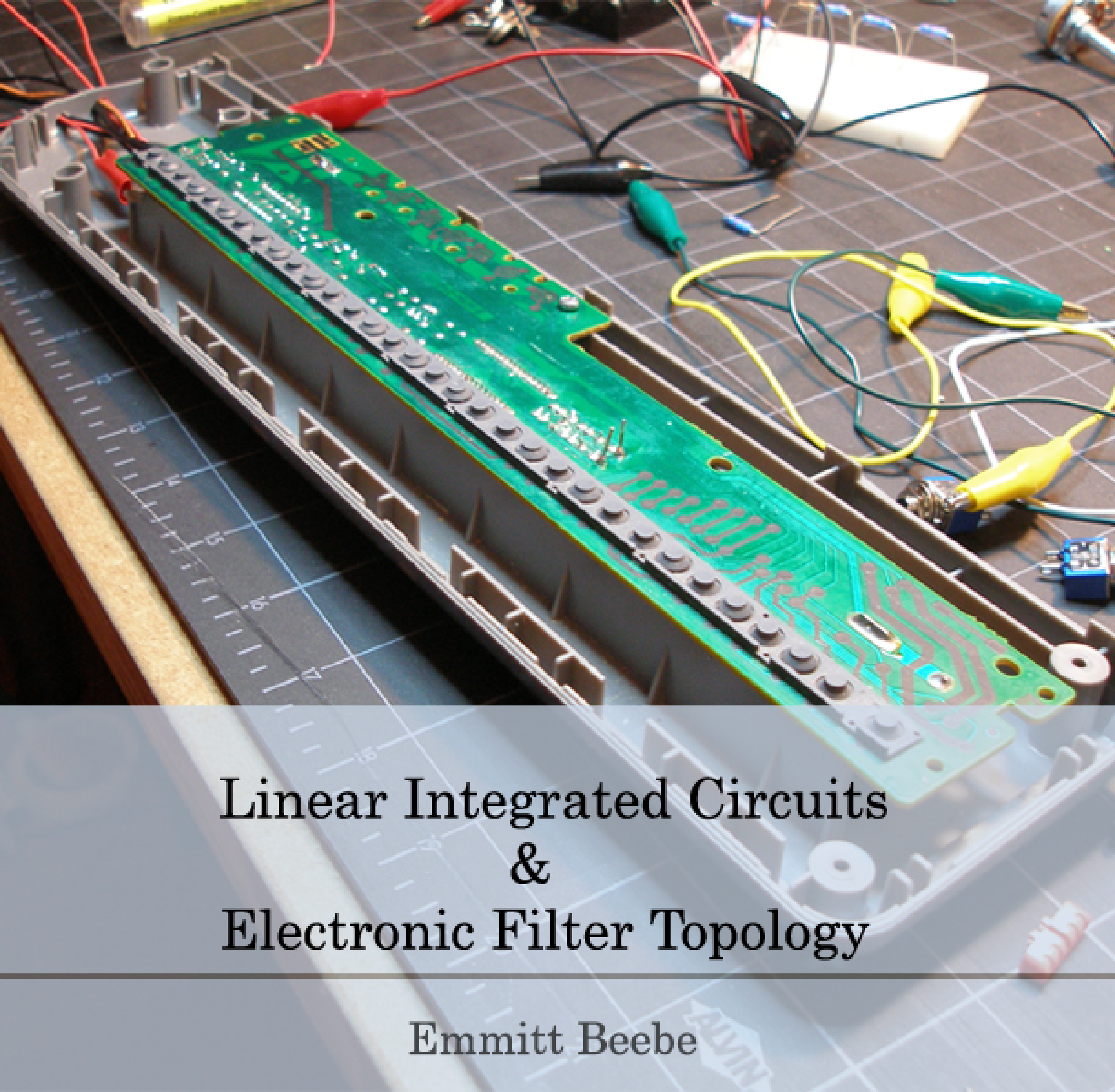


Linear Integrated Circuits & Electronic Filter Topology

Emmitt Beebe

First Edition, 2012

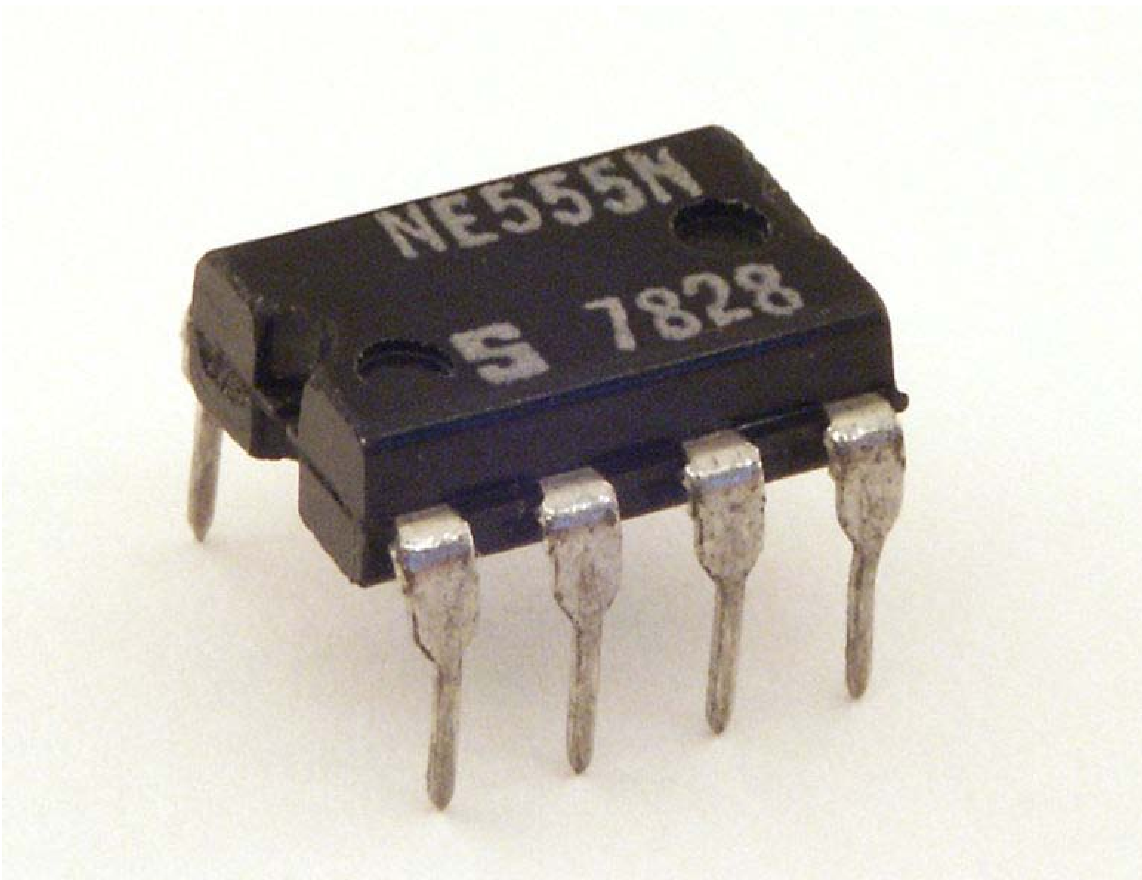
ISBN 978-81-323-3079-0

Published by:
Research World
4735/22 Prakashdeep Bldg,
Ansari Road, Darya Ganj,
Delhi - 110002
Email: info@wtbooks.com

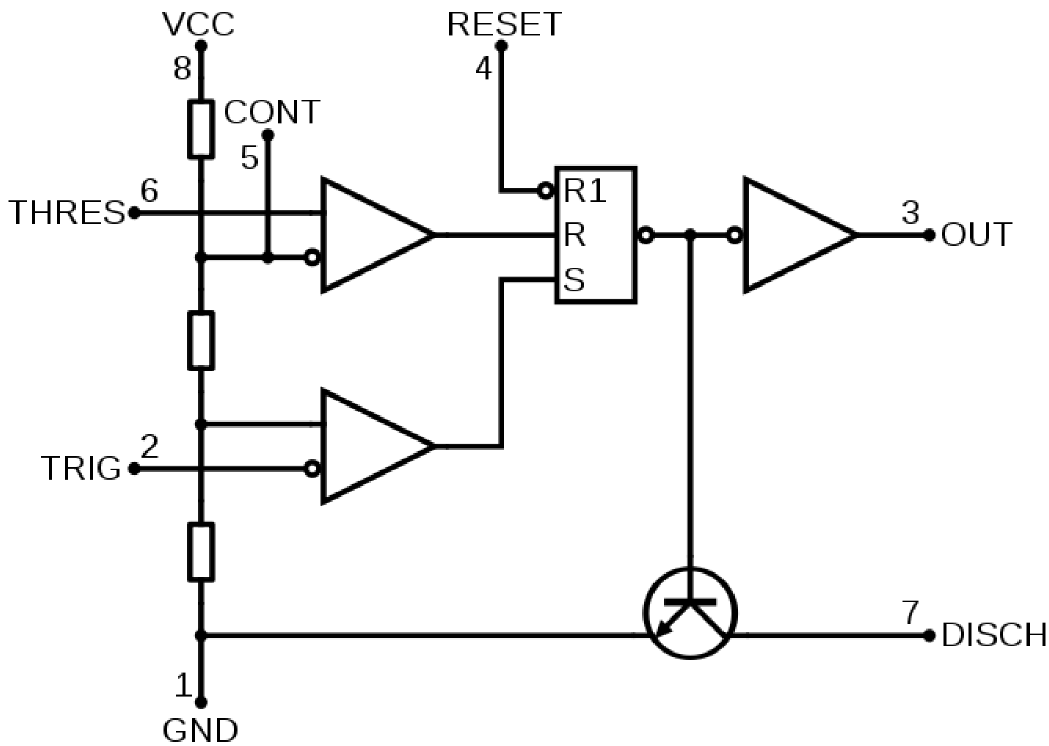
Table of Contents

Chapter 1

555 Timer IC



NE555 from Signetics in dual-in-line package



Internal block diagram

The **555 Timer IC** is an integrated circuit (chip) used in a variety of timer, pulse generation and oscillator applications. The IC was designed by Hans R. Camenzind in 1970 and brought to market in 1971 by Signetics (later acquired by Philips). The original name was the SE555 (metal can)/**NE555** (plastic DIP) and the part was described as "The IC Time Machine". It has been claimed that the 555 gets its name from the three 5 kΩ resistors used in typical early implementations, but Hans Camenzind has stated that the number was arbitrary. The part is still in wide use, thanks to its ease of use, low price and good stability. As of 2003, it is estimated that 1 billion units are manufactured every year.

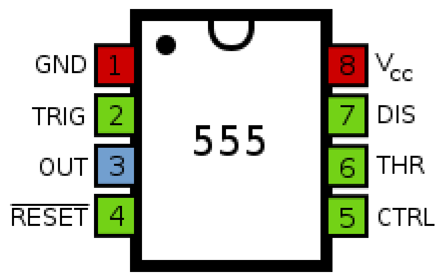
Depending on the manufacturer, the standard 555 package includes over 20 transistors, 2 diodes and 15 resistors on a silicon chip installed in an 8-pin mini dual-in-line package (DIP-8). Variants available include the 556 (a 14-pin DIP combining two 555s on one chip), and the 558 (a 16-pin DIP combining four slightly modified 555s with DIS & THR connected internally, and TR falling edge sensitive instead of level sensitive).

Ultra-low power versions of the 555 are also available, such as the 7555 and TLC555. The 7555 is designed to cause less supply glitching than the classic 555 and the manufacturer claims that it usually does not require a "control" capacitor and in many cases does not require a power supply bypass capacitor.

The 555 has three operating modes:

- Monostable mode: in this mode, the 555 functions as a "one-shot". Applications include timers, missing pulse detection, bouncefree switches, touch switches, frequency divider, capacitance measurement, pulse-width modulation (PWM) etc
- Astable - free running mode: the 555 can operate as an oscillator. Uses include LED and lamp flashers, pulse generation, logic clocks, tone generation, security alarms, pulse position modulation, etc.
- Bistable mode or Schmitt trigger: the 555 can operate as a flip-flop, if the DIS pin is not connected and no capacitor is used. Uses include bouncefree latched switches, etc.

Usage

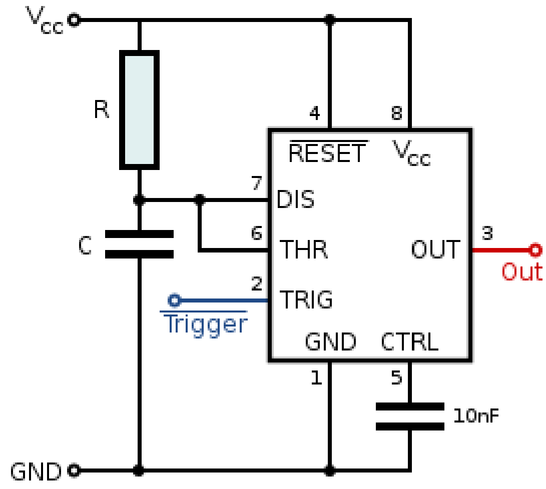


Pinout diagram

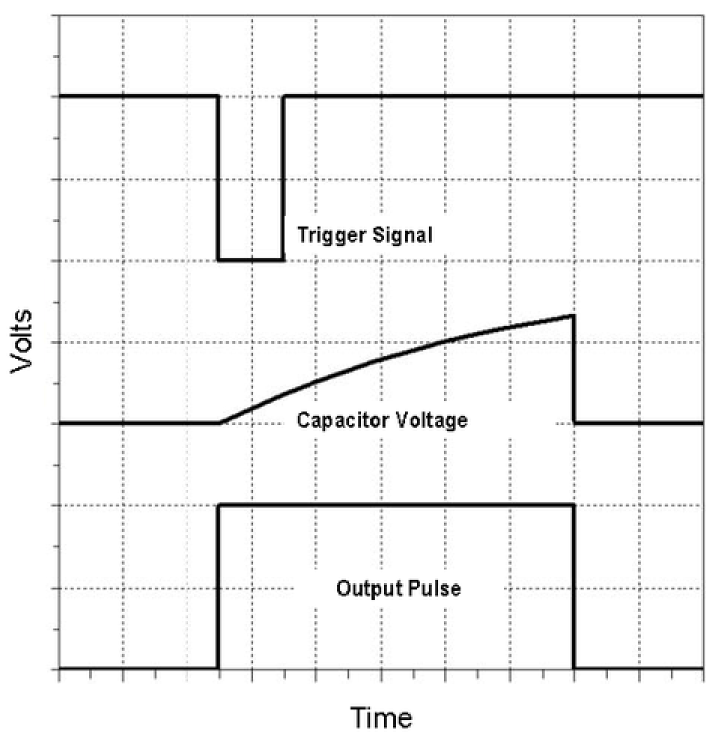
The connection of the pins is as follows:

Pin	Name	Purpose
1	GND	Ground, low level (0 V)
2	TRIG	OUT rises, and interval starts, when this input falls below 1/3 V_{CC}.
3	OUT	This output is driven to $+V_{CC}$ or GND.
4	RESET	A timing interval may be interrupted by driving this input to GND.
5	CTRL	"Control" access to the internal voltage divider (by default, 2/3 V_{CC}).
6	THR	The interval ends when the voltage at THR is greater than at CTRL.
7	DIS	Open collector output; may discharge a capacitor between intervals.
8	$V+$, V_{CC}	Positive supply voltage is usually between 3 and 15 V.

Monostable mode



Schematic of a 555 in monostable mode



The relationships of the trigger signal, the voltage on C and the pulse width in monostable mode

In the monostable mode, the 555 timer acts as a “one-shot” pulse generator. The pulse begins when the 555 timer receives a signal at the trigger input that falls below a third of the voltage supply. The width of the output pulse is determined by the time constant of an RC network, which consists of a capacitor (C) and a resistor (R). The output pulse ends when the charge on the C equals 2/3 of the supply voltage. The output pulse width can be lengthened or shortened to the need of the specific application by adjusting the values of R and C.

The output pulse width of time *t*, which is the time it takes to charge C to 2/3 of the supply voltage, is given by

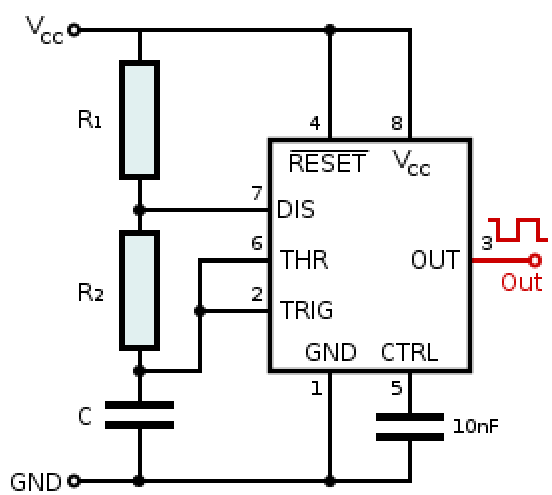
$$t = RC\ln(3) \approx 1.1RC$$

where t is in seconds, R is in ohms and C is in farads.

Bistable Mode

In bistable mode, the 555 timer acts as a basic flip-flop. The trigger and reset inputs (pins 2 and 4 respectively on a 555) are held high via Pull-up resistors while the threshold input (pin 6) is simply grounded. Thus configured, pulling the trigger momentarily to ground acts as a 'set' and transitions the output pin (pin 3) to Vcc (high state). Pulling the reset input to ground acts as a 'reset' and transitions the output pin to ground (low state). No capacitors are required in a bistable configuration. Pins 5 and 7 (control and discharge) are left floating.

Astable mode



Standard 555 Astable Circuit

In astable mode, the 555 timer puts out a continuous stream of rectangular pulses having a specified frequency. Resistor R_1 is connected between V_{CC} and the discharge pin (pin 7) and another resistor (R_2) is connected between the discharge pin (pin 7), and the trigger (pin 2) and threshold (pin 6) pins that share a common node. Hence the capacitor is charged through R_1 and R_2, and discharged only through R_2, since pin 7 has low impedance to ground during output low intervals of the cycle, therefore discharging the capacitor.

In the astable mode, the frequency of the pulse stream depends on the values of R_1, R_2 and C:

$$f = \frac{1}{\ln(2) \cdot C \cdot (R_1 + 2R_2)}$$

The high time from each pulse is given by

$$\text{high} = \ln(2) \cdot (R_1 + R_2) \cdot C$$

and the low time from each pulse is given by

$$\text{low} = \ln(2) \cdot R_2 \cdot C$$

where R_1 and R_2 are the values of the resistors in ohms and C is the value of the capacitor in farads.

note: power of R_1 must be greater than $\frac{V_{cc}^2}{R_1}$

To achieve a duty cycle of less than 50% a diode can be added in parallel with R_2 towards the capacitor. This bypasses R_2 during the high part of the cycle so that the high interval depends only on R_1 and C.

Specifications

These specifications apply to the NE555. Other 555 timers can have different specifications depending on the grade (military, medical, etc).

Supply voltage (V_{CC})	4.5 to 15 V
Supply current (V_{CC} = +5 V)	3 to 6 mA
Supply current (V_{CC} = +15 V)	10 to 15 mA
Output current (maximum)	200 mA
Maximum Power dissipation	600 mW
Power Consumption (minimum operating)	30 mW@5V, 225 mW@15V
Operating temperature	0 to 70 °C

Derivatives

Many pin-compatible variants, including CMOS versions, have been built by various companies. Bigger packages also exist with two or four timers on the same chip. The 555 is also known under the following type numbers:

Manufacturer	Model	Remark
Avago Technologies	Av-555M	
Custom Silicon Solutions	CSS555/CSS555C	CMOS from 1.2 V, IDD < 5 μA
ECG Philips	ECG955M	
Exar	XR-555	
Fairchild Semiconductor	NE555/KA555	
Harris	HA555	
IK Semicon	ILC555	CMOS from 2 V
Intersil	SE555/NE555	
Intersil	ICM7555	CMOS
Lithic Systems	LC555	
Maxim	ICM7555	CMOS from 2 V
Motorola	MC1455/MC1555	
National Semiconductor	LM1455/LM555/LM555C	
National Semiconductor	LMC555	CMOS from 1.5 V
NTE Sylvania	NTE955M	
Raytheon	RM555/RC555	
RCA	CA555/CA555C	
STMicroelectronics	NE555N/ K3T647	
Texas Instruments	SN52555/SN72555	
Texas Instruments	TLC555	CMOS from 2 V
USSR	K1006ВИ1	
Zetex	ZSCT1555	down to 0.9 V
NXP Semiconductors	ICM7555	CMOS
HFO / East Germany	B555	

Dual timer 556

The dual version is called 556. It features two complete 555s in a 14 pin DIL package.

Quad timer 558

The quad version is called 558 and has 16 pins. To fit four 555s into a 16 pin package the control, voltage, and reset lines are shared by all four modules. Also for each module the discharge and threshold are internally wired together and called *timing*.

Example applications

Joystick interface circuit using quad timer 558

The Apple II microcomputer used a quad timer 558 in monostable (or "one-shot") mode to interface up to four "game paddles" or two joysticks to the host computer.

A similar circuit was used in the IBM personal computer. In the joystick interface circuit of the IBM PC, the capacitor (C) of the RC network was generally a 10 nF capacitor. The resistor (R) of the RC network consisted of the potentiometer inside the joystick along with an external resistor of 2.2 kilohms. The joystick potentiometer acted as a variable resistor. By moving the joystick, the resistance of the joystick increased from a small value up to about 100 kilohms. The joystick operated at 5 V.

Software running in the host computer started the process of determining the joystick position by writing to a special address (ISA bus I/O address 201h). This would result in a trigger signal to the quad timer, which would cause the capacitor (C) of the RC network to begin charging and cause the quad timer to output a pulse. The width of the pulse was determined by how long it took the C to charge up to 2/3 of 5 V (or about 3.33 V), which was in turn determined by the joystick position.

Software running in the host computer measured the pulse width to determine the joystick position. A wide pulse represented the full-right joystick position, for example, while a narrow pulse represented the full-left joystick position.

Atari Punk Console

One of Forrest M. Mims III's many books was dedicated to the 555 timer. In it, he first published the "Stepped Tone Generator" circuit which has been adopted as a popular circuit, known as the Atari Punk Console, by circuit benders for its distinctive low-fi sound similar to classic Atari games.

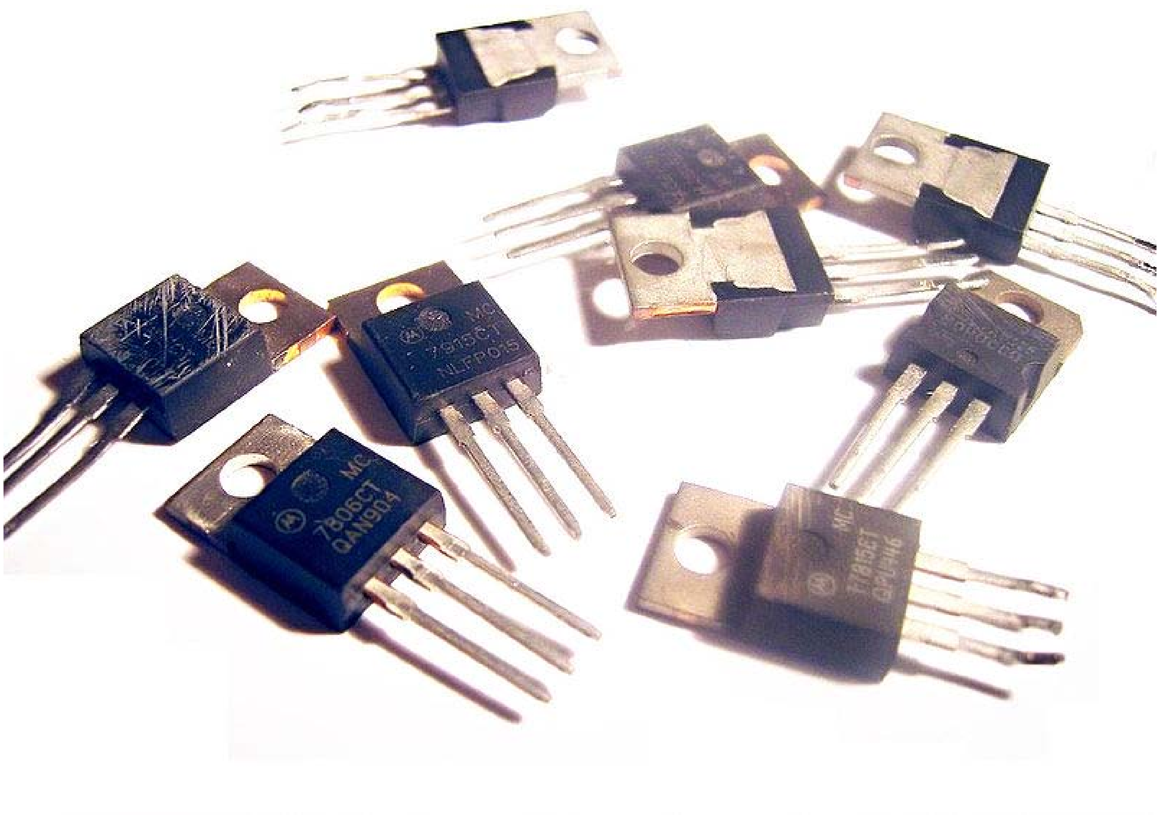
Pulse-width modulation

The 555 can be used to generate a variable PWM signal using a few external components. The chip alone can drive small external loads or an amplifying transistor for larger loads.

Chapter 2

78xx & ZN414

78xx



An assortment of 78XX ICs

The **78xx** (sometimes **LM78xx**) is a family of self-contained fixed linear voltage regulator integrated circuits. The 78xx family is commonly used in electronic circuits requiring a regulated power supply due to their ease-of-use and low cost. For ICs within the family, the *xx* is replaced with two digits, indicating the output voltage (for example, the 7805 has a 5 volt output, while the 7812 produces 12 volts). The 78xx line are positive voltage regulators: they produce a voltage that is positive relative to a common ground. There is a related line of **79xx** devices which are complementary negative voltage regulators. 78xx and 79xx ICs can be used in combination to provide positive and negative supply voltages in the same circuit.

78xx ICs have three terminals and are commonly found in the TO220 form factor, although smaller surface-mount and larger TO3 packages are available. These devices support an input voltage anywhere from a couple of volts over the intended output voltage, up to a maximum of 35 or 40 volts, and typically provide 1 or 1.5 amps of current (though smaller or larger packages may have a lower or higher current rating).

Advantages

- 78xx series ICs do not require additional components to provide a constant, regulated source of power, making them easy to use, as well as economical and efficient uses of space. Other voltage regulators may require additional components to set the output voltage level, or to assist in the regulation process. Some other designs (such as a switching power supply) may need substantial engineering expertise to implement.

- 78xx series ICs have built-in protection against a circuit drawing too much power. They have protection against overheating and short-circuits, making them quite robust in most applications. In some cases, the current-limiting features of the 78xx devices can provide protection not only for the 78xx itself, but also for other parts of the circuit.

Disadvantages

- The input voltage must always be higher than the output voltage by some minimum amount (typically 2 volts). This can make these devices unsuitable for powering some devices from certain types of power sources (for example, powering a circuit that requires 5 volts using 6-volt batteries will not work using a 7805).

- As they are based on a linear regulator design, the input current required is always the same as the output current. As the input voltage must always be higher than the output voltage, this means that the total power (voltage multiplied by current) going into the 78xx will be more than the output power provided. The extra input power is dissipated as heat. This means both that for some applications an adequate heatsink must be provided, and also that a (often substantial) portion of the input power is wasted during the process, rendering them less efficient than some other types of power supplies. When the input voltage is significantly higher than the regulated output voltage (for example, powering a 7805 using a 24 volt power source), this inefficiency can be a significant issue.

- Even in larger packages, 78xx integrated circuits cannot supply as much power as many designs which use discrete components, and are generally inappropriate for applications requiring more than a few amps of current.

Individual Devices in the Series

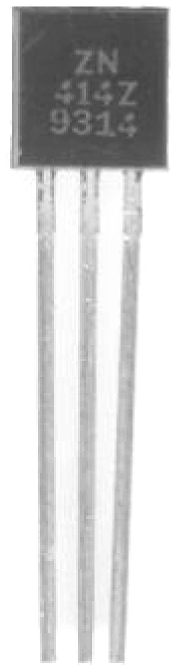
There are common configurations for 78xx ICs, including 7805 (5 volt), 7806 (6 volt), 7808 (8 volt), 7809 (9 volt), 7810 (10 volt), 7812 (12 volt), 7815 (15 volt), 7818 (18 volt), and 7824 (24 volt) versions. The 7805 is common, as its regulated 5 volt supply provides a convenient power source for most TTL components.

Less common are lower-power versions such as the LM78Mxx series (500mA) and LM78Lxx series (100mA) from National Semiconductor. Some devices provide slightly different voltages than usual, such as the LM78L62 (6.2 volts) and LM78L82 (8.2 volts).

Unrelated Devices

The LM78S40 from National Semiconductor is not part of the 78xx family, and does not use the same design. It is a component in switching regulator designs, and is not a linear regulator like other 78xx devices. The 7803SR from Datel is a full switching power supply module (designed as a drop-in replacement for 78xx chips), and not a linear regulator like the 78xx ICs.

ZN414

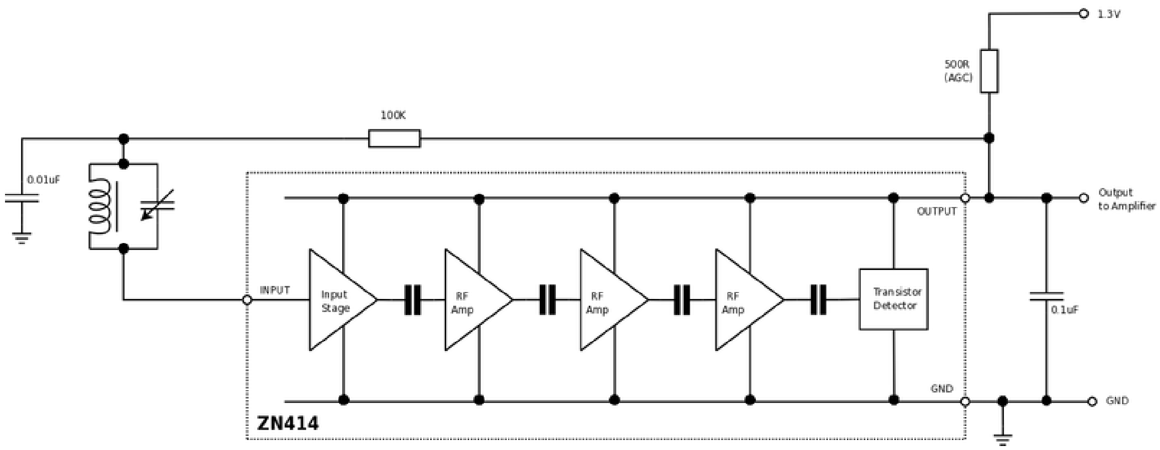


ZN414 in TO92 package

The ZN414 was a low cost, single-chip AM radio integrated circuit. Launched in 1972, the part was designed and supplied by Ferranti, but was also available from GEC-Plessey. The ZN414 was popular amongst hobbyists as a fully working AM radio could be made with just a few external components, a crystal earpiece and a 1.5 V cell.

