

Electronic Systems and Technologies

Lecture Notes

Claudio Sansoè and Francesco Gregoretti

April 4, 2022

Contents

1 Amplifiers	1-1
1.1 The amplifier as black box	1-1
1.1.1 Power and voltage gains	1-3
1.1.2 Voltage amplifier	1-4
1.1.3 Other amplifiers	1-6
1.1.4 Amplifier parameters	1-10
1.1.5 Cascaded amplifiers	1-12
1.2 Frequency behavior	1-14
1.2.1 Bode diagram	1-15
1.2.2 High pass filter	1-22
1.2.3 Low pass filter	1-25
1.2.4 Band pass filter	1-27
1.2.5 Amplifier with complex impedance	1-30
1.2.6 Amplifier bandwidth	1-34
1.3 Linearity	1-35
1.3.1 Gain and offset errors	1-35
1.3.2 Non linear amplifiers	1-36
1.3.3 Saturation	1-37
1.4 Recap questions	1-38
1.5 Exercises	1-39
2 The Operational Amplifier	2-1
2.1 Introductory Analysis	2-1
2.1.1 Ideal operational amplifiers	2-1
2.1.2 Non ideal operational amplifiers	2-4
2.1.3 Voltage follower	2-8
2.1.4 Transresistance Amplifier	2-9
2.1.5 Current Amplifier	2-9
2.1.6 Transconductance Amplifier	2-10
2.1.7 Inverting Amplifier	2-11
2.2 Differential and common mode signals and gain	2-12
2.3 Real Operational Amplifier	2-16
2.3.1 Current and voltage offsets	2-16
2.3.2 Common mode input range	2-17
2.3.3 Output voltage range	2-17
2.3.4 Input Impedance	2-17
2.3.5 Differential gain	2-18
2.3.6 Common mode gain	2-18

2.4	Sizing the components of an amplifier	2-19
2.4.1	Specifications	2-19
2.4.2	Design	2-19
2.5	Frequency behavior	2-25
2.5.1	Transfer function	2-26
2.5.2	Loop gain	2-26
2.5.3	Gain-Bandwidth product	2-27
2.5.4	Slew Rate	2-29
2.6	Exercises	2-32
3	Operational Amplifier applications	3-1
3.1	Inverting configurations	3-1
3.1.1	Integrator	3-2
3.1.2	Differentiator	3-4
3.1.3	Inverting adder	3-6
3.2	Differential amplifier	3-7
3.3	Non inverting adder	3-10
3.4	Generalized adder	3-12
3.5	Instrumentation amplifiers	3-16
3.5.1	Introduction	3-16
3.5.2	Electromagnetic compatibility	3-16
3.5.3	Instrumentation amplifiers circuits	3-17
3.6	Audio amplifiers	3-21
3.7	Single supply circuits	3-24
3.8	Exercises	3-28
4	Non linear Amplifiers	4-1
4.1	Precision half wave rectifier	4-1
4.1.1	Shift of the transfer characteristic	4-3
4.2	Full wave rectifier	4-4
4.2.1	Extended full wave rectifier	4-7
4.2.2	Design example	4-11
4.2.3	Frequency behaviour	4-12
4.2.4	Changing the diode polarity	4-13
4.3	Exercises	4-14
5	Use of the Operational Amplifier out of linearity	5-1
5.1	Voltage comparators	5-1
5.1.1	Inverting comparator	5-2
5.1.2	Non inverting voltage comparator	5-3
5.1.3	Comparators and binary decision	5-3
5.1.4	Noise sensitivity	5-3
5.1.5	Comparators with hysteresis	5-4
5.1.6	Circuit implementation of voltage comparators	5-7
5.2	Astable Multivibrator	5-8
5.2.1	Circuit analysis	5-8
5.2.2	Quantitative analysis	5-10
5.3	Triangular wave generator	5-11
5.3.1	Quantitative analysis	5-11
5.3.2	Design example	5-13

5.4 Exercises	5-21
6 Silicon diodes, BJTs and MOSFETs	6-1
6.1 Basic concepts on semiconductors	6-2
6.1.1 Band model	6-3
6.1.2 Generation and recombination	6-4
6.1.3 Doped semiconductors	6-4
6.1.4 Semiconductor currents	6-5
6.2 Diode	6-7
6.2.1 pn junction	6-7
6.2.2 The diode	6-8
6.2.3 Diode static characteristic	6-10
6.2.4 Using diodes	6-12
6.2.5 Zener diodes	6-18
6.2.6 Diode applications	6-21
6.3 Bipolar transistor	6-25
6.3.1 The BJT as amplifier	6-29
6.4 MOSFET devices	6-42
6.4.1 MOS capacitor	6-42
6.4.2 MOSFET transistor	6-43
6.4.3 p channel MOSFET	6-46
6.4.4 Breakdown of MOSFET transistors	6-48
6.4.5 MOSFET symbols	6-49
6.4.6 The MOSFET as amplifier	6-49
6.5 Exercises	6-54
7 Transistors in switching Mode	7-1
7.1 BJT switches	7-2
7.1.1 Operation in Saturation	7-3
7.1.2 Design example	7-4
7.2 MOSFET switches	7-5
7.2.1 Open and close switch	7-7
7.2.2 Design example	7-7
7.3 Dynamic behaviour	7-9
7.3.1 Dynamic behaviour of diodes	7-9
7.3.2 Dynamic behaviour of BJTs	7-11
7.3.3 MOSFET dynamic behaviour	7-12
7.4 Exercises	7-16
8 Logic Circuits	8-1
8.1 Logic gates	8-2
8.1.1 Electrical parameters of logic gates	8-2
8.1.2 Logic gate compatibility	8-6
8.1.3 Noise margins	8-7
8.1.4 Fan-out	8-7
8.1.5 Logic families	8-7
8.1.6 Logic gate inputs	8-9
8.1.7 Gate output circuits	8-11
8.2 CMOS gates	8-14
8.2.1 Complementary static logic	8-14

8.2.2	nMOS-like logic	8-20
8.3	Examples	8-21
8.3.1	Fan out computation	8-21
8.3.2	LED driving	8-23
8.3.3	Wired-or	8-24
8.4	Dynamic behaviour of logic gates	8-26
8.4.1	Definitions	8-26
8.4.2	Propagation delay	8-27
8.4.3	Power analysis	8-29
8.5	Sequential circuits	8-31
8.5.1	Bistable element	8-31
8.5.2	S-R latch with NAND gates	8-31
8.5.3	SR latch with NOR gates	8-33
8.5.4	Oscillatory Metastability	8-33
8.5.5	SR Transparent Latch	8-34
8.5.6	D-latch	8-35
8.5.7	Synchronous and asynchronous circuits	8-36
8.5.8	D Flip-flop	8-37
8.5.9	Dynamic behaviour of flip flops	8-38
8.6	Exercises	8-40
9	Passive devices	9-1
9.1	Resistors	9-1
9.1.1	Nominal value and tolerance	9-2
9.1.2	Power dissipation and packaging	9-3
9.1.3	Maximum voltage and temperature coefficient	9-4
9.1.4	Material and parasitic parameters	9-4
9.1.5	Noise	9-5
9.2	Capacitors	9-5
9.2.1	Parasitic parameters	9-6
9.2.2	Ceramic capacitors	9-6
9.2.3	Plastic film capacitors	9-8
9.2.4	Electrolytic capacitors	9-9
9.3	Inductors	9-11
9.3.1	Parasitic parameters	12
A	Solution to exercises	A-1
A.1	Solution to chapter 1 exercises	A-1
A.1.1	Solution to exercise 1.1	A-1
A.1.2	Solution to exercise 1.2	A-1
A.1.3	Solution to exercise 1.3	A-2
A.1.4	Solution to exercise 1.4	A-3
A.1.5	Solution to exercise 1.5	A-4
A.1.6	Solution to exercise 1.6	A-5
A.2	Solution to chapter 2 exercises	A-6
A.2.1	Solution to exercise 2.1	A-6
A.2.2	Solution to exercise 2.2	A-7
A.2.3	Solution to exercise 2.3	A-8
A.2.4	Solution to exercise 2.4	A-9
A.2.5	Solution to exercise 2.5	A-10

A.2.6	Solution to exercise 2.6	A-13
A.2.7	Solution to exercise 2.7	A-17
A.2.8	Solution to exercise 2.8	A-18
A.3	Solution to chapter 3 exercises	A-20
A.3.1	Solution to exercise 3.1	A-20
A.3.2	Solution to exercise 3.2	A-22
A.3.3	Solution to exercise 3.3	A-24
A.3.4	Solution to exercise 3.4	A-25
A.3.5	Solution to exercise 3.5	A-26
A.3.6	Solution to exercise 3.6	A-27
A.3.7	Solution to exercise 3.7	A-28
A.3.8	Solution to exercise 3.8	A-29
A.3.9	Solution to exercise 3.9	A-30
A.4	Solution to chapter 4 exercises	A-32
A.4.1	Solution to exercise 4.1	A-32
A.4.2	Solution to exercise 4.2	A-33
A.4.3	Solution to exercise 4.3	A-35
A.5	Solution to chapter 5 exercises	A-36
A.5.1	Solution to exercise 5.1	A-36
A.5.2	Solution to exercise 5.2	A-37
A.5.3	Solution to exercise 5.3	A-39
A.6	Solution to chapter 6 exercises	A-40
A.6.1	Solution to exercise 6.1	A-40
A.6.2	Solution to exercise 6.2	A-40
A.6.3	Solution to exercise 6.3	A-41
A.6.4	Solution to exercise 6.4	A-43
A.6.5	Solution to exercise 6.5	A-45
A.6.6	Solution to exercise 6.6	A-47
A.6.7	Solution to exercise 6.7	A-49
A.7	Solution to chapter 7 exercises	A-51
A.7.1	Solution to exercise 7.1	A-51
A.7.2	Solution to exercise 7.2	A-52
A.7.3	Solution to exercise 7.3	A-53
A.8	Solution to chapter 8 exercises	A-54
A.8.1	Solution to exercise 8.1	A-54
A.8.2	Solution to exercise 8.2	A-54
A.8.3	Solution to exercise 8.3	A-55
A.8.4	Solution to exercise 8.4	A-56
A.8.5	Solution to exercise 8.5	A-56

Chapter 1

Amplifiers

Contents

1.1	The amplifier as black box	1-1
1.1.1	Power and voltage gains	1-3
1.1.2	Voltage amplifier	1-4
1.1.3	Other amplifiers	1-6
1.1.4	Amplifier parameters	1-10
1.1.5	Cascaded amplifiers	1-12
1.2	Frequency behavior	1-14
1.2.1	Bode diagram	1-15
1.2.2	High pass filter	1-22
1.2.3	Low pass filter	1-25
1.2.4	Band pass filter	1-27
1.2.5	Amplifier with complex impedance	1-30
1.2.6	Amplifier bandwidth	1-34
1.3	Linearity	1-35
1.3.1	Gain and offset errors	1-35
1.3.2	Non linear amplifiers	1-36
1.3.3	Saturation	1-37
1.4	Recap questions	1-38
1.5	Exercises	1-39

THIS chapter summarizes concepts that the students should already know from Circuit Theory. We will analyze generic amplifiers, using the double port equivalent. First we will work in DC, then we will investigate the frequency behavior of the circuits. To be able to understand the concepts introduced in this chapter, the student should be proficient with KVL and KCL circuit analysis and should also know the behavior of RC networks, both in time and frequency domain. Bode plots will be recalled and used in the analysis.

1.1 The amplifier as black box

A generic amplifier is an electronic circuit with an input port, an output port and a power supply port. A signal is provided to the input port. The output generates a signal with an associated power greater than the power absorbed

from the input port. The power supply port provides power to the output signal and to the internal circuitry of the amplifier.

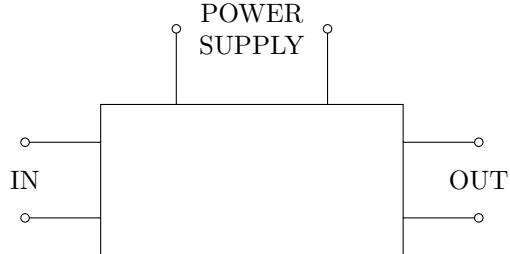


Figure 1.1: Amplifier

Graphically we can represent an amplifier as a box, like in figure 1.1.

From the above description we can define the amplifier as a linear block that generates an output signal proportional to the input signal. The key concept is that the amplifier must provide power gain. An example of a double port device which is not an amplifier, even if it provides an output signal proportional to the input one, is the transformer. The transformer does not increase the power of the output signal with respect to the one of the input signal (a real transformer has internal power losses, the output power is less than the input one).

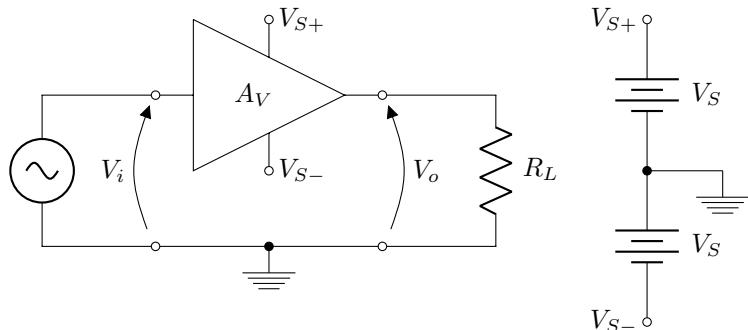


Figure 1.2: Voltage amplifier

Often the input and output ports are referenced to the same potential, normally 0 V, so the input and output ports have one common terminal. Moreover, to easily identify a double port device as an amplifier, we normally prefer to represent it graphically as a triangle. If the input and output signals are voltages, then we can draw the amplifier like in figure 1.2.

In the schematic, A_V is the voltage gain of the amplifier, V_{S+} and V_{S-} are the power supply voltages. Very often, amplifiers need positive and negative voltages from the power supply, to be able to cope with bipolar signals. When representing a simplified schematic of a circuit, it is common practice to omit the indication of the power supply connections. This is done to focus the attention on non-trivial parts of the schematic and does not mean that the amplifier can work without them...

