (b) 15.00 V and -17.00 V

18.3

(a) 25 Ω (b) 1.28 mJ

18.4

(a) $^{14}\!/_{125}$ H, $^{1}\!/_{70}$ F (b) 3.630 J

18.5

12.89 kV

18.6

$200\left(e^{-5000t} - e^{-15000t}\right)$ V, $t > 0$

18.7

-9.992 kV

18.8

(a) 500 J in *L*, 80 J in *C* (b) 335.4 J in *L*, 62.21 J in *C*

18.9

$20 + 250\left(e^{-100t} - e^{-150t}\right)$ V, $^{1}\!/_{24}\left(3e^{-150t} - 2e^{-100t}\right)$ A

18.10

(a) $-50 + {}^{25}\!/_{3}\left(16e^{-2000t} - e^{-8000t}\right)$ V, $t > 0$

(b) $-2 + {}^{5}\!/_{3}\left(4e^{-2000t} - e^{-8000t}\right)$ A, $t > 0$

18.11

(a) 0 A (b) 1.748 A (c) -1.073 A

20.1

(a) $1000\ \text{rads}^{-1}$, 5 (b) $120\ \text{krads}^{-1}$, 60 (c) $602.1\ \text{rads}^{-1}$, 6.021

20.2

(a) $4\ \text{krads}^{-1}$, 40, $80\cos(4000t)$ V

(b) $2\cos(4000t)$ mA, $400\sin(4000t)$ mA, $-400\sin(4000t)$ mA

(c) 20 mW, 4 mJ

20.3

$115.5\ \text{rads}^{-1}$

20.4

(a) $4.999\ \text{Mrads}^{-1}$ (b) 49.99 (c) $100.0\ \text{krads}^{-1}$

(d) $4.949\ \text{Mrads}^{-1}$ (e) $5.049\ \text{Mrads}^{-1}$ (f) $40\angle -88.85°$ mV

20.5

$5\ \text{k}\Omega$, $2.360\ \mu\text{H}$, $4.237\ \text{nF}$

20.6

(a) $5\ \text{Mrads}^{-1}$ (b) $20.59\ \text{krads}^{-1}$

20.7

$4.472\ \text{krads}^{-1}$, 22.36

20.8

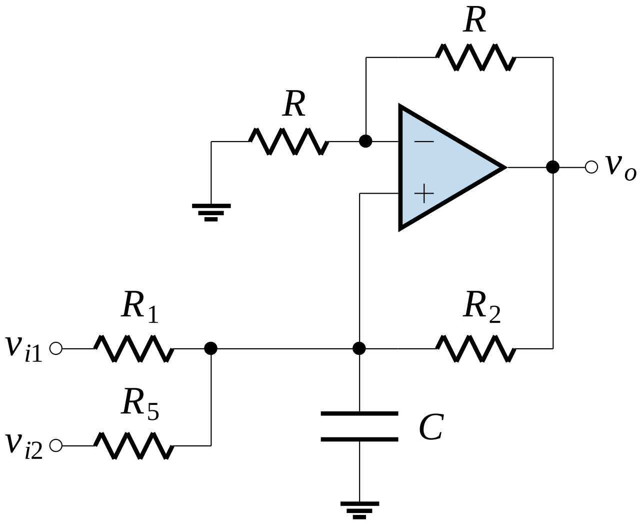
(a) $5\ \text{krads}^{-1}$, 40.01 (b) $5\ \text{krads}^{-1}$, 20.03

20.9

$10\ \Omega$, 514.3 mH, $875\ \mu\text{F}$

21.1

(a) If we analyse the circuit:



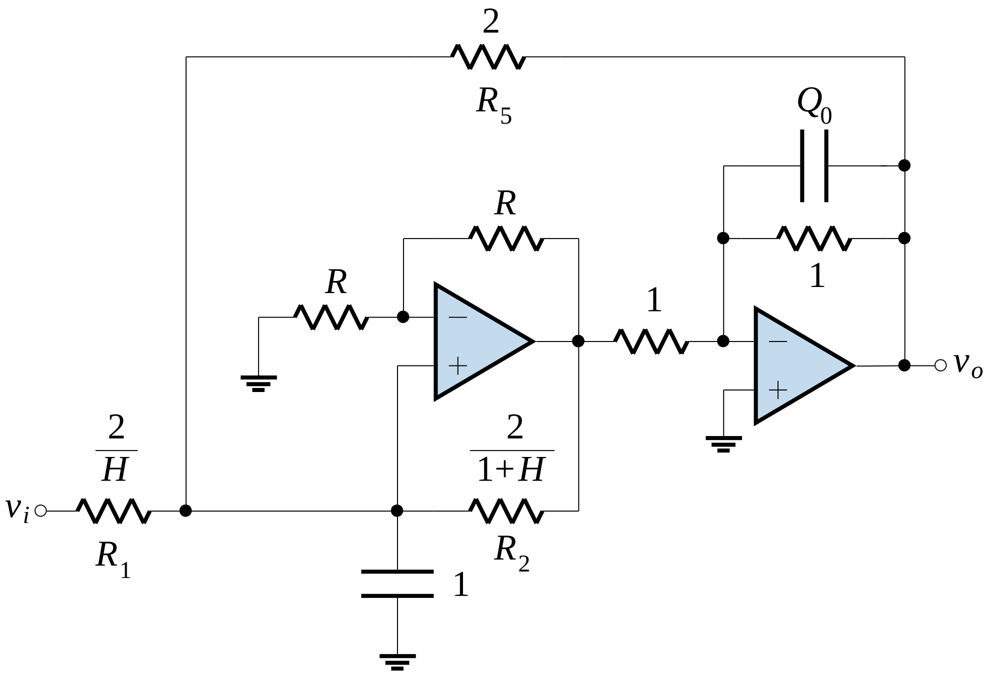
then the output voltage phasor is found to be:

$$\mathbf{V}_o = \frac{2R_2\left(R_5\mathbf{V}_{i1} + R_1\mathbf{V}_{i2}\right)}{j\omega C R_1 R_2 R_5 + R_1 R_2 + R_2 R_5 - R_1 R_5}$$

Note that we must have $R_2 = R_1 \| R_5$ for the circuit to be a pure integrator. If $R_1 = R_5 = 2R$ then $R_2 = R$ and we have:

$$\mathbf{V}_o = \frac{2\mathbf{V}_{i1}}{j\omega C R_1} + \frac{2\mathbf{V}_{i2}}{j\omega C R_5} = \frac{1}{j\omega C R}\mathbf{V}_{i1} + \frac{1}{j\omega C R}\mathbf{V}_{i2}$$

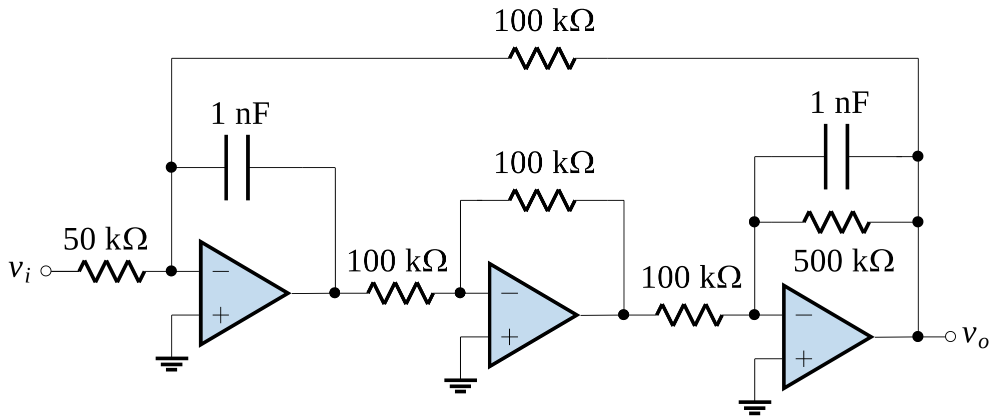
(b) The circuit is a noninverting integrator, so that it accomplishes the same objective as cascading an inverting integrator and an inverter. A normalized version of the biquad with the noninverting integrator is shown below:



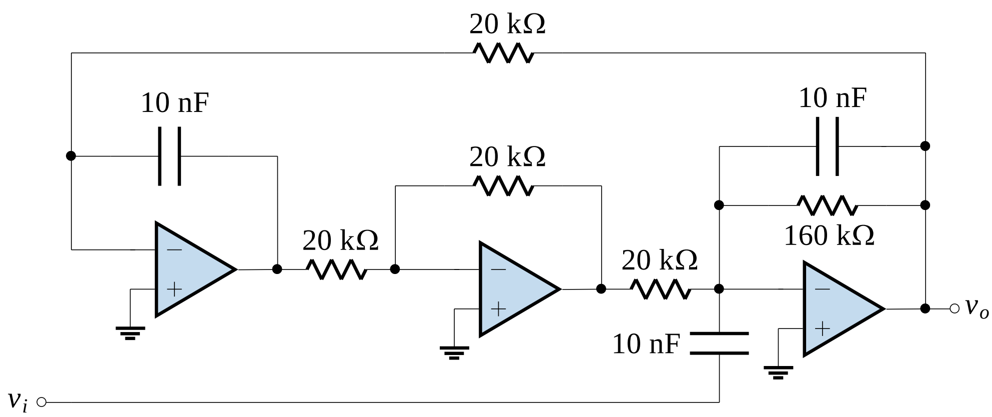
Note that we must have $R_2 = R_1 \parallel R_5$ for the noninverting integrator to be a pure integrator, and if $H = 1$ then $R_1 = R_5 = 2$ and $R_2 = 1$.

21.2

(a) One possible solution is:



(b) One possible solution is:



21.3

$$\mathbf{T}(j\omega) = \frac{1}{1-\omega^2 + j\omega(1/Q_0)}$$

21.4

$$\mathbf{T}(j\omega) = \frac{-j\omega 2Q_0}{1-\omega^2 + j\omega(1/Q_0)}$$

22.1

(a) -10, -40 (b) $-16 \pm j12$

22.2

(a) $5.754\angle -58.50°$ mA (b) $7.211\angle -33.69°$ mA

22.3

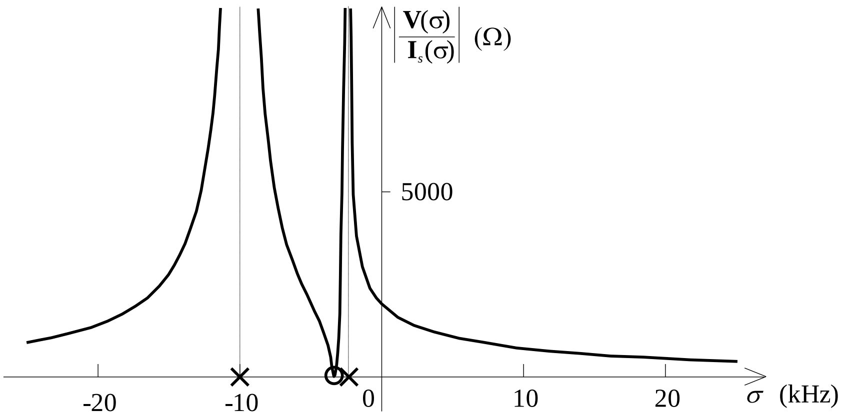
(a) 10.00 A (b) -1.995 A (c) -97.01 mA (d) 53.90 mA