The original ZN414 chip from Ferranti was supplied in a 3-pin, metal TO-18 'transistor' package whereas the GEC part and later Ferranti ones (ZN414Z) used the plastic TO-92 encapsulation. Later variants, the ZN415 and ZN416, came in 8-pin DIL packages and included a built-in amplifier that could drive headphones and small speakers directly.

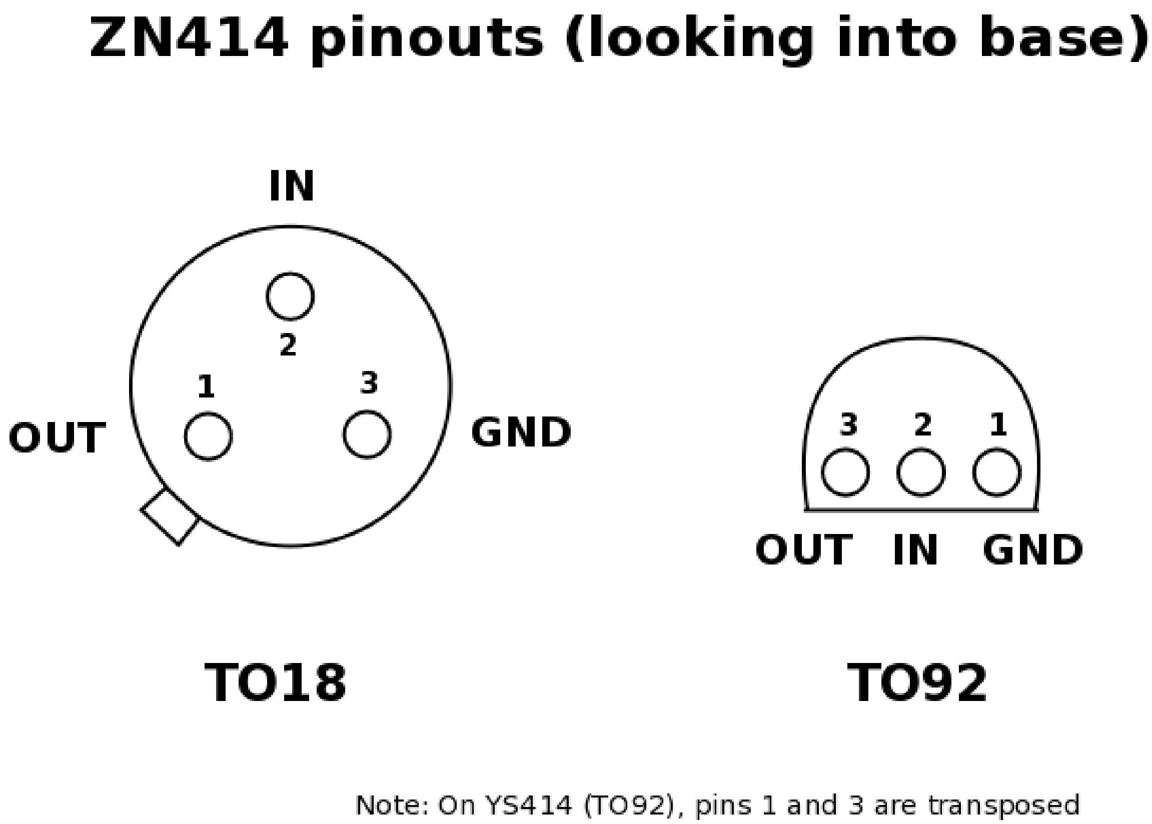
The radio circuit inside the ZN414 was based on a design known as Tuned Radio Frequency (TRF). The TRF design is much simpler than the popular, but more complex, superheterodyne radio circuit often used in modern AM receivers. It was principally the use of the TRF circuit that allowed almost a whole radio to be fitted into one small, three-pin package.



ZN414 in basic functional circuit

The manufacturing process for the ZN414 chip used a relatively new (for the time) technique known as Collector Diffusion Isolation (CDI). CDI was invented by engineers at Bell Telephone Laboratories and subsequently developed into a commercial process by Ferranti in the UK.

The original ZN41x family have not been manufactured for some time, but modern equivalents to the original 3-pin ZN414 are available, with part codes of MK484, TA7642 and (mainly in India, the Far East & Australasia) YS414 and LMF501T. Note that on the YS414 part, pins 1 (output) and 3 (ground/earth) are transposed.



ZN414 pinouts

Chapter 3

Operational Amplifier Applications

Practical considerations

Input offset problems

It is important to note that the equations shown below, pertaining to each type of circuit, assume that an ideal op amp is used.

Resistors used in practical solid-state op-amp circuits are typically in the kΩ range. Resistors much greater than 1 MΩ cause excessive thermal noise and make the circuit operation susceptible to significant errors due to bias or leakage currents.

Practical operational amplifiers draw a small current from each of their inputs due to bias requirements and leakage. These currents flow through the resistances connected to the inputs and produce small voltage drops across those resistances. In AC signal applications this seldom matters. If high-precision DC operation is required, however, these voltage drops need to be considered. The design technique is to try to ensure that these voltage drops are equal for both inputs, and therefore cancel. If these voltage drops are equal and the common-mode rejection ratio of the operational amplifier is good, there will be considerable cancellation and improvement in DC accuracy.

If the input currents into the operational amplifier are equal, to reduce offset voltage the designer must ensure that the DC resistance looking out of each input is also matched. In general input currents differ, the difference being called the *input offset current*, I_{os}. Matched external input resistances R_{in} will still produce an input voltage error of $R_{in} \cdot I_{os}$. Most manufacturers provide a method for tuning the operational amplifier to balance the input currents (e.g., "offset null" or "balance" pins that can interact with an external voltage source attached to a potentiometer). Otherwise, a tunable external voltage can be added to one of the inputs in order to balance out the offset effect. In cases where a design calls for one input to be short-circuited to ground, that short circuit can be replaced with a variable resistance that can be tuned to mitigate the offset problem.

Note that many operational amplifiers that have MOSFET-based input stages have input leakage currents that will truly be negligible to most designs.

Power supply effects

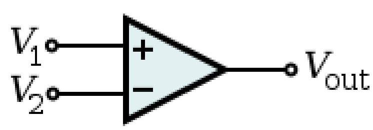
Although the power supplies are not shown in the operational amplifier designs below, they can be critical in operational amplifier design.

Power supply imperfections (e.g., power signal ripple, non-zero source impedance) may lead to noticeable deviations from ideal operational amplifier behavior. For example, operational amplifiers have a specified power supply rejection ratio that indicates how well the output can reject signals that appear on the power supply inputs. Power supply inputs are often noisy in large designs because the power supply is used by nearly every component in the design, and inductance effects prevent current from being instantaneously delivered to every component at once. As a consequence, when a component requires large injections of current (e.g., a digital component that is frequently switching from one state to another), nearby components can experience sagging at their connection to the power supply. This problem can be mitigated with copious use of bypass capacitors placed connected across each power supply pin and ground. When bursts of current are required by a component, the component can *bypass* the power supply by receiving the current directly from the nearby capacitor (which is then slowly charged by the power supply).

Additionally, current drawn into the operational amplifier from the power supply can be used as inputs to external circuitry that augment the capabilities of the operational amplifier. For example, an operational amplifier may not be fit for a particular high-gain application because its output would be required to generate signals outside of the safe range generated by the amplifier. In this case, an external push–pull amplifier can be controlled by the current into and out of the operational amplifier. Thus, the operational amplifier may itself operate within its factory specified bounds while still allowing the negative feedback path to include a large output signal well outside of those bounds.

Circuit applications

Comparator

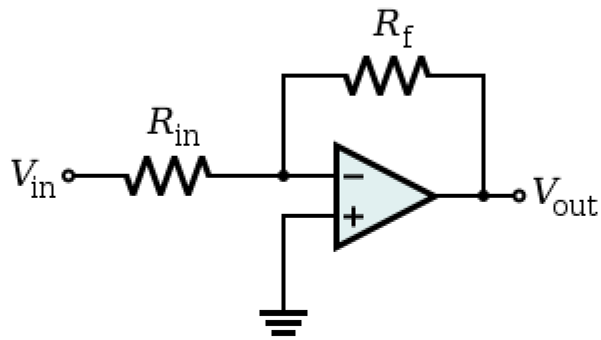


Compares two voltages and switches its output to indicate which voltage is larger.

- $$V_{\text{out}} = \begin{cases} V_{\text{S+}} & V_1 > V_2 \\ V_{\text{S-}} & V_1 < V_2 \end{cases}$$

(where V_s is the supply voltage and the opamp is powered by $+V_s$ and $-V_s$.)

Inverting amplifier



An inverting amplifier uses negative feedback to invert and amplify a voltage. The R_{in},R_f resistor network allows some of the output signal to be returned to the input. Since the output is 180° out of phase, this amount is effectively subtracted from the input, thereby reducing the input into the operational amplifier. This reduces the overall gain of the amplifier and is dubbed negative feedback.

$$V_{\mathrm{out}} = -\frac{R_{\mathrm{f}}}{R_{\mathrm{in}}} V_{\mathrm{in}}$$

- $Z_{in} = R_{in}$ (because V_{-} is a virtual ground)
- A third resistor, of value $R_{\mathrm{f}} \| R_{\mathrm{in}} \triangleq R_{\mathrm{f}} R_{\mathrm{in}} / (R_{\mathrm{f}} + R_{\mathrm{in}})$, added between the non-inverting input and ground, while not necessary, minimizes errors due to input bias currents.

The gain of the amplifier is determined by the ratio of R_f to R_{in}. That is:

$$A = -\frac{R_f}{R_{in}}$$

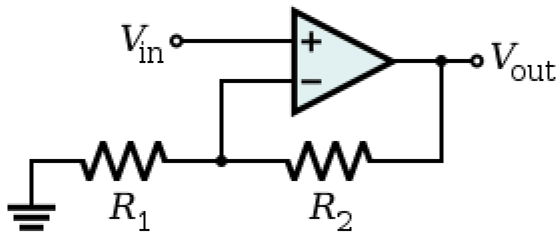
The presence of the negative sign is a convention indicating that the output is inverted. For example, if R_f is 10 000 Ω and R_{in} is 1 000 Ω, then the gain would be -10 000Ω/1 000Ω, which is -10.

Theory of operation: An Ideal Operational Amplifier has 2 characteristics that imply the operation of the inverting amplifier: Infinite input impedance, and infinite differential gain. Infinite input impedance implies there is no current in either of the input pins because current cannot flow through an infinite impedance. Infinite differential gain implies that both the (+) and (-) input pins are at the same voltage because the output is equal to infinity times (V+ - V-). As the output approaches any arbitrary finite voltage, then the term (V+ - V-) approaches 0, thus the two input pins are at the same voltage for any finite output.

To begin analysis, first it is noted that with the (+) pin grounded, the (-) must also be at 0 volts potential due to implication 2. with the (-) at 0 volts, the current through Rin (from left to right) is given by I = Vin/Rin by Ohm's law. Second, since no current is flowing into the op amp through the (-) pin due to implication 1, all the current through Rin must also be flowing through Rf. Therefore, with V- = 0 volts and I(Rf) = Vin/Rin the output voltage given by Ohm's law is -Vin*Rf/Rin.

Real op amps have both finite input impedance and differential gain, however both are high enough as to induce error that is considered negligible in most applications.

Non-inverting amplifier



Amplifies a voltage (multiplies by a constant greater than 1)

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

- Input impedance $Z_{\mathbf{in}} \approx \infty$
    - The input impedance is *at least* the impedance between non-inverting (+) and inverting (−) inputs, which is typically 1 MΩ to 10 TΩ, plus the impedance of the path from the inverting (−) input to ground (i.e., R_1 in parallel with R_2).
    - Because negative feedback ensures that the non-inverting and inverting inputs match, the input impedance is actually **much higher.**
- Although this circuit has a large input impedance, it suffers from error of input bias current.
    - The non-inverting (+) and inverting (−) inputs draw small leakage currents into the operational amplifier.
    - These input currents generate voltages that act like unmodeled input offsets. These unmodeled effects can lead to noise on the output (e.g., offsets or drift).
    - Assuming that the two leaking currents are **matched,** their effect can be mitigated by ensuring the DC impedance looking **out** of each input is the same.
        - The voltage produced by each bias current is equal to the product of the bias current with the equivalent DC impedance looking out of each input. Making those impedances equal makes the offset

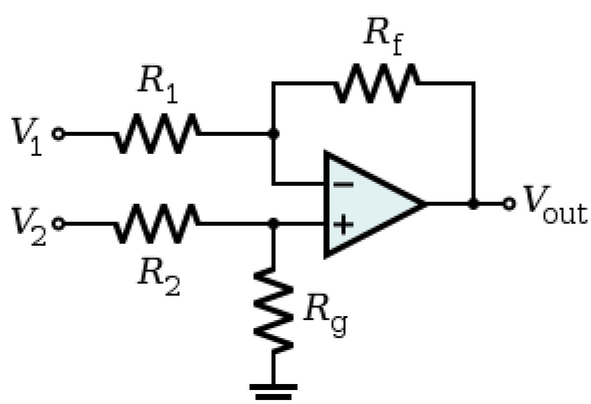
voltage at each input equal, and so the non-zero bias currents will have no impact on the **difference** between the two inputs.

- A resistor of value

$$R_1 \| R_2 \triangleq \left(\frac{1}{R_1} + \frac{1}{R_2} \right)^{-1} = \frac{R_1 R_2}{R_1 + R_2},$$

- which is the equivalent resistance of R_1 in parallel with R_2, between the V_{in} source and the non-inverting (+) input will ensure the impedances looking **out** of each input will be matched.
  - The matched bias currents will then generate matched offset voltages, and their effect will be hidden to the operational amplifier (which acts on the difference between its inputs) so long as the CMRR is good.
- Very often, the input currents are *not* matched.
  - Most operational amplifiers provide some method of balancing the two input currents (e.g., by way of an external potentiometer).
  - Alternatively, an external offset can be added to the operational amplifier input to nullify the effect.
  - Another solution is to insert a variable resistor between the V_{in} source and the non-inverting (+) input. The resistance can be tuned until the offset voltages at each input are matched.
  - Operational amplifiers with MOSFET-based input stages have input currents that are so small that they often can be neglected.

Differential amplifier



The circuit shown is used for finding the difference of two voltages each multiplied by some constant (determined by the resistors).

$$V_{out} = \frac{(R_f + R_1) R_g}{(R_g + R_2) R_1} V_2 - \frac{R_f}{R_1} V_1$$

- Differential Z_{in} (between the two input pins) = $R_1 + R_2$ (Note: this is approximate)

For common-mode rejection, anything done to one input must be done to the other. The addition of a compensation capacitor in parallel with Rf, for instance, must be balanced by an equivalent capacitor in parallel with Rg.

The "instrumentation amplifier", which is also shown on this page, is another form of differential amplifier that also provides high input impedance.

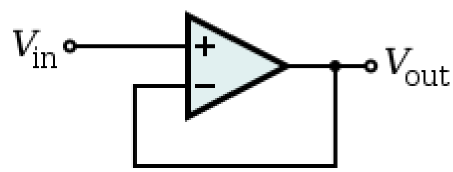
Whenever $R_1 = R_2$ and $R_\mathrm{f} = R_\mathrm{g}$, the differential gain is

$$V_\mathrm{out} = A(V_2 - V_1) \text{ and } A \triangleq \frac{R_\mathrm{f}}{R_1}$$

When $R_1 = R_\mathrm{f}$ and $R_2 = R_\mathrm{g}$ the differential gain is A = 1 and the circuit acts as a differential follower:

$$V_\mathrm{out} = V_2 - V_1$$

Voltage follower

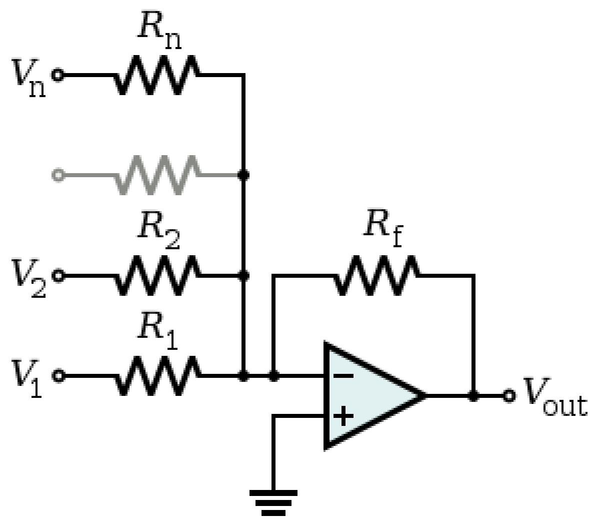


Used as a buffer amplifier to eliminate loading effects (e.g., connecting a device with a high source impedance to a device with a low input impedance).

$$V_\mathrm{out} = V_\mathrm{in}$$
$Z_\mathrm{in} = \infty$ (realistically, the differential input impedance of the op-amp itself, 1 MΩ to 1 TΩ)

Due to the strong (i.e., unity gain) feedback and certain non-ideal characteristics of real operational amplifiers, this feedback system is prone to have poor stability margins. Consequently, the system may be unstable when connected to sufficiently capacitive loads. In these cases, a lag compensation network (e.g., connecting the load to the voltage follower through a resistor) can be used to restore stability. The manufacturer data sheet for the operational amplifier may provide guidance for the selection of components in external compensation networks. Alternatively, another operational amplifier can be chosen that has more appropriate internal compensation.

Summing amplifier



A summing amplifer sums several (weighted) voltages:

$$V_{\text{out}} = -R_{\text{f}}\left(\frac{V_1}{R_1} + \frac{V_2}{R_2} + \cdots + \frac{V_n}{R_n}\right)$$

- When $R_1 = R_2 = \cdots = R_n$, and R_{f} independent

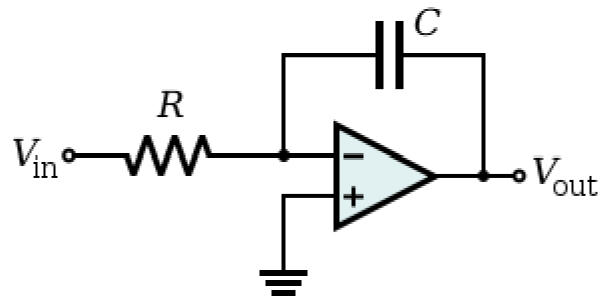
$$V_{\text{out}} = -\frac{R_{\text{f}}}{R_1}(V_1 + V_2 + \cdots + V_n)$$

- When $R_1 = R_2 = \cdots = R_n = R_{\text{f}}$

$$V_{\text{out}} = -(V_1 + V_2 + \cdots + V_n)$$

- Output is inverted
- Input impedance of the ***n*th** input is $Z_n = R_n$ (V_- is a virtual ground)

Inverting integrator



Integrates the (inverted) signal over time

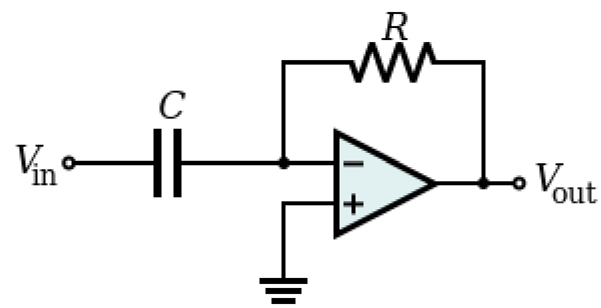
$$V_{\text{out}} = -\int_0^t \frac{V_{\text{in}}}{RC}\,\mathrm{d}t + V_{\text{initial}}$$

(where V_{in} and V_{out} are functions of time, V_{initial} is the output voltage of the integrator at time $t = 0$.)

- Note that this can also be viewed as a low-pass electronic filter. It is a filter with a single pole at DC (i.e., where $\omega = 0$) and gain.
- There are several potential problems with this circuit.
  - It is usually assumed that the input V_{in} has zero DC component (i.e., has a zero average value). Otherwise, unless the capacitor is periodically discharged, the output will drift outside of the operational amplifier's operating range.
  - Even when V_{in} has no offset, the leakage or bias currents into the operational amplifier inputs can add an unexpected offset voltage to V_{in} that causes the output to drift. Balancing input currents **and** replacing the non-inverting (+) short-circuit to ground with a resistor with resistance R can reduce the severity of this problem.
  - Because this circuit provides no DC feedback (i.e., the capacitor appears like an open circuit to signals with $\omega = 0$), the offset of the output may not agree with expectations (i.e., V_{initial} may be out of the designer's control with the present circuit).

  Many of these problems can be made less severe by adding a *large* resistor R_F in parallel with the feedback capacitor. At significantly high frequencies, this resistor will have negligible effect. However, at low frequencies where there are drift and offset problems, the resistor provides the necessary feedback to hold the output steady at the correct value. In effect, this resistor reduces the DC gain of the "integrator" – it goes from infinite to some finite value R_F / R.

Inverting differentiator

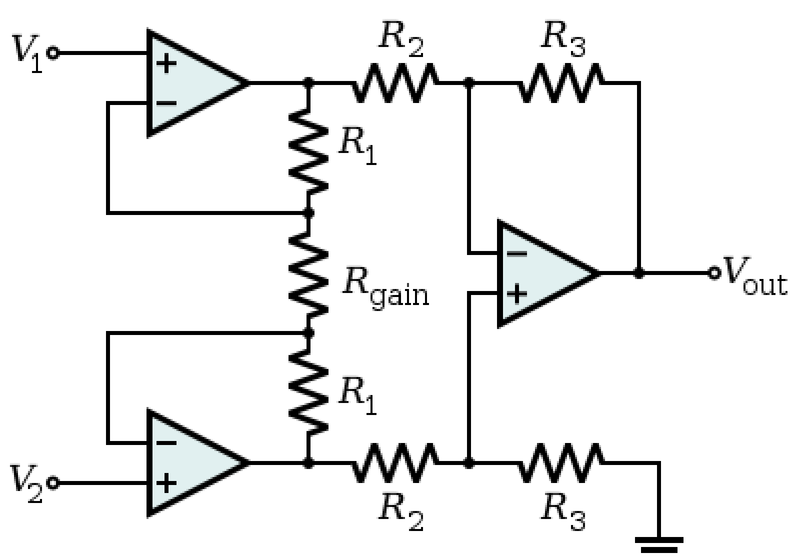


Differentiates the (inverted) signal over time.

$$V_{\text{out}} = -RC\,\frac{\mathrm{d}V_{\text{in}}}{\mathrm{d}t} \qquad \text{where } V_{\text{in}} \text{ and } V_{\text{out}} \text{ are functions of time.}$$

- Note that this can also be viewed as a high-pass electronic filter. It is a filter with a single zero at DC (i.e., where ω = 0) and gain. The high pass characteristics of a differentiating amplifier can lead to unstable behavior when the circuit is used in an analog servo loop. For this reason the system function would be re-formulated to use integrators.

Instrumentation amplifier

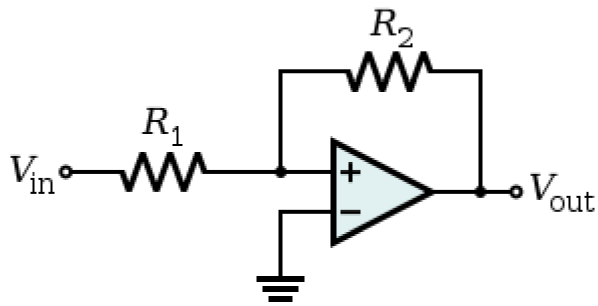


Combines very high input impedance, high common-mode rejection, low DC offset, and other properties used in making very accurate, low-noise measurements

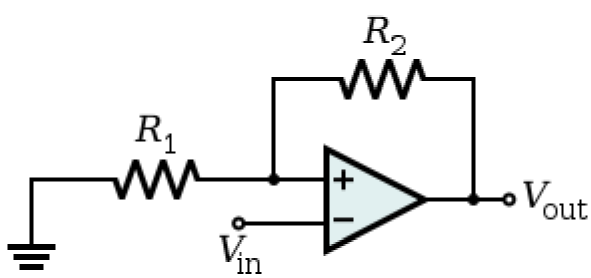
- Is made by adding a non-inverting buffer to each input of the differential amplifier to increase the input impedance.

Schmitt trigger

A bistable multivibrator implemented as a comparator with hysteresis.

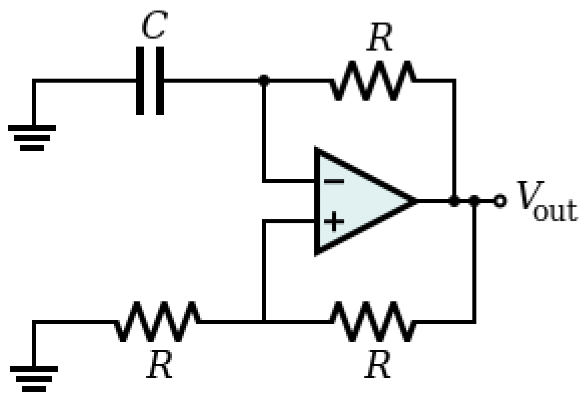


In this configuration, the input voltage is applied through the resistor R_1 (which may be the source internal resistance) to the non-inverting input and the inverting input is grounded or referenced. The hysteresis curve is non-inverting and the switching thresholds are $\pm\frac{R_1}{R_2}V_{\mathbf{sat}}$ where V_{sat} is the greatest output magnitude of the operational amplifier.



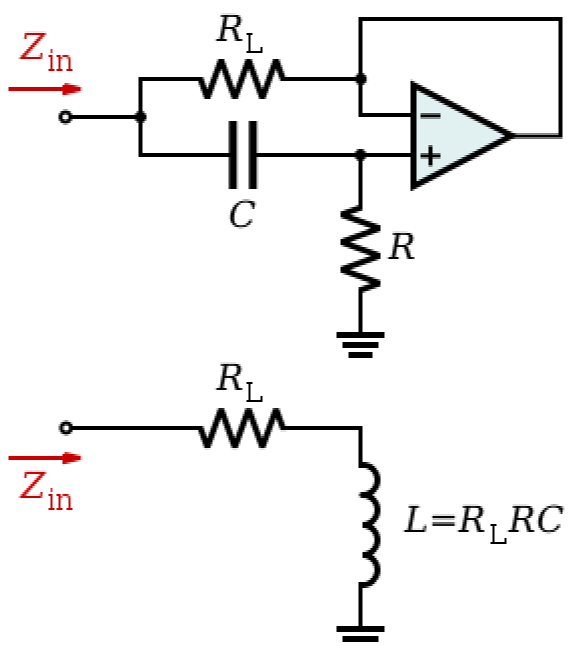
Alternatively, the input source and the ground may be swapped. Now the input voltage is applied directly to the inverting input and the non-inverting input is grounded or referenced. The hysteresis curve is inverting and the switching thresholds are $\pm\frac{R_1}{R_1+R_2}V_{\mathbf{sat}}$. Such a configuration is used in the relaxation oscillator shown below.

Relaxation oscillator



By using an RC network to add slow negative feedback to the inverting Schmitt trigger, a relaxation oscillator is formed. The feedback through the RC network causes the Schmitt trigger output to oscillate in an endless symmetric square wave (i.e., the Schmitt trigger in this configuration is an astable multivibrator).

Inductance gyrator



Simulates an inductor (i.e., provides inductance without the use of a possibly costly inductor). The circuit exploits the fact that the current flowing through a capacitor behaves through time as the voltage across an inductor. The capacitor used in this circuit is smaller than the inductor it simulates and its capacitance is less subject to changes in value due to environmental changes.

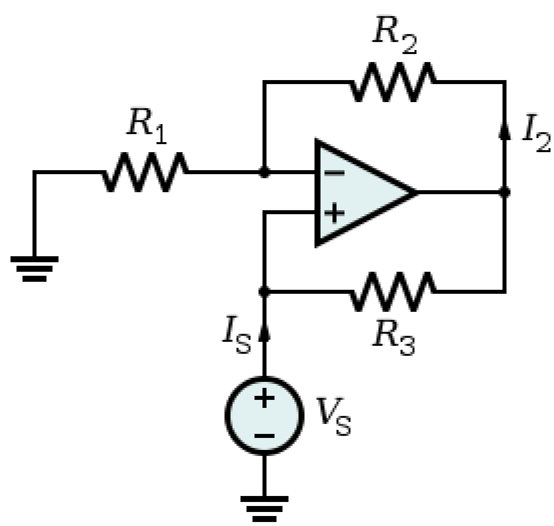
This circuit is unsuitable for applications relying on the back EMF property of an inductor as this will be limited in a gyrator circuit to the voltage supplies of the op-amp.

Zero level detector

Voltage divider reference

- Zener sets reference voltage.
- Acts as a comparator with one input tied to ground.
- When input is at zero, op-amp output is zero (assuming split supplies.)

Negative impedance converter (NIC)



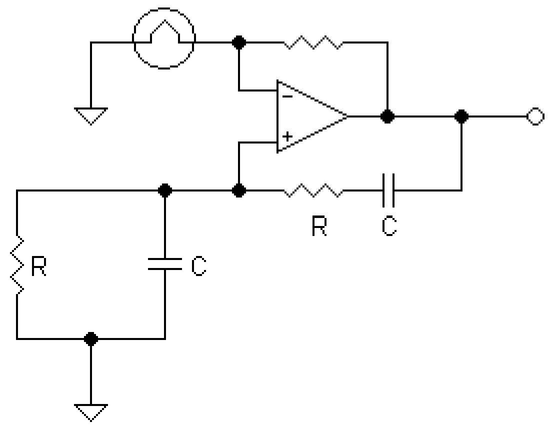
Creates a resistor having a negative value for any signal generator

- In this case, the ratio between the input voltage and the input current (thus the input resistance) is given by:

$$R_{\text{in}} = -R_3 \frac{R_1}{R_2}$$

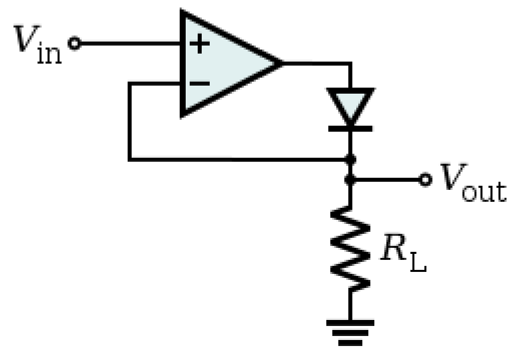
In general, the components R_1, R_2, and R_3 need not be resistors; they can be any component that can be described with an impedance.