1.1.1 Power and voltage gains

The power gain of an amplifier is defined as the ratio between the power delivered to the output and the power received from the input:

$$K_p = \frac{P_o}{P_i}$$

It is common practice to measure the magnitude of the power gain of an amplifier in decibel.

$$G = |K_p|_{\text{dB}} = 10 \log_{10} \left(\left| \frac{P_o}{P_i} \right| \right)$$

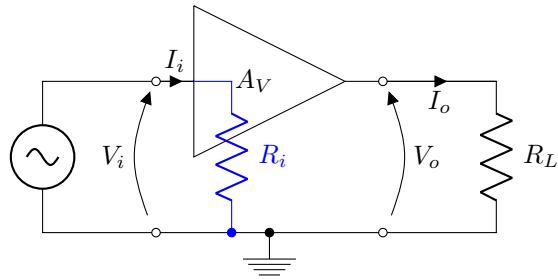


Figure 1.3: Input and output power

K_p	G
1	0 dB
2	3 dB
10	10 dB
100	20 dB
1000	30 dB

Table 1.1: Power gain in dB

The relationship between K_p and G for some meaningful values is shown in table 1.1. If we consider input and output power more in detail, we can see that, referring to figure 1.3, obviously the input power is the product of the input voltage and the input current, while the output power is the product between the output voltage and the output current. This means that we need to know the equivalent input resistance of the amplifier to be able to calculate the input power and we need to know the magnitude of the load resistance to compute the output power.

$$P_i = V_i \cdot I_i = \frac{V_i^2}{R_i}$$

$$P_o = V_o \cdot I_o = \frac{V_o^2}{R_L}$$

If the input resistance and the load resistance are the same ($R_L = R_i$), the

power gain in decibel can also be written as:

$$G = 10 \log_{10} \left(\frac{V_o^2}{V_i^2} \right) = 20 \log_{10} \left(\frac{|V_o|}{|V_i|} \right)$$

We can extend this notation neglecting R_i and R_L and express, even if $R_i \neq R_L$, the voltage gain in decibel as:

$$|A_V|_{\text{dB}} = 20 \log_{10} \left(\frac{|V_o|}{|V_i|} \right)$$

Table 1.2 lists some useful values.

$ A_V $	$ A_V _{\text{dB}}$
1	0 dB
2	6 dB
10	20 dB
100	40 dB
1000	60 dB

Table 1.2: Voltage gain in dB

As already specified, the voltage is not the only parameter defining an amplifier's gain: one should always refer primarily to the power gain. As an example, we can have an amplifier which has a voltage gain of one, but, if the input impedance is $R_i = 100 \text{ k}\Omega$ and the load resistance is $R_L = 1 \text{ k}\Omega$, we have that the current in the load is one hundred times the current in the amplifier's input. The power gain is 100 and so is the current gain.

1.1.2 Voltage amplifier

An amplifier can be modeled as a double port device, neglecting the always necessary power supply connections. If we want to amplify the voltage from a voltage source, then the best amplifier is a device that senses the input voltage, produces a proportional output voltage and is able to drive any load. What we really need is a voltage controlled voltage source, like in figure 1.4.

This circuit is an ideal voltage to voltage amplifier because it has infinite input impedance, so it does not alter the input voltage, and the output is an ideal voltage source, so that the output voltage is not modified by the load resistor.

In practice this circuit can only be approximated by real word devices. A more realistic model includes:

- Amplifier input finite resistance, modeled by a resistor R_i .
- Amplifier output resistance, modeled by a resistor R_o .
- Signal generator equivalent resistance, modeled by a resistor R_g .

The new circuit is displayed in figure 1.5.

When considering input and output impedance of the amplifier, the circuit voltage gain, V_o/V_g , is lower than the no-load amplifier gain A_V .

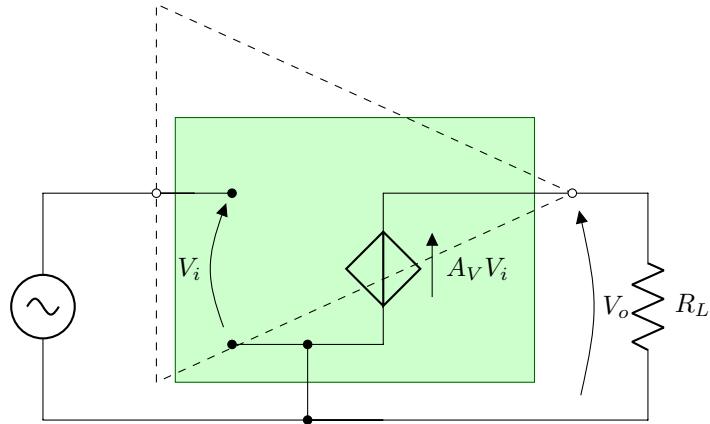


Figure 1.4: Ideal voltage amplifier

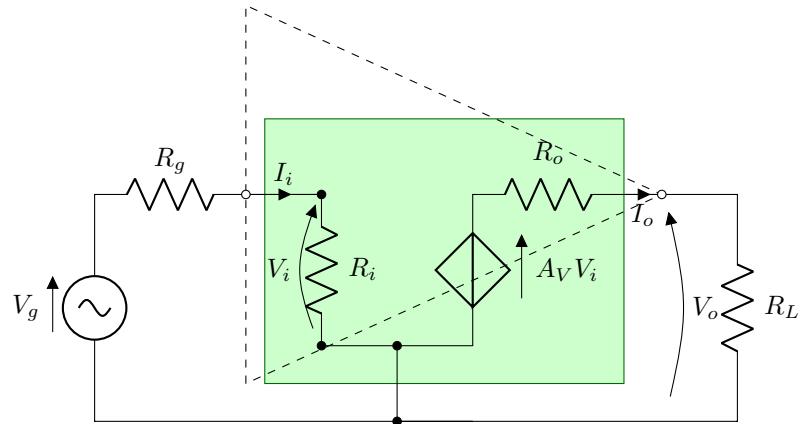


Figure 1.5: Real voltage amplifier

R_g and R_i introduce an input voltage division, so that:

$$V_i = \frac{R_i}{R_i + R_g} V_g$$

At the same time, R_o and R_L introduce an output voltage division:

$$V_o = A_V V_i \frac{R_L}{R_L + R_o}$$

Combining the two effects, we finally have:

$$\frac{V_o}{V_g} = \frac{R_i}{R_i + R_g} \frac{R_L}{R_L + R_o} A_V$$

Numerical example

Let us consider a real amplifier modeled like in figure 1.5, with the following parameters:

- Amplifier:

- $A_V = 120$;
- $R_i = 10 \text{ k}\Omega$;
- $R_o = 300 \Omega$.

- External devices:

- $R_g = 600 \Omega$;
- $R_L = 1.5 \text{ k}\Omega$.

We can split the computation in two steps. First we compute the voltage at the input of the amplifier:

$$V_i = \frac{R_i}{R_i + R_g} V_g = \frac{10 \text{ k}\Omega}{10.6 \text{ k}\Omega} V_g = 0.943 \cdot V_g$$

We can then compute the output voltage as function of V_i and the gain A_V :

$$V_o = A_V V_i \frac{R_L}{R_L + R_o} = 120 \cdot \frac{1.5 \text{ k}\Omega}{1.8 \text{ k}\Omega} V_i = 120 \cdot 0.833 \cdot V_i = 100 \cdot V_i$$

The total gain of the circuit is therefore:

$$\frac{V_o}{V_g} = 100 \cdot 0.943 = 94.3$$

We can identify

- Input losses: due to finite input impedance R_i , only 94.3 % of the generator voltage is passed to the amplifier.
- Output losses: due to finite output impedance R_o , only 83.3 % of the unloaded output voltage is transferred to the load.

These results confirm what we already noted: a good voltage amplifier should have very high input impedance and very low output impedance.

1.1.3 Other amplifiers

Even if in most cases the signal sources can be modeled as a voltage source, there are times in which the input to our system is a current generator.

Example of voltage sources are microphones or logic circuits. Example of current sources are optical sensors (photodiodes) or chemical sensors.

When we consider the load of our system, we normally drive it by imposing an output voltage, but there are cases in which we need to supply a specific current to the load instead of a specific voltage. Example of voltage loads are logic circuits, light bulbs and some kind of motors, while current loads are required in battery chargers, electromagnetic actuators, LEDs, magnetic needle instruments.

Depending on the kind of source and load, we can then classify our amplifiers in four classes:

- Voltage amplifier, amplifies the input voltage and produces a proportional output voltage: $A_V = V_o/V_i$
- Current amplifier, amplifies the input current and generates a proportional output current: $A_I = I_o/I_i$
- Transconductance amplifier, amplifies the input voltage and produces a proportional output current: $G_m = I_o/V_i$
- Transresistance amplifier, amplifies the input current and generates a proportional output voltage: $R_m = V_o/I_i$

Current at the input

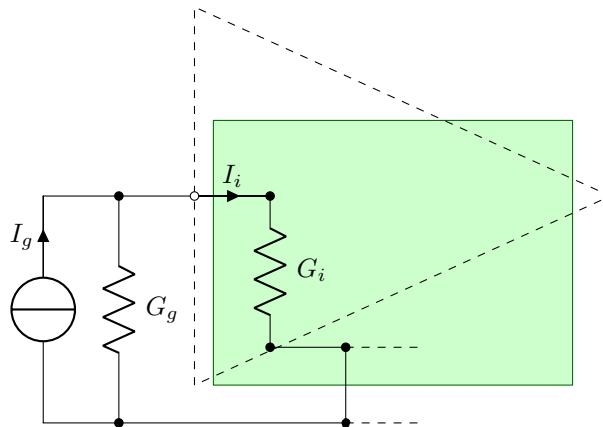


Figure 1.6: Current input amplifier

If the input source can be modeled as a current source with possibly an equivalent parallel conductance G_g , then our amplifier should receive most of the input current, independently from G_g . The circuit is depicted in figure 1.6.

The ideal situation to be able to transfer all of the generator current to the input of the amplifier is to have $G_i \rightarrow \infty$, that is $R_i \rightarrow 0$.

If the amplifier has to be connected to a current source, its input impedance must be low. If the input impedance is not zero, then the input current is:

$$I_i = I_g \frac{G_i}{G_i + G_g} = I_g \frac{R_g}{R_i + R_g}$$

Current at the output

If the load of the amplifier requires a specific current instead of a specific voltage, then we need a current amplifier, that is an amplifier whose output can be modeled as a controlled current source, like in figure 1.7.

We want to avoid current division at the output, so that we can transfer as much current as possible to the load. To obtain this, we need G_o as low as possible, that is $R_o \rightarrow \infty$.

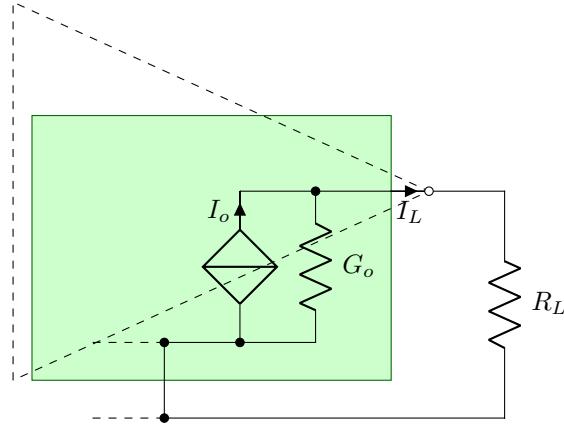


Figure 1.7: Current output amplifier

Every time the output of an amplifier can be modeled as a controlled current source and we are interested in maximizing the current transfer to the load, then the output impedance should be as high as possible.

We are now ready to study the schematic of the three amplifiers other than the voltage amplifier, listed at the beginning of this section.

Current amplifier

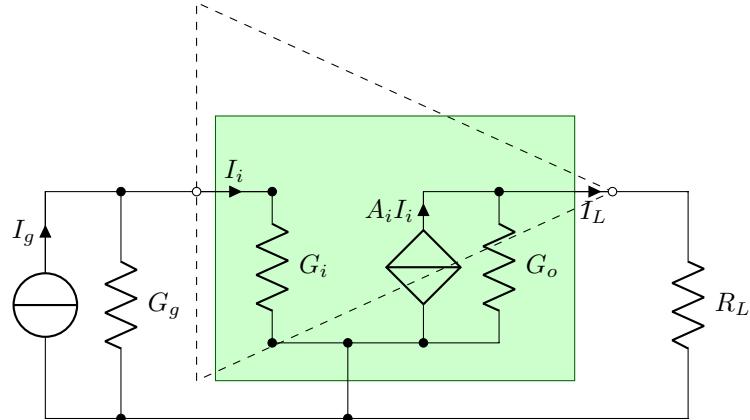


Figure 1.8: Current amplifier

The current amplifier receives its input from a current source and drives a load that needs a specific current, for example a battery that needs to be recharged at a specific current or a LED.

The schematic is in figure 1.8. The ideal current amplifier has $R_i = 0$ (or $G_i \rightarrow \infty$) and $R_o \rightarrow \infty$ (or $G_o = 0$). If this is not the case, there will be some current partition at the input and at the output:

$$I_i = I_g \frac{G_i}{G_i + G_g} = I_g \frac{R_g}{R_i + R_g}$$

$$I_L = A_i I_i \frac{G_L}{G_L + G_o} = A_i I_i \frac{R_o}{R_L + R_o}$$

The current gain will be less than A_i :

$$\frac{I_L}{I_g} = A_i \frac{R_g}{R_i + R_g} \frac{R_o}{R_L + R_o}$$

Transconductance amplifier

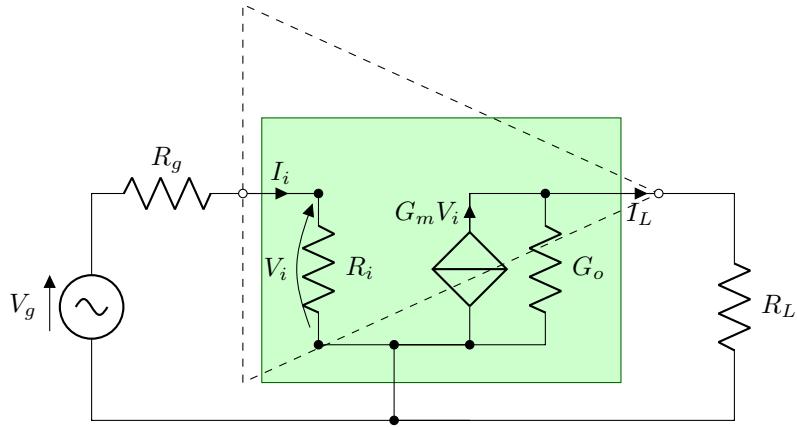


Figure 1.9: Transconductance amplifier

The transconductance amplifier accepts a voltage at the input and delivers a current to the output. Its gain is expressed as $G_m = I_o/V_i$, therefore it is not a-dimensional like in voltage or current amplifiers but it is a conductance, hence the name.