22.4

2.508 J

22.5

Zeros: $\mathbf{s} = -3333, \infty$; poles $\mathbf{s} = -2500, -10\,000$



22.6

(a) 100 Ω (b) 12.5 H (c) 689.7 μF

22.7

(a) $53.85\angle 21.80°$ from zero at $\mathbf{s} = -50$,

$53.85\angle 68.20°$ from pole at $\mathbf{s} = -20 - j30$,

$22.36\angle -26.57°$ from pole at $\mathbf{s} = -20 + j30$

(b) $1.137\angle 42.13°$

23.1

(a) $v_{o1} = \left(1 + \frac{R_1}{R_G}\right)v_{i1} - \frac{R_1}{R_G}v_{i2}$

$v_{o2} = \left(1 + \frac{R_2}{R_G}\right)v_{i2} - \frac{R_2}{R_G}v_{i1}$

(b) $A_d = 201$

(c) 1.005 V

(d) $200.2\,\Omega$

23.2

$A_d = 10$

24.1

(a), (b) and (c) -2 and -5

24.2

$\left(15/8\sqrt{2}e^{-t}\cos(2t+45°)+9/8e^{-3t}\right)u(t)$ V

24.3

$\left(6/5+4/5e^{-25t}-2e^{-10t}\right)u(t)$ A

24.4

(a) 0 (b) $10/(\mathbf{s}+17.5)\,\Omega$ (c) 0 (d) $Ae^{-17.5t}$ V

24.5

(a) $\mathbf{I}/\mathbf{V}_s=5/\left[(\mathbf{s}+2)(\mathbf{s}^2+2\mathbf{s}+5)\right]$, $\mathbf{s}=-2,-1\pm j2$

(b) $i_f(t)=1\,\text{A}$

(c) $i_n(t)=Ae^{-2t}+Be^{-t}\cos(2t)+Ce^{-t}\sin(2t)$

(d) $i(t)=\left(1-e^{-2t}-e^{-t}\sin(2t)\right)u(t)\,\text{A}$

24.6

(a) $$\begin{bmatrix} R_1+\mathbf{s}L & -R_1 \\ -R_1 & R_1+R_2+\dfrac{1}{\mathbf{s}C}\end{bmatrix}\begin{bmatrix}\mathbf{I}_1\\ \mathbf{I}_2\end{bmatrix}=\begin{bmatrix}\mathbf{V}_s\\ 0\end{bmatrix}$$

(b) $$\frac{\mathbf{I}_2}{\mathbf{V}_s}=\frac{\dfrac{R_1}{(R_1+R_2)L}\mathbf{s}}{\mathbf{s}^2+\left(\dfrac{R_1R_2}{(R_1+R_2)L}+\dfrac{1}{(R_1+R_2)C}\right)\mathbf{s}+\dfrac{R_1}{(R_1+R_2)LC}}$$

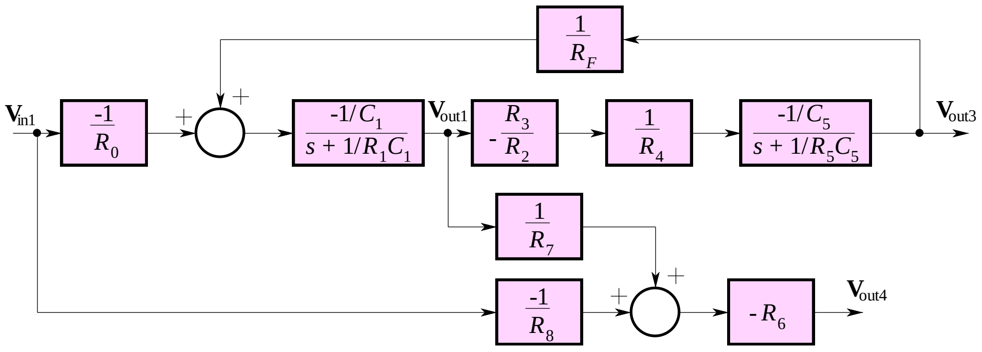
(c) $$i_2(t)=\frac{3}{500}\left(e^{-1000/3t}-e^{-500t}\right)u(t)$$