Wien bridge oscillator



Produces a very low distortion sine wave. Uses negative temperature compensation in the form of a light bulb or diode.

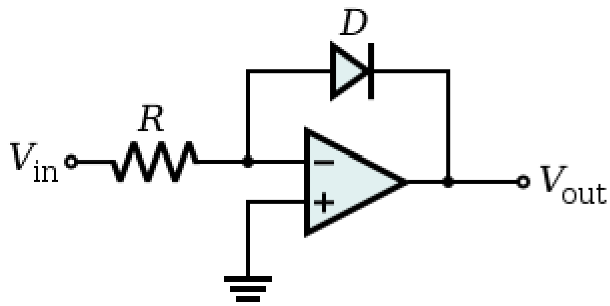
Precision rectifier



The voltage drop V_F across the forward biased diode in the circuit of a passive rectifier is undesired. In this active version, the problem is solved by connecting the diode in the negative feedback loop. The op-amp compares the output voltage across the load with the input voltage and increases its own output voltage with the value of V_F. As a result, the voltage drop V_F is compensated and the circuit behaves very nearly as an ideal (*super*) diode with $V_F = 0$ V.

The circuit has speed limitations at high frequency because of the slow negative feedback and due to the low slew rate of many non-ideal op-amps.

Logarithmic output



- The relationship between the input voltage v_{in} and the output voltage v_{out} is given by:

  $$v_{\mathrm{out}} = -V_{\mathrm{T}} \ln\left(\frac{v_{\mathrm{in}}}{I_{\mathrm{S}}\, R}\right)$$

  where I_S is the *saturation current* and V_T is the *thermal voltage*.

- If the operational amplifier is considered ideal, the negative pin is virtually grounded, so the current flowing into the resistor from the source (and thus through the diode to the output, since the op-amp inputs draw no current) is:

  $$\frac{v_{\mathrm{in}}}{R} = I_{\mathrm{R}} = I_{\mathrm{D}}$$

  where I_D is the current through the diode. As known, the relationship between the current and the voltage for a diode is:

  $$I_{\mathrm{D}} = I_{\mathrm{S}}\left(e^{\frac{V_{\mathrm{D}}}{V_{\mathrm{T}}}} - 1\right).$$

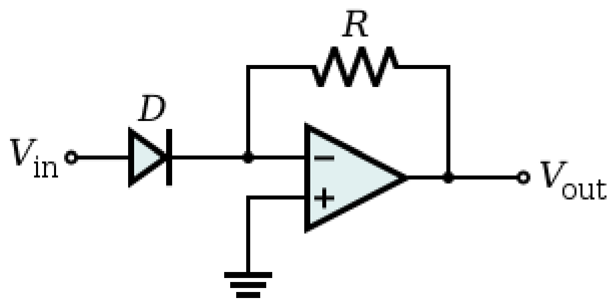
  This, when the voltage is greater than zero, can be approximated by:

  $$I_{\mathrm{D}} \simeq I_{\mathrm{S}} e^{\frac{V_{\mathrm{D}}}{V_{\mathrm{T}}}}.$$

  Putting these two formulae together and considering that the output voltage is the negative of the voltage across the diode ($V_{out} = -V_D$), the relationship is proven.

Note that this implementation does not consider temperature stability and other non-ideal effects.

Exponential output



- The relationship between the input voltage v_{in} and the output voltage v_{out} is given by:

$$v_{\text{out}} = -RI_{\text{S}}e^{\frac{v_{\text{in}}}{V_{\text{T}}}}$$

where I_S is the *saturation current* and V_T is the *thermal voltage*.

- Considering the operational amplifier ideal, then the negative pin is virtually grounded, so the current through the diode is given by:

$$I_{\text{D}} = I_{\text{S}}\left(e^{\frac{V_{\text{D}}}{V_{\text{T}}}} - 1\right)$$

when the voltage is greater than zero, it can be approximated by:

$$I_{\text{D}} \simeq I_{\text{S}}e^{\frac{V_{\text{D}}}{V_{\text{T}}}}.$$

The output voltage is given by:

$$v_{\text{out}} = -RI_{\text{D}}.$$

Chapter 4

Operational Amplifier

A Signetics µa741 operational amplifier, one of the most successful op-amps.

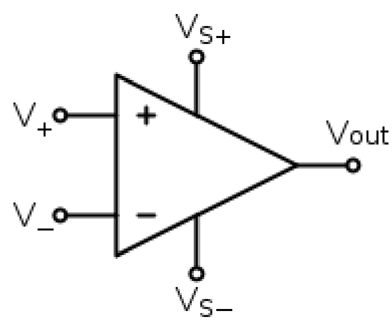
An **operational amplifier** ("op-amp") is a DC-coupled high-gain electronic voltage amplifier with a differential input and, usually, a single-ended output. An op-amp produces an output voltage that is typically hundreds of thousands times larger than the voltage *difference* between its input terminals.

Operational amplifiers are important building blocks for a wide range of electronic circuits. They had their origins in analog computers where they were used in many linear, non-linear and frequency-dependent circuits. Their popularity in circuit design largely stems from the fact that characteristics of the final elements (such as their gain) are set by external components with little dependence on temperature changes and manufacturing variations in the op-amp itself.

Op-amps are among the most widely used electronic devices today, being used in a vast array of consumer, industrial, and scientific devices. Many standard IC op-amps cost only a few cents in moderate production volume; however some integrated or hybrid operational amplifiers with special performance specifications may cost over $100 US in small quantities. Op-amps may be packaged as components, or used as elements of more complex integrated circuits.

The op-amp is one type of differential amplifier. Other types of differential amplifier include the fully differential amplifier (similar to the op-amp, but with two outputs), the instrumentation amplifier (usually built from three op-amps), the isolation amplifier (similar to the instrumentation amplifier, but with tolerance to common-mode voltages that would destroy an ordinary op-amp), and negative feedback amplifier (usually built from one or more op-amps and a resistive feedback network).

Circuit notation



Circuit diagram symbol for an op-amp

The circuit symbol for an op-amp is shown to the right, where:

- V_{+}: non-inverting input
- V_{-}: inverting input
- V_{out}: output
- V_{S+}: positive power supply
- V_{S-}: negative power supply

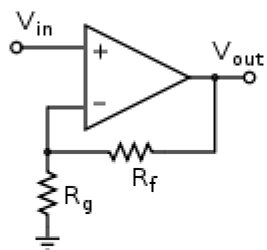
The power supply pins (V_{S+} and V_{S-}) can be labeled in different ways. Despite different labeling, the function remains the same — to provide additional power for amplification of the signal. Often these pins are left out of the diagram for clarity, and the power configuration is described or assumed from the circuit.

Operation

The amplifier's differential inputs consist of a V_+input and a V_-input, and ideally the op-amp amplifies only the difference in voltage between the two, which is called the *differential input voltage*. The output voltage of the op-amp is given by the equation,

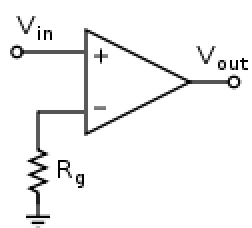
$$V_{\text{out}} = (V_+ - V_-)\, A_{OL}$$

where $V_+$is the voltage at the non-inverting terminal, $V_-$is the voltage at the inverting terminal and A_{OL} is the open-loop gain of the amplifier. (The term "open-loop" refers to the absence of a feedback loop from the output to the input).



Typically the op-amp's very large gain is controlled by negative feedback, which largely determines the magnitude of its output ("closed-loop") voltage gain in amplifier applications, or the transfer function required (in analog computers). Without negative feedback, and perhaps with positive feedback for regeneration, an op-amp acts as a comparator. High input impedance at the input terminals and low output impedance at the output terminal(s) are important typical characteristics.

With no negative feedback, the op-amp acts as a comparator. The inverting input is held at ground (0 V) by the resistor, so if the V_{in} applied to the non-inverting input is positive, the output will be maximum positive, and if V_{in} is negative, the output will be maximum negative. Since there is no feedback from the output to either input, this is an *open loop* circuit. The circuit's gain is just the G_{OL} of the op-amp.



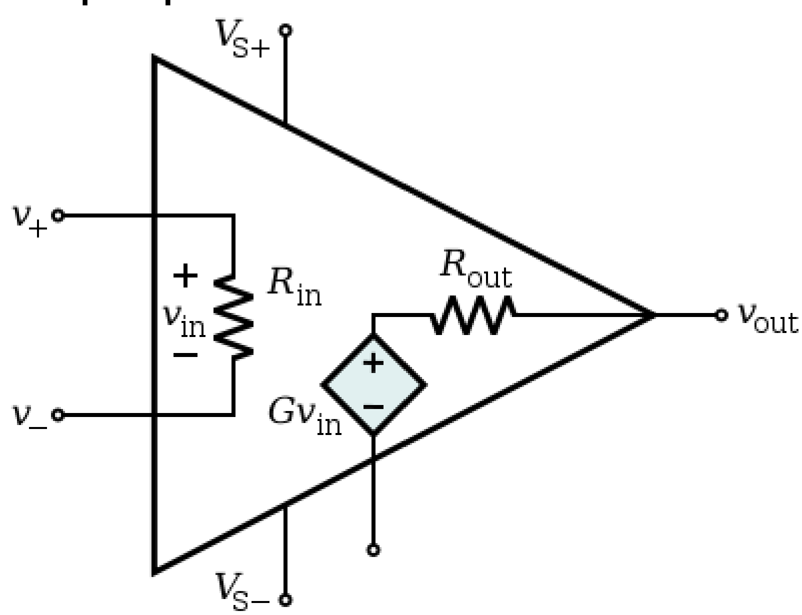
Adding negative feedback via the voltage divider R_f,R_g reduces the gain. Equilibrium will be established when V_{out} is just sufficient to reach around and "pull" the inverting input to the same voltage as V_{in}. As a simple example, if V_{in} = 1 V and $R_f = R_g$, V_{out} will be 2 V, the amount required to keep V_- at 1 V. Because of the feedback provided by R_f, R_g this is a *closed loop* circuit. Its over-all gain V_{out} / V_{in} is called the *closed-loop gain* A_{CL}. Because the feedback is negative, in this case A_{CL} is less than the A_{OL} of the op-amp.

The magnitude of A_{OL} is typically very large—10,000 or more for integrated circuit op-amps—and therefore even a quite small difference between V_+ and V_- drives the amplifier output nearly to the supply voltage. This is called *saturation* of the amplifier. The magnitude of A_{OL} is not well controlled by the manufacturing process, and so it is impractical to use an operational amplifier as a stand-alone differential amplifier. If predictable operation is desired, negative feedback is used, by applying a portion of the output voltage to the inverting input. The *closed loop* feedback greatly reduces the gain of the amplifier. If negative feedback is used, the circuit's overall gain and other parameters become determined more by the feedback network than by the op-amp itself. If the feedback network is made of components with relatively constant, stable values, the unpredictability and inconstancy of the op-amp's parameters do not seriously affect the circuit's performance.

If no negative feedback is used, the op-amp functions as a switch or comparator.

Positive feedback may be used to introduce hysteresis or oscillation.

Ideal and real op-amps



An equivalent circuit of an operational amplifier that models some resistive non-ideal parameters.

An ideal op-amp is usually considered to have the following properties, and they are considered to hold for all input voltages:

- Infinite open-loop gain (when doing theoretical analysis, a limit may be taken as open loop gain A_{OL} goes to infinity).
- Infinite voltage range available at the output (v_{out}) (in practice the voltages available from the output are limited by the supply voltages $V_{S+}$and V_{S-}). The power supply sources are called rails.
- Infinite bandwidth (i.e., the frequency magnitude response is considered to be flat everywhere with zero phase shift).
- Infinite input impedance (so, in the diagram, $R_{\text{in}} = \infty$, and zero current flows from $v_+$to v_-).
- Zero input current (i.e., there is assumed to be no leakage or bias current into the device).
- Zero input offset voltage (i.e., when the input terminals are shorted so that $v_+ = v_-$, the output is a virtual ground or $v_{\text{out}} = 0$).
- Infinite slew rate (i.e., the rate of change of the output voltage is unbounded) and power bandwidth (full output voltage and current available at all frequencies).
- Zero output impedance (i.e., $R_{\text{out}} = 0$, so that output voltage does not vary with output current).
- Zero noise.
- Infinite Common-mode rejection ratio (CMRR).
- Infinite Power supply rejection ratio for both power supply rails.

These ideals can be summarized by the two "golden rules":

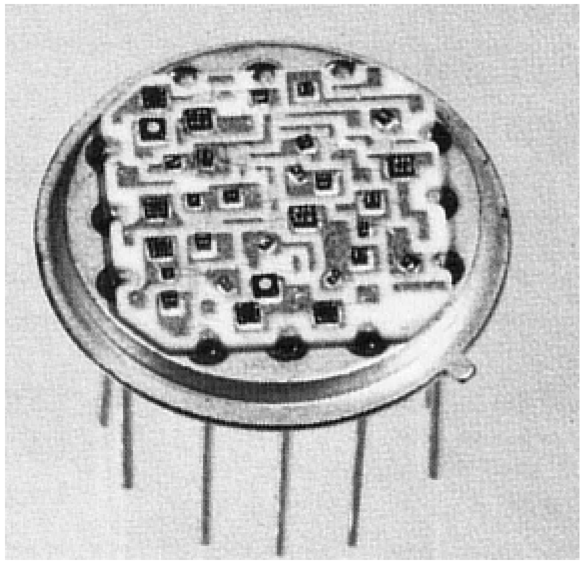
> I. The output attempts to do whatever is necessary to make the voltage difference between the inputs zero.
> II. The inputs draw no current.[:177]

The first rule only applies in the usual case where the op-amp is used in a closed-loop design (negative feedback, where there is a signal path of some sort feeding back from the output to the inverting input). These rules are commonly used as a good first approximation for analyzing or designing op-amp circuits.[:177]

In practice, none of these ideals can be perfectly realized, and various shortcomings and compromises have to be accepted. Depending on the parameters of interest, a real op-amp may be modeled to take account of some of the non-infinite or non-zero parameters using equivalent resistors and capacitors in the op-amp model. The designer can then include the effects of these undesirable, but real, effects into the overall performance of the final circuit. Some parameters may turn out to have negligible effect on the final design while others represent actual limitations of the final performance, that must be evaluated.

History

GAP/R's K2-W: a vacuum-tube op-amp (1953)



ADI's HOS-050: a high speed hybrid IC op-amp (1979)

An op-amp in a modern DIP

1941: First (vacuum tube) op-amp

An op-amp, defined as a general-purpose, DC-coupled, high gain, inverting feedback amplifier, is first found in U.S. Patent 2,401,779 "Summing Amplifier" filed by Karl D. Swartzel Jr. of Bell labs in 1941. This design used three vacuum tubes to achieve a gain of 90 dB and operated on voltage rails of ±350 V. It had a single inverting input rather than differential inverting and non-inverting inputs, as are common in today's op-amps. Throughout World War II, Swartzel's design proved its value by being liberally used in the M9 artillery director designed at Bell Labs. This artillery director worked with the SCR584 radar system to achieve extraordinary hit rates (near 90%) that would not have been possible otherwise.

1947: First op-amp with an explicit non-inverting input

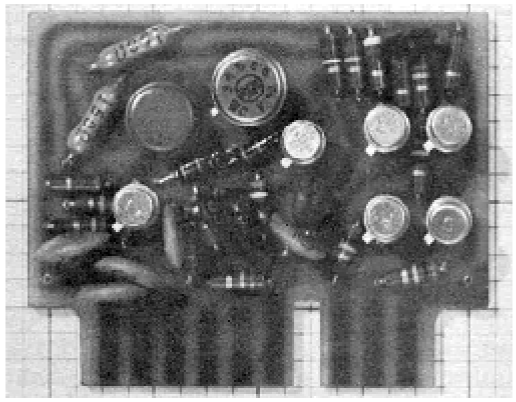
In 1947, the operational amplifier was first formally defined and named in a paper by Professor John R. Ragazzini of Columbia University. In this same paper a footnote mentioned an op-amp design by a student that would turn out to be quite significant. This op-amp, designed by Loebe Julie, was superior in a variety of ways. It had two major innovations. Its input stage used a long-tailed triode pair with loads matched to reduce drift in the output and, far more importantly, it was the first op-amp design to have two inputs (one inverting, the other non-inverting). The differential input made a whole range of new functionality possible, but it would not be used for a long time due to the rise of the chopper-stabilized amplifier.

1949: First chopper-stabilized op-amp

In 1949, Edwin A. Goldberg designed a chopper-stabilized op-amp. This set-up uses a normal op-amp with an additional AC amplifier that goes alongside the op-amp. The chopper gets an AC signal from DC by switching between the DC voltage and ground at a fast rate (60 Hz or 400 Hz). This signal is then amplified, rectified, filtered and fed into the op-amp's non-inverting input. This vastly improved the gain of the op-amp while significantly reducing the output drift and DC offset. Unfortunately, any design that used a chopper couldn't use their non-inverting input for any other purpose. Nevertheless, the much improved characteristics of the chopper-stabilized op-amp made it the dominant way to use op-amps. Techniques that used the non-inverting input regularly would not be very popular until the 1960s when op-amp ICs started to show up in the field.

In 1953, vacuum tube op-amps became commercially available with the release of the model K2-W from George A. Philbrick Researches, Incorporated. The designation on the devices shown, GAP/R, is an acronym for the complete company name. Two nine-pin 12AX7 vacuum tubes were mounted in an octal package and had a model K2-P chopper add-on available that would effectively "use up" the non-inverting input. This op-amp was based on a descendant of Loebe Julie's 1947 design and, along with its successors, would start the widespread use of op-amps in industry.

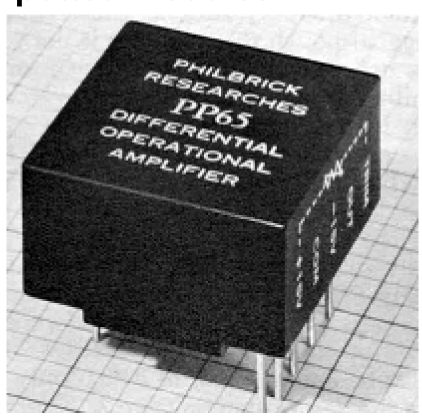
1961: First discrete IC op-amps



GAP/R's model P45: a solid-state, discrete op-amp (1961).

With the birth of the transistor in 1947, and the silicon transistor in 1954, the concept of ICs became a reality. The introduction of the planar process in 1959 made transistors and ICs stable enough to be commercially useful. By 1961, solid-state, discrete op-amps were being produced. These op-amps were effectively small circuit boards with packages such as edge-connectors. They usually had hand-selected resistors in order to improve things such as voltage offset and drift. The P45 (1961) had a gain of 94 dB and ran on ±15 V rails. It was intended to deal with signals in the range of ±10 V.

1962: First op-amps in potted modules



GAP/R's model PP65: a solid-state op-amp in a potted module (1962)

By 1962, several companies were producing modular potted packages that could be plugged into printed circuit boards. These packages were crucially important as they made the operational amplifier into a single black box which could be easily treated as a component in a larger circuit.

1963: First monolithic IC op-amp

In 1963, the first monolithic IC op-amp, the µA702 designed by Bob Widlar at Fairchild Semiconductor, was released. Monolithic ICs consist of a single chip as opposed to a chip and discrete parts (a discrete IC) or multiple chips bonded and connected on a circuit board (a hybrid IC). Almost all modern op-amps are monolithic ICs; however, this first IC did not meet with much success. Issues such as an uneven supply voltage, low gain and a small dynamic range held off the dominance of monolithic op-amps until 1965 when the µA709 (also designed by Bob Widlar) was released.

1966: First varactor bridge op-amps

Since the 741, there have been many different directions taken in op-amp design. Varactor bridge op-amps started to be produced in the late 1960s. They were designed to have extremely small input current and are still amongst the best op-amps available in terms of common-mode rejection with the ability to correctly deal with hundreds of volts at their inputs.

1968: Release of the µA741

The popularity of monolithic op-amps was further improved upon the release of the LM101 in 1967, which solved a variety of issues, and the subsequent release of the µA741 in 1968. The µA741 was extremely similar to the LM101 except that Fairchild's facilities allowed them to include a 30 pF compensation capacitor inside the chip instead of requiring external compensation. This simple difference has made the 741 *the* canonical op-amp and many modern amps base their pinout on the 741s. The µA741 is still in production, and has become ubiquitous in electronics—many manufacturers produce a version of this classic chip, recognizable by part numbers containing *741*. The same part is manufactured by several companies.

1970: First high-speed, low-input current FET design

In the 1970s high speed, low-input current designs started to be made by using FETs. These would be largely replaced by op-amps made with MOSFETs in the 1980s. During the 1970s single sided supply op-amps also became available.

1972: Single sided supply op-amps being produced

A single sided supply op-amp is one where the input and output voltages can be as low as the negative power supply voltage instead of needing to be at least two volts above it. The result is that it can operate in many applications with the negative supply pin on the op-amp being connected to the signal ground, thus eliminating the need for a separate negative power supply.

The LM324 (released in 1972) was one such op-amp that came in a quad package (four separate op-amps in one package) and became an industry standard. In addition to packaging multiple op-amps in a single package, the 1970s also saw the birth of op-amps in hybrid packages. These op-amps were generally improved versions of existing monolithic op-amps. As the properties of monolithic op-amps improved, the more complex hybrid ICs were quickly relegated to systems that are required to have extremely long service lives or other specialty systems.

Recent trends

Recently supply voltages in analog circuits have decreased (as they have in digital logic) and low-voltage op-amps have been introduced reflecting this. Supplies of ±5 V and increasingly 3.3 V (sometimes as low as 1.8 V) are common. To maximize the signal range modern op-amps commonly have rail-to-rail output (the ouput signal can range from the lowest supply voltage to the highest) and sometimes rail-to-rail inputs.

Classification

Op-amps may be classified by their construction:

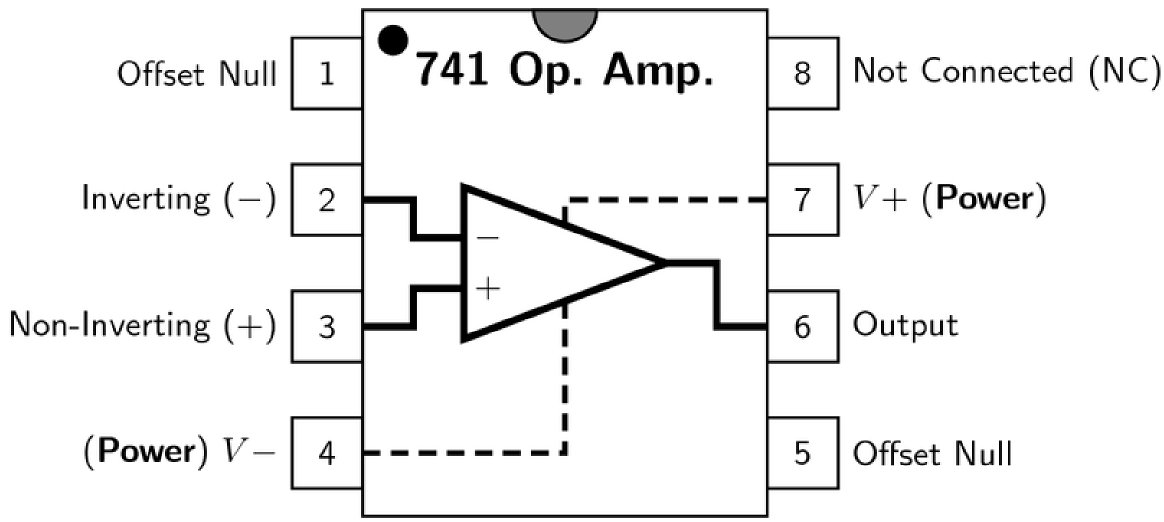
- discrete (built from individual transistors or tubes/valves)
- IC (fabricated in an Integrated circuit) - most common
- hybrid

IC op-amps may be classified in many ways, including:

- Military, Industrial, or Commercial grade (for example: the LM301 is the commercial grade version of the LM101, the LM201 is the industrial version). This may define operating temperature ranges and other environmental or quality factors.
- Classification by package type may also affect environmental hardiness, as well as manufacturing options; DIP, and other through-hole packages are tending to be replaced by Surface-mount devices.
- Classification by internal compensation: op-amps may suffer from high frequency instability in some negative feedback circuits unless a small compensation capacitor modifies the phase- and frequency- responses; op-amps with capacitor built in are termed "*compensated*", or perhaps compensated for closed-loop gains down to (say) 5, others: uncompensated.
- Single, dual and quad versions of many commercial op-amp IC are available, meaning 1, 2 or 4 operational amplifiers are included in the same package.
- Rail-to-rail input (and/or output) op-amps can work with input (and/or output) signals very close to the power supply rails.
- CMOS op-amps (such as the CA3140E) provide extremely high input resistances, higher than JFET-input op-amps, which are normally higher than bipolar-input op-amps.
- other varieties of op-amp include programmable op-amps (simply meaning the quiescent current, gain, bandwidth and so on can be adjusted slightly by an external resistor).

- manufacturers often tabulate their op-amps according to purpose, such as low-noise pre-amplifiers, wide bandwidth amplifiers, and so on.

Applications



DIP pinout for 741-type operational amplifier

Use in electronics system design

The use of op-amps as circuit blocks is much easier and clearer than specifying all their individual circuit elements (transistors, resistors, etc.), whether the amplifiers used are integrated or discrete. In the first approximation op-amps can be used as if they were ideal differential gain blocks; at a later stage limits can be placed on the acceptable range of parameters for each op-amp.

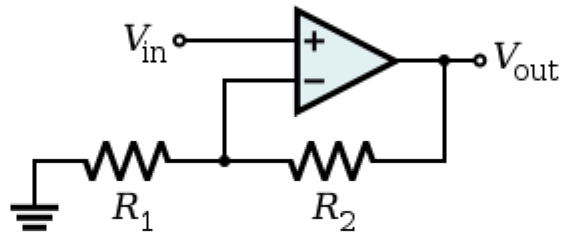
Circuit design follows the same lines for all electronic circuits. A specification is drawn up governing what the circuit is required to do, with allowable limits. For example, the gain may be required to be 100 times, with a tolerance of 5% but drift of less than 1% in a specified temperature range; the input impedance not less than one megohm; etc.

A basic circuit is designed, often with the help of circuit modeling (on a computer). Specific commercially available op-amps and other components are then chosen that meet the design criteria within the specified tolerances at acceptable cost. If not all criteria can be met, the specification may need to be modified.

A prototype is then built and tested; changes to meet or improve the specification, alter functionality, or reduce the cost, may be made.

Basic single stage amplifiers

Non-inverting amplifier



An op-amp connected in the non-inverting amplifier configuration

In a non-inverting amplifier, the output voltage changes in the same direction as the input voltage.

The gain equation for the op-amp is:

$$V_{\mathrm{out}} = (V_+ - V_-)\, A_{OL}$$

However, in this circuit V_- is a function of V_{out} because of the negative feedback through the R_1R_2 network. R_1 and R_2 form a voltage divider, and as V_- is a high-impedance input, it does not load it appreciably. Consequently:

$$V_- \;= \beta \cdot V_{\mathrm{out}}$$

where

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

Substituting this into the gain equation, we obtain:

$$V_{\mathrm{out}} = (V_{\mathrm{in}} - \beta \cdot V_{\mathrm{out}}) \cdot A_{OL}$$

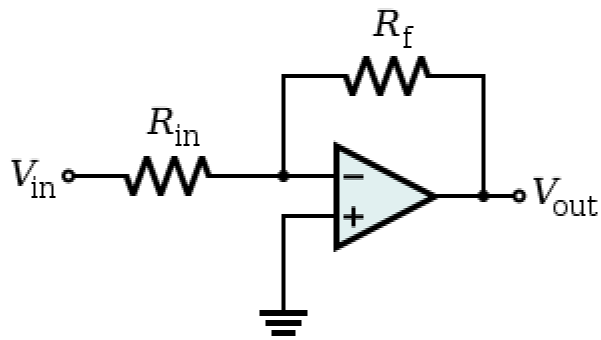
Solving for V_{out}:

$$V_{\mathrm{out}} = V_{\mathrm{in}} \cdot \left(\frac{1}{\beta + 1/A_{OL}}\right)$$

If A_{OL} is very large, this simplifies to

$$V_{\mathrm{out}} \approx \frac{V_{\mathrm{in}}}{\beta} = \frac{V_{\mathrm{in}}}{\frac{R_1}{R_1+R_2}} = V_{\mathrm{in}}(1+\frac{R_2}{R_1})$$.

Inverting amplifier



An op-amp connected in the inverting amplifier configuration

In an inverting amplifier, the output voltage changes in an opposite direction to the input voltage.

As with the non-inverting amplifier, we start with the gain equation of the op-amp:

$$V_{\mathrm{out}} = (V_+ - V_-)\,A_{OL}$$

This time, V_- is a function of both V_{out} and V_{in} due to the voltage divider formed by R_{f} and R_{in}. Again, the op-amp input does not apply an appreciable load, so:

$$V_- \ = \frac{1}{R_{\mathrm{f}} + R_{\mathrm{in}}}\left(R_{\mathrm{f}} V_{\mathrm{in}} + R_{\mathrm{in}} V_{\mathrm{out}}\right)$$

Substituting this into the gain equation and solving for V_{out}:

$$V_{\mathrm{out}} = -V_{\mathrm{in}} \cdot \frac{A_{OL} R_{\mathrm{f}}}{R_{\mathrm{f}} + R_{\mathrm{in}} + A_{OL} R_{\mathrm{in}}}$$

If A_{OL} is very large, this simplifies to

$$V_{\mathrm{out}} \approx -V_{\mathrm{in}} \frac{R_{\mathrm{f}}}{R_{\mathrm{in}}}.$$

A resistor is often inserted between the non-inverting input and ground (so both inputs "see" similar resistances), reducing the input offset voltage due to different voltage drops due to bias current, and may reduce distortion in some op-amps.

A DC-blocking capacitor may be inserted in series with the input resistor when a frequency response down to DC is not needed and any DC voltage on the input is unwanted. That is, the capacitive component of the input impedance inserts a DC zero and a low-frequency pole that gives the circuit a bandpass or high-pass characteristic.

The potentials at the operational amplifier inputs remain virtually constant (near ground) in the inverting configuration. The constant operating potential typically results in distortion levels that are lower than those attainable with the non-inverting topology.

Positive feedback configurations

Another typical configuration of op-amps is with positive feedback, which takes a fraction of the output signal back to the non-inverting input. An important application of it is the comparator with hysteresis, the Schmitt trigger.

Positive voltage level detector

A positive reference voltage V_{ref} is applied to one of the op-amp's inputs. This means that the op-amp is set up as a comparator to detect a positive voltage. If the voltage to be sensed, E_i, is applied to op amp's (+) input, the result is a noninverting positive-level detector. When E_i is above V_{ref}, V_O equals $+V_{sat}$. When E_i is below V_{ref}, V_O equals $-V_{sat}$.