Its schematic is shown in figure 1.9.

The ideal transconductance amplifier has very high input impedance and very high output impedance, that is $R_i \rightarrow \infty$ and $G_o = 0$.

The transconductance of a real amplifier is:

$$V_i = \frac{R_i}{R_i + R_g} V_g$$

$$I_L = G_m V_i \frac{G_L}{G_L + G_o} = G_m V_i \frac{R_o}{R_L + R_o}$$

$$\frac{I_L}{V_g} = G_m \frac{R_i}{R_i + R_g} \frac{R_o}{R_L + R_o}$$

Transresistance amplifier

The last configuration we study accepts a current at the input and produces a voltage at the output. Its gain is expressed as $R_m = V_o/I_i$, therefore it is a resistance. This is why it is called transresistance amplifier, figure 1.10.

An ideal transresistance amplifier has $R_i = 0$ and $R_o = 0$.

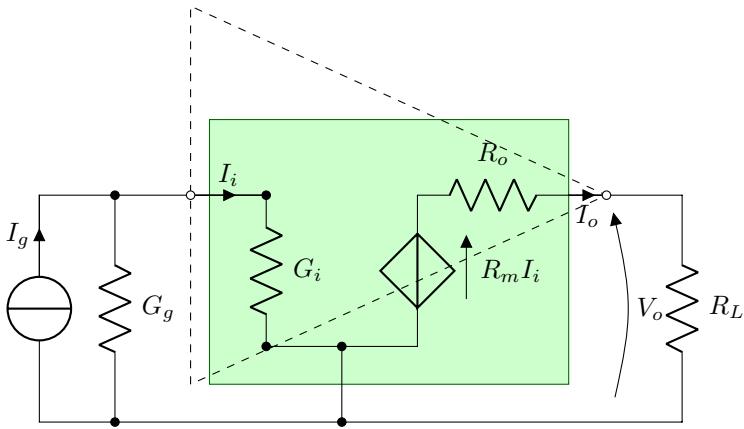


Figure 1.10: Transresistance amplifier

1.1.4 Amplifier parameters

Given a circuit with voltage or current dependent sources, it is useful to transform it in a double port equivalent amplifier.

To be able to do so, we need to identify the parameters of the equivalent double port amplifier. Let us for example transform a circuit in an equivalent voltage amplifier. We need to find what is the amplifier's input resistance, R_i , output resistance, R_o , and its no-load gain.

We can find the three parameters by using the following procedure.

- **Input resistance.** To find the input resistance we apply a voltage generator V_x at the input of the circuit and find the corresponding input current I_x . Input resistance will be:

$$R_i = \frac{V_x}{I_x}$$

- **Output resistance.** To find the output resistance we apply a voltage generator V_x at the output of the circuit, after short-circuiting all the independent voltage generators at the input. We find the current drawn from V_x , I_x , and we have:

$$R_o = \frac{V_x}{I_x}$$

- **No-load gain.** To find the gain of the amplifier we connect an ideal voltage generator V_i at the input, we leave the output in open-circuit and we compute the output voltage V_o . The gain is therefore

$$A_V = \frac{V_o}{V_i}$$

If we want to model a circuit as any of the other amplifier types studied in the previous section, we need to modify how we compute the gain.

- If we want to model the output of the amplifier as a current source, we have to compute the output current when the load is a short-circuit and not an open circuit.

- If we want to model the input to accept input current, then the independent generator should be an ideal current source.

Example

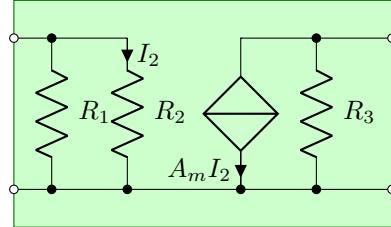
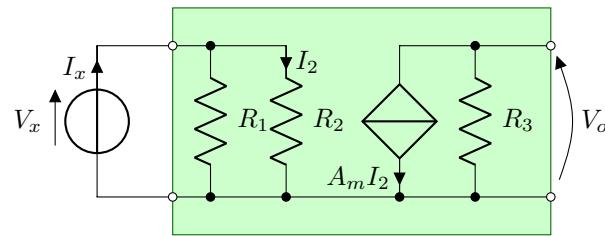
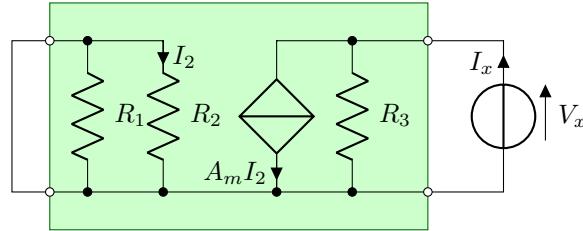


Figure 1.11: Example network.

To clarify the procedure, let us try to model as a voltage amplifier the simple circuit of figure 1.11.



a) input impedance and gain



b) output impedance

Figure 1.12: Probing input and output impedance.

Input impedance

To compute the input impedance we connect an ideal voltage source V_x at the input and we measure the current I_x drained by the voltage source. The procedure is shown in figure 1.12a. In this simple case, we see that I_x is shared by the two parallel resistors R_1 and R_2 , so that the input impedance can be identified without further calculations as

$$R_i = R_1 \parallel R_2$$

Output impedance

It is also easy to compute the output impedance. If we short-circuit the input voltage generator, like in figure 1.12b, The current of the current dependent current generator is 0. By inserting a voltage generator at the output, the only current is through R_3 , so the output impedance of this circuit is

$$R_o = R_3$$

Gain

To compute the no-load gain A_V we use an ideal voltage generator V_i as input and we leave open the output of the circuit, like in figure 1.12a.

The current I_2 can be expressed as:

$$I_2 = \frac{V_i}{R_2}$$

The output voltage is the voltage drop across R_3 due to the current dependent generator's current.

$$V_o = -A_m I_2 R_3 = -V_i A_m \frac{R_3}{R_2}$$

Finally, the gain is:

$$A_V = \frac{V_o}{V_i} = -A_m \frac{R_3}{R_2}$$

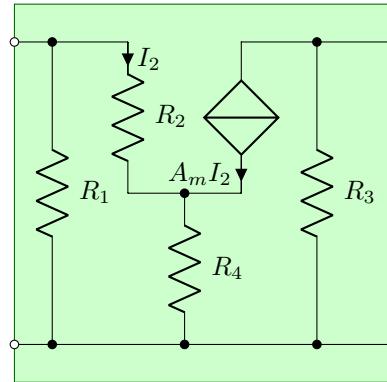


Figure 1.13: Example network 2.

The above example is deliberately simple. A slightly more complex circuit, like the one shown in figure 1.13 requires more computations to obtain the equivalent double port voltage amplifier. The solution of this circuit is left to the student.

1.1.5 Cascaded amplifiers

Frequently, analog circuits can be modeled as a sequence of amplifiers or double port circuits. In this case it is possible to model the complete chain as a single equivalent double port.

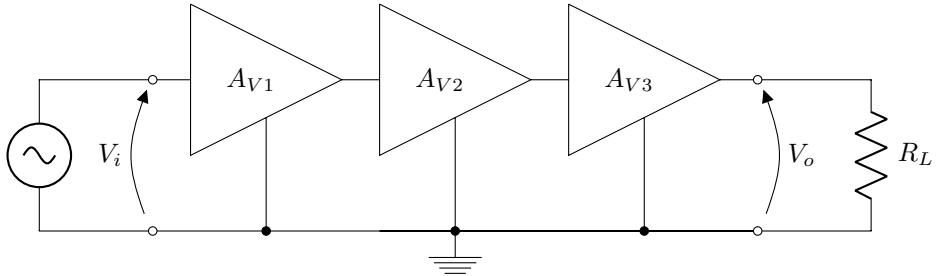


Figure 1.14: Cascaded amplifiers

Figure 1.14 depicts this situation. If there is no feedback among the different stages, the input impedance of the equivalent double port is that of the first element, while the output impedance is that of the last one. To compute the gain, one has to inspect the elements and keep into account the interaction between the output impedance of the previous stage and input impedance of the next one.

Example

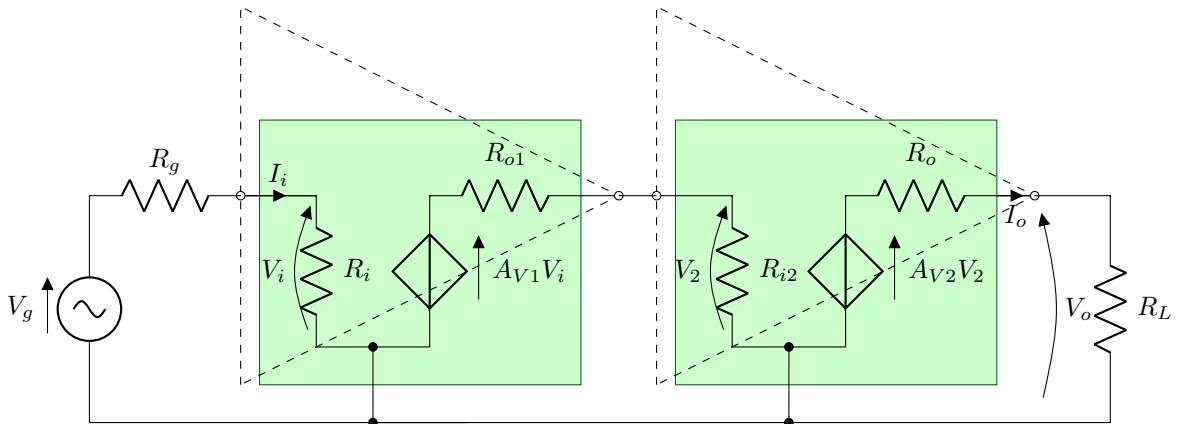


Figure 1.15: Two cascaded voltage amplifiers

Let us compute the equivalent double port of two cascaded voltage amplifiers, drawn in figure 1.15.

There is no feedback from the output of the second amplifier to the input of the first, so the input resistance of the equivalent amplifier is R_i and the output resistance R_o .

To compute the equivalent no-load gain, we disconnect R_L and connect an ideal voltage source at the input (or we consider $R_g = 0 \Omega$ in the schematic). First we find V_2

$$V_2 = A_{V1}V_i \frac{R_{i2}}{R_{i2} + R_{o1}}$$

We can then easily find A_V

$$A_V = \frac{V_o}{V_i} = A_{V1}A_{V2} \frac{R_{i2}}{R_{i2} + R_{o1}}$$

We can observe that the gain is always less than the product of the gains of the chained amplifiers unless their output impedance is much less than the input impedance of the next element.

1.2 Frequency behavior

When we apply a time-varying signal at the input of an amplifier, we can study the behavior of the output, called dynamic behavior, using different techniques.

If we consider the Fourier transform of the signal, we can study how the different frequency components of the signal are modified by the amplifier. The main instrument we use is the so called Bode plot.

Alternatively we can study the behavior of the amplifier in the time domain. In this case, we can perform transient analysis and step response.

With respect to the static analysis, which gives us the DC response of the amplifier, we now have to take into account the reactive parameters of the circuit, that is capacitors and inductors.

When we work in the frequency domain, we find the dependency of the gain of our amplifier from the frequency in hertz (f , Hz) or the angular frequency (ω) in radians per second. It is easy to transform angular frequency in frequency: $f = \omega/(2\pi)$. The gain of an amplifier will be written as function of either ω or f . We will then have $A_V = A_V(j\omega)$ or $A_V = A_V(jf)$. We normally prefer to use the Laplace transform instead of the Fourier transform. We just need to recall that the Fourier transform is equivalent to the Laplace transform computed on the imaginary axis. To study the frequency behavior of a transfer function expressed in Laplace domain we just need to substitute in the expression the complex variable s with $j\omega$. The Bode plot is a graph of the magnitude of a transfer function in the frequency domain. There is also a Bode plot of the phase of a transfer function, but we will not use this second graph in our course, even if the phase versus frequency behavior will be very important for other more advanced topics in electronics.

Conversely, we can work in the time domain. In this domain we study the output voltage of the amplifier as function of time and of an input stimulus. The input is normally a step input and this type of analysis is called transient response. The two analyses are complementary. The transient response can take into account also non linear behavior of the amplifier while it is much easier to obtain the frequency response for linear circuits.

In the following, we will apply frequency and time domain analyses to simple circuits, to recall what we already studied in circuit theory. In both types of analysis, we will first obtain what we call asymptotic behavior, and then the detailed behavior in frequency or time. Asymptotic behavior in the frequency domain means to evaluate the transfer function in DC (or for $\omega \rightarrow 0$) and in high frequency (for $\omega \rightarrow \infty$). In time domain analysis it means to find the behavior for $t = 0$, replacing capacitors by short-circuits, and for $t \rightarrow \infty$, replacing capacitors by open-circuit. Detailed behavior in frequency domain means to find the position of poles and zeros of the circuit and to draw the Bode diagram, in

time domain means to find the time constants of the circuit and to draw the $V(t)$ graph.

1.2.1 Bode diagram

Every network function of a linear circuit can be expressed as the ratio of two polynomials in the Laplace transform complex variable s :

$$H(s) = \frac{N(s)}{D(s)}$$

The two polynomials can in turn be expressed as the product of their roots, which can be either real first order roots or complex conjugates second order roots. We will not study second order roots even if they are very important in filter design and will be used in more advanced electronics topics.