25.1

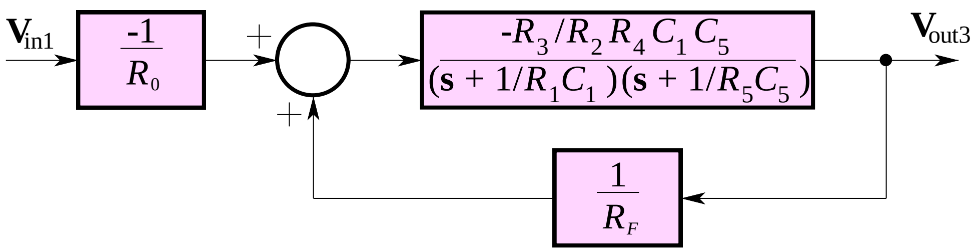
(b) $v_o = -\frac{V_B}{R_1 + R_2}\Delta R$

26.1

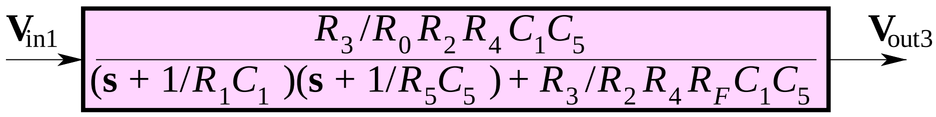
(a)



(b) Considering just the output $\mathbf{V}_{\text{out3}}$, this can be reduced to:



Further reduction gives:



Expansion of the denominator results in the given transfer function.

(c)

$$K_1 = \frac{R_1R_3R_5R_F}{R_0\left(R_1R_3R_5 + R_2R_4R_F\right)}$$

$$\alpha = \frac{1}{2}\left(\frac{1}{R_1C_1} + \frac{1}{R_5C_5}\right)$$

$$\omega_0 = \sqrt{\frac{R_3}{R_2R_4R_FC_1C_5} + \frac{1}{R_1R_5C_1C_5}}$$

(d)

$$\alpha = \frac{1}{2}\left(\frac{1}{R_1C_1} + \frac{1}{R_1C_1}\right) = \frac{1}{2}\frac{2}{R_1C_1} = \frac{1}{R_1C_1}$$

$$\omega_d = \sqrt{\frac{R_3}{R_2R_4^2C_1^2} + \frac{1}{R_1^2C_1^2} - \frac{1}{R_1^2C_1^2}} = \sqrt{\frac{R_3}{R_2}}\frac{1}{R_4C_1}$$

(e) From the block diagram, we can see that:

$$\frac{\mathbf{V}_{\text{out3}}}{\mathbf{V}_{\text{out1}}} = \frac{R_3}{R_2R_4C_5}\frac{1}{\mathbf{s} + 1/R_5C_5}$$

Then:

$$\frac{\mathbf{V}_{\text{out1}}}{\mathbf{V}_{\text{in1}}} = \frac{\mathbf{V}_{\text{out3}}}{\mathbf{V}_{\text{in1}}}\Bigg/\frac{\mathbf{V}_{\text{out3}}}{\mathbf{V}_{\text{out1}}} = \frac{\dfrac{\mathbf{s} + 1/R_5C_5}{R_0C_1}}{\mathbf{s}^2 + \left(\dfrac{1}{R_1C_1} + \dfrac{1}{R_5C_5}\right)\mathbf{s} + \dfrac{R_3}{R_2R_4R_FC_1C_5} + \dfrac{1}{R_1R_5C_1C_5}}$$

For the special case of $R_1 = R_5$, $C_1 = C_5$ and $R_F = R_4$, this reduces to:

$$\frac{\mathbf{V}_{\text{out1}}}{\mathbf{V}_{\text{in1}}} = \frac{1}{R_0C_1}\frac{(\mathbf{s} + \alpha)}{\mathbf{s}^2 + 2\alpha\mathbf{s} + \omega_0^2} = K_2\frac{(\mathbf{s} + \alpha)}{\mathbf{s}^2 + 2\alpha\mathbf{s} + \omega_0^2}$$