If E_i, is applied to the inverting input, the circuit is an inverting positive-level detector: When E_i is above V_{ref}, V_O equals $-V_{sat}$.

Negative voltage level detector

A negative voltage detector is a circuit that detects when input signal E_i crosses the negative voltage $-V_{ref}$. When E_i is above $-V_{ref}$, V_O equals $+V_{sat}$. When $E_i$is below $-V_{ref}$, V_O equals $-V_{sat}$.When E_i is above $-V_{ref}$, V_O equals $-V_{sat}$, and when $E_i$is below $-V_{ref}$, V_O equals $+V_{sat}$.

Sine to square wave converter

The zero detector will convert the output of a sine-wave from a function generator into a variable-frequency square wave. If E_i is a sine wave, triangular wave, or wave of any other shape that is symmetrical around zero, the zero-crossing detector's output will be square.

Because of the wide slew-range and lack of positive feedback, the response of all the level detectors described above will be relatively slow. Using a general-purpose op-amp, for example, the frequency of E_i for the sine to square wave converter should probably be below 100 Hz.

Other applications

- audio- and video-frequency pre-amplifiers and buffers
- voltage comparators
- differential amplifiers
- differentiators and integrators
- filters
- precision rectifiers
- precision peak detectors
- voltage and current regulators
- analog calculators
- analog-to-digital converters
- digital-to-analog converter
- voltage clamps
- oscillators and waveform generators

Most single, dual and quad op-amps available have a standardized pin-out which permits one type to be substituted for another without wiring changes. A specific op-amp may be chosen for its open loop gain, bandwidth, noise performance, input impedance, power consumption, or a compromise between any of these factors.

Limitations of real op-amps

Real op-amps differ from the ideal model in various respects.

DC imperfections

Real operational amplifiers suffer from several non-ideal effects:

Finite gain

> Open-loop gain is infinite in the ideal operational amplifier but finite in real operational amplifiers. Typical devices exhibit open-loop DC gain ranging from 100,000 to over 1 million. So long as the loop gain (i.e., the product of open-loop and feedback gains) is very large, the circuit gain will be determined entirely by the amount of negative feedback (i.e., it will be independent of open-loop gain). In cases where closed-loop gain must be very high, the feedback gain will be very low, and the low feedback gain causes low loop gain; in these cases, the operational amplifier will cease to behave ideally.

Finite input impedances

> The *differential input impedance* of the operational amplifier is defined as the impedance *between* its two inputs; the *common-mode input impedance* is the impedance from each input to ground. MOSFET-input operational amplifiers often have protection circuits that effectively short circuit any input differences greater than a small threshold, so the input impedance can appear to be very low in some tests. However, as long as these operational amplifiers are used in a typical high-gain negative feedback application, these protection circuits will be

inactive. The input bias and leakage currents described below are a more important design parameter for typical operational amplifier applications.

Non-zero output impedance

Low output impedance is important for low-impedance loads; for these loads, the voltage drop across the output impedance of the amplifier will be significant. Hence, the output impedance of the amplifier limits the maximum power that can be provided. In a negative-feedback configuration, the output impedance of the amplifier is effectively lowered; thus, in linear applications, op-amps usually exhibit a very low output impedance indeed. Negative feedback can not, however, reduce the limitations that R_{load} in conjunction with R_{out} place on the maximum and minimum possible output voltages; it can only reduce output errors *within* that range.

Low-impedance outputs typically require high quiescent (i.e., idle) current in the output stage and will dissipate more power, so low-power designs may purposely sacrifice low output impedance.

Input current

Due to biasing requirements or leakage, a small amount of current (typically ~10 nanoamperes for bipolar op-amps, tens of picoamperes for JFET input stages, and only a few pA for MOSFET input stages) flows into the inputs. When large resistors or sources with high output impedances are used in the circuit, these small currents can produce large unmodeled voltage drops. If the input currents are matched, *and* the impedance looking *out* of *both* inputs are matched, then the voltages produced at each input will be equal. Because the operational amplifier operates on the *difference* between its inputs, these matched voltages will have no effect (unless the operational amplifier has poor CMRR, which is described below). It is more common for the input currents (or the impedances looking out of each input) to be slightly mismatched, and so a small *offset voltage* can be produced. This offset voltage can create offsets or drifting in the operational amplifier. It can often be nulled externally; however, many operational amplifiers include *offset null* or *balance* pins and some procedure for using them to remove this offset. Some operational amplifiers attempt to nullify this offset automatically.

Input offset voltage

This voltage, which is what is required across the op-amp's input terminals to drive the output voltage to zero, is related to the mismatches in input bias current. In the perfect amplifier, there would be no input offset voltage. However, it exists in actual op-amps because of imperfections in the differential amplifier that constitutes the input stage of the vast majority of these devices. Input offset voltage creates two problems: First, due to the amplifier's high voltage gain, it virtually assures that the amplifier output will go into saturation if it is operated without negative feedback, even when the input terminals are wired together. Second, in a closed loop, negative feedback configuration, the input offset voltage is amplified along with the signal and this may pose a problem if high precision DC amplification is required or if the input signal is very small.

Common mode gain

A perfect operational amplifier amplifies only the voltage difference between its two inputs, completely rejecting all voltages that are common to both. However, the differential input stage of an operational amplifier is never perfect, leading to the amplification of these identical voltages to some degree. The standard measure of this defect is called the common-mode rejection ratio (denoted CMRR). Minimization of common mode gain is usually important in non-inverting amplifiers (described below) that operate at high amplification.

Temperature effects

All parameters change with temperature. Temperature drift of the input offset voltage is especially important.

Power-supply rejection

The output of a perfect operational amplifier will be completely independent from ripples that arrive on its power supply inputs. Every real operational amplifier has a specified power supply rejection ratio (PSRR) that reflects how well the op-amp can reject changes in its supply voltage. Copious use of bypass capacitors can improve the PSRR of many devices, including the operational amplifier.

Drift

Real op-amp parameters are subject to slow change over time and with changes in temperature, input conditions, etc.

Noise

Amplifiers generate random voltage at the output even when there is no signal applied. This can be due to thermal noise and flicker noise of the devices. For applications with high gain or high bandwidth, noise becomes a very important consideration.

AC imperfections

The op-amp gain calculated at DC does not apply at higher frequencies. To a first approximation, the gain of a typical op-amp is inversely proportional to frequency. This means that an op-amp is characterized by its gain-bandwidth product. For example, an op-amp with a gain bandwidth product of 1 MHz would have a gain of 5 at 200 kHz, and a gain of 1 at 1 MHz. This low-pass characteristic is introduced deliberately, because it tends to stabilize the circuit by introducing a dominant pole. This is known as frequency compensation.

Typical low cost, general purpose op-amps exhibit a gain bandwidth product of a few megahertz. Specialty and high speed op-amps can achieve gain bandwidth products of hundreds of megahertz. For very high-frequency circuits, a completely different form of op-amp called the current-feedback operational amplifier is often used.

Other imperfections include:

Finite bandwidth

All amplifiers have a finite bandwidth. This creates several problems for op amps. First, associated with the bandwidth limitation is a phase difference between the input signal and the amplifier output that can lead to oscillation in some feedback circuits. The internal frequency compensation used in some op amps to increase the gain or phase margin intentionally reduces the bandwidth even further to maintain output stability when using a wide variety of feedback networks. Second, reduced bandwidth results in lower amounts of feedback at higher frequencies, producing higher distortion, noise, and output impedance and also reduced output phase linearity as the frequency increases.

Input capacitance

Most important for high frequency operation because it further reduces the open loop bandwidth of the amplifier.

Non-linear imperfections

Saturation

output voltage is limited to a minimum and maximum value close to the power supply voltages. Saturation occurs when the output of the amplifier reaches this value and is usually due to:

- In the case of an op-amp using a bipolar power supply, a voltage gain that produces an output that is more positive or more negative than that maximum or minimum; or
- In the case of an op-amp using a single supply voltage, either a voltage gain that produces an output that is more positive than that maximum, or a signal so close to ground that the amplifier's gain is not sufficient to raise it above the lower threshold.

Slewing

the amplifier's output voltage reaches its maximum rate of change. Measured as the slew rate, it is usually specified in volts per microsecond. When slewing occurs, further increases in the input signal have no effect on the rate of change of the output. Slewing is usually caused by internal capacitances in the amplifier, especially those used to implement its frequency compensation.

Non-linear input-output relationship

The output voltage may not be accurately proportional to the difference between the input voltages. It is commonly called distortion when the input signal is a waveform. This effect will be very small in a practical circuit if substantial negative feedback is used.

Power considerations

Limited output current

The output current must be finite. In practice, most op-amps are designed to limit the output current so as not to exceed a specified level — around 25 mA for a type 741 IC op-amp — thus protecting the op-amp and associated circuitry from damage. Modern designs are electronically more rugged than earlier implementations and some can sustain direct short circuits on their outputs without damage.

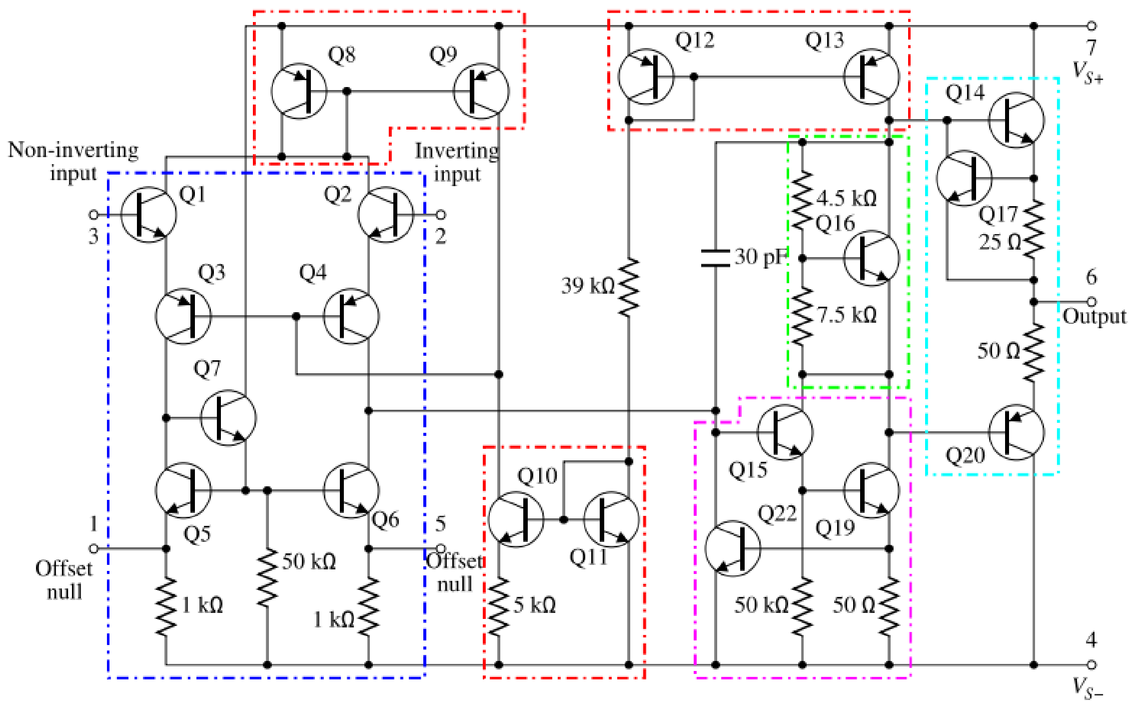
Limited dissipated power

The output current flows through the op-amp's internal output impedance, dissipating heat. If the op-amp dissipates too much power, then its temperature will increase above some safe limit. The op-amp may enter thermal shutdown, or it may be destroyed.

Modern integrated FET or MOSFET op-amps approximate more closely the ideal op-amp than bipolar ICs when it comes to input impedance and input bias and offset currents. Bipolars are generally better when it comes to input *voltage* offset, and often have lower noise. Generally, at room temperature, with a fairly large signal, and limited bandwidth, FET and MOSFET op-amps now offer better performance.

Internal circuitry of 741 type op-amp

Though designs vary between products and manufacturers, all op-amps have basically the same internal structure, which consists of three stages:



A component level diagram of the common 741 op-amp. Dotted lines outline: current mirrors (red); differential amplifier (blue); class A gain stage (magenta); voltage level shifter (green); output stage (cyan).

1. Differential amplifier – provides low noise amplification, high input impedance, usually a differential output.
2. Voltage amplifier – provides high voltage gain, a single-pole frequency roll-off, usually single-ended output.
3. Output amplifier – provides high current driving capability, low output impedance, current limiting and short circuit protection circuitry.

Input stage

Constant-current stabilization system

The input stage DC conditions are stabilized by a high-gain negative feedback system whose main parts are the two current mirrors on the left of the figure, outlined in red. The main purpose of this negative feedback system—to supply the differential input stage with a stable constant current—is realized as follows.

The current through the 39 kΩ resistor acts as a current reference for the other bias currents used in the chip. The voltage across the resistor is equal to the voltage across the supply rails ($V_{S+} - V_{S-}$) minus two transistor diode drops (i.e., from Q11 and Q12), and so the current has value $I_{\text{ref}} = (V_{S+} - V_{S-} - 2V_{be})/(39\ \text{k}\Omega)$. The Widlar current mirror built by Q10, Q11, and the 5 kΩ resistor produces a very small fraction of I_{ref} at the Q10 collector. This small constant current through Q10's collector supplies the base currents for Q3 and Q4 as well as the Q9 collector current. The Q8/Q9 current mirror tries to make Q9's collector current the same as the Q3 and Q4 *collector* currents. Thus Q3 and Q4's combined base currents (which are of the same order as the overall chip's input currents) will be a small fraction of the already small Q10 current.

So, if the input stage current increases for any reason, the Q8/Q9 current mirror will draw current away from the bases of Q3 and Q4, which reduces the input stage current, and vice versa. The feedback loop also isolates the rest of the circuit from common-mode signals by making the base voltage of Q3/Q4 follow tightly $2V_{be}$ below the higher of the two input voltages.

Differential amplifier

The blue outlined section is a differential amplifier. Q1 and Q2 are input emitter followers and together with the common base pair Q3 and Q4 form the differential input stage. In addition, Q3 and Q4 also act as level shifters and provide voltage gain to drive the class A amplifier. They also help to increase the reverse V_{be} rating on the input transistors (the emitter-base junctions of the NPN transistors Q1 and Q2 break down at around 7 V but the PNP transistors Q3 and Q4 have breakdown voltages around 50 V).

The differential amplifier formed by Q1–Q4 drives a current mirror active load formed by transistors Q5–Q7 (actually, Q6 is the very active load). Q7 increases the accuracy of the current mirror by decreasing the amount of signal current required from Q3 to drive the bases of Q5 and Q6. This configuration provides differential to single ended conversion as follows:

The signal current of Q3 is the input to the current mirror while the output of the mirror (the collector of Q6) is connected to the collector of Q4. Here, the signal currents of Q3 and Q4 are summed. For differential input signals, the signal currents of Q3 and Q4 are equal and opposite. Thus, the sum is twice the individual signal currents. This completes the differential to single ended conversion.

The open circuit signal voltage appearing at this point is given by the product of the summed signal currents and the paralleled collector resistances of Q4 and Q6. Since the collectors of Q4 and Q6 appear as high resistances to the signal current, the open circuit voltage gain of this stage is very high.

The base current at the inputs is not zero and the effective (differential) input impedance of a 741 is about 2 MΩ. The "offset null" pins may be used to place external resistors in parallel with the two 1 kΩ resistors (typically in the form of the two ends of a potentiometer) to adjust the balancing of the Q5/Q6 current mirror and thus indirectly control the output of the op-amp when zero signal is applied between the inputs.

Class A gain stage

The section outlined in magenta is the class A gain stage. The top-right current mirror Q12/Q13 supplies this stage by a constant current load, via the collector of Q13, that is largely independent of the output voltage. The stage consists of two NPN transistors in a Darlington configuration and uses the output side of a current mirror as its collector load to achieve high gain. The 30 pF capacitor provides frequency selective negative feedback around the class A gain stage as a means of frequency compensation to stabilise the amplifier in feedback configurations. This technique is called Miller compensation and functions in a similar manner to an op-amp integrator circuit. It is also known as 'dominant pole compensation' because it introduces a dominant pole (one which masks the effects of other poles) into the open loop frequency response. This pole can be as low as 10 Hz in a 741 amplifier and it introduces a −3 dB loss into the open loop response at this frequency. This internal compensation is provided to achieve unconditional stability of the amplifier in negative feedback configurations where the feedback network is non-reactive and the closed loop gain is unity or higher. Hence, the use of the operational amplifier is simplified because no external compensation is required for unity gain stability; amplifiers without this internal compensation may require external compensation or closed loop gains significantly higher than unity.

Output bias circuitry

The green outlined section (based on Q16) is a voltage level shifter or rubber diode (i.e., a V_{BE} multiplier); a type of voltage source. In the circuit as shown, Q16 provides a constant voltage drop between its collector and emitter regardless of the current through the circuit. If the base current to the transistor is assumed to be zero, and the voltage between base and emitter (and across the 7.5 kΩ resistor) is 0.625 V (a typical value for a BJT in the active region), then the current through the 4.5 kΩ resistor will be the same as that through the 7.5 kΩ, and will produce a voltage of 0.375 V across it. This keeps the voltage across the transistor, and the two resistors at 0.625 + 0.375 = 1 V. This serves to bias the two output transistors slightly into conduction reducing crossover distortion. In some discrete component amplifiers this function is achieved with (usually two) silicon diodes.

Output stage

The output stage (outlined in cyan) is a Class AB push-pull emitter follower (Q14, Q20) amplifier with the bias set by the V_{be} multiplier voltage source Q16 and its base resistors. This stage is effectively driven by the collectors of Q13 and Q19. Variations in the bias with temperature, or between parts with the same type number, are common so crossover distortion and quiescent current may be subject to significant variation. The output range of the amplifier is about one volt less than the supply voltage, owing in part to V_{be} of the output transistors Q14 and Q20.

The 25 Ω resistor in the output stage acts as a current sense to provide the output current-limiting function which limits the current in the emitter follower Q14 to about 25 mA for the 741. Current limiting for the negative output is done by sensing the voltage across Q19's emitter resistor and using this to reduce the drive into Q15's base. Later versions of this amplifier schematic may show a slightly different method of output current limiting. The output resistance is not zero, as it would be in an ideal op-amp, but with negative feedback it approaches zero at low frequencies.

Note: while the 741 was historically used in audio and other sensitive equipment, such use is now rare because of the improved noise performance of more modern op-amps. Apart from generating noticeable hiss, 741s and other older op-amps may have poor common-mode rejection ratios and so will often introduce cable-borne mains hum and other common-mode interference, such as switch 'clicks', into sensitive equipment.

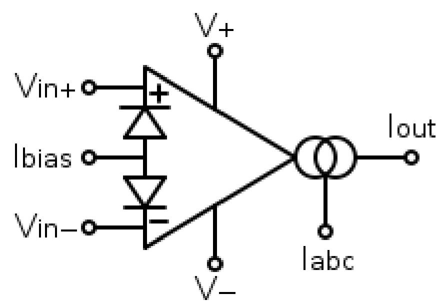
The "741" has come to often mean a generic op-amp IC (such as uA741, LM301, 558, LM324, TBA221 - or a more modern replacement such as the TL071). The description of the 741 output stage is qualitatively similar for many other designs (that may have quite different input stages), except:

- Some devices (uA748, LM301, LM308) are not internally compensated (require an external capacitor from output to some point within the operational amplifier, if used in low closed-loop gain applications).
- Some modern devices have rail-to-rail output capability (output can be taken to positive or negative power supply rail within a few millivolts).

Chapter 5

Operational Transconductance Amplifier & Low-Dropout Regulator

Operational Transconductance Amplifier



Schematic symbol for the OTA. Like the standard operational amplifier, it has both inverting (−) and noninverting (+) inputs; power supply lines (V+ and V−); and a single output. Unlike the traditional op-amp, it has two additional biasing inputs, I_{abc} and I_{bias}, explained in **Basic operation** and **Subsequent improvements**, below.

The **operational transconductance amplifier** (**OTA**) is an amplifier whose differential input voltage produces an output current. Thus, it is a voltage controlled current source (VCCS). There is usually an additional input for a current to control the amplifier's transconductance. The OTA is similar to a standard operational amplifier in that it has a high impedance differential input stage and that it may be used with negative feedback.

The first commercially available integrated circuit units were produced by RCA in 1969 (before being acquired by General Electric), in the form of the CA3080, and they have been improved since that time. Although most units are constructed with bipolar transistors, field effect transistor units are also produced. The OTA is not as useful by itself in the vast majority of standard op-amp functions as the ordinary op-amp because its output is a current. One of its principal uses is in implementing electronically controlled applications such as variable frequency oscillators and filters and variable gain amplifier stages which are more difficult to implement with standard op-amps.

Principal differences from standard operational amplifiers

- Its output of a *current* contrasts to that of standard operational amplifier whose output is a *voltage*.
- It is usually used "open-loop"; without negative feedback in linear applications. This is possible because the magnitude of the resistance attached to its output controls its output voltage. Therefore a resistance can be chosen that keeps the output from going into saturation, even with high differential input voltages.

Basic operation

In the ideal OTA, the output current is a linear function of the differential input voltage, calculated as follows:

$$I_{\mathrm{out}} = (V_{\mathrm{in+}} - V_{\mathrm{in-}}) \cdot g_{\mathrm{m}}$$

where $V_{\mathrm{in+}}$ is the voltage at the non-inverting input, $V_{\mathrm{in-}}$ is the voltage at the inverting input and g_m is the transconductance of the amplifier.

The amplifier's output voltage is the product of its output current and its load resistance:

$$V_{\mathrm{out}} = I_{\mathrm{out}} \cdot R_{\mathrm{load}}$$

The voltage gain is then the output voltage divided by the differential input voltage:

$$G_{\mathrm{voltage}} = \frac{V_{\mathrm{out}}}{(V_{\mathrm{in+}} - V_{\mathrm{in-}})} = R_{\mathrm{load}} \cdot g_{\mathrm{m}}$$

The transconductance of the amplifier is usually controlled by an input current, denoted I_{abc} ("amplifier bias current"). The amplifier's transconductance is directly proportional to this current. This is the feature that makes it useful for electronic control of amplifier gain, etc.

Non-ideal characteristics

As with the standard op-amp, practical OTA's have some non-ideal characteristics. These include:

- Input stage non-linearity at higher differential input voltages due to the characteristics of the input stage transistors. In the early devices, such as the CA 3080, the input stage consisted of two bipolar transistors connected in the differential amplifier configuration. The transfer characteristics of this connection are approximately linear for differential input voltages of 20 mV or less. This is an important limitation when the OTA is being used open loop as there is no negative feedback to linearize the output. One scheme to improve this parameter is mentioned below.

- Temperature sensitivity of transconductance.

- Variation of input and output impedance, input bias current and input offset voltage with the transconductance control current I_{abc}.

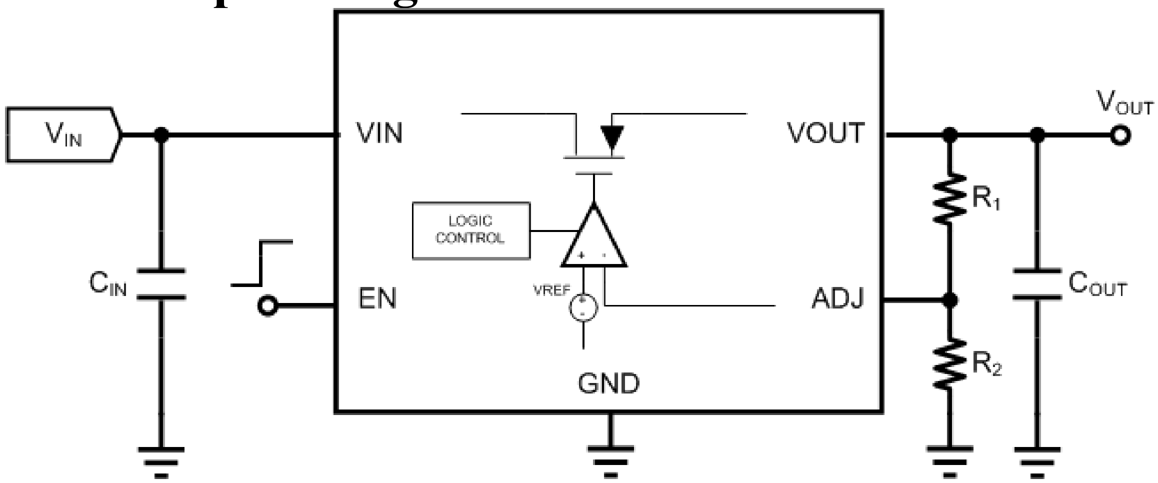
Subsequent improvements

Earlier versions of the OTA had neither the I_{bias} terminal shown in the diagram nor the diodes shown adjacent to it. They were all added in later versions. As depicted in the diagram, the anodes of the diodes are attached together and the cathode of one is attached to the non inverting input (Vin+) and the cathode of the other to the inverting input (Vin−). The diodes are biased at the anodes by a current (I_{bias}) that is injected into the I_{bias} terminal. These additions make two substantial improvements to the OTA. First, when used with input resistors, the diodes distort the differential input voltage to offset a significant amount of input stage non linearity at higher differential input voltages. According to National Semiconductor, the addition of these diodes increases the linearity of the input stage by a factor of 4. That is, using the diodes, the signal distortion level at 80 mV of differential input is the same as that of the simple differential amplifier at a differential input of 20 mV. Second, the action of the biased diodes offsets much of the temperature sensitivity of the OTA's transconductance.

A second improvement is the integration of an optional-use output buffer amplifier to the chip on which the OTA resides. This is actually a convenience to a circuit designer rather than an improvement to the OTA itself; dispensing with the need to employ a separate buffer. It also allows the OTA to be used as a traditional op-amp, if desired, by converting its output current to a voltage.

An example of a chip combining both of these features is the National Semiconductor LM13600 and its successor, the LM13700, the data sheet for which can be found here:

Low-Dropout Regulator



Schematic of a low-dropout regulator

A **low-dropout** or **LDO** regulator is a DC linear voltage regulator which can operate with a very small input–output differential voltage. The advantages of a low dropout voltage include a lower minimum operating voltage, higher efficiency operation and lower heat dissipation.

The main components are a power FET and a differential amplifier (error amplifier). One input of the differential amplifier monitors the fraction of the output determined by the resistor ratio of **R1** and **R2**. The second input to the differential amplifier is from a stable voltage reference (bandgap reference). If the output voltage rises too high relative to the reference voltage, the drive to the power FET changes to maintain a constant output voltage.

Overview

The adjustable low-dropout regulator debuted on April 12, 1977 in an Electronic Design article entitled "*Break Loose from Fixed IC Regulators*". The article was written by Robert Dobkin, an IC designer then working for National Semiconductor. Because of this, National Semiconductor claims the title of "*LDO inventor*". Dobkin later left National in 1981 to found Linear Technology where he is currently chief technology officer.

Current advances in semiconductor technology made possible to significantly reduce voltage drop on all linear voltage regulators, so producers today advertise almost everything as low-dropout. For example, the AS1117 is advertised as low drop out even if it uses same topology as 78xx series.

Regulation

Low-dropout (LDO) regulators work in the same way as all linear voltage regulators. The main difference between LDO and non-LDO regulators is their schematic topology. Instead of an emitter follower topology, Low-dropout regulators utilize open collector or open drain topology. This enables transistor saturation, which limits the voltage drop to only the saturation voltage.

If a bipolar transistor is used however, significant additional power is lost to control it, when non-LDO regulators take that power from voltage drop itself. For high voltages under very low In-Out difference there will be significant power loss in control circuit.

Because the power control element is working as an inverter, another inverting amplifier is required to control it, which increases schematic complexity compared to a simple voltage stabilizer.

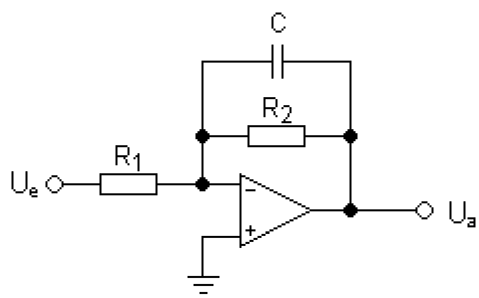
Power FETs may be preferable to reduce power consumption, but this significantly increases price and poses problems when the regulator is used for low input voltage, because FETs usually require 10 to 5V to open completely.

Quiescent current

Among other important characteristics is the quiescent current (the current flowing through the system when no load is present), which creates a difference between the input and output currents. The series pass element, topologies, and ambient temperature are the primary contributors to quiescent current. Quiescent current and input–output drop limit the efficiency of LDO regulators and should thus be minimized.

Chapter 6

Electronic Filter Topology



An elementary filter topology introduces a capacitor into the feedback path of an op-amp to achieve an unbalanced active implementation of a low-pass transfer function

Electronic filter topology defines electronic filter circuits without taking note of the values of the components used but only the manner in which those components are connected.

Filter design characterises filter circuits primarily by their transfer function rather than their topology. Transfer functions may be linear or nonlinear. Common types of linear filter transfer function are; high-pass, low-pass, bandpass, band-reject or notch and all-pass. Once the transfer function for a filter is chosen the particular topology to implement such a prototype filter can be selected so that, for example, one might choose to design a Butterworth filter using the Sallen–Key topology.

Filter topologies may be divided into passive and active types. Passive topologies do not contain a generator of energy, either in reality or, due to non-linearity, in their equivalent circuit, but only capacitors, inductors and also, in some topologies, resistors: particularly, variable resistors are often included in order to control the depth of filtering. Active topologies include components that require power. Further, topologies may be implemented either in unbalanced form or else in balanced form when employed in balanced circuits. Implementations such as electronic mixers and stereo sound may require arrays of identical circuits.

Passive topologies

Passive filters have been long in development and use. Most are built from simple two-port networks called "sections". There is no formal definition of a section except that it must have at least one series component and one shunt component. Sections are invariably connected in a "cascade" or "daisy-chain" topology, consisting of either repeats of the same section or of completely different sections. Impedance would combine two sections consisting only of series components or shunt components into a single section.

Some passive filters, consisting of only one or two filter sections, are given special names including the L-section, T-section and Π-section, which are unbalanced filters, and the C-section, H-section and box-section, which are balanced. All are built upon a very simple "ladder" topology (see below). The chart at the bottom of the page shows these various topologies in terms of general constant k filters.

Filters designed using network synthesis usually repeat the simplest form of L-section topology though component values may change in each section. Image designed filters, on the other hand, keep the same basic component values from section to section though the topology may vary and tend to make use of more complex sections.

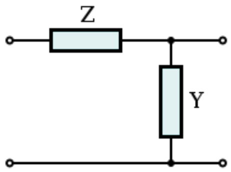
L-sections are never symmetrical but two L-sections back-to-back form a symmetrical topology and many other sections are symmetrical in form.