Therefore, our network function will always be in the form:

$$H(s) = K \frac{s^{n_z} \left(1 + \frac{s}{z_1}\right) \left(1 + \frac{s}{z_2}\right) \cdots}{s^{n_p} \left(1 + \frac{s}{p_1}\right) \left(1 + \frac{s}{p_2}\right) \cdots}$$

Our functions will have zeros, that is roots of the numerator, on the real axis in the left plane (that is $s < 0$) and/or in the origin. The functions will also have poles, that is roots of the denominator, on the real axis in the left plane and/or in the origin. Zeros or poles in the right plane usually mean instability in the circuit but we will only study stable circuits.

The Bode plot of the magnitude of a transfer function is in the frequency domain ($s \rightarrow j\omega$) and is drawn using logarithmic axes. The x axis represents the logarithm of the frequency, while the y axis is the logarithm of the magnitude of the transfer function, in decibel. The reason for using log axes is very simple. If we use the properties of the logarithms on the magnitude of the above written network function, using the definition of decibel, we obtain:

$$\begin{aligned} |H(j\omega)|(\text{dB}) &= 20 \log_{10}(|H(j\omega)|) = 20 \log_{10}(|K|) + (n_z - n_p)20 \log_{10}(\omega) + \\ &+ \sum_i 20 \log_{10} \left(\left| 1 + \frac{j\omega}{z_i} \right| \right) - \sum_j 20 \log_{10} \left(\left| 1 + \frac{j\omega}{p_j} \right| \right) \end{aligned}$$

What we learn from the above equation is that a Bode plot can be constructed by summing together the contribution of the single poles and zeros. It is easy to graphically perform this operation.

Let us then study each type of contribution.

Single pole in the origin

If in our network function we have the term

$$H(s) = \frac{1}{s}$$

we have a pole in the origin, that is the magnitude is infinity for $s = 0$, that is $j\omega = 0$, or in DC.

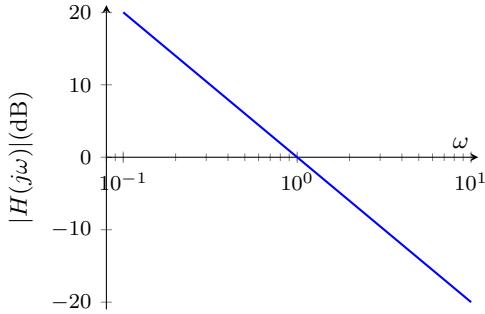


Figure 1.16: Bode plot for single pole in the origin.

The magnitude of $H(j\omega)$ is 1, that is 0 dB, for $\omega = 1$ and decreases by a factor of ten (20 dB) when we multiply the frequency by ten.

$$\begin{aligned}|H(0.1j)| &= 20 \log_{10}(10) = 20 \text{ dB} \\ |H(j)| &= 20 \log_{10}(1) = 0 \text{ dB} \\ |H(10j)| &= 20 \log_{10}(0.1) = -20 \text{ dB}\end{aligned}$$

We have a straight line with slope of -20 dB/decade , crossing the 0 dB axis at $\omega = 1$. Figure 1.16 shows the plot.

Single zero in the origin

If in our network function we have the term

$$H(s) = s$$

we have a zero in the origin, that is the magnitude is zero for $s = 0$, that is $j\omega = 0$, or in DC. This case creates a Bode plot mirrored with respect to the previous one.

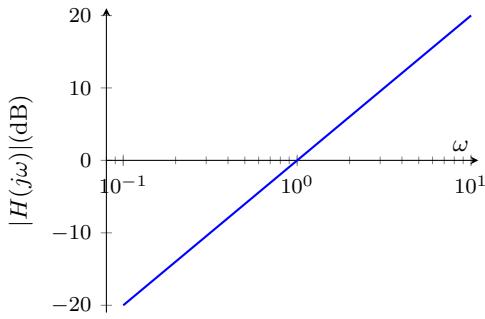


Figure 1.17: Bode plot for single zero in the origin.

The magnitude of $H(j\omega)$ is 1, that is 0 dB, for $\omega = 1$ and increases by a

factor of ten (20 dB) when we multiply the frequency by ten.

$$|H(0.1j)| = 20 \log_{10}(0.1) = -20 \text{ dB}$$

$$|H(j)| = 20 \log_{10}(1) = 0 \text{ dB}$$

$$|H(10j)| = 20 \log_{10}(10) = 20 \text{ dB}$$

We have a straight line with slope of 20 dB/decade, crossing the 0 dB axis at $\omega = 1$. Figure 1.17 shows the plot.

Single zero at angular frequency z_i

If in our network function we have the term

$$H(s) = 1 + \frac{s}{z_i}$$

we have a zero at angular frequency z_i . That means that, in the Fourier domain, at angular frequency z_i the imaginary part of the term equals 1.

To draw the Bode plot we first analyze the asymptotic behavior of the function:

$$\lim_{\omega \rightarrow 0} (H(j\omega)) = 1$$

That means that for low frequency the Bode plot is a flat line coincident with the 0 dB axis.

$$\lim_{\omega \rightarrow \infty} (H(j\omega)) = \frac{j\omega}{z_i}$$

If we compute the magnitude in decibel of the asymptotic function we obtain

$$\left| \frac{j\omega}{z_i} \right| (\text{dB}) = 20 \log_{10}(\omega) - 20 \log_{10}(z_i)$$

That means that for high frequency the Bode plot is a straight line with slope of 20 dB/decade, crossing the 0 dB axis at $\omega = z_i$.

The asymptotic behavior is correct for low and high frequency, but the function will be not coincident with the asymptote for frequency close to z_i . We should then calculate the exact function:

$$|H(j\omega)|(\text{dB}) = 20 \log_{10} \left(\left| 1 + \frac{j\omega}{z_i} \right| \right) = 20 \log_{10} \left(\sqrt{1 + \frac{\omega^2}{z_i^2}} \right)$$

At angular frequency z_i , the magnitude is

$$|H(jz_i)|(\text{dB}) = 20 \log_{10} \left(\left| 1 + \frac{jz_i}{z_i} \right| \right) = 20 \log_{10} (\sqrt{2}) = 3 \text{ dB}$$

Figure 1.18 shows both the asymptotic and the exact behavior of the term, assuming $z_i = 1 \text{ rad s}^{-1}$.

When studying the frequency response of a circuit, the asymptotic behavior of a zero or pole can be used as first approximation, the exact function is then used for more precise analysis.

When examining a circuit in the laboratory, we can identify the frequency of the zero as the frequency where we find an increase of 3 dB in the magnitude. Other methods include observation of the phase shift between input and output signal.

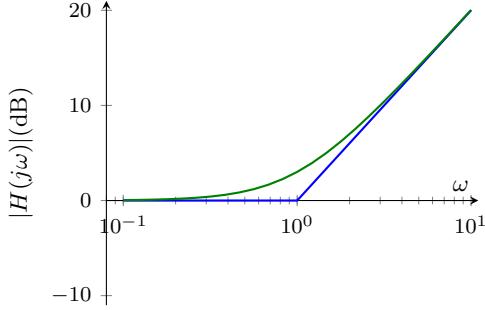


Figure 1.18: Bode plot for zero at $\omega = 1 \text{ rad s}^{-1}$. Asymptotic behavior in blue, exact behavior in green

Single pole at angular frequency p_j

If in our network function we have the term

$$H(s) = \frac{1}{1 + \frac{s}{p_j}}$$

we have a pole at angular frequency p_j . That means that, in the Fourier domain, at angular frequency p_j the imaginary part of the term equals 1.

To draw the Bode plot we first analyze the asymptotic behavior of the function:

$$\lim_{\omega \rightarrow 0} (H(j\omega)) = 1$$

That means that for low frequency the Bode plot is a flat line coincident with the 0 dB axis.

$$\lim_{\omega \rightarrow \infty} (H(j\omega)) = \frac{1}{j\omega} = \frac{p_j}{j\omega}$$

If we compute the magnitude in decibel of the asymptotic function we obtain

$$\left| \frac{p_j}{j\omega} \right| (\text{dB}) = -20 \log_{10}(\omega) + 20 \log_{10}(p_j)$$

That means that for high frequency the Bode plot is a straight line with slope of -20 dB/decade , crossing the 0 dB axis at $\omega = p_j$.

The asymptotic behavior is correct for low and high frequency, but the function will be not coincident with the asymptote for frequency close to p_j . We should then compute the true function:

$$|H(j\omega)|(\text{dB}) = -20 \log_{10} \left(\left| 1 + \frac{j\omega}{p_j} \right| \right) = -20 \log_{10} \left(\sqrt{1 + \frac{\omega^2}{p_j^2}} \right)$$

At angular frequency p_j , the magnitude is

$$|H(jz_i)|(\text{dB}) = -20 \log_{10} \left(\left| 1 + \frac{jp_j}{p_j} \right| \right) = -20 \log_{10} (\sqrt{2}) = -3 \text{ dB}$$

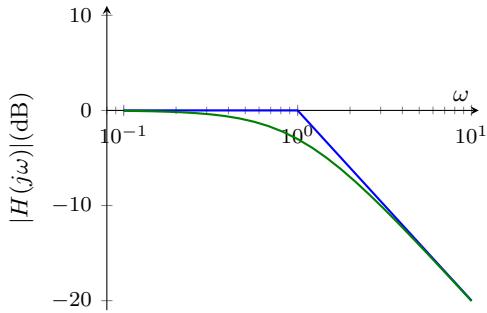


Figure 1.19: Bode plot for pole at $\omega = 1 \text{ rad s}^{-1}$. Asymptotic behavior in blue, exact behavior in green

Figure 1.19 shows both the asymptotic and the exact behavior of the term, assuming $p_j = 1 \text{ rad s}^{-1}$.

Examples with poles and zeros

The frequency response of a complex circuit exhibits several poles and zeros.

As previously noted, it is easy to analyze the frequency response using the Bode plot, because the contribution of every pole or zero is simply summed to obtain the overall behavior. Let us work on two typical functions

Example 1 Let us analyze the following function:

$$H(s) = \frac{1 + 0.1s}{1 + 5 \cdot 10^{-3}s}$$

Before considering zeros and poles, it is always a good idea to find the DC and high frequency behavior.

$$A_{V0} = \lim_{j\omega \rightarrow 0} (H(j\omega)) = 1$$

Therefore the magnitude in DC expressed in decibel is 0 dB and the graph is a flat line.

$$A_{V\infty} = \lim_{j\omega \rightarrow \infty} (H(j\omega)) = 20$$

Again, this is a finite value, so we expect that the transfer function is flat and its value in decibel is 26 dB.

We have one zero and one pole.

The angular frequency of the zero is $\omega_z = 10 \text{ rad s}^{-1}$. We are normally more interested in the frequency which is $f_z = \omega_z/(2\pi) = 1.59 \text{ Hz}$.

The angular frequency of the pole is $\omega_p = 200 \text{ rad s}^{-1}$, the frequency is $f_p = 31.8 \text{ Hz}$.

To create the Bode plot of the magnitude we can draw separately the asymptotic graphs of the zero and the pole and then sum them.

Figure 1.20 shows the resulting Bode plot. Let us note something more on this function: from the DC and high frequency analyses we computed the two values A_{V0} and $A_{V\infty}$. The two flat areas are joined by a line with slope of

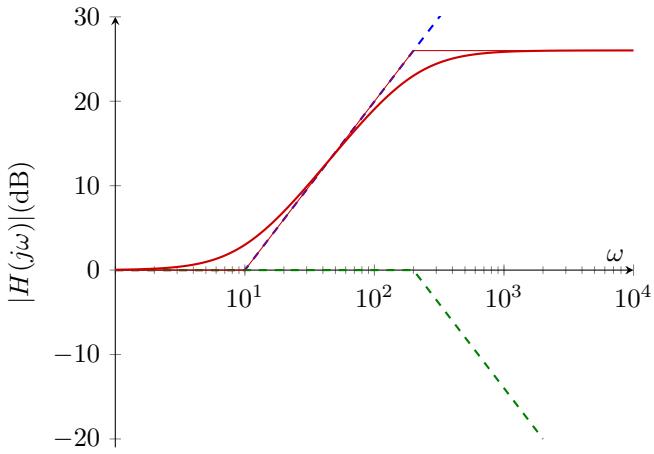


Figure 1.20: Bode plot for the example transfer function. Asymptotic behavior of the zero in blue, of the pole in green. Complete transfer function in red (both asymptotic and exact)

20 dB/decade. Such a slope means that in that region the ratio between the magnitude corresponding to two points equals the ratio of their frequencies. This allows us to write:

$$f_p = f_z \frac{A_{V\infty}}{A_{V0}}$$

In some occasions we will find this relationship very useful: in a circuit it is normally easy to compute the behavior in DC and at very high frequency. There are circuits that exhibit one pole and one zero, where the frequency of one of the two depends only on the value of a single capacitor and resistor, while the other is much more complex to compute. Instead of computing it, we can use the above relationship.

Example 2 Let us now write a transfer function starting from specifications, that is we now want to determine which transfer function is appropriate for a certain circuit.

The circuit we want to build is an audio amplifier. Our amplifier should block the DC component of the input signal, that is $H(0) = 0$, should have a flat response in the audible frequencies range, with a certain gain, and should attenuate higher frequencies.

We define as audible frequencies range the interval from $f_L = 20\text{ Hz}$ to $f_H = 20\text{ kHz}$. We set the gain of the amplifier in this range to $A_V = 10$ or 20 dB.