(f) The output voltage $\mathbf{V}_{\text{out4}}$ is given by:

$$\mathbf{V}_{\text{out4}} = \frac{R_6}{R_8}\mathbf{V}_{\text{in1}} - \frac{R_6}{R_7}\mathbf{V}_{\text{out1}}$$

Then substituting for $\mathbf{V}_{\text{out1}}$ gives:

$$\mathbf{V}_{\text{out4}} = \frac{R_6}{R_8}\mathbf{V}_{\text{in1}} - \frac{R_6}{R_7}\frac{1}{R_0C_1}\frac{\mathbf{s} + \alpha}{\mathbf{s}^2 + 2\alpha\mathbf{s} + \omega_0^2}V_{\text{in1}}$$

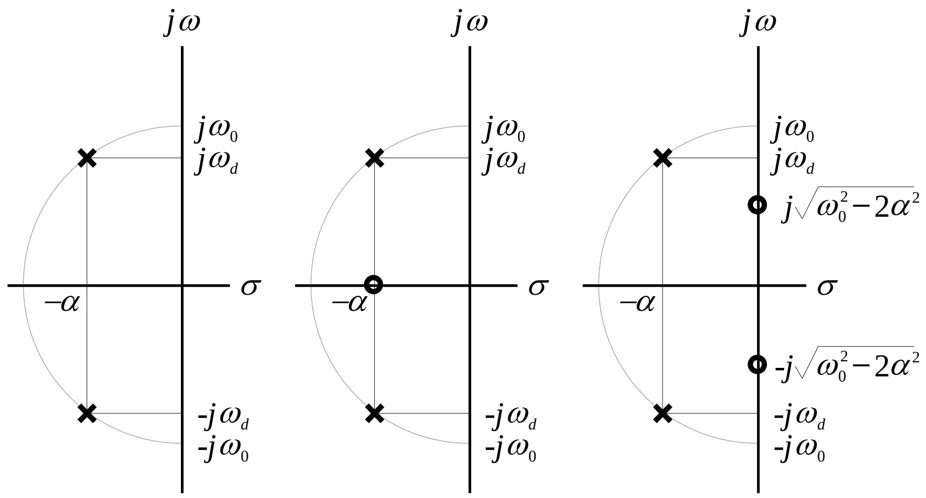
The transfer function is then:

$$\mathbf{T}_3(s) = \frac{R_6}{R_8}\left[\frac{\mathbf{s}^2 + 2\alpha\mathbf{s} + \omega_0^2}{\mathbf{s}^2 + 2\alpha\mathbf{s} + \omega_0^2} - \frac{R_8}{R_0 R_7 C_1}\frac{\mathbf{s} + \alpha}{\mathbf{s}^2 + 2\alpha\mathbf{s} + \omega_0^2}\right]$$

Substituting the special conditions on the values gives:

$$\mathbf{T}_3(s) = \frac{\mathbf{s}^2 + \left(2\alpha - \frac{R_8}{R_0 R_7 C_1}\right)\mathbf{s} + \omega_0^2 - \frac{\alpha R_8}{R_0 R_7 C_1}}{\mathbf{s}^2 + 2\alpha\mathbf{s} + \omega_0^2} = \frac{\mathbf{s}^2 + \omega_0^2 - 2\alpha^2}{\mathbf{s}^2 + 2\alpha\mathbf{s} + \omega_0^2}$$

(g) The pole-zero plots for $\mathbf{T}_1(s)$, $\mathbf{T}_2(s)$ and $\mathbf{T}_3(s)$ are respectively:



(h) Lowpass, lowpass, notch.

Index

A

B

E

F

G

Q

R

S

T

U

V

W

Z