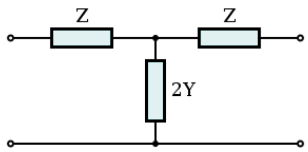
Ladder topologies

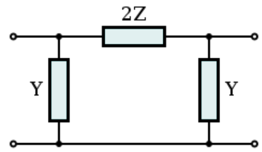
Ladder topology, often called **Cauer topology** after Wilhelm Cauer (inventor of the Elliptical filter), was in fact first used by George Campbell (inventor of the Constant k filter). Campbell published in 1922 but had clearly been using the topology for some time before this. Cauer first picked up on ladders (published 1926) inspired by the work of Foster (1924). There are two forms of basic ladder topologies; unbalanced and balanced. Cauer topology is usually thought of as an unbalanced ladder topology.

A ladder network consists of cascaded asymmetrical L-sections (unbalanced) or C-sections (balanced). In low pass form the topology would consist of series inductors and shunt capacitors. Other bandforms would have an equally simple topology transformed from the lowpass topology. The transformed network will have shunt admittances that are dual networks of the series impedances if they were duals in the starting network - which is the case with series inductors and shunt capacitors.

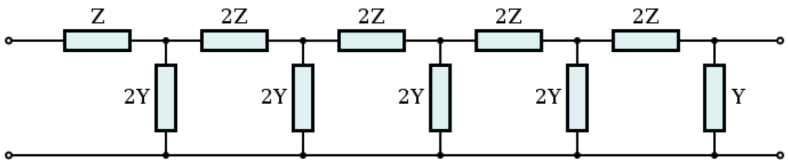
Image filter sections		
Unbalanced		
L Half section	T Section	Π Section



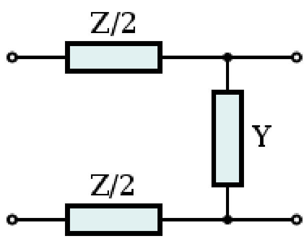
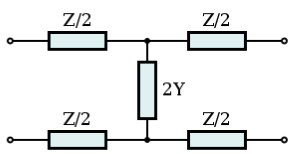
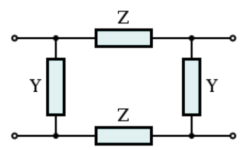




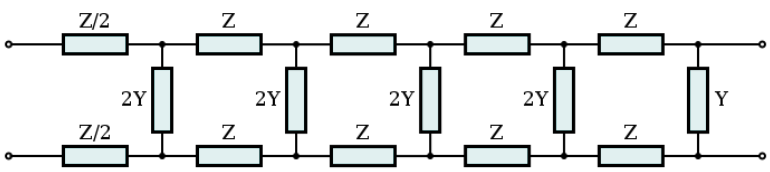
Ladder network

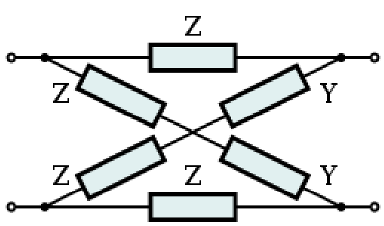
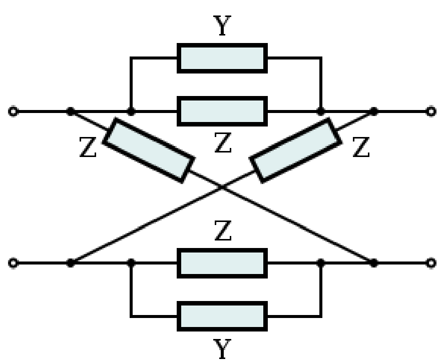


Balanced

C Half-section	H Section	Box Section
		

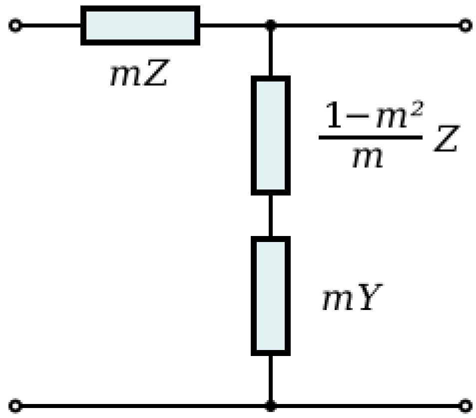
Ladder network



X Section (mid-T-Derived)	X Section (mid-Π-Derived)
	

N.B. Textbooks and design drawings usually show the unbalanced implementations, but in telecoms it is often required to convert the design to the balanced implementation when used with balanced lines.

Modified ladder topologies



series m-derived topology

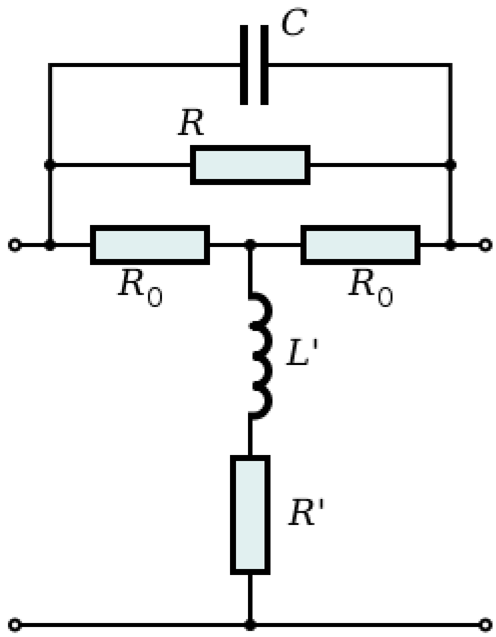
Image filter design commonly uses modifications of the basic ladder topology. These topologies, invented by Otto Zobel, have the same passbands as the ladder on which they are based but their transfer functions are modified to improve some parameter such as impedance matching, stopband rejection or passband-to-stopband transition steepness. Usually the design applies some transform to a simple ladder topology: the resulting topology is ladder-like but no longer obeys the rule that shunt admittances are the dual network of series impedances: it invariably becomes more complex with higher component count. Such topologies include;

- m-derived filter
- mm'-type filter
- General m_n-type filter

The m-type (m-derived) filter is by far the most commonly used modified image ladder topology. There are two m-type topologies for each of the basic ladder topologies; the series-derived and shunt-derived topologies. These have identical transfer functions to each other but different image impedances. Where a filter is being designed with more than one passband, the m-type topology will result in a filter where each passband has an analogous frequency-domain response. It is possible to generalise the m-type topology for filters with more than one passband using parameters m_1, m_2, m_3 etc., which are not equal to each other resulting in general m_n-type filters which have bandforms that can differ in different parts of the frequency spectrum.

The mm'-type topology can be thought of as a double m-type design. Like the m-type it has the same bandform but offers further improved transfer characteristics. It is, however, a rarely used design due to increased component count and complexity as well as its normally requiring basic ladder and m-type sections in the same filter for impedance matching reasons. It is normally only found in a composite filter.

Bridged-T topologies



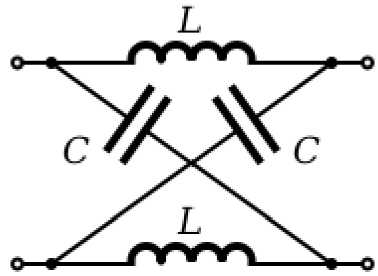
Typical bridged-T Zobel network equaliser used to correct high end roll-off

Zobel constant resistance filters use a topology that is somewhat different from other filter types, distinguished by having a constant input resistance at all frequencies and in that they use resistive components in the design of their sections. The higher component and section count of these designs usually limits their use to equalisation applications. Topologies usually associated with constant resistance filters are the bridged-T and its variants, all described in the Zobel network article;

- Bridged-T topology
- Balanced bridged-T topology
- Open-circuit L-section topology
- Short-circuit L-section topology
- Balanced open-circuit C-section topology
- Balanced short-circuit C-section topology

The bridged-T topology is also used in sections intended to produce a signal delay but in this case no resistive components are used in the design.

Lattice topology



Lattice topology X-section phase correction filter

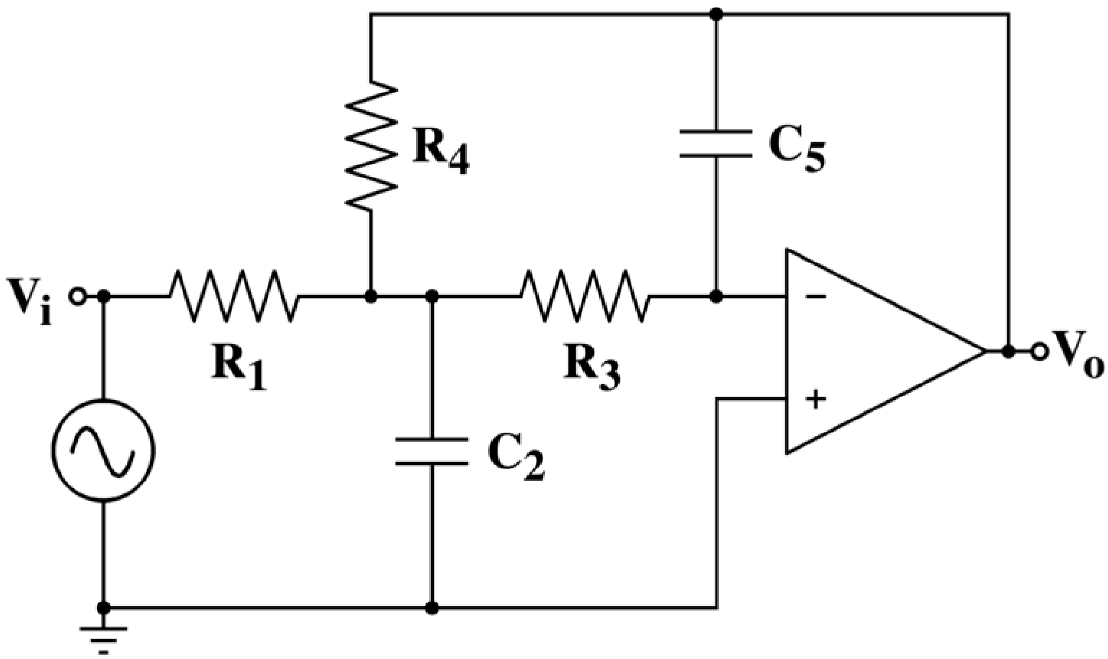
Both the T-section (from ladder topology) and the bridge-T (from Zobel topology) can be transformed into a lattice topology filter section but in both cases this results in high component count and complexity. The most common application of lattice filters (X-sections) is in all-pass filters used for phase equalisation.

Although T and bridged-T sections can always be transformed into X-sections the reverse is not always possible because of the possibility of negative values of inductance and capacitance arising in the transform.

Lattice topology is identical to the more familiar **bridge topology**, the difference being merely the drawn representation on the page rather than any real difference in topology, cicuitry or function.

Active topologies

Multiple feedback topology



Multiple feedback topology circuit.

Multiple feedback topology is an electronic filter topology which is used to implement an electronic filter by adding two poles to the transfer function. A diagram of the circuit topology for a second order low pass filter is shown in the figure on the right.

The transfer function of the multiple feedback topology circuit, like all second-order linear filters, is:

$$H(s) = \frac{V_o}{V_i} = -\frac{1}{As^2 + Bs + C} = \frac{K\omega_0^2}{s^2 + \frac{\omega_0}{Q}s + \omega_0^2}.$$

In an MF filter,

$$A = (R_1 R_3 C_2 C_5)$$
$$B = R_3 C_5 + R_1 C_5 + R_1 R_3 C_5 / R_4$$
$$C = R_1 / R_4$$
$$Q = \frac{\sqrt{R_3 R_4 C_2 C_5}}{(R_4 + R_3 + |K| R_3) C_5}$$ is the Q factor.

$$K = -R_4 / R_1$$ is the DC voltage gain

$\omega_0 = 2\pi f_0 = 1/\sqrt{R_3 R_4 C_2 C_5}$ is the corner frequency

Biquad filter

A **biquad filter** is a type of linear filter that implements a transfer function that is the ratio of two quadratic functions. The name *biquad* is short for *biquadratic*.

Biquad filters are typically active and implemented with a **single-amplifier biquad (SAB)** or **two-integrator-loop** topology.

- The SAB topology uses feedback to generate complex poles and possibly complex zeros. In particular, the feedback moves the real poles of an RC circuit in order to generate the proper filter characteristics.
- The two-integrator-loop topology is derived from rearranging a biquadratic transfer function. The rearrangement will equate one signal with the sum of another signal, its integral, and the integral's integral. In other words, the rearrangement reveals a state variable filter structure. By using different states as outputs, any kind of second-order filter can be implemented.

The SAB topology is sensitive to component choice and can be more difficult to adjust. Hence, usually the term **biquad** refers to the two-integrator-loop state variable filter topology.

Tow-Thomas Biquad Example

For example, the basic configuration in Figure 1 can be used as either a low-pass or bandpass filter depending on where the output signal is taken from.

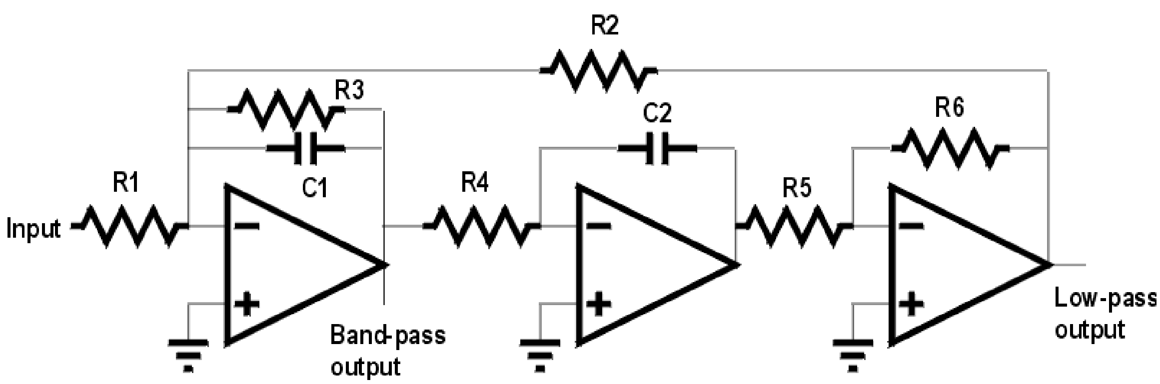


Figure 1: The common Tow-Thomas biquad filter topology.

The second-order low-pass transfer function is given by

$$H(s) = \frac{G_{lpf}\omega_0^2}{s^2 + \frac{\omega_0}{Q}s + \omega_0^2}$$

where low-pass gain $G_{lpf} = R_2 / R_1$. The second-order bandpass transfer function is given by

$$H(s) = \frac{G_{bpf} \frac{\omega_0}{Q} s}{s^2 + \frac{\omega_0}{Q} s + \omega_0^2}.$$

with bandpass gain $G_{bpf} = - R_4 / R_2$. In both cases, the

- Natural frequency is $\omega_0 = 1/\sqrt{R_2 R_4 C_1 C_2}$.
- Quality factor is $Q = \sqrt{\frac{R_3^2 C_1}{R_2 R_4 C_2}}$.

The bandwidth is approximated by $B = \omega_0 / Q$, and Q is sometimes expressed as a damping constant $\zeta = 1 / 2Q$. If a noninverting low-pass filter is required, the output can be taken at the output of the second operational amplifier. If a noninverting bandpass filter is required, the order of the second integrator and the inverter can be switched, and the output taken at the output of the inverter's operational amplifier.

Chapter 7

RC Circuit

A **resistor–capacitor circuit (RC circuit)**, or **RC filter** or **RC network**, is an electric circuit composed of resistors and capacitors driven by a voltage or current source. A first order RC circuit is composed of one resistor and one capacitor and is the simplest type of RC circuit.

RC circuits can be used to filter a signal by blocking certain frequencies and passing others. The four most common RC filters are the high-pass filter, low-pass filter, band-pass filter, and band-stop filter.

Introduction

There are three basic, linear passive lumped analog circuit components: the resistor (R), capacitor (C) and inductor (L). These may be combined in: the RC circuit, the RL circuit, the LC circuit and the RLC circuit with the abbreviations indicating which components are used. These circuits, between them, exhibit a large number of important types of behaviour that are fundamental to much of analog electronics. In particular, they are able to act as passive filters.

Natural response

The simplest RC circuit is a capacitor and a resistor in series. When a circuit consists of only a charged capacitor and a resistor, the capacitor will discharge its stored energy through the resistor. The voltage across the capacitor, which is time dependent, can be found by using Kirchhoff's current law, where the current through the capacitor must equal the current through the resistor. This results in the linear differential equation

$$C\frac{dV}{dt} + \frac{V}{R} = 0.$$

Solving this equation for V yields the formula for exponential decay:

$$V(t) = V_0 e^{-\frac{t}{RC}} \ ,$$

where V_0 is the capacitor voltage at time $t = 0$.

The time required for the voltage to fall to $\frac{V_0}{e}$ is called the RC time constant and is given by

$$\tau = RC \ .$$

Complex impedance

The complex impedance, Z_C (in ohms) of a capacitor with capacitance C (in farads) is

$$Z_C = \frac{1}{sC}$$

The complex frequency s is, in general, a complex number,

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

where

- j represents the imaginary unit:

  $j^2 = -1$

- σ is the exponential decay constant (in radians per second), and
- ω is the sinusoidal angular frequency (also in radians per second).

Sinusoidal steady state

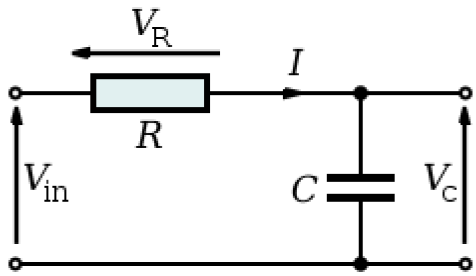
Sinusoidal steady state is a special case in which the input voltage consists of a pure sinusoid (with no exponential decay). As a result,

$$\sigma \ = \ 0$$

and the evaluation of s becomes

$$s \ = \ j\omega$$

Series circuit



Series RC circuit

By viewing the circuit as a voltage divider, the voltage across the capacitor is:

$$V_C(s) = \frac{1/Cs}{R + 1/Cs} V_{in}(s) = \frac{1}{1 + RCs} V_{in}(s)$$

and the voltage across the resistor is:

$$V_R(s) = \frac{R}{R + 1/Cs} V_{in}(s) = \frac{RCs}{1 + RCs} V_{in}(s).$$

Transfer functions

The transfer function for the capacitor is

$$H_C(s) = \frac{V_C(s)}{V_{in}(s)} = \frac{1}{1 + RCs}.$$

Similarly, the transfer function for the resistor is

$$H_R(s) = \frac{V_R(s)}{V_{in}(s)} = \frac{RCs}{1 + RCs}.$$

Poles and zeros

Both transfer functions have a single pole located at

$$s = -\frac{1}{RC}.$$

In addition, the transfer function for the resistor has a zero located at the origin.

Gain and phase angle

The magnitude of the gains across the two components are:

$$G_C = |H_C(j\omega)| = \left|\frac{V_C(j\omega)}{V_{in}(j\omega)}\right| = \frac{1}{\sqrt{1+(\omega RC)^2}}$$

and

$$G_R = |H_R(j\omega)| = \left|\frac{V_R(j\omega)}{V_{in}(j\omega)}\right| = \frac{\omega RC}{\sqrt{1+(\omega RC)^2}},$$

and the phase angles are:

$$\phi_C = \angle H_C(j\omega) = \tan^{-1}(-\omega RC)$$

and

$$\phi_R = \angle H_R(j\omega) = \tan^{-1}\left(\frac{1}{\omega RC}\right).$$

These expressions together may be substituted into the usual expression for the phasor representing the output:

$$\begin{aligned} V_C &= G_C V_{in} e^{j\phi_C} \\ V_R &= G_R V_{in} e^{j\phi_R}. \end{aligned}$$

Current

The current in the circuit is the same everywhere since the circuit is in series:

$$I(s) = \frac{V_{in}(s)}{R+\frac{1}{Cs}} = \frac{Cs}{1+RCs} V_{in}(s)$$

Impulse response

The impulse response for each voltage is the inverse Laplace transform of the corresponding transfer function. It represents the response of the circuit to an input voltage consisting of an impulse or Dirac delta function.

The impulse response for the capacitor voltage is

$$h_C(t) = \frac{1}{RC}e^{-t/RC}u(t) = \frac{1}{\tau}e^{-t/\tau}u(t)$$

where $u(t)$ is the Heaviside step function and

$$\tau = RC$$

is the time constant.

Similarly, the impulse response for the resistor voltage is

$$h_R(t) = \delta(t) - \frac{1}{RC}e^{-t/RC}u(t) = \delta(t) - \frac{1}{\tau}e^{-t/\tau}u(t)$$

where $\delta(t)$ is the Dirac delta function

Frequency-domain considerations

These are frequency domain expressions. Analysis of them will show which frequencies the circuits (or filters) pass and reject. This analysis rests on a consideration of what happens to these gains as the frequency becomes very large and very small.

As $\omega \to \infty$:

$$G_C \to 0$$
$$G_R \to 1.$$

As $\omega \to 0$:

$$G_C \to 1$$
$$G_R \to 0.$$

This shows that, if the output is taken across the capacitor, high frequencies are attenuated (rejected) and low frequencies are passed. Thus, the circuit behaves as a *low-pass filter*. If, though, the output is taken across the resistor, high frequencies are passed and low frequencies are rejected. In this configuration, the circuit behaves as a *high-pass filter*.

The range of frequencies that the filter passes is called its bandwidth. The point at which the filter attenuates the signal to half its unfiltered power is termed its cutoff frequency. This requires that the gain of the circuit be reduced to

$$G_C = G_R = \frac{1}{\sqrt{2}}.$$

Solving the above equation yields

$$\omega_c = \frac{1}{RC} \text{ rad/s}$$

or

$$f_c = \frac{1}{2\pi RC} \text{ Hz}$$

which is the frequency that the filter will attenuate to half its original power.

Clearly, the phases also depend on frequency, although this effect is less interesting generally than the gain variations.

As $\omega \to 0$:

$$\begin{aligned} \phi_C &\to 0 \\ \phi_R &\to 90^\circ = \pi/2^c \end{aligned}.$$

As $\omega \to \infty$:

$$\begin{aligned} \phi_C &\to -90^\circ = -\pi/2^c \\ \phi_R &\to 0 \end{aligned}$$

So at DC (0 Hz), the capacitor voltage is in phase with the signal voltage while the resistor voltage leads it by 90°. As frequency increases, the capacitor voltage comes to have a 90° lag relative to the signal and the resistor voltage comes to be in-phase with the signal.

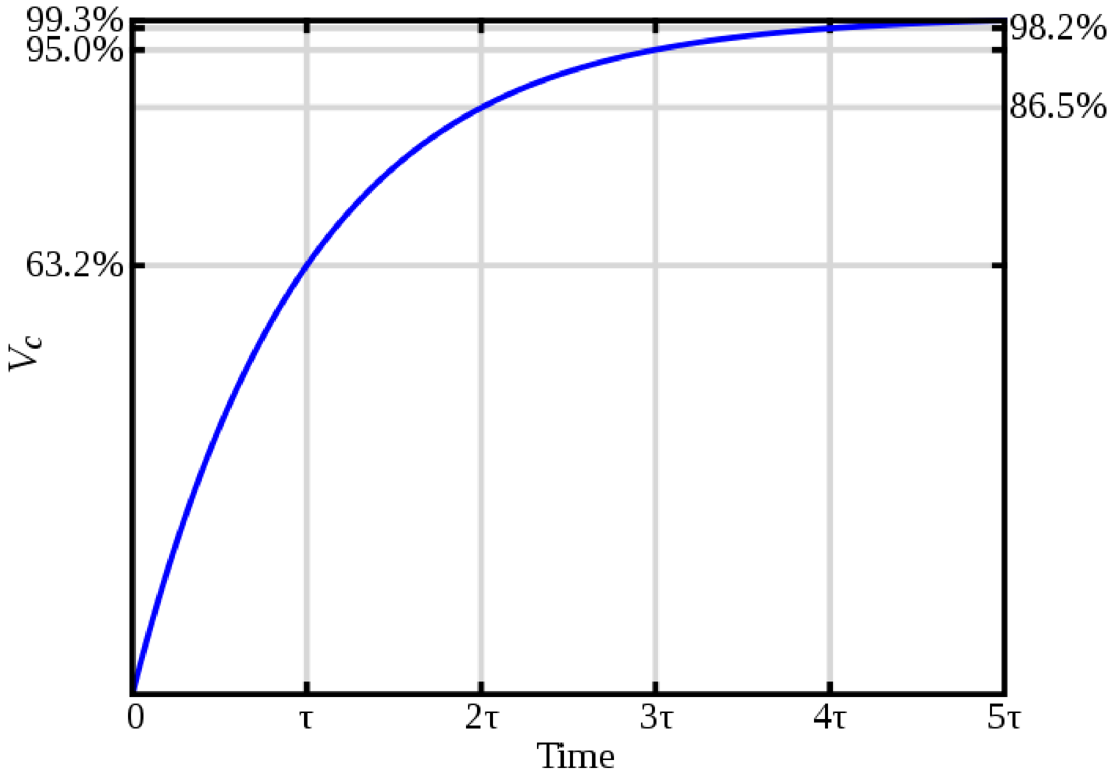
Time-domain considerations

The most straightforward way to derive the time domain behaviour is to use the Laplace transforms of the expressions for V_C and V_R given above. This effectively transforms $j\omega \to s$. Assuming a step input (i.e. $V_{in} = 0$ before $t = 0$ and then $V_{in} = V$ afterwards):

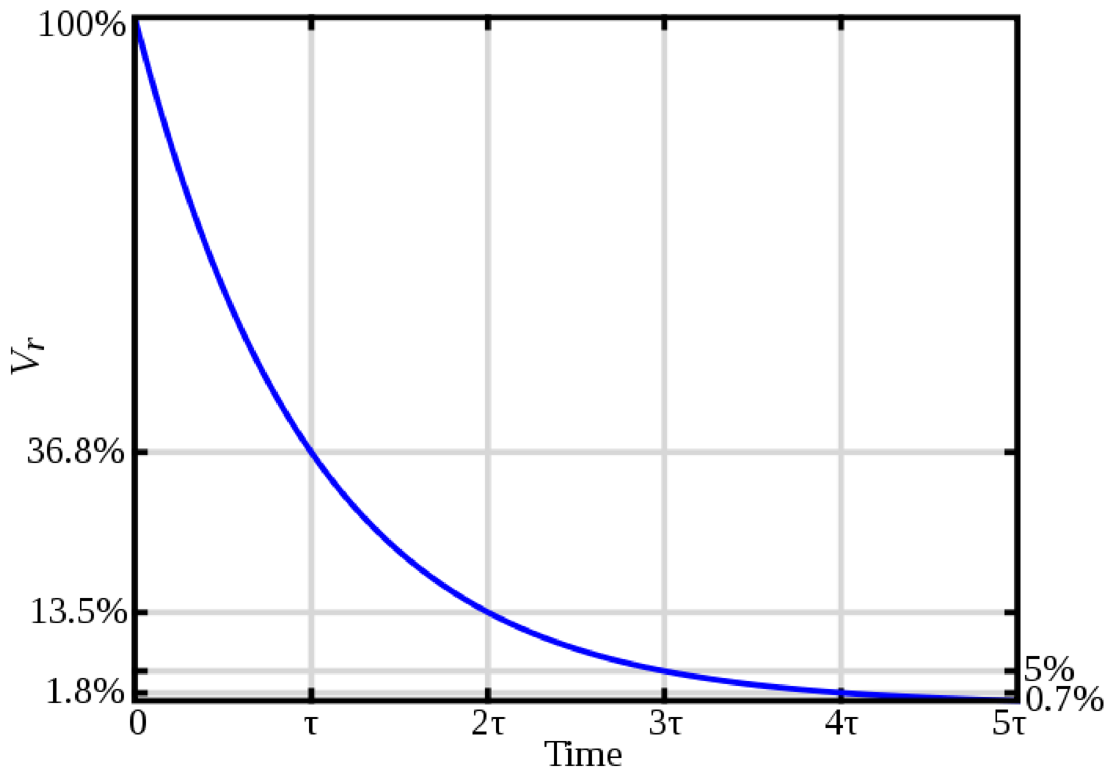
$$\begin{aligned} V_{in}(s) &= V\frac{1}{s} \\ V_C(s) &= V\frac{1}{1+sRC}\frac{1}{s} \end{aligned}$$

and

$$V_R(s) = V\frac{sRC}{1+sRC}\frac{1}{s}.$$



Capacitor voltage step-response.



Resistor voltage step-response.

Partial fractions expansions and the inverse Laplace transform yield:

$$V_C(t) = V\left(1 - e^{-t/RC}\right)$$
$$V_R(t) = Ve^{-t/RC}.$$

These equations are for calculating the voltage across the capacitor and resistor respectively while the capacitor is charging; for discharging, the equations are vice-versa. These equations can be rewritten in terms of charge and current using the relationships C=Q/V and V=IR.

Thus, the voltage across the capacitor tends towards V as time passes, while the voltage across the resistor tends towards 0, as shown in the figures. This is in keeping with the intuitive point that the capacitor will be charging from the supply voltage as time passes, and will eventually be fully charged and form an open circuit.

These equations show that a series RC circuit has a time constant, usually denoted $\tau = RC$ being the time it takes the voltage across the component to either rise (across C) or fall (across R) to within $1 / e$ of its final value. That is, τ is the time it takes V_C to reach $V(1 - 1 / e)$ and V_R to reach $V(1 / e)$.

The rate of change is a *fractional* $\left(1 - \frac{1}{e}\right)$ per τ. Thus, in going from $t = N\tau$ to $t = (N + 1)\tau$, the voltage will have moved about 63.2 % of the way from its level at $t = N\tau$ toward its final value. So C will be charged to about 63.2 % after τ, and essentially fully charged (99.3 %) after about 5τ. When the voltage source is replaced with a short-circuit, with C fully charged, the voltage across C drops exponentially with *t* from *V* towards 0. C will be discharged to about 36.8 % after τ, and essentially fully discharged (0.7 %) after about 5τ. Note that the current, *I*, in the circuit behaves as the voltage across R does, via Ohm's Law.

These results may also be derived by solving the differential equations describing the circuit:

$$\frac{V_{in} - V_C}{R} = C\frac{dV_C}{dt}$$

and

$$V_R = V_{in} - V_C.$$

The first equation is solved by using an integrating factor and the second follows easily; the solutions are exactly the same as those obtained via Laplace transforms.