How do we build this transfer function? Having $H(0) = 0$ means that our transfer function needs a zero in the origin, so we insert the term $H_0(s) = s$. In this case at low frequencies we will have a line with slope of 20 dB/decade. Our function should be flat for frequencies above 20 Hz. That means that we need to insert a pole at 20 Hz, or at $\omega_{p1} = 2\pi \cdot 20\text{ rad s}^{-1} = 125.7\text{ rad s}^{-1}$. So we

introduce a term:

$$H_1(s) = \frac{1}{1 + 7.96 \times 10^{-3}s} = \frac{1}{1 + \frac{s}{2\pi \cdot 20}}$$

If we combine the two terms, we can easily find that our transfer function is now

$$H(s) = H_0(s)H_1(s) = \frac{s}{1 + 7.96 \times 10^{-3}s}$$

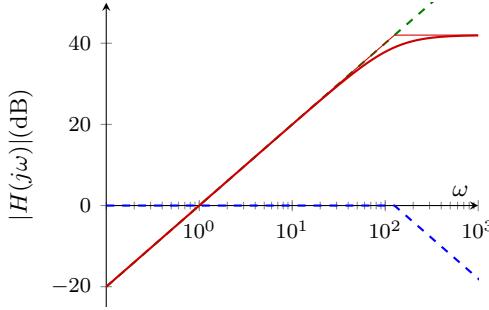


Figure 1.21: Bode plot for the example transfer function. Asymptotic behavior of the pole in blue, of the zero in origin in green. Complete transfer function in red (both asymptotic and exact)

We show the resulting Bode plot in figure 1.21. If we observe the figure we find out that it is still different from our goal, because the gain in the flat region is not 20 dB, but 42 dB. We also note that high frequencies are not attenuated, but that will come in a subsequent step.

Let us set the correct gain. We compute

$$A_{V\infty} = \lim_{j\omega \rightarrow \infty} \left(\frac{j\omega}{1 + 7.96 \times 10^{-3}j\omega} \right) = 125.7$$

Instead, we need $A_{V\infty} = 10$. We can obtain this by introducing a constant $K = 10/125.7 = 7.96 \times 10^{-2}$. Our new transfer function is:

$$H(s) = K \cdot H_0(s)H_1(s) = \frac{7.96 \times 10^{-2}s}{1 + 7.96 \times 10^{-3}s}$$

The last problem is that our transfer function is flat from 20 Hz to infinity, while we want to attenuate the frequencies above 20 kHz. We then need to insert a second pole at 20 kHz, or at $\omega_{p2} = 2\pi \cdot 20 \text{ krad s}^{-1} = 125.7 \times 10^3 \text{ rad s}^{-1}$.

The additional term in our transfer function will be:

$$H_2(s) = \frac{1}{1 + \frac{s}{2\pi \cdot 20 \times 10^3}}$$

The final transfer function is

$$H(s) = K \cdot H_0(s)H_1(s)H_2(s) = \frac{7.96 \times 10^{-2}s}{\left(1 + \frac{s}{2\pi \cdot 20}\right)\left(1 + \frac{s}{2\pi \cdot 20 \times 10^3}\right)}$$

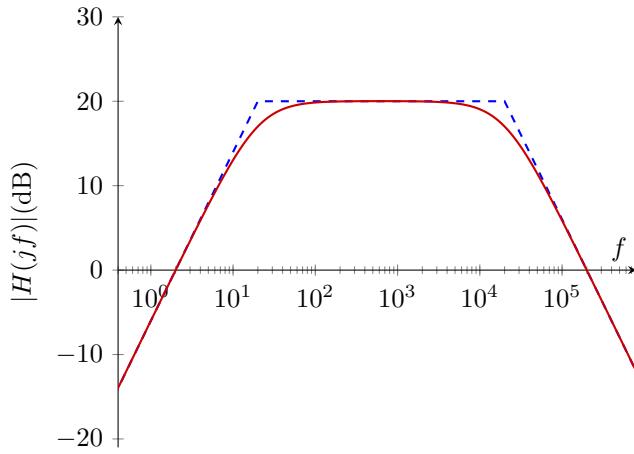


Figure 1.22: Final Bode plot for example 2. The x axis is now scaled in frequency (hertz)

Figure 1.22 shows the Bode plot of the magnitude of our function. In the figure we scaled the x axis in frequency instead of angular frequency, for better readability.

1.2.2 High pass filter

It is often necessary to exclude the DC component from a signal at the input of an amplifier. This is normally done by inserting a capacitor in series with the amplifier input.

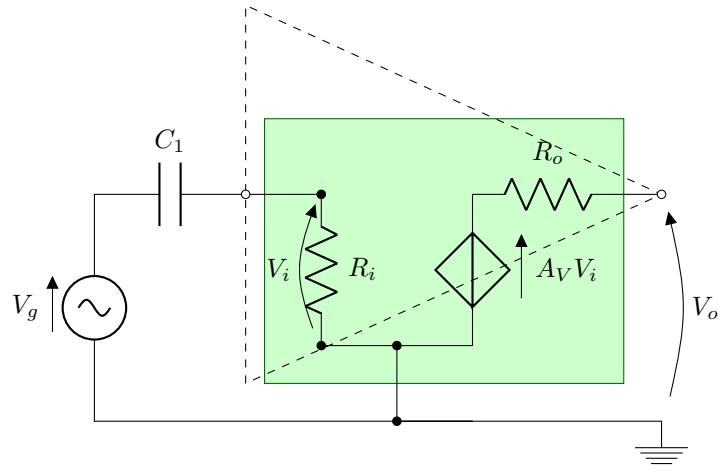


Figure 1.23: High pass filter at the input of an amplifier

The circuit is displayed in figure 1.23. We can analyze the behavior of this circuit either in frequency or time domain.

Frequency domain

In the frequency domain, we can immediately find the asymptotic behavior in DC, considering the capacitor as an open circuit, and at high frequency, considering the capacitor a short circuit. Let us observe the input voltage V_i versus generator voltage V_g in those conditions, that is to compute the gain $A_1 = V_i/V_g$:

- DC: if C_1 is an open circuit, there is no connection between V_g and V_i , so we can say that $V_i = 0$ V. Then we have a zero in DC, or $A_1(\omega \rightarrow 0) = 0$
- High frequency: if C_1 is a short circuit, $V_i = V_g$. We can then say that $A_1(\omega \rightarrow \infty) = 1$.

Even without computing the Bode plot, we know that we have a zero in DC, so the diagram will have a slope of 20 dB/decade for low frequency. Given that the high frequency asymptote is flat, we should then have a pole somewhere. Let us compute the true transfer function between V_g and V_i . We have a voltage divider:

$$V_i = V_g \frac{R_i}{R_i + \frac{1}{sC_1}} = V_g \frac{sR_i C_1}{1 + sR_i C_1}$$

Thus our transfer function is:

$$A_1 = \frac{V_i}{V_g} = \frac{sR_i C_1}{1 + sR_i C_1}$$

We can confirm what we noted in the asymptotic analysis: we have a zero in DC and a pole at the frequency:

$$f_p = \frac{1}{2\pi R_i C_1}$$

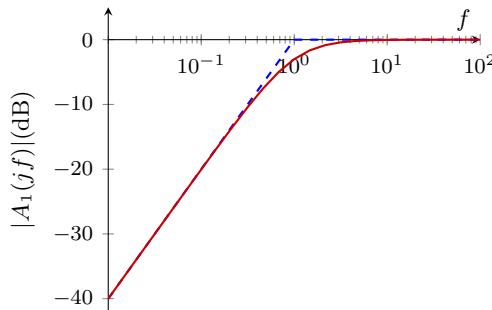


Figure 1.24: Bode plot of high pass filter for $f_p = 1$ Hz.

The Bode plot of the function is drawn in figure 1.24.

Time domain analysis

We perform the time domain analysis by assuming a 1 V step input as V_g . Again, we can start with asymptotic behavior, by calculating the value of V_i immediately after the step (time 0+) and at infinity.

- At time $t = 0+$ we consider the capacitor a short circuit, so the value we expect at the input of the amplifier is $V_{i0} = V_g(0+) = 1\text{ V}$.
- At time $t \rightarrow \infty$ the capacitor is an open circuit, so the voltage at the input of the amplifier is $V_{i\infty} = 0\text{ V}$.

So we easily know the initial and final value of V_i . To perform the exact analysis, we recall from circuit theory that in this case the voltage at V_i follows an exponential equation with time constant $\tau = R_i C_1$.

To write the exponential equation, we know that the general formula is:

$$V_i(t) = (V_{i0} - V_{i\infty})e^{-t/\tau} + V_{i\infty}$$

In our case:

$$V_i(t) = 1\text{ V}e^{-t/R_i C_1}$$

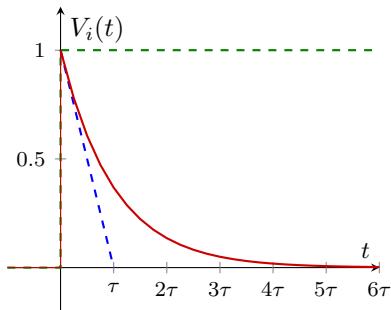


Figure 1.25: V_i time domain step response. V_g in green, V_i in red.

The waveforms are in figure 1.25. By analyzing the time domain formula we can also state that the initial slope of the decaying exponential is $1/\tau = 1/R_i C_1$. After τ seconds the voltage has decayed by 63.2% of the total. After about 5τ seconds the transient is almost finished.

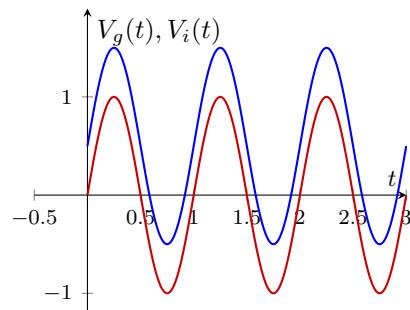


Figure 1.26: V_i time domain response to sine wave. V_g in blue, V_i in red.

If we use as input signal V_g a sine wave with some DC offset and frequency at least one decade higher than f_p , the resulting V_i will be the same wave but with DC offset null (figure 1.26). For lower frequencies, we will have attenuation and phase shift.

1.2.3 Low pass filter

Another common situation in amplifiers is the addition of a low pass filter to limit the bandwidth of the amplifier to that of the intended input signal and to avoid amplifying high frequency noise.

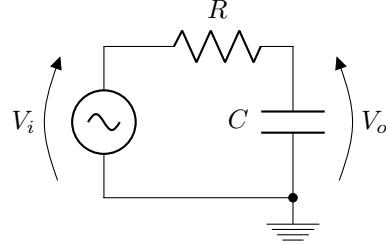


Figure 1.27: Passive low pass filter

Let us study the frequency and time behavior of a passive first order low pass filter. The simple circuit is shown in figure 1.27. We will analyze the behavior of the circuit both in frequency and time domain.

Frequency domain We note that the asymptotic behavior at low frequency is $V_o = V_i$, that is $A_v = 1$, because in DC the capacitor is an open circuit so that no current flows in R . The Bode plot is flat at low frequency and coincides with the 0 dB axis. At high frequency the capacitor is a short circuit. That means that $\lim_{\omega \rightarrow \infty} (V_o(j\omega)) = 0$. The Bode plot goes to minus infinity at high frequency.

Let us now perform the exact analysis in the Laplace domain:

$$A_v = \frac{V_o}{V_i} = \frac{\frac{1}{sC}}{\frac{1}{sC} + R} = \frac{1}{sRC + 1}$$

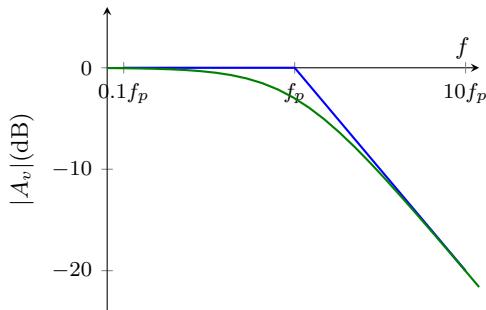


Figure 1.28: Bode plot of low pass filter. Asymptotic behavior in blue, exact behavior in green

The equation shows that we have a single pole at frequency $f_p = 1/(2\pi RC)$. The Bode plot of the magnitude is in figure 1.28.

Time domain For time domain analysis, we note that, given a unity step input on V_i , $V_o(0+) = 0\text{ V}$ because the capacitor acts as a short circuit for very high frequency, and $\lim_{t \rightarrow \infty}(V_o(t)) = 1\text{ V}$. We know that the transient is exponential with time constant $\tau = RC$, so the exact solution is:

$$V_o(t) = 1\text{ V} \left(1 - e^{-t/\tau}\right)$$

Figure 1.29 shows the circuit response to a step input. We note that after τ

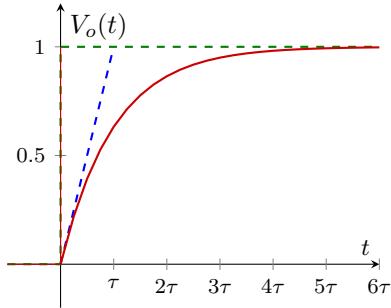


Figure 1.29: V_o time domain step response. V_i in green, V_o in red.

seconds the output voltage has changed by 63.2% of the total. This information can be used to find the time constant in laboratory.

Example Let us introduce a low pass filter at the output of an amplifier and see how the frequency behavior is modified.

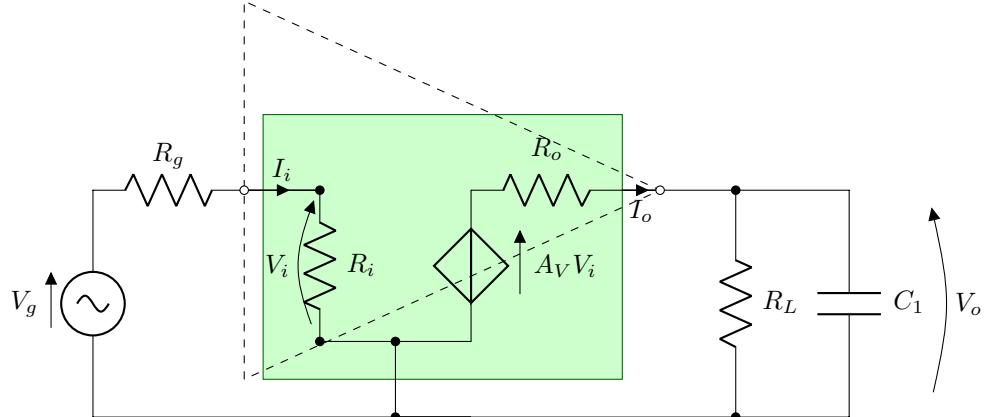


Figure 1.30: Example circuit

Figure 1.30 shows the schematic of the circuit. Let us set the following values: $R_g = 1\text{ k}\Omega$, $R_i = 1\text{ k}\Omega$, $R_o = 100\text{ }\Omega$, $R_L = 2\text{ k}\Omega$, $C_1 = 47\text{ nF}$, $A_V = 100$.