Integrator

Consider the output across the capacitor at *high* frequency i.e.

$$\omega \gg \frac{1}{RC}.$$

This means that the capacitor has insufficient time to charge up and so its voltage is very small. Thus the input voltage approximately equals the voltage across the resistor. To see this, consider the expression for *I* given above:

$$I = \frac{V_{in}}{R + 1/j\omega C}$$

but note that the frequency condition described means that

$$\omega C \gg \frac{1}{R}$$

so

$$I \approx \frac{V_{in}}{R}$$ which is just Ohm's Law.

Now,

$$V_C = \frac{1}{C}\int_0^t I dt$$

so

$$V_C \approx \frac{1}{RC}\int_0^t V_{in} dt,$$

which is an integrator *across the capacitor*.

Differentiator

Consider the output across the resistor at *low* frequency i.e.,

$$\omega \ll \frac{1}{RC}.$$

This means that the capacitor has time to charge up until its voltage is almost equal to the source's voltage. Considering the expression for *I* again, when

$$R \ll \frac{1}{\omega C},$$

so

$$I \approx \frac{V_{in}}{1/j\omega C}$$
$$V_{in} \approx \frac{I}{j\omega C} \approx V_C$$

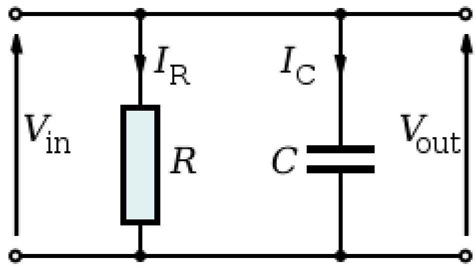
Now,

$$V_R = IR = C\frac{dV_C}{dt}R$$
$$V_R \approx RC\frac{dV_{in}}{dt}$$

which is a differentiator *across the resistor*.

More accurate integration and differentiation can be achieved by placing resistors and capacitors as appropriate on the input and feedback loop of operational amplifiers.

Parallel circuit



Parallel RC circuit

The parallel RC circuit is generally of less interest than the series circuit. This is largely because the output voltage V_{out} is equal to the input voltage V_{in} — as a result, this circuit does not act as a filter on the input signal unless fed by a current source.

With complex impedances:

$$I_R = \frac{V_{in}}{R}$$

and

$$I_C = j\omega C V_{in}.$$

This shows that the capacitor current is 90° out of phase with the resistor (and source) current. Alternatively, the governing differential equations may be used:

$$I_R = \frac{V_{in}}{R}$$

and

$$I_C = C\frac{dV_{in}}{dt}.$$

When fed by a current source, the transfer function of a parallel RC circuit is

$$\frac{V_{out}}{I_{in}} = \frac{R}{1 + sRC}.$$

Chapter 8

RL Circuit

A **resistor-inductor circuit (RL circuit)**, or **RL filter** or **RL network**, is one of the simplest analogue infinite impulse response electronic filters. It consists of a resistor and an inductor, either in series or in parallel, driven by a voltage source.

Introduction

The fundamental passive linear circuit elements are the resistor (R), capacitor (C) and inductor (L). These circuit elements can be combined to form an electrical circuit in four distinct ways: the RC circuit, the RL circuit, the LC circuit and the RLC circuit with the abbreviations indicating which components are used. These circuits exhibit important types of behaviour that are fundamental to analogue electronics. In particular, they are able to act as passive filters.

In practice, however, capacitors (and RC circuits) are usually preferred to inductors since they can be more easily manufactured and are generally physically smaller, particularly for higher values of components.

Complex Impedance

The complex impedance Z_L (in ohms) of an inductor with inductance L (in henries) is

$$Z_L = Ls$$

The complex frequency s is a complex number,

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

where

- j represents the imaginary unit:

  $j^2 = -1$

- σ is the exponential decay constant (in radians per second), and

- ω is the angular frequency (in radians per second).

Eigenfunctions

The complex-valued **eigenfunctions** of ANY linear time-invariant (LTI) system are of the following forms:

$V(t) = \mathbf{A}e^{st} = \mathbf{A}e^{(\sigma+j\omega)t}$, or letting $\mathbf{A} = Ae^{j\phi}$ and rewriting; $= Ae^{j\phi}e^{(\sigma+j\omega)t}$, and collecting terms is $= Ae^{\sigma t}e^{j(\omega t+\phi)}$

From Euler's formula, the **real-part** of these eigenfunctions are exponentially-decaying sinusoids:

$$v(t) = \mathrm{Re}\{V(t)\} = Ae^{\sigma t}\cos(\omega t + \phi)$$

Sinusoidal Steady State

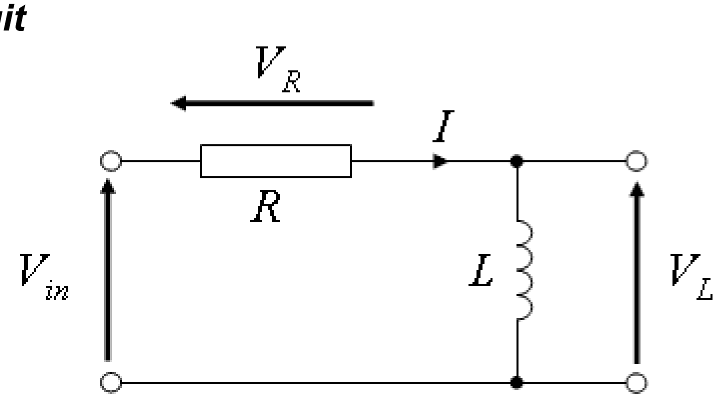
Sinusoidal steady state is a special case in which the input voltage consists of a pure sinusoid (with no exponential decay). As a result,

$$\sigma = 0$$

and the evaluation of s becomes

$$s = j\omega$$

Series circuit



Series RL circuit

By viewing the circuit as a voltage divider, we see that the voltage across the inductor is:

$$V_L(s) = \frac{Ls}{R + Ls} V_{in}(s)$$

and the voltage across the resistor is:

$$V_R(s) = \frac{R}{R + Ls} V_{in}(s)$$.

Current

The current in the circuit is the same everywhere since the circuit is series:

$$I(s) = \frac{V_{in}(s)}{R + Ls}.$$

Transfer functions

The transfer function for the inductor is

$$H_L(s) = \frac{V_L(s)}{V_{in}(s)} = \frac{Ls}{R + Ls} = G_L e^{j\phi_L}$$

Similarly, the transfer function for the resistor is

$$H_R(s) = \frac{V_R(s)}{V_{in}(s)} = \frac{R}{R + Ls} = G_R e^{j\phi_R}$$

Poles and zeros

Both transfer functions have a single pole located at

$$s = -\frac{R}{L}$$

In addition, the transfer function for the inductor has a zero located at the origin.

Gain and phase angle

The gains across the two components are found by taking the magnitudes of the above expressions:

$$G_L = |H_L(s)| = \left|\frac{V_L(s)}{V_{in}(s)}\right| = \frac{\omega L}{\sqrt{R^2 + (\omega L)^2}}$$

and

$$G_R = |H_R(s)| = \left|\frac{V_R(s)}{V_{in}(s)}\right| = \frac{R}{\sqrt{R^2 + (\omega L)^2}},$$

and the phase angles are:

$$\phi_L = \angle H_L(s) = \tan^{-1}\left(\frac{R}{\omega L}\right)$$

and

$$\phi_R = \angle H_R(s) = \tan^{-1}\left(-\frac{\omega L}{R}\right).$$

Phasor notation

These expressions together may be substituted into the usual expression for the phasor representing the output:

$$V_L = G_L V_{in} e^{j\phi_L}$$
$$V_R = G_R V_{in} e^{j\phi_R}.$$

Impulse Response

The impulse response for each voltage is the inverse Laplace transform of the corresponding transfer function. It represents the response of the circuit to an input voltage consisting of an impulse or Dirac delta function.

The impulse response for the inductor voltage is

$$h_L(t) = \delta(t) - \frac{R}{L}e^{-tR/L}u(t) = \delta(t) - \frac{1}{\tau}e^{-t/\tau}u(t)$$

where $u(t)$ is the Heaviside step function and

$$\tau = \frac{L}{R}$$

is the time constant.

Similarly, the impulse response for the resistor voltage is

$$h_R(t) = \frac{R}{L} e^{-tR/L} u(t) = \frac{1}{\tau} e^{-t/\tau} u(t)$$

Zero input response (ZIR)

The **Zero input response**, also called the **natural response**, of an RL circuit describes the behavior of the circuit after it has reached constant voltages and currents and is disconnected from any power source. It is called the zero-input response because it requires no input.

The ZIR of an RL circuit is:

$$i(t) = i(0) e^{-(R/L)t} = i(0) e^{-t/\tau}.$$

Frequency domain considerations

These are frequency domain expressions. Analysis of them will show which frequencies the circuits (or filters) pass and reject. This analysis rests on a consideration of what happens to these gains as the frequency becomes very large and very small.

As $\omega \to \infty$:

$$G_L \to 1$$
$$G_R \to 0.$$

As $\omega \to 0$:

$$G_L \to 0$$
$$G_R \to 1.$$

This shows that, if the output is taken across the inductor, high frequencies are passed and low frequencies are attenuated (rejected). Thus, the circuit behaves as a *high-pass filter*. If, though, the output is taken across the resistor, high frequencies are rejected and low frequencies are passed. In this configuration, the circuit behaves as a *low-pass filter*. Compare this with the behaviour of the resistor output in an RC circuit, where the reverse is the case.

The range of frequencies that the filter passes is called its bandwidth. The point at which the filter attenuates the signal to half its unfiltered power is termed its cutoff frequency. This requires that the gain of the circuit be reduced to

$$G_L = G_R = \frac{1}{\sqrt{2}}.$$

Solving the above equation yields

$$\omega_c = \frac{R}{L} \text{ rad/s}$$

or

$$f_c = \frac{R}{2\pi L} \text{ Hz}$$

which is the frequency that the filter will attenuate to half its original power.

Clearly, the phases also depend on frequency, although this effect is less interesting generally than the gain variations.

As $\omega \to 0$:

$$\phi_L \to 90^\circ = \pi/2^c$$
$$\phi_R \to 0.$$

As $\omega \to \infty$:

$$\phi_L \to 0$$
$$\phi_R \to -90^\circ = -\pi/2^c$$

So at DC (0 Hz), the resistor voltage is in phase with the signal voltage while the inductor voltage leads it by 90°. As frequency increases, the resistor voltage comes to have a 90° lag relative to the signal and the inductor voltage comes to be in-phase with the signal.

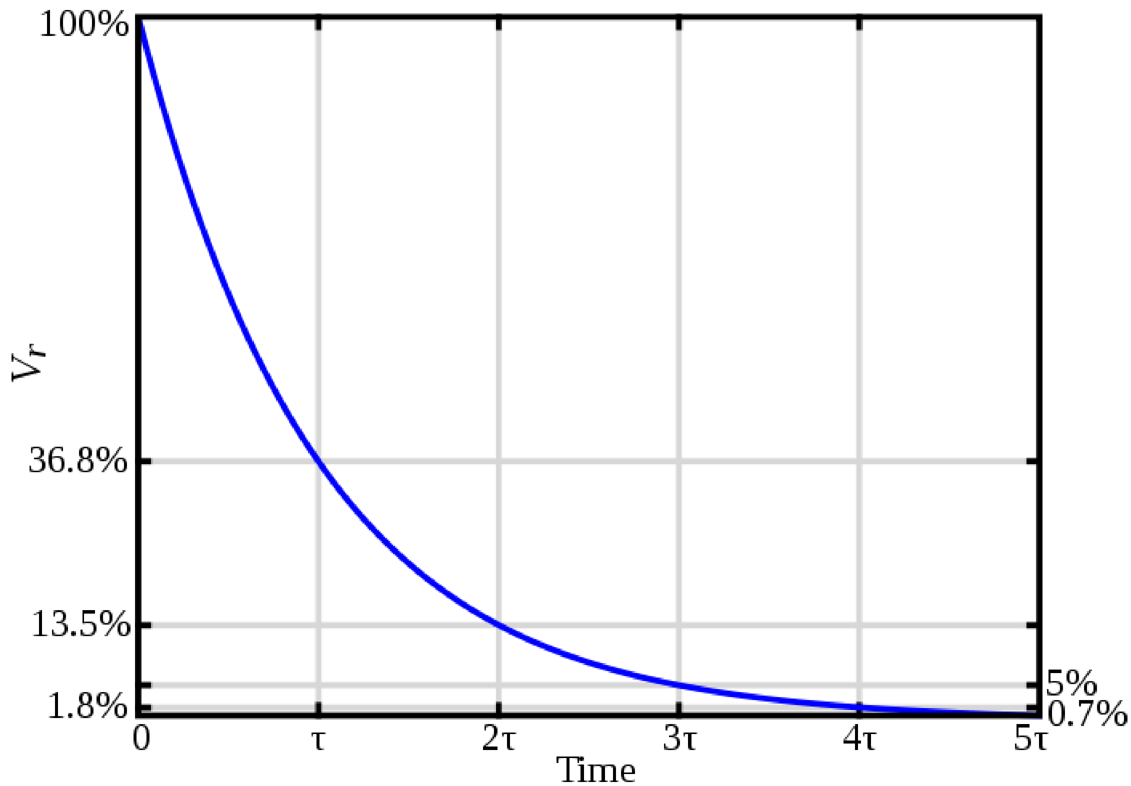
Time domain considerations

The most straightforward way to derive the time domain behaviour is to use the Laplace transforms of the expressions for V_L and V_R given above. This effectively transforms $j\omega \to s$. Assuming a step input (i.e. $V_{in} = 0$ before $t = 0$ and then $V_{in} = V$ afterwards):

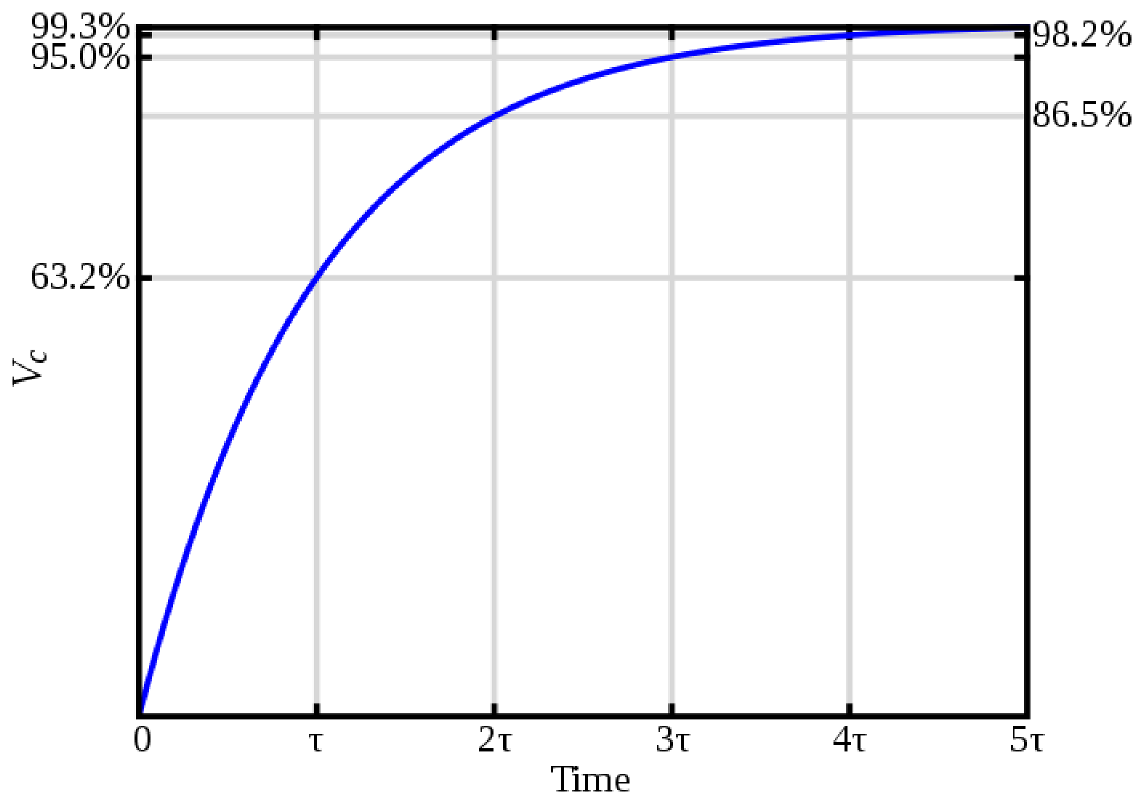
$$V_{in}(s) = V\frac{1}{s}$$
$$V_L(s) = V\frac{sL}{R + sL}\frac{1}{s}$$

and

$$V_R(s) = V\frac{R}{R+sL}\frac{1}{s}.$$



Inductor voltage step-response.



Resistor voltage step-response.

Partial fractions expansions and the inverse Laplace transform yield:

$$V_L(t) = Ve^{-tR/L}$$
$$V_R(t) = V\left(1 - e^{-tR/L}\right).$$

Thus, the voltage across the inductor tends towards 0 as time passes, while the voltage across the resistor tends towards V, as shown in the figures. This is in keeping with the intuitive point that the inductor will only have a voltage across as long as the current in the circuit is changing — as the circuit reaches its steady-state, there is no further current change and ultimately no inductor voltage.

These equations show that a series RL circuit has a time constant, usually denoted $\tau = L / R$ being the time it takes the voltage across the component to either fall (across L) or rise (across R) to within $1 / e$ of its final value. That is, τ is the time it takes V_L to reach $V(1 / e)$ and V_R to reach $V(1 - 1 / e)$.

The rate of change is a *fractional* $\left(1 - \frac{1}{e}\right)$ per τ. Thus, in going from $t = N\tau$ to $t = (N + 1)\tau$, the votage will have moved about 63% of the way from its level at $t = N\tau$ toward its final value. So the voltage across L will have dropped to about 37% after τ, and essentially to zero (0.7%) after about 5τ. Kirchhoff's voltage law implies that the voltage across the resistor will *rise* at the same rate. When the voltage source is then replaced with a short-circuit, the voltage across R drops exponentially with *t* from *V* towards 0. R will be discharged to about 37% after τ, and essentially fully discharged (0.7%) after about 5τ. Note that the current, *I*, in the circuit behaves as the voltage across R does, via Ohm's Law.

The delay in the rise/fall time of the circuit is in this case caused by the back-EMF from the inductor which, as the current flowing through it tries to change, prevents the current (and hence the voltage across the resistor) from rising or falling much faster than the time-constant of the circuit. Since all wires have some self-inductance and resistance, all circuits have a time constant. As a result, when the power supply is switched on, the current does not instantaneously reach its steady-state value, *V* / *R*. The rise instead takes several time-constants to complete. If this were not the case, and the current were to reach steady-state immediately, extremely strong inductive electric fields would be generated by the sharp change in the magnetic field — this would lead to breakdown of the air in the circuit and electric arcing, probably damaging components (and users).

These results may also be derived by solving the differential equation describing the circuit:

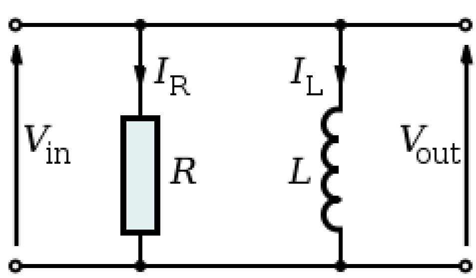
$$V_{in} = IR + L\frac{dI}{dt},$$

and

$$V_R = V_{in} - V_L.$$

The first equation is solved by using an integrating factor and yields the current which must be differentiated to give V_L; the second equation is straightforward. The solutions are exactly the same as those obtained via Laplace transforms.

Parallel circuit



Parallel RL circuit

The parallel RL circuit is generally of less interest than the series circuit unless fed by a current source. This is largely because the output voltage V_{out} is equal to the input voltage V_{in} — as a result, this circuit does not act as a filter for a voltage input signal.

With complex impedances:

$$I_R = \frac{V_{in}}{R}$$

and

$$I_L = \frac{V_{in}}{j\omega L} = -\frac{jV_{in}}{\omega L}.$$

This shows that the inductor lags the resistor (and source) current by 90°.

The parallel circuit is seen on the output of many amplifier circuits, and is used to isolate the amplifier from capacitive loading effects at high frequencies. Because of the phase shift introduced by capacitance, some amplifiers become unstable at very high frequencies, and tend to oscillate. This affects sound quality and component life (especially the transistors), and is to be avoided.

Chapter 9

Sallen–Key Topology

The **Sallen–Key topology** is an electronic filter topology used to implement second-order active filters that is particularly valued for its simplicity. It is a degenerate form of a **voltage-controlled voltage-source (VCVS) filter topology**. A VCVS filter uses a super-unity-gain voltage amplifier with practically infinite input impedance and zero output impedance to implement a 2-pole (12 dB/octave) low-pass, high-pass, or bandpass response. The super-unity-gain amplifier allows for very high Q factor and passband gain without the use of inductors. A Sallen–Key filter is a variation on a VCVS filter that uses a unity-gain amplifier (i.e., a pure buffer amplifier with 0 dB gain). It was introduced by R.P. Sallen and E. L. Key of MIT Lincoln Laboratory in 1955.

Because of its high input impedance and easily selectable gain, an operational amplifier in a conventional non-inverting configuration is often used in VCVS implementations. Implementations of Sallen–Key filters often use an operational amplifier configured as a voltage follower; however, emitter or source followers are other common choices for the buffer amplifier.

VCVS filters are relatively resilient to component tolerance, but obtaining high Q factor may require extreme component value spread or high amplifier gain. Higher-order filters can be obtained by cascading two or more stages.

Generic Sallen–Key topology

The generic unity-gain Sallen–Key filter topology implemented with a unity-gain operational amplifier is shown in Figure 1. The following analysis is based on the assumption that the operational amplifier is ideal.

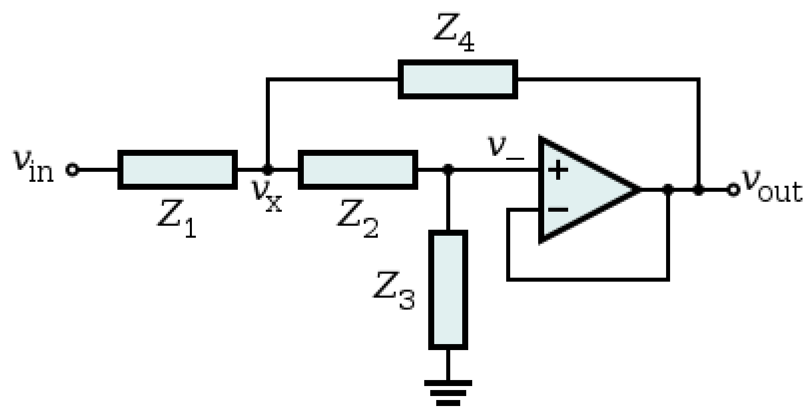


Figure 1: The generic Sallen–Key filter topology.

Because the operational amplifier (OA) is in a negative-feedback configuration, its v_+ and v_- inputs must match (i.e., $v_+ = v_-$). However, the inverting input v_- is connected directly to the output v_{out}, and so

$$v_+ = v_- = v_{out}. \tag{1}$$

By Kirchhoff's current law (KCL) applied at the v_x node,

$$\frac{v_{in} - v_x}{Z_1} = \frac{v_x - v_{out}}{Z_4} + \frac{v_x - v_-}{Z_2}. \tag{2}$$

By combining Equations (1) and (2),

$$\frac{v_{in} - v_x}{Z_1} = \frac{v_x - v_{out}}{Z_4} + \frac{v_x - v_{out}}{Z_2}$$

Applying KCL again at the OA's non-inverting input v_+ (= v_- = v_{out}) gives

$$\frac{v_x - v_{out}}{Z_2} = \frac{v_{out}}{Z_3},$$

which means that

$$v_x = v_{out}\left(\frac{Z_2}{Z_3} + 1\right).$$

$$(3)$$

Combining Equations (2) and (3) gives

$$\frac{v_{\text{in}} - v_{\text{out}}\left(\frac{Z_2}{Z_3} + 1\right)}{Z_1} = \frac{v_{\text{out}}\left(\frac{Z_2}{Z_3} + 1\right) - v_{\text{out}}}{Z_4} + \frac{v_{\text{out}}\left(\frac{Z_2}{Z_3} + 1\right) - v_{\text{out}}}{Z_2}. \tag{4}$$

Rearranging Equation (4) gives the transfer function

$$\frac{v_{\text{out}}}{v_{\text{in}}} = \frac{Z_3 Z_4}{Z_1 Z_2 + Z_4(Z_1 + Z_2) + Z_3 Z_4}, \tag{5}$$

which typically describes a second-order LTI system.

Interpretation

If the Z_4component were connected to ground, the filter would be a voltage divider composed of the $Z_1$and Z_4components cascaded with another voltage divider composed of the $Z_2$and Z_3components. The buffer bootstraps the "bottom" of the Z_4component to the output of the filter, which will improve upon the simple two divider case. This interpretation is the reason why Sallen–Key filters are often drawn with the operational amplifier's non-inverting input below the inverting input, thus emphasizing the similarity between the output and ground.

Example applications

By choosing different passive components (e.g., resistors and capacitors) for Z_1, Z_2, Z_3, and Z_4, the filter can be made with low-pass, bandpass, and high-pass characteristics. In the examples below, recall that a resistor with resistance Rhas impedance $Z_R$of

$$Z_R = R\,,$$

and a capacitor with capacitance Chas impedance $Z_C$of

$$Z_C = \frac{1}{sC},$$

where $s = j\omega = \left(\sqrt{-1}\right) 2\pi f$ and f is a frequency of a pure sine wave input. That is, a capacitor's impedance is frequency dependent and a resistor's impedance is not.

Example: Low-pass filter

An example of a unity-gain low-pass configuration is shown in Figure 2.

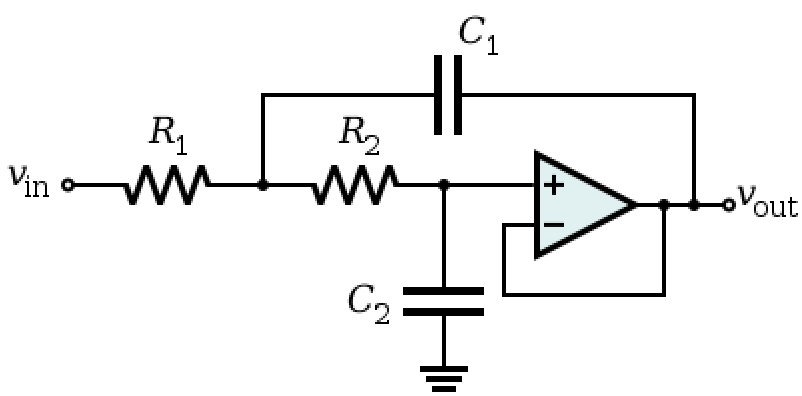


Figure 2: A unity-gain low-pass filter implemented with a Sallen–Key topology.

An operational amplifier is used as the buffer here, although an emitter follower is also effective. This circuit is equivalent to the generic case above with

$$Z_1 = R_1, \quad Z_2 = R_2, \quad Z_3 = \frac{1}{sC_2}, \quad \text{and} \quad Z_4 = \frac{1}{sC_1}.$$

The transfer function for this second-order unity-gain low-pass filter is

$$H(s) = \frac{\overbrace{(2\pi f_c)^2}^{\omega_c^2}}{s^2 + \underbrace{2\pi \frac{f_c}{Q}}_{\frac{\omega_c}{Q} = 2\zeta\omega_c} s + \underbrace{(2\pi f_c)^2}_{\omega_c^2}}$$

where the cutoff frequency f_c and Q factor Q (i.e., damping ratio ζ) are given by

$$f_c = \frac{1}{2\pi\sqrt{R_1 R_2 C_1 C_2}}$$

and

$$2\zeta = \frac{1}{Q} = \frac{\sqrt{R_1 R_2 C_1 C_2}}{C_1}\left(\frac{1}{R_1} + \frac{1}{R_2}\right).$$

So,

$$Q = \frac{\sqrt{R_1 R_2 C_1 C_2}}{C_2\left(R_1 + R_2\right)}$$

The Q factor determines the height and width of the peak of the frequency response of the filter. As this parameter increases, the filter will tend to "ring" at a single resonant frequency near f_c.

A designer must choose the Q and f_c appropriate for his application. For example, a second-order Butterworth filter, which has maximally flat passband frequency response, has a Q of $1/\sqrt{2}$. Because there are two parameters and four unknowns, the design procedure typically fixes one resistor as a ratio of the other resistor and one capacitor as a ratio of the other capacitor. One possibility is to set the ratio between C_1 and C_2 as n and the ratio between R_1 and R_2 as m. So,

$$\begin{aligned} R_1 &= mR, \\ R_2 &= R, \\ C_1 &= nC, \\ C_2 &= C. \end{aligned}$$

Therefore, the f_c and Q expressions are

$$f_c = \frac{1}{2\pi\ RC\sqrt{mn}},$$

and

$$Q = \frac{\sqrt{mn}}{m+1}.$$

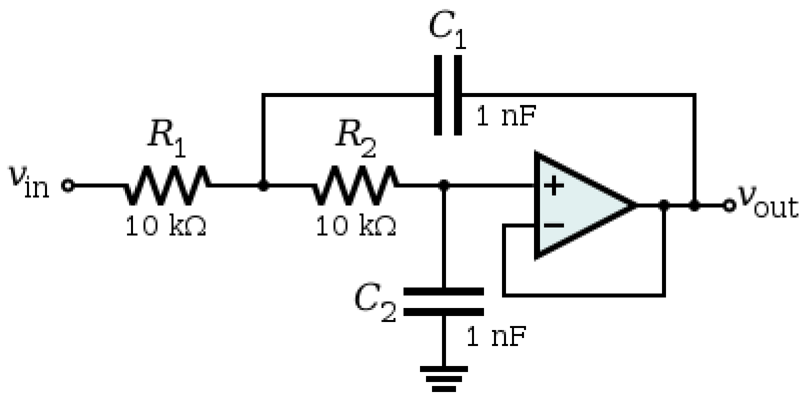


Figure 3: A low-pass filter, which is implemented with a Sallen–Key topology, with f_c=15.9 kHz and Q = 0.5.

For example, the circuit in Figure 3 has an f_c of $15.9\ \mathrm{kHz}$ and a Q of 0.5. The transfer function is given by

$$H(s) = \frac{1}{1 + \underbrace{C_2(R_1+R_2)}_{\frac{2\zeta}{\omega_c} = \frac{1}{\omega_c Q}} s + \underbrace{C_1 C_2 R_1 R_2}_{\frac{1}{\omega_c^2}} s^2},$$

and, after substitution, this expression is equal to

$$H(s) = \frac{1}{1 + \underbrace{RC(m+1)}_{\frac{2\zeta}{\omega_c} = \frac{1}{\omega_c Q}} s + \underbrace{mnR^2C^2}_{\frac{1}{\omega_c^2}} s^2}$$

which shows how every (R, C) combination comes with some (m, n) combination to provide the same f_c and Q for the low-pass filter. A similar design approach is used for the other filters below.

Example: High-pass filter

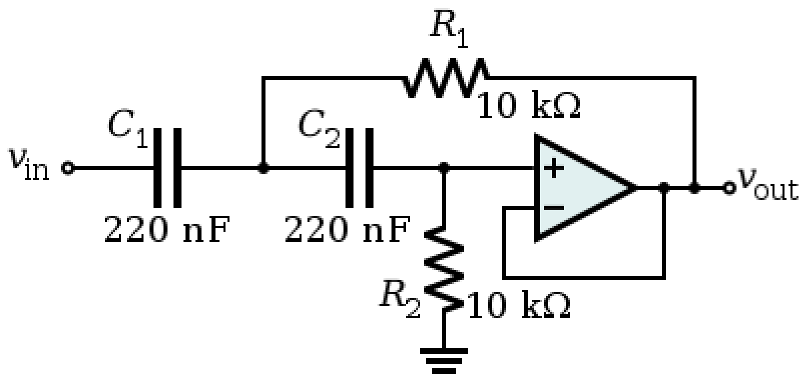


Figure 4: A specific Sallen–Key high-pass filter with f_c=72 Hz and Q = 0.5.

A second-order unity-gain high-pass filter with f_c of $72\ \mathrm{Hz}$ and Q of 0.5 is shown in Figure 4.

A second-order unity-gain high-pass filter has the transfer function

$$H(s) = \frac{s^2}{s^2 + \underbrace{2\pi(\frac{f_c}{Q})}_{2\zeta\omega_c = \frac{\omega_c}{Q}} s + \underbrace{(2\pi f_c)^2}_{\omega_c^2}},$$

where cutoff frequency f_c and Q factor are discussed above in the low-pass filter discussion. The circuit above implements this transfer function by the equations

$$f_c = \frac{1}{2\pi\sqrt{R_1 R_2 C_1 C_2}}$$

(as before), and

$$\frac{1}{2\zeta} = Q = \frac{\sqrt{R_1 R_2 C_1 C_2}}{R_1(C_1 + C_2)}.$$

So

$$2\zeta f_c = \frac{f_c}{Q} = \frac{C_1 + C_2}{2\pi R_2 C_1 C_2}.$$

Follow an approach similar to the one used to design the low-pass filter above.

VCVS Example: Bandpass configuration

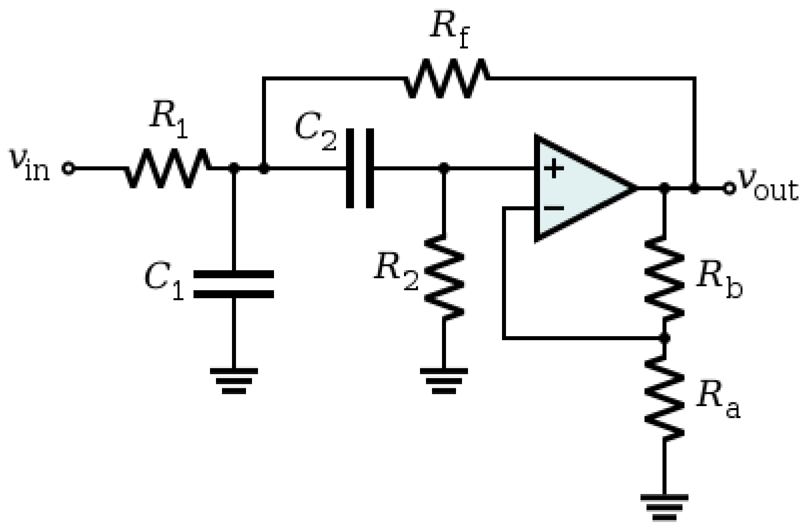


Figure 5: A bandpass filter realized with a VCVS topology.

An example of a non-unity-gain bandpass filter implemented with a VCVS filter is shown in Figure 5. Although it uses a different topology and an operational amplifier configured to provide non-unity-gain, it can be analyzed using similar methods as with the generic Sallen–Key topology. Its transfer function is given by:

$$H(s) = \frac{\overbrace{\left(1+\frac{R_b}{R_a}\right)}^{G} \frac{s}{R_1 C_1}}{s^2 + \underbrace{\left(\frac{1}{R_1 C_1} + \frac{1}{R_2 C_1} + \frac{1}{R_2 C_2} - \frac{R_b}{R_a R_f C_1}\right)}_{2\zeta\omega_0 = \frac{\omega_0}{Q}} s + \underbrace{\frac{R_1 + R_f}{R_1 R_f R_2 C_1 C_2}}_{\omega_0^2 = (2\pi f_0)^2}}$$

The center frequency f_0 (i.e., the frequency where the magnitude response has its *peak*) is given by:

$$f_0 = \frac{1}{2\pi}\sqrt{\frac{R_f + R_1}{C_1 C_2 R_1 R_2 R_f}}$$

The voltage divider in the negative feedback loop controls the gain. The "inner gain" G provided by the operational amplifier is given by

$$G = 1 + \frac{R_b}{R_a}$$

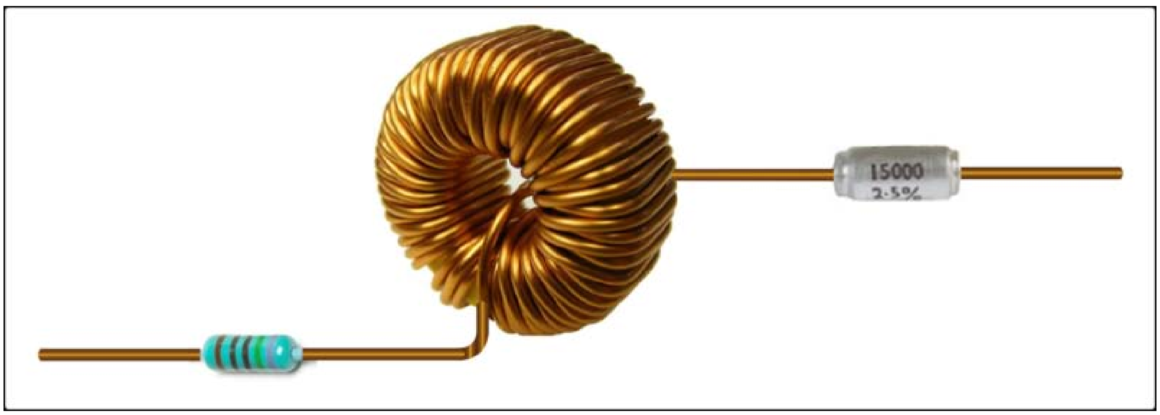
while the amplifier gain at the peak frequency is given by:

$$A = \frac{G}{3 - G}$$

It can be seen that G must be kept below 3 or else the filter will oscillate. The filter is usually optimized by selecting $R_2 = 2R_1$ and $C_1 = C_2$.

Chapter 10

RLC Circuit



A series RLC circuit: a resistor, inductor, and a capacitor

An **RLC circuit** (or **LCR circuit**) is an electrical circuit consisting of a resistor, an inductor, and a capacitor, connected in series or in parallel. The RLC part of the name is due to those letters being the usual electrical symbols for resistance, inductance and capacitance respectively. The circuit forms a harmonic oscillator for current and will resonate in just the same way as an LC circuit will. The difference that the presence of the resistor makes is that any oscillation induced in the circuit will die away over time if it not kept going by a source. This effect of the resistor is called damping. Some resistance is unavoidable in real circuits, even if a resistor is not specifically included as a component. A pure LC circuit is an ideal which really only exists in theory.

There are many applications for this circuit. They are used in many different types of oscillator circuit. Another important application is for tuning, such as in radio receivers or television sets, where they are used to select a narrow range of frequencies from the ambient radio waves. In this role the circuit is often referred to as a tuned circuit. An RLC circuit can be used as a band-pass filter or a band-stop filter. The tuning application, for instance, is an example of band-pass filtering. The RLC filter is described as a *second-order* circuit, meaning that any voltage or current in the circuit can be described by a second-order differential equation in circuit analysis.

The three circuit elements can be combined in a number of different topologies. All three elements in series or all three elements in parallel are the simplest in concept and the most straightforward to analyse. There are, however, other arrangements, some with practical importance in real circuits. One issue often encountered is the need to take into account inductor resistance. Inductors are typically constructed from coils of wire, the resistance of which is not usually desirable, but it often has a significant effect on the circuit.

Basic concepts

Resonance

An important property of this circuit is its ability to resonate at a specific frequency, the resonance frequency, f_0. Frequencies are measured in units of hertz. However, angular frequency, ω_0, is used which is more mathematically convenient. This is measured in radians per second. They are related to each other by a simple proportion,

$$\omega_0 = 2\pi f_0$$

Resonance occurs because energy is stored in two different ways: in an electric field as the capacitor is charged and in a magnetic field as current flows through the inductor. Energy can be transferred from one to the other within the circuit and this can be oscillatory. A mechanical analogy is a weight suspended on a spring which will oscillate up and down when released. This is no passing metaphor, a weight on a spring is described by exactly the same second order differential equation as an RLC circuit and for all the properties of the one system there will be found an analogous property of the other. The mechanical property answering to the resistor in the circuit is friction in the spring/weight system. Friction will slowly bring any oscillation to a halt if there is no external force driving it. Likewise, the resistance in an RLC circuit will "damp" the oscillation, diminishing it with time if there is no driving AC power source in the circuit.

The resonance frequency is defined as the frequency at which the impedance of the circuit is at a minimum. Equivalently, it can be defined as the frequency at which the impedance is purely real (that is, purely resistive). This occurs because the impedance (reactance) of the inductor and capacitor at resonance are equal but of opposite sign and cancel out. Circuits where L and C are in parallel rather than series actually have a maximum impedance rather than a minimum impedance. For this reason they are often described as antiresonators, it is still usual, however, to name the frequency at which this occurs as the resonance frequency.

Natural frequency

The resonance frequency is defined in terms of the impedance presented to a driving source. It is still possible for the circuit to carry on oscillating (for a time) after the driving source has been removed or it is subjected to a step in voltage (including a step down to zero). This is similar to the way that a tuning fork will carry on ringing after it has been struck, and the effect is often called ringing. This effect is the undriven natural resonance frequency of the circuit and in general is not exactly the same as the driven resonance frequency, although the two will usually be quite close to each other. Various terms are used by different authors to distinguish the two, but resonance frequency unqualified usually means the driven resonance frequency. The driven frequency may be called the undamped resonance frequency or undamped natural frequency and the undriven frequency may be called the damped resonance frequency or the damped natural frequency. The reason for this terminology is that the driven resonance frequency in a series or parallel resonant circuit has the value

$$\omega_0 = \frac{1}{\sqrt{LC}}$$

This is exactly the same as the resonance frequency of an LC circuit, that is, one with no resistor present, that is, it is the same as a circuit in which there is no damping, hence undamped resonance frequency. The undriven resonance frequency, on the other hand, depends on the value of the resistor and hence is described as the damped resonance frequency. A highly damped circuit will fail to resonate at all when undriven. A circuit with a value of resistor that causes it to be just on the edge of ringing is called critically damped. Either side of critically damped are described as underdamped (ringing happens) and overdamped (ringing is suppressed).

Circuits with topologies more complex than straightforward series or parallel have a driven resonance frequency that deviates from $\omega_0 = \frac{1}{\sqrt{LC}}$and for those the undamped resonance frequency, damped resonance frequency and driven resonance frequency can all be different.

Damping

Damping is caused by the resistance in the circuit. It determines whether or not the circuit will resonate naturally (that is, without a driving source). Circuits which will resonate in this way are described as underdamped and those that will not are overdamped. Damping attenuation (symbol α) is measured in nepers per second. However, the unitless damping factor (symbol ζ) is often a more useful measure, which is related to α by

$$\zeta = \frac{\alpha}{\omega_0}$$

The special case of $\zeta = 1$ is called critical damping and represents the case of a circuit that is just on the border of oscillation. It is the minimum damping that can be applied without causing oscillation.

Bandwidth

The resonance effect can be used for filtering, the rapid change in impedance near resonance can be used to pass or block signals close to the resonance frequency. Both band-pass and band-stop filters can be constructed and some filter circuits are shown later here. A key parameter in filter design is bandwidth. The bandwidth is measured between the 3dB-points, that is, the frequencies at which the power passed through the circuit has fallen to half the value passed at resonance. There are two of these half-power frequencies, one above, and one below the resonance frequency

$$\Delta\omega = \omega_2 - \omega_1$$

where $\Delta\omega$is the bandwidth, $\omega_1$is the lower half-power frequency and $\omega_2$is the upper half-power frequency. The bandwidth is related to attenuation by,

$$\Delta\omega = 2\alpha$$

when the units are radians per second and nepers per second respectively. Other units may require a conversion factor. A more general measure of bandwidth is the fractional bandwidth, which expresses the bandwidth as a fraction of the resonance frequency and is given by

$$F_\mathrm{b} = \frac{\Delta\omega}{\omega_0}$$

The fractional bandwidth is also often stated as a percentage. The damping of filter circuits is adjusted to result in the required bandwidth. A narrow band filter, such as a notch filter, requires low damping. A wide band filter requires high damping.

Q factor

The Q factor is a widespread measure used to characterise resonators. It is defined as the peak energy stored in the circuit divided by the average energy dissipated in it per cycle at resonance. Low *Q* circuits are therefore damped and lossy and high *Q* circuits are underdamped. *Q* is related to bandwidth; low *Q* circuits are wide band and high *Q* circuits are narrow band. In fact, it happens that *Q* is the inverse of fractional bandwidth

$$Q = \frac{1}{F_\mathrm{b}} = \frac{\omega_0}{\Delta\omega}$$

Q factor is directly proportional to selectivity, as Q factor depends inversely on bandwidth.

Scaled parameters

The parameters ζ, F_b, and Q are all scaled to ω_0. This means that circuits which have similar parameters share similar characteristics regardless of whether or not they are operating in the same frequency band.

Next gives the analysis for the series RLC circuit in detail. Other configurations are not described in such detail, but the key differences from the series case are given. The general form of the differential equations given in the series circuit section are applicable to all second order circuits and can be used to describe the voltage or current in any element of each circuit.

Series RLC circuit

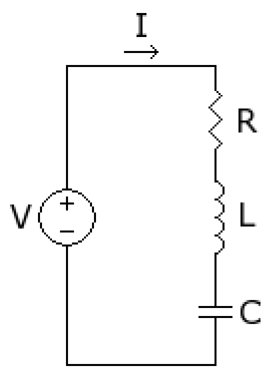


Figure 1. RLC series circuit

V - the voltage of the power source
I - the current in the circuit
R - the resistance of the resistor
L - the inductance of the inductor
C - the capacitance of the capacitor

In this circuit, the three components are all in series with the voltage source. The governing differential equation can be found by substituting into Kirchhoff's voltage law (KVL) the constitutive equation for each of the three elements. From KVL,

$$v_R + v_L + v_C = v(t)$$

where v_R, v_L, v_C are the voltages across R, L and C respectively and $v(t)$ is the time varying voltage from the source. Substituting in the constitutive equations,

$$Ri(t) + L\frac{di}{dt} + \frac{1}{C}\int_{-\infty}^{\tau=t} i(\tau)\,d\tau = v(t)$$

For the case where the source is an unchanging voltage, differentiating and dividing by L leads to the second order differential equation:

$$\frac{d^2i(t)}{dt^2} + \frac{R}{L}\frac{di(t)}{dt} + \frac{1}{LC}i(t) = 0$$

This can usefully be expressed in a more generally applicable form:

$$\frac{d^2i(t)}{dt^2} + 2\alpha\frac{di}{dt} + {\omega_0}^2 i(t) = 0$$

α and ω_0 are both in units of angular frequency. α is called the *neper frequency*, or *attenuation*, and is a measure of how fast the transient response of the circuit will die away after the stimulus has been removed. Neper occurs in the name because the units can also be considered to be nepers per second, neper being a unit of attenuation. ω_0 is the angular resonance frequency and is discussed later.

For the case of the series RLC circuit these two parameters are given by:

$$\alpha = \frac{R}{2L} \text{ and } \omega_0 = \frac{1}{\sqrt{LC}}$$

A useful parameter is the *damping factor*, ζ which is defined as the ratio of these two,

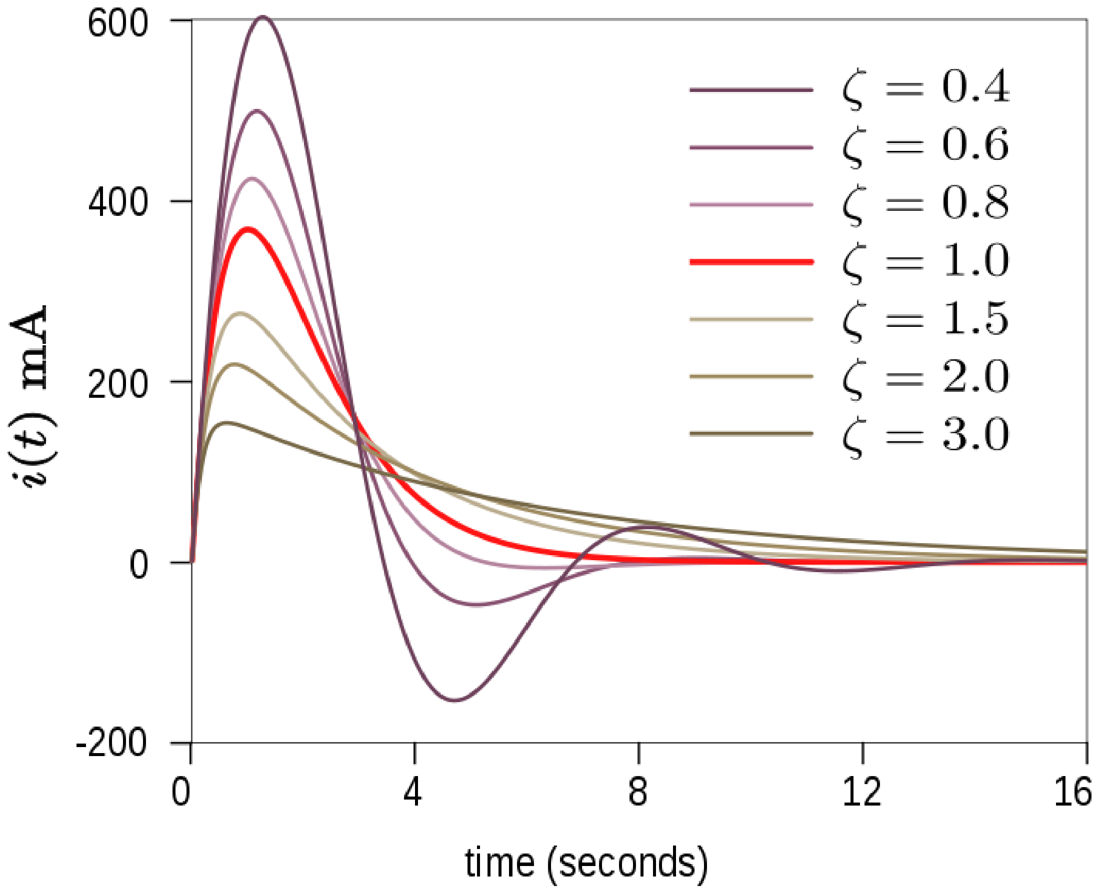
$$\zeta = \frac{\alpha}{\omega_0}$$

In the case of the series RLC circuit, the damping factor is given by,

$$\zeta = \frac{R}{2}\sqrt{\frac{C}{L}}$$

The value of the damping factor determines the type of transient that the circuit will exhibit. Some authors do not use ζ and call α the damping factor.

Transient response



Plot showing underdamped and overdamped responses of a series RLC circuit. The critical damping plot is the bold red curve. The plots are normalised for L=1, C=1 and $\omega_0=1$

The differential equation for the circuit solves in three different ways depending on the value of ζ. These are underdamped ($\zeta<1$), overdamped ($\zeta>1$) and critically damped ($\zeta=1$). The differential equation has the characteristic equation,

$$s^2 + 2\alpha s + {\omega_0}^2 = 0$$

The roots of the equation in *s* are,

$$s_1 = -\alpha + \sqrt{\alpha^2 - {\omega_0}^2}$$
$$s_2 = -\alpha - \sqrt{\alpha^2 - {\omega_0}^2}$$

The general solution of the differential equation is an exponential in either root or a linear superposition of both,

$$i(t) = A_1 e^{s_1 t} + A_2 e^{s_2 t}$$

The coefficients A_1 and A_2 are determined by the boundary conditions of the specific problem being analysed. That is, they are set by the values of the currents and voltages in the circuit at the onset of the transient and the presumed value they will settle to after infinite time.

The overdamped response ($\zeta > 1$) is,

$$i(t) = A_1 e^{-\omega_0(\zeta + \sqrt{\zeta^2 - 1})t} + A_2 e^{-\omega_0(\zeta - \sqrt{\zeta^2 - 1})t}$$

The overdamped response is a decay of the transient current without oscillation.

The underdamped response ($\zeta < 1$) is,

$$i(t) = B_1 e^{-\alpha t} \cos(\omega_d t) + B_2 e^{-\alpha t} \sin(\omega_d t)$$

By applying standard trigonometric identities the two trigonometric functions may be expressed as a single sinusoid with phase shift,

$$i(t) = B_3 e^{-\alpha t} \sin(\omega_d t + \varphi)$$

The underdamped response is a decaying oscillation at frequency ω_d. The oscillation decays at a rate determined by the attenuation α. The exponential in αdescribes the envelope of the oscillation. B_1 and B_2 (or B_3 and the phase shift φin the second form) are arbitrary constants determined by boundary conditions. The frequency $\omega_d$is given by,

$$\omega_d = \sqrt{{\omega_0}^2 - \alpha^2} = \omega_0 \sqrt{1 - \zeta^2}$$

This is called the damped resonance frequency or the damped natural frequency. It is the frequency the circuit will naturally oscillate at if not driven by an external source. The resonance frequency, ω_0, which is the frequency at which the circuit will resonate when driven by an external oscillation, may often be referred to as the undamped resonance frequency to distinguish it.

The critically damped response ($\zeta = 1$) is,

$$i(t) = D_1 t e^{-\alpha t} + D_2 e^{-\alpha t}$$

The critically damped response represents the circuit response that decays in the fastest possible time without going into oscillation. This consideration is important in control systems where it is required to reach the desired state as quickly as possible without overshooting. D_1 and D_2 are arbitrary constants determined by boundary conditions.

Laplace domain

The series RLC can be analyzed for both transient and steady AC state behavior using the Laplace transform. If the voltage source above produces a waveform with Laplace-transformed *V*(*s*) (where *s* is the complex frequency $s = \sigma + i\omega$), KVL can be applied in the Laplace domain:

$$V(s) = I(s)\left(R + Ls + \frac{1}{Cs}\right)$$

where *I*(*s*) is the Laplace-transformed current through all components. Solving for *I*(*s*):

$$I(s) = \frac{1}{R + Ls + \frac{1}{Cs}} V(s)$$

And rearranging, we have that

$$I(s) = \frac{s}{L\left(s^2 + \frac{R}{L}s + \frac{1}{LC}\right)} V(s)$$

Laplace admittance

Solving for the Laplace admittance *Y*(*s*):

$$Y(s) = \frac{I(s)}{V(s)} = \frac{s}{L\left(s^2 + \frac{R}{L}s + \frac{1}{LC}\right)}$$

Simplifying using parameters α and ω_o defined in the previous section, we have

$$Y(s) = \frac{I(s)}{V(s)} = \frac{s}{L\left(s^2 + 2\alpha s + {\omega_0}^2\right)}$$

Poles and zeros

The zeros of *Y*(*s*) are those values of *s* such that *Y*(*s*) = 0:

$s = 0$ and $|s| \rightarrow \infty$

The poles of *Y*(*s*) are those values of *s* such that $Y(s) \rightarrow \infty$. By the quadratic formula, we find

$$s = -\alpha \pm \sqrt{\alpha^2 - {\omega_0}^2}$$

The poles of $Y(s)$ are identical to the roots s_1 and s_2 of the characteristic polynomial of the differential equation in the section above.

General solution

For an arbitrary E(t), the solution obtained by inverse transform of I(s) is:

$$I(t) = \frac{1}{L}\int_0^t E(t-\tau)e^{-\alpha\tau}\left(\cos\omega_d\tau - \frac{\alpha}{\omega_d}\sin\omega_d\tau\right)d\tau$$ in the underdamped case ($\omega_0 > \alpha$)

$$I(t) = \frac{1}{L}\int_0^t E(t-\tau)e^{-\alpha\tau}(1-\alpha\tau)d\tau$$ in the critically damped case ($\omega_0 = \alpha$)

$$I(t) = \frac{1}{L}\int_0^t E(t-\tau)e^{-\alpha\tau}\left(\cosh\omega_r\tau - \frac{\alpha}{\omega_r}\sinh\omega_r\tau\right)d\tau$$ in the overdamped case ($\omega_0 < \alpha$)

where $\omega_r = \sqrt{\alpha^2 - {\omega_0}^2}$, and cosh and sinh are the usual hyperbolic functions.

Sinusoidal steady state

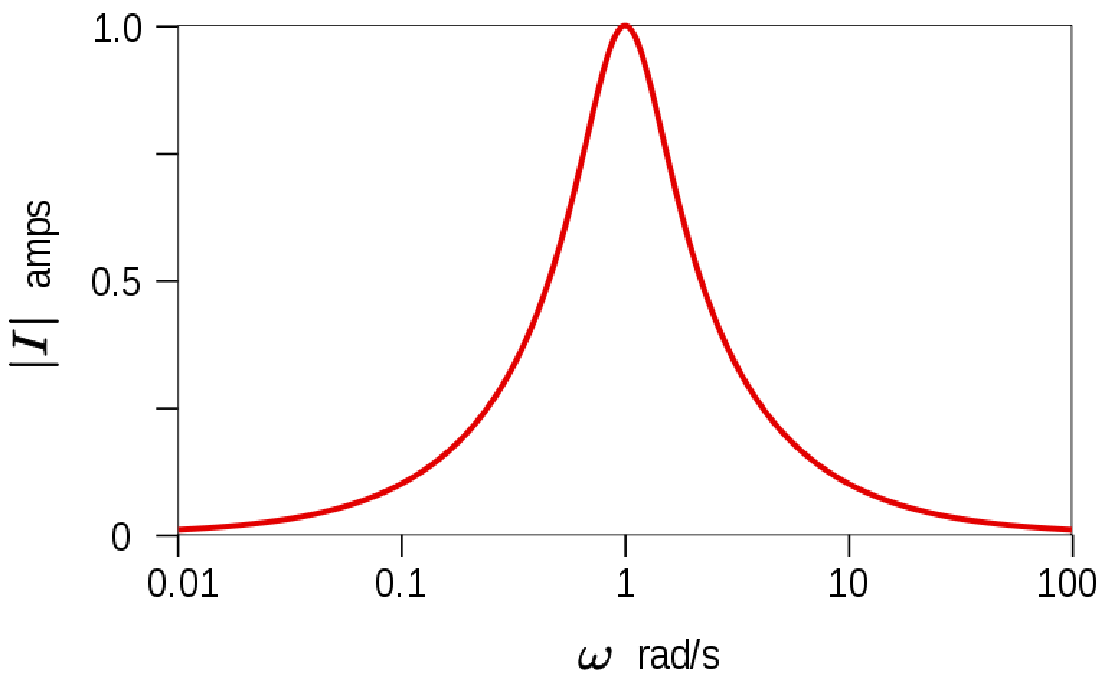


Figure 4. *Sinusoidal steady-state analysis*

normalised to R = 1 ohm, C = 1 farad, L = 1 henry, and V = 1.0 volt

Sinusoidal steady state is represented by letting $s = i\omega$

Taking the magnitude of the above equation with this substitution:

$$|Y(s = i\omega)| = \frac{1}{\sqrt{R^2 + \left(\omega L - \frac{1}{\omega C}\right)^2}}.$$

and the current as a function of ω can be found from

$$|I(i\omega)| = |Y(i\omega)||V(i\omega)|.$$

Note that there is a peak at $i_{mag}(\omega) = 1$. This is known as the resonance frequency. Solving for this value:

$$\omega_0 = \frac{1}{\sqrt{LC}}.$$

Parallel RLC circuit

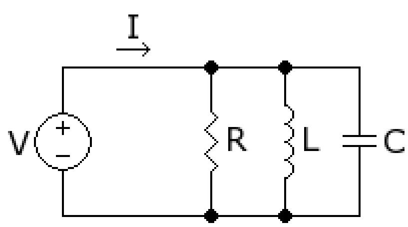


Figure 5. RLC parallel circuit

V - the voltage of the power source
I - the current in the circuit
R - the resistance of the resistor
L - the inductance of the inductor
C - the capacitance of the capacitor

The properties of the parallel RLC circuit can be obtained from the duality relationship of electrical circuits and considering that the parallel RLC is the dual impedance of a series RLC. From this consideration is immediately obtained the result that the differential equations describing this circuit will be identical to the general form of those describing a series RLC.

For the parallel circuit, the attenuation α is given by

$$\alpha = \frac{1}{2RC}$$

and the damping factor is consequently

$$\zeta = \frac{1}{2R}\sqrt{\frac{L}{C}}$$

This is the inverse of the expression for ζ in the series circuit. Likewise, the other scaled parameters, fractional bandwidth and Q are also the inverse of each other. This means that a wide band, low Q circuit in one topology will become a narrow band, high Q circuit in the other topology when constructed from components with identical values. The Q and fractional bandwidth of the parallel circuit are given by

$$Q = R\sqrt{\frac{C}{L}}$$ and $$F_{\mathrm{b}} = \frac{1}{R}\sqrt{\frac{L}{C}}$$

Frequency domain

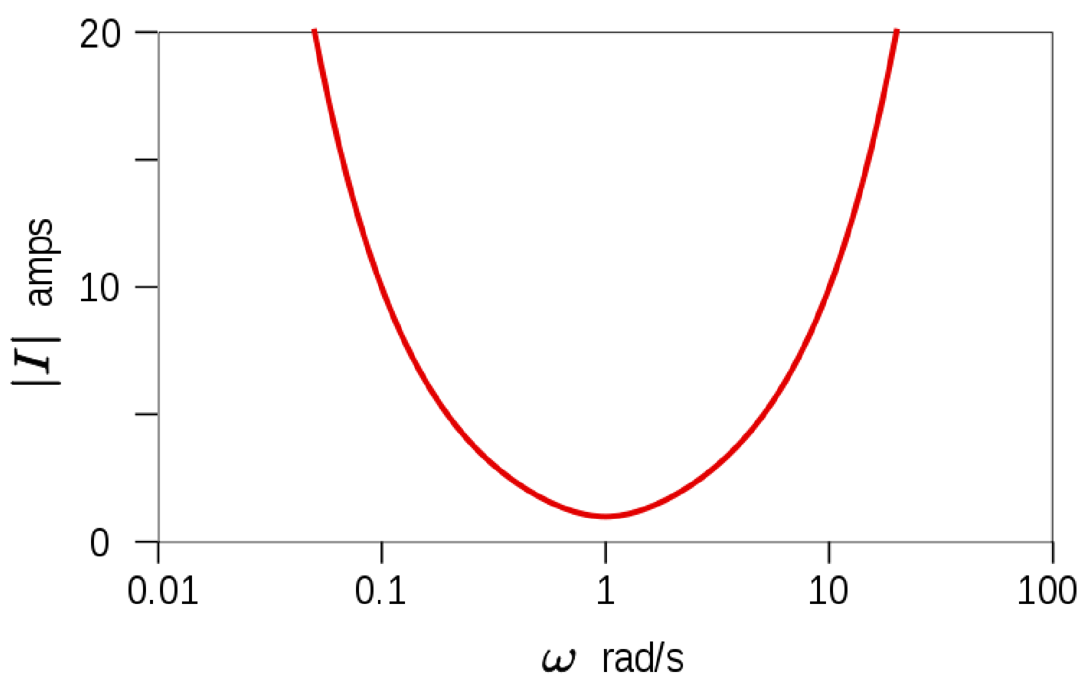


Figure 6. *Sinusoidal steady-state analysis*s

normalised to $R = 1$ ohm, $C = 1$ farad, $L = 1$ henry, and $V = 1.0$ volt

The complex admittance of this circuit is given by adding up the admittances of the components:

$$\frac{1}{Z} = \frac{1}{Z_L} + \frac{1}{Z_C} + \frac{1}{Z_R} = \frac{1}{j\omega L} + j\omega C + \frac{1}{R}$$

The change from a series arrangement to a parallel arrangement results in the circuit having a peak in impedance at resonance rather than a minimum, so the circuit is an antiresonator.