The capacitor shorts the output of the circuit at high frequency, while it is equivalent to open circuit at low frequency.

We want to compute the gain of the amplifier $A = V_o/V_g$. As we did in the previous sections, we can split our calculations in two: first we find $A_i = V_i/V_g$, then $A_o = V_o/V_i$, and finally we concatenate the results.

A_i does not depend on frequency, and we easily find:

$$V_i = \frac{R_i}{R_i + R_g} V_g \implies A_i = \frac{R_i}{R_i + R_g} = 0.5$$

A_o instead depends on frequency. We can find the asymptotic behavior as a first step, then perform the exact analysis. If we consider the capacitor an open circuit, we have:

$$V_o = \frac{R_L}{R_o + R_L} A_V V_i \implies A_o = \frac{R_L}{R_o + R_L} A_V = 95.2$$

As we already noted, the gain at high frequency is zero, so the capacitor inserts a pole in the frequency response. To find the frequency of the pole we first compute the complex impedance Z_L :

$$Z_L = R_L \parallel \frac{1}{sC_1} = \frac{R_L}{sR_L C_1 + 1}$$

The output voltage divider becomes:

$$\frac{V_o}{V_i} = A_V \frac{Z_L}{Z_L + R_o} = A_V \frac{\frac{R_L}{sR_L C_1 + 1}}{R_o + \frac{R_L}{sR_L C_1 + 1}} = A_V \frac{R_L}{(R_L + R_o)} \frac{1}{[1 + s(R_L \parallel R_o)C_1]}$$

The equation shows that, as we expected, the gain at low frequency is the one we computed above and we have a pole at frequency:

$$f_p = \frac{1}{2\pi(R_L \parallel R_o)C_1} = 35.6 \text{ kHz}$$

Finally, by combining the input and output part, we obtain

$$\frac{V_o}{V_g} = \frac{A_i A_o}{1 + \frac{s}{2\pi f_p}} = \frac{47.6}{1 + \frac{s}{2\pi \times 35.6 \text{ kHz}}}$$

To plot the Bode diagram, we need to convert the low frequency gain in dB

$$A_i A_o (\text{dB}) = 20 \log_{10}(47.6) = 33.6 \text{ dB}$$

1.2.4 Band pass filter

In several applications, only a certain range of frequencies should be amplified, while all the others, lower or higher, should be attenuated or ideally canceled. This is done by band pass filtering. We only study first order filters, while a band pass frequency response requires at least a second order polynomial. We will only study band pass filters obtained by chaining a first order high pass filter and a first order low pass filter. This technique has some limitations and band pass filters obtained this way are called wide bandwidth band pass filters.

This kind of band pass filtering does not require more study effort once one knows how to design a low pass filter and a high pass filter. There is only one warning: if the poles of the two filters are not separated by at least one decade of

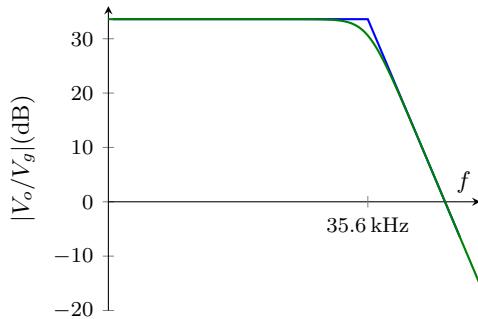


Figure 1.31: Bode plot of example circuit. Asymptotic behavior in blue, exact behavior in green

frequency, that is if $f_{p2} < 10f_{p1}$, then the exact frequency response will be lower than the asymptotic one in the entire bandwidth of the filter and the frequencies at which we have a -3 dB attenuation with respect to the maximum amplitude of the transfer function will be more spaced than the pole frequencies (or in other words, the bandwidth of the filter is larger than the quantity $f_{p2} - f_{p1}$).

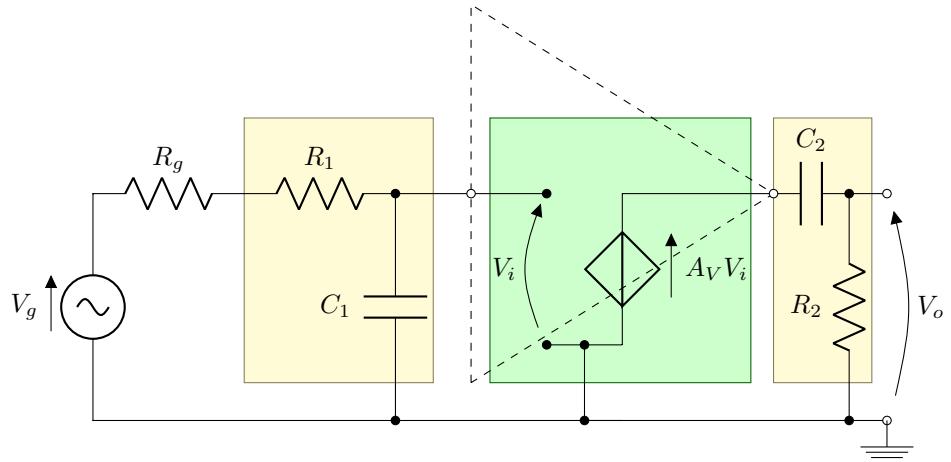


Figure 1.32: Amplifier with band pass filter

Example Figure 1.32 shows an example of amplifier with band pass filtering. With respect to the previous examples, now we swap the position of high pass and low pass filters, to show that both kind of filtering is possible at the input and at the output of the amplifier. In this example, for the sake of simplicity, we neglected the input resistance of the amplifier, assuming $R_i \rightarrow \infty$, and its output resistance, assuming $R_o \rightarrow 0$. Let us use the following values: $R_g = 600 \Omega$, $R_1 = 10 \text{ k}\Omega$, $R_2 = 1.2 \text{ k}\Omega$, $C_1 = 220 \text{ pF}$, $C_2 = 3.3 \mu\text{F}$, $A_V = 15$.

frequency domain As always, we start by evaluating the asymptotic behavior of the circuit. In DC, the gain V_o/V_g is zero because of the presence of C_2

which acts as open circuit. At high frequency, the gain is again zero, because the input of the amplifier is shorted by C_1 , short circuit at high frequency.

Within the bandwidth of the amplifier we can assume that we are above the pole introduced by C_2 and below the one introduced by C_1 . We can therefore say that C_1 is still an open circuit while C_2 is a short circuit. In this region, $V_o/V_i = A_V$.

We can now find the exact solution. As we did in previous examples, let us find V_i/V_g and $V_o/(A_V V_i)$. We then concatenate the two terms to find the complete solution.

$$V_i = V_g \frac{\frac{1}{sC_1}}{\frac{1}{sC_1} + R_g + R_1} = \frac{V_g}{1 + sC_1(R_g + R_1)}$$

$$V_o = A_V V_i \frac{R_2}{\frac{1}{sC_2} + R_2} = A_V V_i \frac{sC_2 R_2}{1 + sC_2 R_2}$$

By combining the two equations, we have:

$$\frac{V_o}{V_g} = A_V \frac{1}{[1 + sC_1(R_g + R_1)]} \cdot \frac{sC_2 R_2}{(1 + sC_2 R_2)}$$

The first frequency dependent term in the equation is a low pass filter, with pole frequency

$$f_{p2} = \frac{1}{2\pi C_1(R_g + R_1)} = 68.2 \text{ kHz}$$

The second term is a high pass filter with a zero in DC and a pole at frequency:

$$f_{p1} = \frac{1}{2\pi C_2 R_2} = 40.2 \text{ Hz}$$

The two poles are sufficiently spaced in frequency, so the asymptotic curve is close enough to the exact one.

Finally, we draw the Bode plot of the transfer function. We know the poles' frequencies, we only need to translate the in-band gain in decibel. $A_V(\text{dB}) = 20 \log_{10}(15) = 23.5 \text{ dB}$. The Bode plot is depicted in figure 1.33.

Time domain This circuit has two time constants, associated with the two capacitors: $\tau_1 = (R_g + R_1)C_1 = 2.2 \mu\text{s}$ and $\tau_2 = R_2C_2 = 3.96 \text{ ms}$. Being τ_1 associated with a low pass RC cell, its effects will be to smooth the rise of the output step, while τ_2 will make the output decay to 0 at infinity.

Figure 1.34 shows the output voltage for a 1 V input step. The graph at left shows the behavior after several milliseconds, where the effect of τ_2 is predominant, while the other zooms what happens in the first few microseconds, where τ_1 is effective.

Another way of performing a time analysis which is easy to implement in a laboratory is to compute the response of the circuit to an input square wave. In figure 1.35, at left we use a 100 kHz square wave with 1 V_p amplitude and we see that the frequency is such that the low pass filtering effect of the amplifier

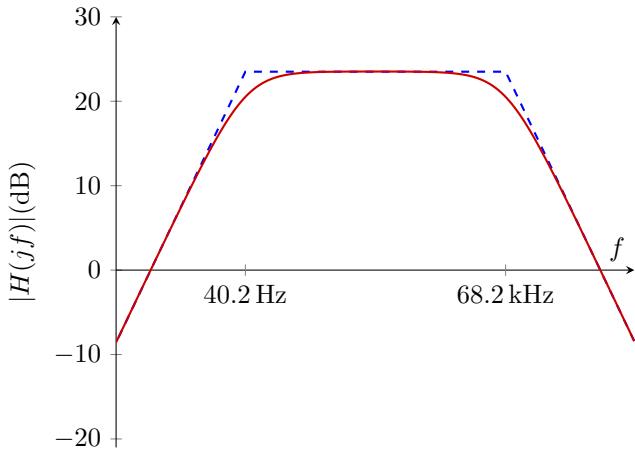


Figure 1.33: Bode plot for band pass circuit.

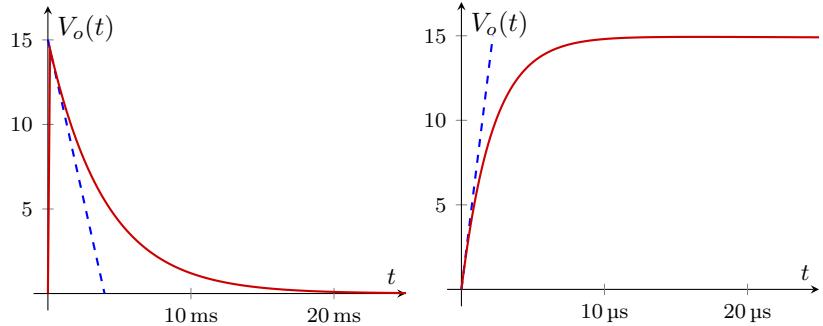


Figure 1.34: V_o time domain step response with two different time scales

is predominant, and the amplifier cannot reach the steady state output before a new transition begins at the input. At right, we changed input to a 40 Hz square wave. At the time scale we are using, the low pass filter effect is not visible, while the high pass filter makes the output come close to 0 V before the new transition is applied. Notice that the peak-to-peak output voltage is close to 60 V because each -1 V to 1 V transition of the square wave is amplified by a factor A_V .

1.2.5 Amplifier with complex impedance

Capacitors (and inductors) can be included in the amplifier model. In this case, the input resistance will become input impedance, output resistance will become output impedance and the gain will be frequency dependent. For a voltage amplifier we have:

- $R_i \rightarrow Z_i(s)$
- $R_o \rightarrow Z_o(s)$
- $A_V \rightarrow A_V(s)$

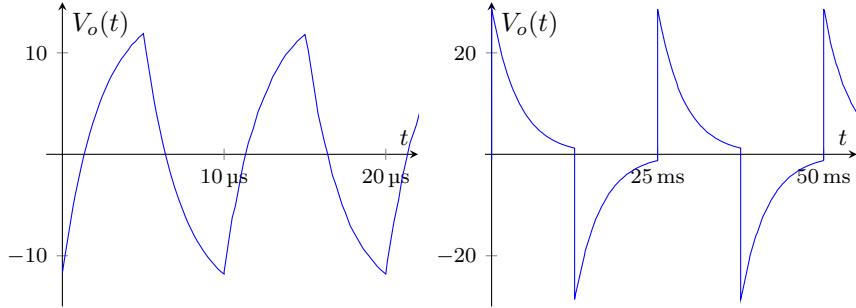


Figure 1.35: V_o time domain step response with two different time scales

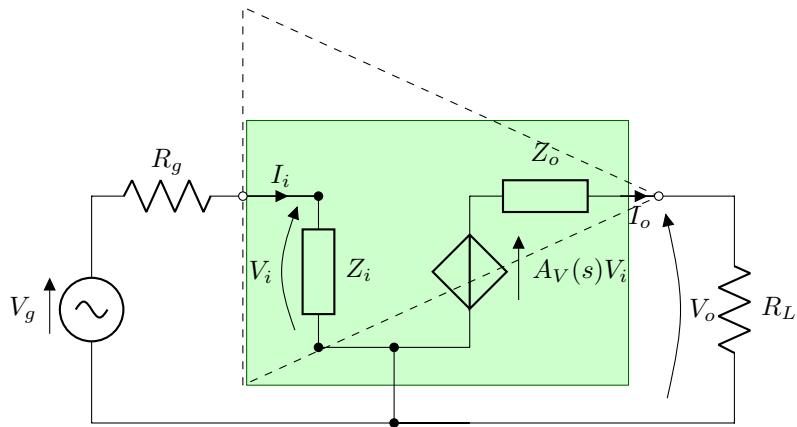


Figure 1.36: Frequency dependent voltage amplifier

Figure 1.36 shows the model of the amplifier. To find gain, input and output impedance of an amplifier, we can use the same techniques we adopted with circuits using only resistors and voltage or current dependent voltage/current sources.