The graph opposite shows that there is a minimum in the frequency response of the current at the resonance frequency $\omega_0 = \frac{1}{\sqrt{LC}}$ when the circuit is driven by a constant voltage. On the other hand, if driven by a constant current, there would be a maximum in the voltage which would follow the same curve as the current in the series circuit.

Other configurations

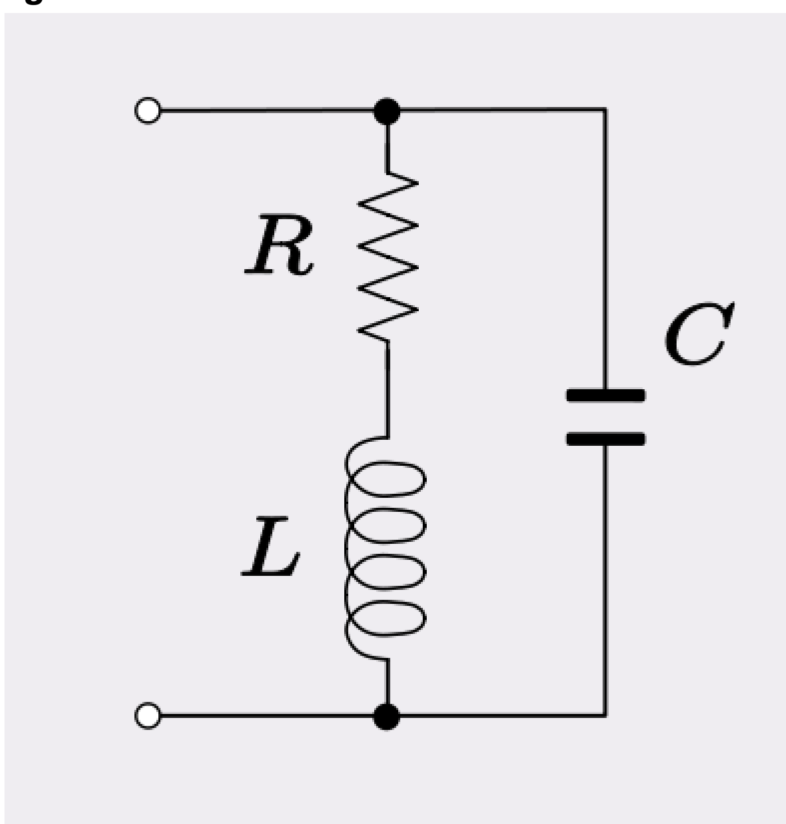


Fig. 7. RLC parallel circuit with resistance in series with the inductor

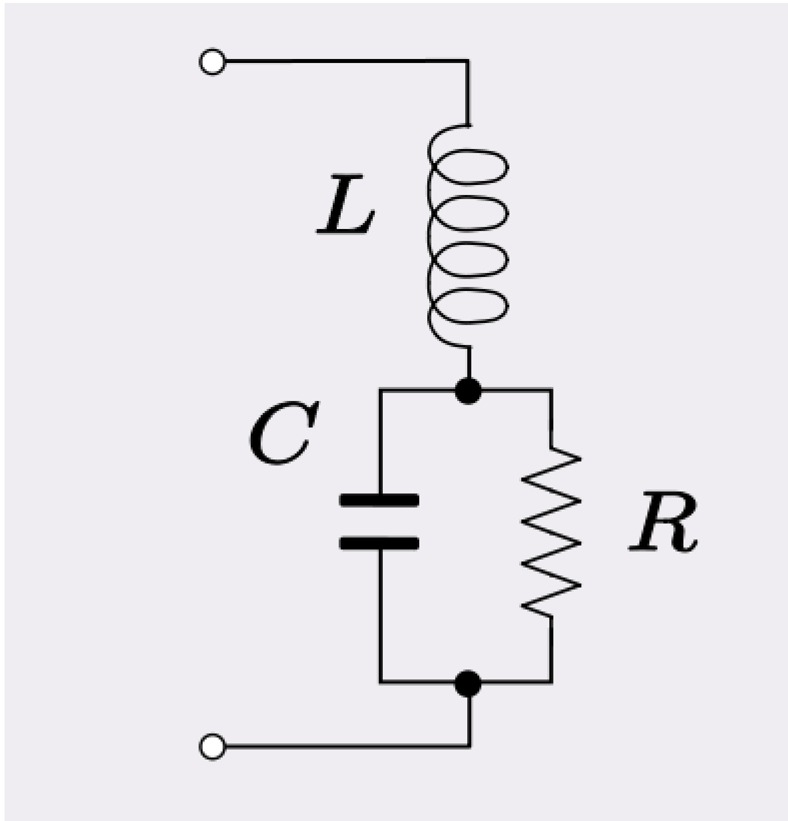


Fig. 8. RLC series circuit with resistance in parallel with the capacitor

A series resistor with the inductor in a parallel LC circuit as shown in figure 7 is a topology commonly encountered where there is a need to take into account the resistance of the coil winding. Parallel LC circuits are frequently used for bandpass filtering and the Q is largely governed by this resistance. The resonant frequency of this circuit is,

$$\omega_0 = \sqrt{\frac{1}{LC} - \left(\frac{R}{L}\right)^2}$$

This is the resonant frequency of the circuit defined as the frequency at which the admittance has zero imaginary part. The frequency that appears in the generalised form of the characteristic equation (which is the same for this circuit as previously)

$$s^2 + 2\alpha s + {\omega'_0}^2 = 0$$

is not the same frequency. In this case it is the undamped resonant frequency

$$\omega'_0 = \sqrt{\frac{1}{LC}}$$

In the same vein, a resistor in parallel with the capacitor in a series LC circuit can be used to represent a capacitor with a lossy dielectric. This configuration is shown in figure 8. The resonant frequency in this case is given by

$$\omega_0 = \sqrt{\frac{1}{LC} - \frac{1}{(RC)^2}}$$

Applications

Variable tuned circuits

A very frequent use of these circuits is in the tuning circuits of analogue radios. Adjustable tuning is commonly achieved with a parallel plate variable capacitor which allows the value of C to be changed and tune to stations on different frequencies. For the IF stage in the radio where the tuning is preset in the factory the more usual solution is an adjustable core in the inductor to adjust L. In this design the core (made of a high permeability material that has the effect of increasing inductance) is threaded so that it can be screwed further in, or screwed further out of the inductor winding as required.

Filters

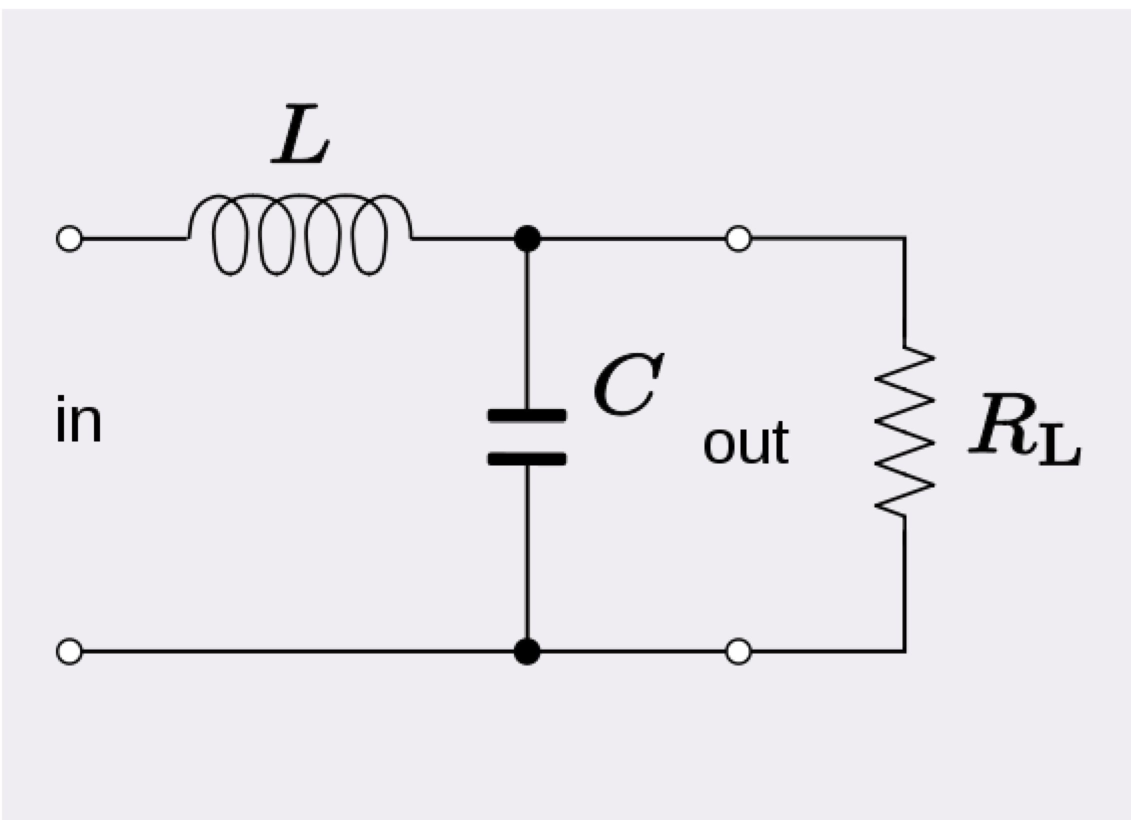


Fig. 9. RLC circuit as a low-pass filter

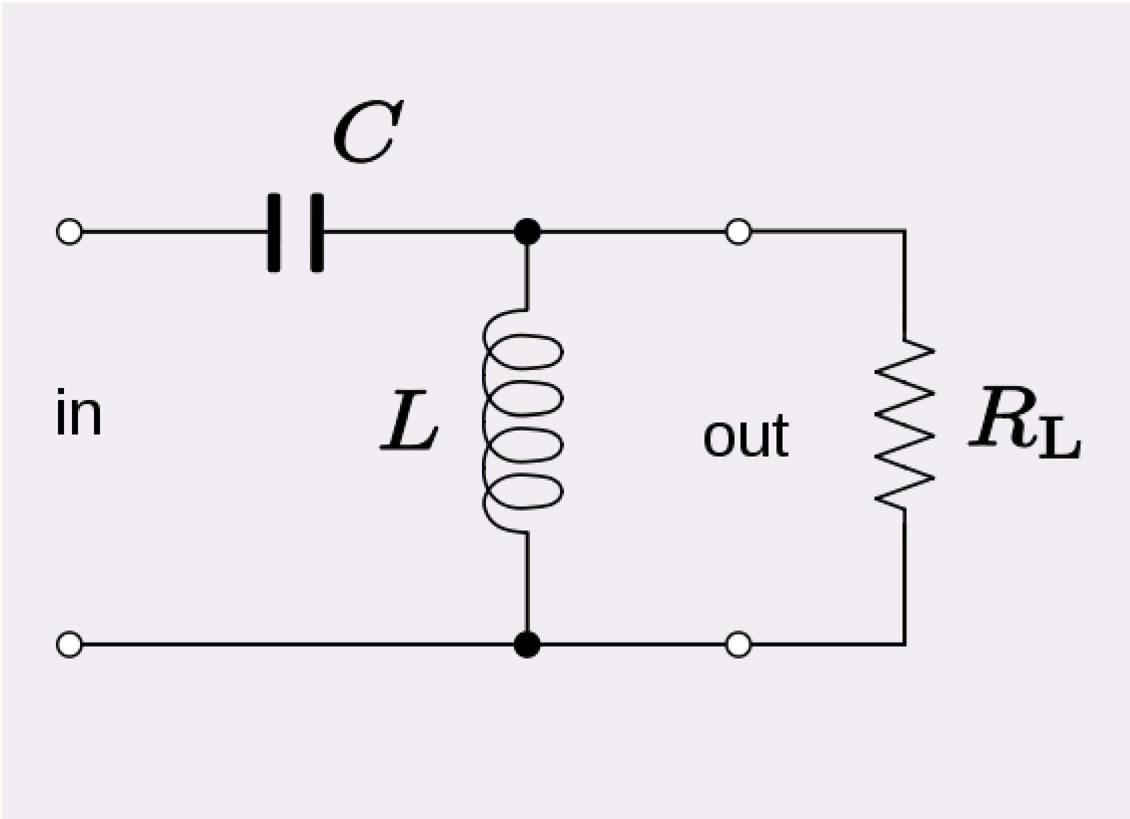


Fig. 10. RLC circuit as a high-pass filter

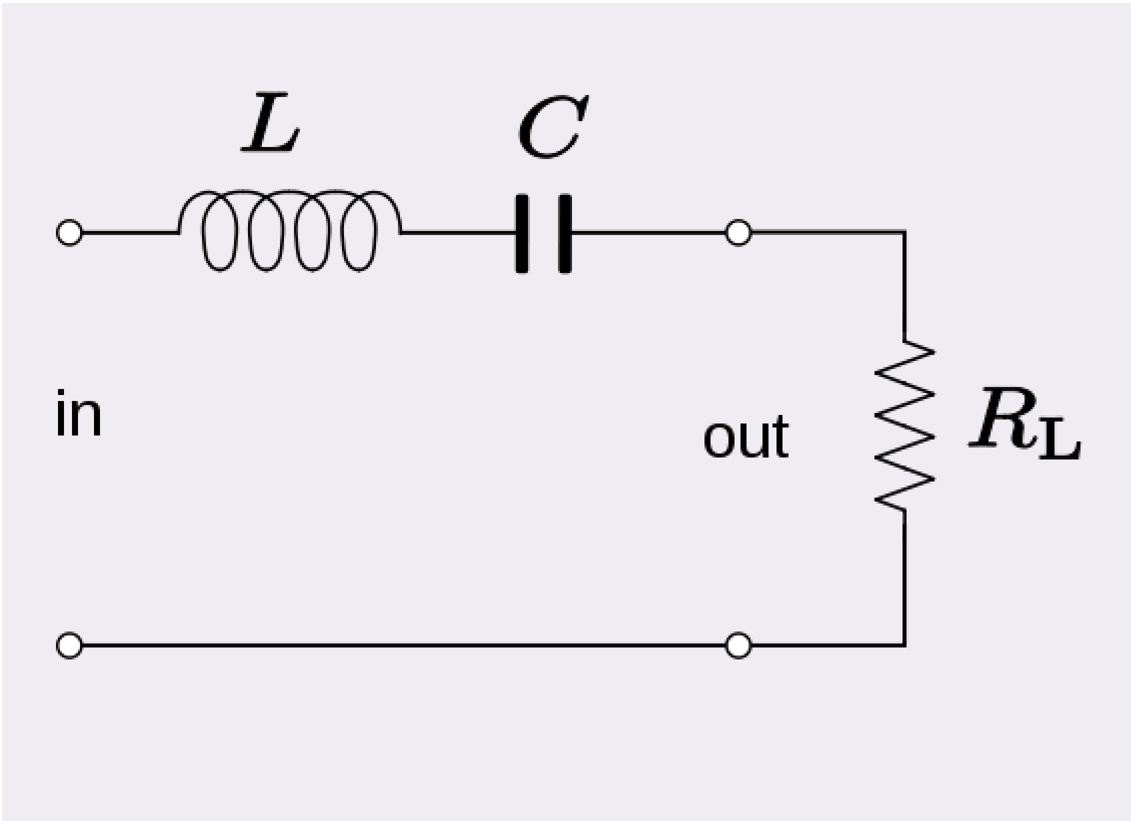


Fig. 11. RLC circuit as a series band-pass filter in series with the line

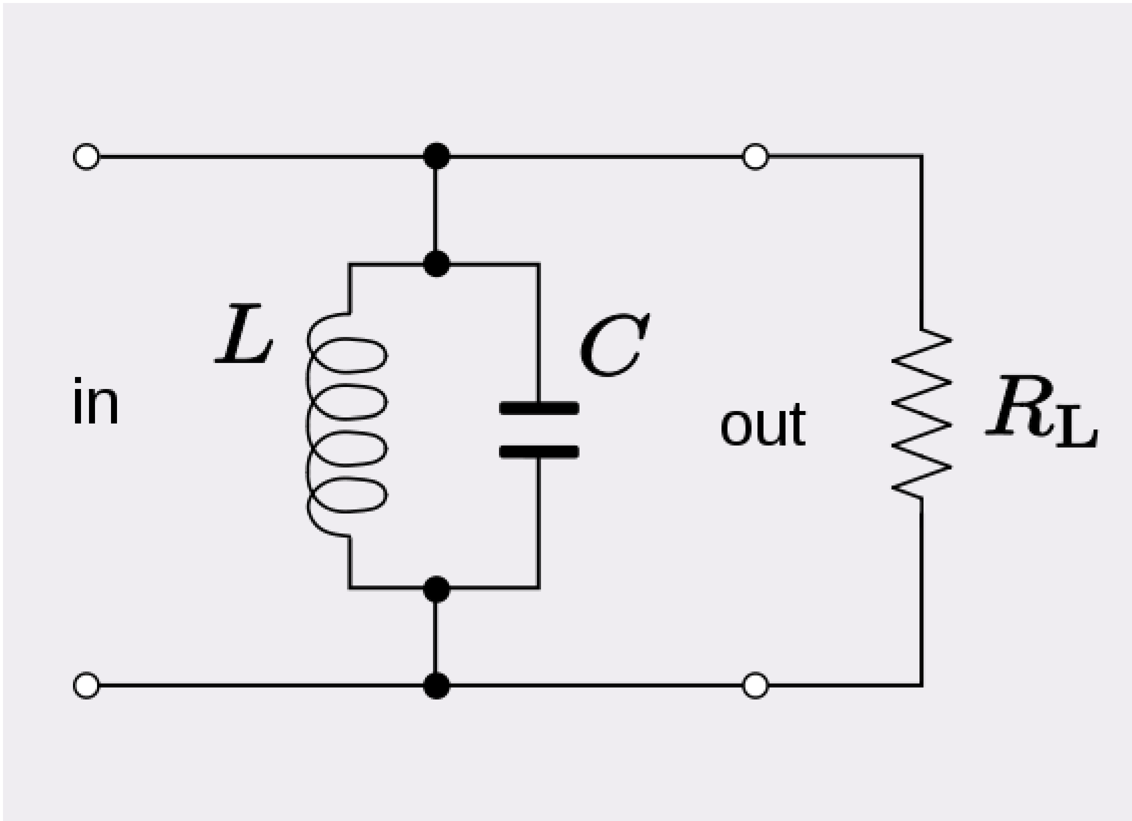


Fig. 12. RLC circuit as a parallel band-pass filter in shunt across the line

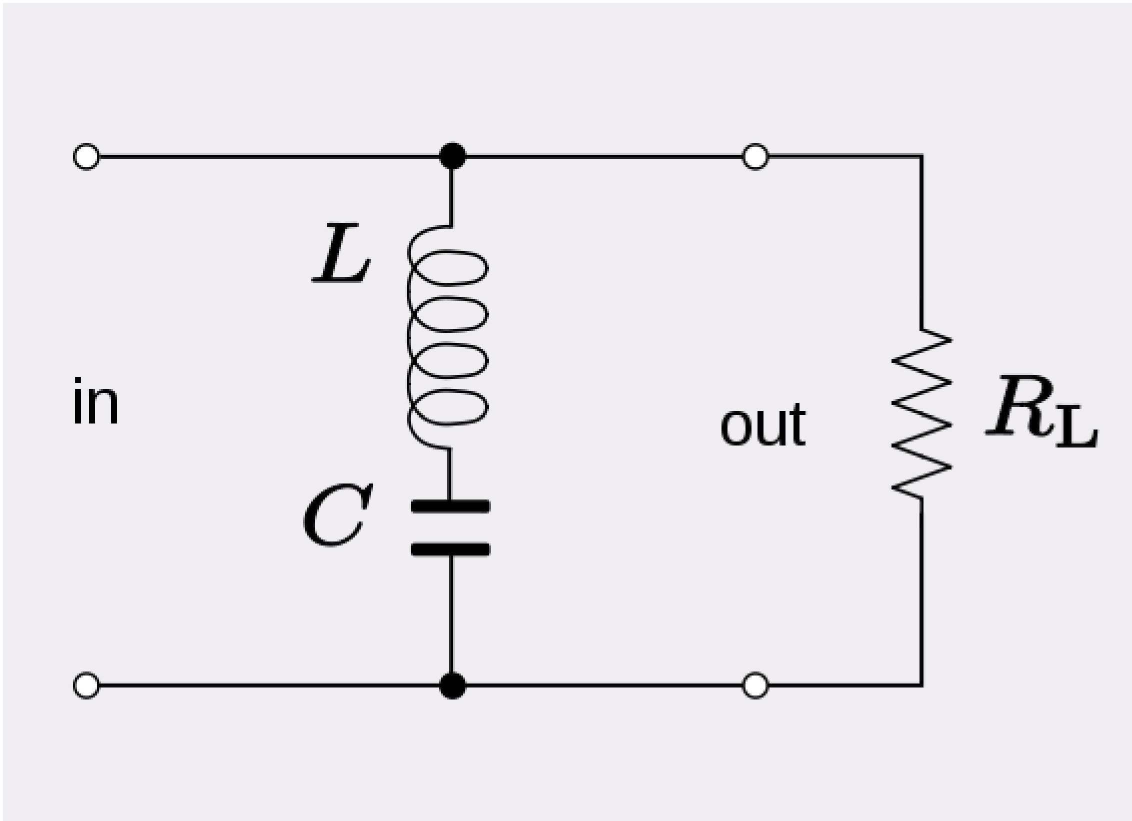


Fig. 13. RLC circuit as a series band-stop filter in shunt across the line

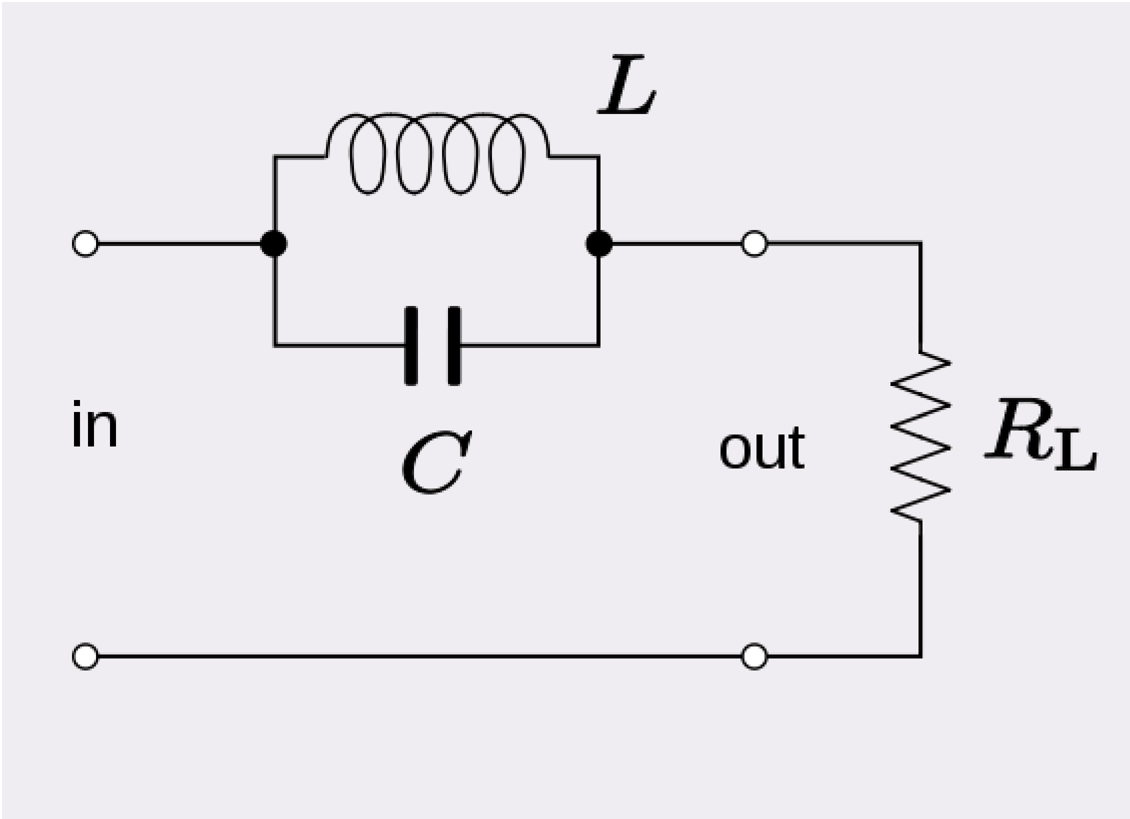


Fig. 14. RLC circuit as a parallel band-stop filter in series with the line

In the filtering application, the resistor *R* becomes the load that the filter is working into. The value of the damping factor is chosen based on the desired bandwidth of the filter. For a wider bandwidth, a larger value of the damping factor is required (and vice versa). The three components give the designer three degrees of freedom. Two of these are required to set the bandwidth and resonant frequency. The designer is still left with one which can be used to scale *R*, *L* and *C* to convenient practical values. Alternatively, *R* may be predetermined by the external circuitry which will use the last degree of freedom.

Low-pass filter

An RLC circuit can be used as a low-pass filter. The circuit configuration is shown in figure 9. The corner frequency, that is, the frequency of the 3dB point, is given by

$$\omega_c = \frac{1}{\sqrt{LC}}$$

This is also the bandwidth of the filter. The damping factor is given by

$$\zeta = \frac{1}{2R_L}\sqrt{\frac{L}{C}}$$

High-pass filter

A high-pass filter is shown in figure 10. The corner frequency is the same as the low-pass filter

$$\omega_{\mathrm{c}} = \frac{1}{\sqrt{LC}}$$

The filter has a stop-band of this width.

Band-pass filter

A band-pass filter can be formed with an RLC circuit by either placing a series LC circuit in series with the load resistor or else by placing a parallel LC circuit in parallel with the load resistor. These arrangements are shown in figures 11 and 12 respectively. The centre frequency is given by

$$\omega_{\mathrm{c}} = \frac{1}{\sqrt{LC}}$$

and the bandwidth for the series circuit is

$$\Delta\omega = \frac{R_{\mathrm{L}}}{L}$$

The shunt version of the circuit is intended to be driven by a high impedance source, that is, a constant current source. Under those conditions the bandwidth is

$$\Delta\omega = \frac{1}{CR_{\mathrm{L}}}$$

Band-stop filter

Figure 13 shows a band-stop filter formed by a series LC circuit in shunt across the load. Figure 14 is a band-stop filter formed by a parallel LC circuit in series with the load. The first case requires a high impedance source so that the current is diverted into the resonator when it becomes low impedance at resonance. The second case requires a low impedance source so that the voltage is dropped across the antiresonator when it becomes high impedance at resonance.

Oscillators

For applications in oscillator circuits, it is generally desirable to make the attenuation (or equivalently, the damping factor) as small as possible. In practice, this objective requires making the circuit's resistance *R* as small as physically possible for a series circuit, or alternatively increasing *R* to as much as possible for a parallel circuit. In either case, the *RLC circuit* becomes a good approximation to an ideal LC circuit. However, for very low attenuation circuits (high Q-factor) circuits, issues such as dielectric losses of coils and capacitors can become important.

In an oscillator circuit

$$\alpha \ll \omega_0.$$

or equivalently

$$\zeta \ll 1.$$

As a result

$$\omega_d \approx \omega_0.$$

Voltage multiplier

In a series RLC circuit at resonance, the current is limited only by the resistance of the circuit

$$I = \frac{V}{R}$$

If R is small, consisting only of the inductor winding resistance say, then this current will be large. It will drop a voltage across the inductor of

$$V_{\mathrm{L}} = \frac{V}{R}\omega_0 L$$

An equal magnitude voltage will also be seen across the capacitor but in antiphase to the inductor. If R can be made sufficiently small, these voltages can be several times the input voltage. The voltage ratio is, in fact, the *Q* of the circuit,

$$\frac{V_{\mathrm{L}}}{V} = Q$$

A similar effect is observed with currents in the parallel circuit. Even though the circuit appears as high impedance to the external source, there is a large current circulating in the internal loop of the parallel inductor and capacitor.

Pulse discharge circuit

An overdamped series RLC circuit can be used as a pulse discharge circuit. Often it is useful to know the values of components that could be used to produce a waveform this is described by the form:

$$I(t) = I_0(e^{-\alpha t} - e^{-\beta t})$$

Such a circuit could consist of an energy storage capacitor, a load in the form of a resistance, some circuit inductance and a switch - all in series. The initial conditions are that the capacitor is at voltage V_0 and there is no current flowing in the inductor. If the inductance L is known, then the remaining parameters are given by the following - Capacitance:

$$C = \frac{1}{L\alpha\beta}$$

Resistance (total of circuit and load):

$$R = L(\alpha + \beta)$$

Initial terminal voltage of capacitor:

$$V_0 = -I_0 L\alpha\beta\left(\frac{1}{\beta} - \frac{1}{\alpha}\right)$$

Rearranging for the case where R is known - Capacitance:

$$C = \frac{(\alpha + \beta)}{R\alpha\beta}$$

Inductance (total of circuit and load):

$$L = \frac{R}{(\alpha + \beta)}$$

Initial terminal voltage of capacitor:

$$V_0 = \frac{-I_0 R\alpha\beta}{(\alpha + \beta)}\left(\frac{1}{\beta} - \frac{1}{\alpha}\right)$$