Example

Figure 1.43 shows the circuit of one of previous examples (figure 1.13) with one added capacitor. Let us compute the equivalent amplifier parameters, using the following numeric values: $R_1 = 50 \text{ k}\Omega$, $R_2 = 1.2 \text{ k}\Omega$, $R_3 = 10 \text{ k}\Omega$, $R_4 = 180 \Omega$, $C = 15 \mu\text{F}$, $A_m = 100$.

We start computing the input impedance. Let us analyze the asymptotic behavior of the network. For high frequency, capacitor C is a short circuit and the input impedance is easily found as:

$$Z_i(s \rightarrow \infty) = R_1 \parallel R_2 = 1.17 \text{ k}\Omega$$

At low frequency we can use the same procedure of section 1.1.4, considering the capacitor an open circuit. We connect a test voltage generator V_x at the input port and we compute the corresponding current I_x . We note that the current is the sum of the currents I_1 flowing in R_1 and I_2 flowing in R_2 .

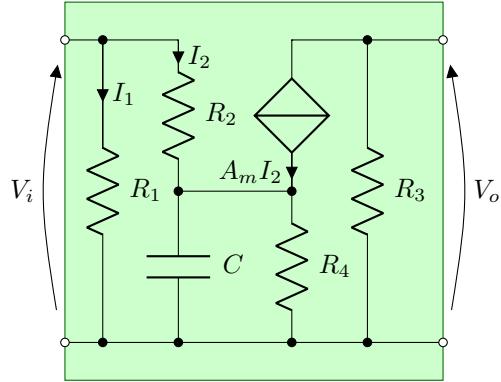


Figure 1.37: Example network with capacitor.

We can find I_2 by KVL:

$$V_x = I_2 R_2 + (1 + A_m) I_2 R_4 = I_2 [R_2 + (1 + A_m) R_4]$$

If we define to define $R_e = R_2 + (1 + A_m) R_4$ we can write the above expression as:

$$\frac{V_x}{I_2} = R_e = 19.4 \text{ k}\Omega$$

$$I_x = I_1 + I_2 = \frac{V_x}{R_1} + \frac{V_x}{R_e} \implies Z_i(s \rightarrow 0) = R_1 \parallel R_e = 14 \text{ k}\Omega$$

To find the impedance at intermediate frequency, we need to take C into account.

It is better to merge the capacitor C and the resistor R_4 in a single impedance Z_4 :

$$Z_4 = \frac{1}{sC + \frac{1}{R_4}} = \frac{R_4}{sR_4 C + 1}$$

We can note that the current that flows in Z_4 is the sum of I_2 and the current in the current dependent current generator, $A_m I_2$. We now write the following KVL equation:

$$V_x = I_2 R_2 + I_2 (1 + A_m) Z_4 = I_2 [R_2 + (1 + A_m) Z_4]$$

If we define $Z_e = R_2 + (1 + A_m) Z_4$, we can see that the input impedance can be written as the parallel of R_1 and Z_e .

$$Z_e = \frac{R_2 + (1 + A_m) R_4 + s R_2 R_4 C}{s R_4 C + 1}$$

The expression of Z_i becomes complex, so it is better to insert the previously defined $R_e = R_2 + (1 + A_m) R_4$. In this way we can rewrite the expression of Z_e :

$$Z_e = \frac{R_e \left(1 + s \frac{R_2 R_4 C}{R_e} \right)}{s R_4 C + 1}$$

Finally we can compute

$$Z_i = R_1 \parallel Z_e = \frac{R_1 R_e \left(1 + s \frac{R_2 R_4 C}{R_e} \right)}{(s R_4 C + 1) \left[R_1 + \frac{R_e \left(1 + s \frac{R_2 R_4 C}{R_e} \right)}{s R_4 C + 1} \right]}$$

After a few maths we can write

$$Z_i = (R_1 \parallel R_e) \frac{1 + s \frac{R_2 R_4 C}{R_e}}{1 + \frac{s(R_1 + R_2) R_4 C}{R_1 + R_e}} = (R_1 \parallel R_e) \frac{1 + \frac{s}{\omega_p}}{1 + \frac{\omega_z}{\omega_p}}$$

Where $\omega_z = R_e / (R_2 R_4 C) = 5988 \text{ rad s}^{-1} = 953 \text{ Hz}$ and $\omega_p = (R_1 + R_e) / [(R_1 + R_2) R_4 C] = 502 \text{ rad s}^{-1} = 80 \text{ Hz}$.

Now we can compute the gain $A_V(s)$. Again, we need to connect a test voltage generator V_x at the input and obtain the corresponding output voltage. We easily see that

$$V_o = -R_3 A_m I_2$$

We already found the relationship between I_2 and V_x :

$$V_x = I_2 Z_e \implies I_2 = \frac{V_x}{Z_e}$$

by substituting in the first equation we obtain

$$V_o = -R_3 A_m \frac{V_x}{Z_e} \implies \frac{V_o}{V_x} = A_V = -\frac{R_3 A_m}{Z_e}$$

We can now expand back Z_e

$$A_V = -\frac{R_3 A_m}{R_e} \frac{1 + s R_4 C}{1 + s \frac{R_2 R_4 C}{R_e}}$$

We can plot the magnitude of the gain by observing that

$$A_V(0) = -\frac{R_3 A_m}{R_e} = -51.5 \implies |A_V|(\text{dB}) = 34.2 \text{ dB}$$

$$A_V(s \rightarrow \infty) = -\frac{R_3 A_m}{R_2} = -833 \implies |A_V|(\text{dB}) = 58.4 \text{ dB}$$

The frequency of the zero is

$$f_z = \frac{1}{2\pi R_4 C} = 59 \text{ Hz}$$

and the frequency of the pole is the same as the frequency of the zero in the input impedance

$$f_p = \frac{R_e}{2\pi R_2 R_4 C} = 953 \text{ Hz}$$

Figure 1.38 shows the resulting Bode plot.

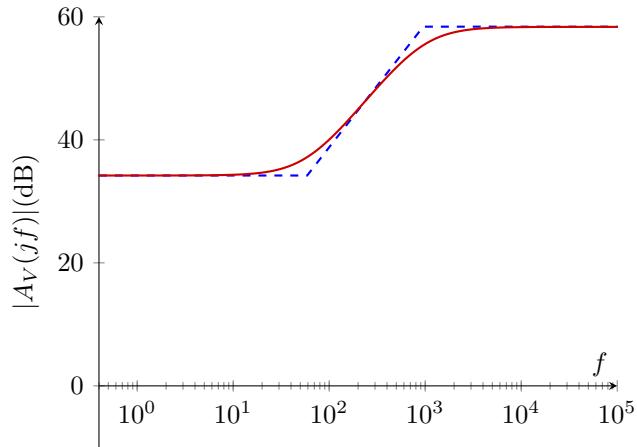


Figure 1.38: Bode plot of A_V

1.2.6 Amplifier bandwidth

The bandwidth of an amplifier is normally defined as the range of frequencies in which the gain of the amplifier is above a threshold which is normally set as -3 dB below its maximum value.

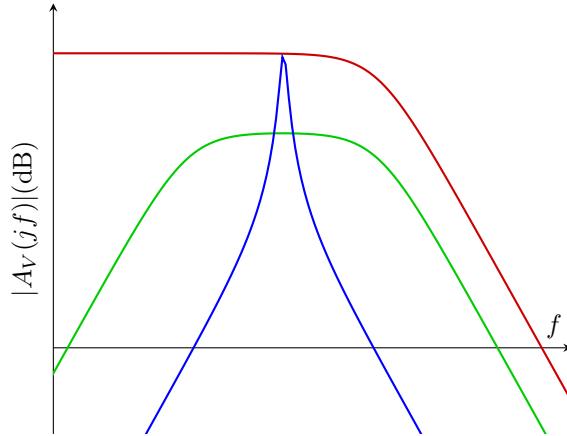


Figure 1.39: DC amplifier (red), AC amplifier (green), narrow band amplifier (blue)

It is important to tailor the bandwidth of the amplifier to the bandwidth of the input signal. Obviously the bandwidth of the amplifier has to be large enough to amplify all the frequency components of the signal, but, if it is too large, it will also amplify unwanted signals mixed to the input signal, such as noise or interference.

In general we can classify the amplifiers in at least three categories, as shown in figure 1.39:

- DC amplifiers: the amplifier amplifies the DC components of the signal. The bandwidth of the amplifier is from 0 to a certain maximum frequency

f_h . This is the normal situation for amplifiers connected to sensors.

- wide bandwidth AC amplifiers: the bandwidth of the amplifier has a lower limit f_l and an higher limit f_h . The two limits are separated by decades of frequency. Examples are voice amplifier ($f_l = 300 \text{ Hz}$, $f_h = 3200 \text{ Hz}$) or hi-fi audio amplifier ($f_l = 20 \text{ Hz}$, $f_h = 20 \text{ kHz}$)
- narrow bandwidth AC amplifiers: the bandwidth is very small compared with the center frequency. If we define the center frequency as $f_c = \sqrt{(f_l \times f_h)}$, we can compute the factor of merit of the amplifier as:

$$Q = \frac{f_c}{f_h - f_l}$$

Amplifiers with high Q are narrow band amplifiers and are used mainly at radio frequency.

1.3 Linearity

For the time being we assumed that an amplifier is a linear circuit. This means that the output signal is proportional to the input signal. The ratio between output and input signal is called gain of the amplifier. If we call K the gain of the system, we have:

$$O(I) = K \times I$$

where O is the output of the amplifier and I is the input.

Linear systems are easy to deal with, because the superposition principle holds:

$$O(A + B) = O(A) + O(B) \implies K \times (A + B) = K \times A + K \times B$$

Where A and B are two different values of the input of the amplifier.

1.3.1 Gain and offset errors

Real systems only approximate these conditions and for a limited range of input signals. The first problem is due to tolerances of the value of the components that build up the amplifier. In this case we can talk about nominal gain and effective gain of the system.

We know what the term K should be, that is its nominal value, but the real value will differ of an amount that we can call ΔK , or *gain error*. It is not normally possible to know the exact value of ΔK , but only its maximum. That means that the effective gain of our amplifier will be somewhere in the range $[K - \Delta K_M, K + \Delta K_M]$, where we indicate as ΔK_M the maximum value of ΔK .

The parameters of a linear amplifier descent from the mathematical parameters of a straight line. If the slope of the line is called gain when we deal with linear amplifiers, should we bother with the intercept too? The answer is yes. So far we assumed that the output of an amplifier can be found by multiplying the input signal by a gain factor. This means that the output of the amplifier when the input is null should be null. This is not always true, because some amplifier adds some DC value to the output, that is:

$$O(I) = K \times I + O_{\text{off}}$$

where O_{off} is a constant, not input dependent. We define O_{off} as the *offset* of the amplifier. In some cases the offset is a wanted quantity, but in most cases it derives from asymmetries in the internal structure of the amplifier. Again, the offset of the real amplifier can only be estimated statistically as a deviation from the nominal one.

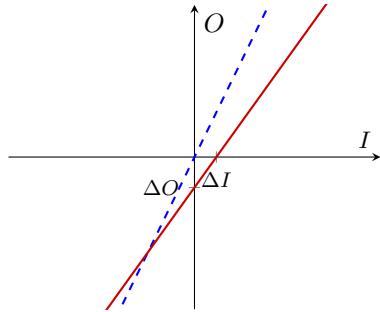


Figure 1.40: Nominal and real amplifier transfer function

Figure 1.40 shows the nominal input/output transfer function of an hypothetical amplifier compared to that of a real device.

The DC transfer function of an amplifier can be expressed in terms of gain error and offset error:

$$O(I) = (K + \Delta K) \times I + \Delta O$$

In this case we say that we refer the offset error to the output: ΔO is the value that O assumes when $I = 0$. In some cases it will be better to refer the offset to the input, that is to consider as error the value that we have to add to the input to obtain $O = 0$. We can then rewrite the function as:

$$O(I) = (K + \Delta K) \times (I - \Delta I)$$

Obviously we can compute the output referred offset from the input referred offset:

$$\Delta O = -\Delta I \times (K + \Delta K)$$

1.3.2 Non linear amplifiers

Some amplifiers are inherently non linear because they are based on devices which have a non linear transfer function.

An example of non linear function (in this case logarithmic) is in figure 1.41. Non linear amplifiers are more difficult to use, because for example the superposition effect does not hold. If we have $O = g(I)$ with $g(x)$ a non linear function:

$$O(A + B) \neq O(A) + O(B) \implies g(A + B) \neq g(A) + g(B)$$

For example, if $g(x) = x^2$ we have

$$(A + B)^2 \neq A^2 + B^2$$

Another problem of non linear amplifier resides in distortion: the shape of the output waveform is different from the input one. This is acceptable in some cases,

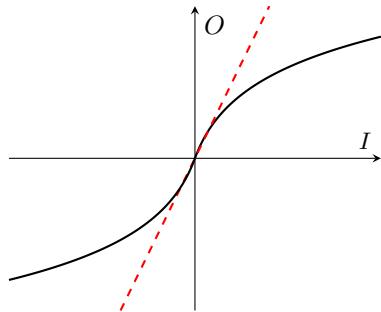


Figure 1.41: Example of nonlinear transfer function

for example many radio frequency amplifiers are nonlinear but the distortion of the output waveform can be recovered by means of simple filtering. In other cases distortion is a big problem and non linearity has to be avoided or corrected, for example in high fidelity audio applications.

If we input a periodic signal like a sine wave $v_i = A \sin(2\pi f_0 t)$ to a non linear amplifier, the output will be a periodic signal but not a pure sine wave. In terms of frequency, the power of the input sine wave is concentrated at the frequency f_0 , while in the output signal we will have contributions also at the harmonics of f_0 , that is at $2f_0, 3f_0, \dots$. The linearity of the amplifier can be measured by the ratio between the power of the harmonics and the power of the fundamental frequency. This measure is called THD or Total Harmonic Distortion.

Even if the global transfer function is nonlinear, in most cases it is possible to approximate a segment of the function with a linear function, like the dashed line in figure 1.41, which is a valid approximation in the range when it is very close to the original function. We will go into deeper details when dealing with transistors and the so called *small signal analysis*.

1.3.3 Saturation

Each amplifier can work linearly only in a well defined input/output dynamic range. In most cases, the output of the amplifier cannot be larger than the power supply voltage. When the signal exits the linearity range, it cannot grow anymore and we identify this condition as *saturation*.

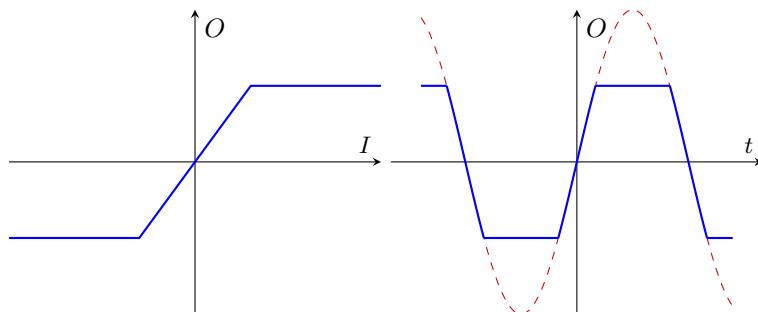


Figure 1.42: Example of saturation: transfer function and effects on a large sine wave

Figure 1.42 shows a transfer function with saturation and the effect of applying a sine wave exceeding the saturation limits.

We name *Output dynamic range* the range of voltages that the amplifier can provide exploiting the nominal transfer function (e.g. without saturation). As noted above, this range is normally a subset of the range defined by the power supply voltages.

We also have an *Input dynamic range* which is the range of voltages accepted by the amplifier without creating abnormal current flow or other conditions that can damage the device. Normally, the range of input voltages tolerated by an amplifier is larger than the range that guarantees operation without distortion or saturation.

1.4 Recap questions

- An amplifier has a power gain of 8. What is its value in dB?
- A voltage amplifier has $A_v = 20$. What is the value of A_v (dB)?
- Draw a table of the best values of the parameters input resistance (R_i) and output resistance (R_o) for the four types of amplifiers you know. Use 0 or ∞ as appropriate.
- What is the procedure to compute the input impedance of a circuit that we want to model as an amplifier?
- Draw the Bode plot of a transfer function with:
 - a pole in DC and a zero at frequency f_z .
 - a zero in DC and a pole at frequency f_p .
- a passive high pass filter has a time constant $\tau = 1$ ms. Draw in the time domain the output voltage as response to a 1 V input step. The time axis should extend from 0 to 6 ms.
- can you create a band pass filter by just concatenating a passive low pass filter and a passive high pass filter? If yes, find zeros and poles of the transfer function of the circuit.
- explain the concept of saturation of an amplifier

1.5 Exercises

Exercise 1.1. Voltage Amplifier

In the real voltage amplifier of figure 1.5, we have the following parameters:

$$R_g = 1.2 \text{ k}\Omega \quad R_i = 10 \text{ k}\Omega \quad R_o = 100 \Omega \quad R_L = 1 \text{ k}\Omega \quad A_V = 100$$

1. Compute the voltage gain V_o/V_g , in number and in dB
2. Compute the power gain P_o/P_g , in number and in dB

Exercise 1.2. Current Amplifier

In the real current amplifier of figure 1.8, we have the following parameters:

$$R_g = 50 \text{ k}\Omega \quad R_i = 1 \text{ k}\Omega \quad R_o = 100 \text{ k}\Omega \quad R_L = 1 \text{ k}\Omega \quad A_i = 20$$

1. Compute the current gain I_o/I_g , in number and in dB

Exercise 1.3. Transconductance Amplifier

In the transconductance amplifier of figure 1.9, we have the following parameters:

$$R_g = 5 \text{ k}\Omega \quad R_i = 10 \text{ k}\Omega \quad R_o = 100 \text{ k}\Omega \quad R_L = 2 \text{ k}\Omega \quad G_m = 50 \text{ mA V}^{-1}$$

1. Compute the transconductance I_L/V_g

Exercise 1.4. Voltage Amplifier parameters

Model the circuit of figure 1.43 as a voltage amplifier, considering the following parameters:

$$R_1 = 1 \text{ k}\Omega \quad R_b = 10 \text{ k}\Omega \quad R_3 = 20 \text{ k}\Omega \quad R_e = 1 \text{ k}\Omega \quad G_m = 10 \text{ mA V}^{-1}$$

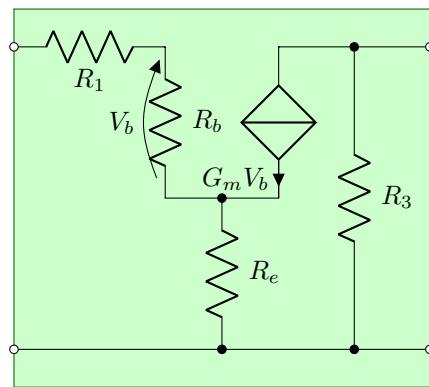


Figure 1.43: Circuit to be modeled as amplifier.

Calculate:

1. Input resistance R_i
2. Gain A_V
3. Output resistance R_o

Exercise 1.5. Low pass filter

The circuit of figure 1.44 has the following parameters:

$$\begin{aligned} R_g &= 5 \text{ k}\Omega & R_i &= 10 \text{ k}\Omega & R_o &= 100 \Omega & R_L &= 2 \text{ k}\Omega \\ A_V &= 10 & C_1 &= 100 \text{ pF} \end{aligned}$$

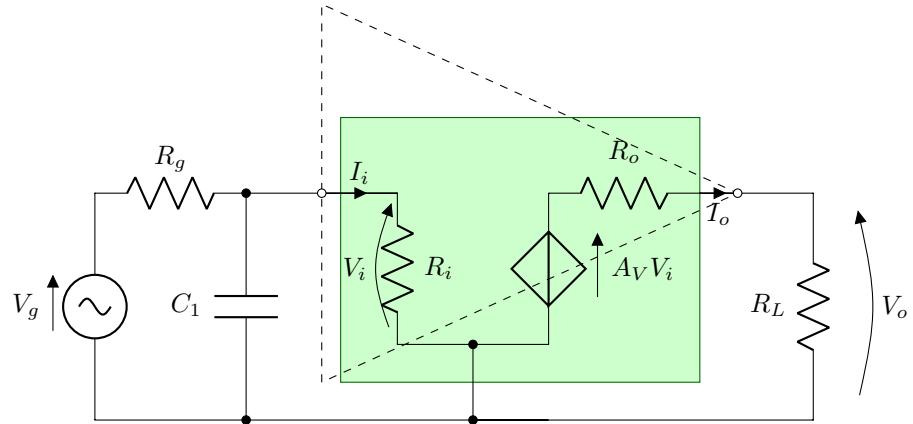


Figure 1.44: Low pass filter

Compute:

- Asymptotic behavior in DC and high frequency of V_o/V_g
- Transfer function $V_o(s)/V_g(s)$
- Bode diagram of $V_o(s)/V_g(s)$

Exercise 1.6. High pass filter

The circuit of figure 1.45 has the following parameters:

$$\begin{aligned} R_g &= 10 \text{ k}\Omega & R_i &= 100 \text{ k}\Omega & R_o &= 100 \Omega & R_L &= 5 \text{ k}\Omega \\ A_V &= 50 & C_1 &= 100 \text{ nF} \end{aligned}$$

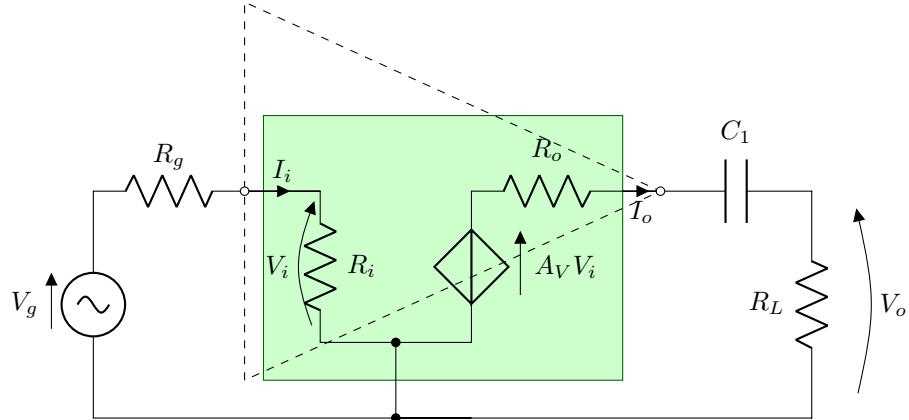


Figure 1.45: High pass filter

Compute:

- Asymptotic behavior in DC and high frequency of V_o/V_g
- Transfer function $V_o(s)/V_g(s)$
- Bode diagram of $V_o(s)/V_g(s)$

Chapter 2

The Operational Amplifier

Contents

2.1	Introductory Analysis	2-1
2.1.1	Ideal operational amplifiers	2-1
2.1.2	Non ideal operational amplifiers	2-4
2.1.3	Voltage follower	2-8
2.1.4	Transresistance Amplifier	2-9
2.1.5	Current Amplifier	2-9
2.1.6	Transconductance Amplifier	2-10
2.1.7	Inverting Amplifier	2-11
2.2	Differential and common mode signals and gain .	2-12
2.3	Real Operational Amplifier	2-16
2.3.1	Current and voltage offsets	2-16
2.3.2	Common mode input range	2-17
2.3.3	Output voltage range	2-17
2.3.4	Input Impedance	2-17
2.3.5	Differential gain	2-18
2.3.6	Common mode gain	2-18
2.4	Sizing the components of an amplifier	2-19
2.4.1	Specifications	2-19
2.4.2	Design	2-19
2.5	Frequency behavior	2-25
2.5.1	Transfer function	2-26
2.5.2	Loop gain	2-26
2.5.3	Gain-Bandwidth product	2-27
2.5.4	Slew Rate	2-29
2.6	Exercises	2-32

2.1 Introductory Analysis

2.1.1 Ideal operational amplifiers

THE simplest approach to the analysis of operational amplifiers based circuits is to use a “black box” model which hides the internal implementation of the device itself.

The most common graphical representation of an operational amplifier is a "triangle" with two input pins, two power supply pins (often omitted in circuits) and one output pin; input pins, characterized by symbols "+" and "-" (named also 'non inverting input' and 'inverting input'), are the input for the signals which the amplifier will amplify; power supply pins power and bias the internal amplifier circuits to the correct operating point.

The basic equations of an "ideal" operational amplifier are:

$$\begin{cases} i_+ = i_- \rightarrow 0 \\ v_d = v_+ - v_- \rightarrow 0 \end{cases}$$

These equations are the very basis for the analysis of any circuit containing one or more operational amplifiers functioning in linearity. Since the gain of the operational amplifier is close to infinity it is easy to understand that if the output is limited, that is if the product of the input voltage v_d (voltage between + and - pins) and the differential gain A_d is finite, then we must have $v_d \rightarrow 0$. Therefore in the ideal operational amplifier operating in linearity the voltage difference between the input pins is assumed to be negligible and therefore the input current is also almost zero provided that the differential resistance between the pins is finite. In real cases this resistance is large and, to simplify the analysis, we can assume the differential resistance $r_d \rightarrow \infty$.

In summary the main characteristics of the **ideal** operational amplifier are:

- Differential gain close to infinity;
- Differential input resistance close to infinity (Simplifies the analysis);
- Output resistance close to 0;
- Differential input voltage close to 0;
- Input currents close to 0.

Example 1. Let us now consider the example circuit of figure 2.1 and analyze it by using the just defined characteristics of the ideal amplifier.

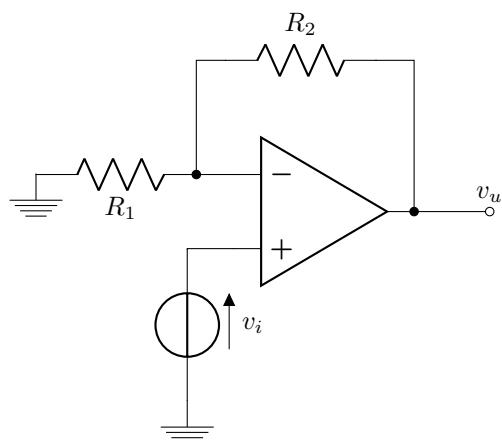


Figure 2.1: Non Inverting amplifiers

This circuit is a non-inverting amplifier, that is it amplifies a signal without modifying its phase. Its gain is defined as the ratio between the output voltage v_u and the input voltage v_i . Considering that the current i_- in the inverting pin is null we can easily write:

$$v_- = v_u \cdot \frac{R_1}{R_1 + R_2}$$

Since $v_+ = v_- = v_i$ we have:

$$\frac{v_u}{v_i} = \frac{R_1 + R_2}{R_1} = \left(1 + \frac{R_2}{R_1}\right)$$

Let us now make a few further comments about the example which has been just presented:

- In the first part of these notes the operational amplifier will be used in a feedback loop. The negative feedback will modify the input and output impedances, the pass band and other parameters.
- When the feedback connection is to the "-" pin of operational amplifier, then it is *negative* because the signal which is fed back is opposed to the input signal and reduces the amplifier input. When feedback is connected to the non inverting input it is *positive*.
- In automatic control theory negative feedback systems are often modeled with an amplifier block A and a feedback block β as shown in figure 2.2. When considering operational amplifier circuits it is often simple to distinguish block A from block β ; block β is the circuit (a passive network in our example) which "feeds" a fraction of the output signal "back" to the input. In the example the signal which is fed back to input is:

$$v_u \cdot \frac{R_1}{R_1 + R_2} = v_- = v_+$$

We can than say that:

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

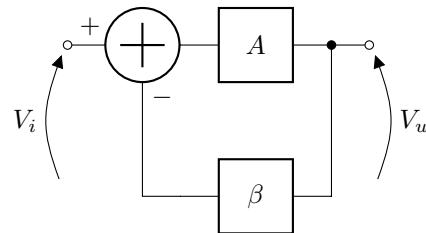


Figure 2.2: Negative feedback system

2.1.2 Non ideal operational amplifiers

We have seen that we can implement a non inverting amplifier with a known gain by setting the ratio of the resistors of the feedback block to a given value. The question is whether the absolute resistor value is important. In other words is it the same to use resistors in the 1Ω e 9Ω range or in the $1 M\Omega$ e $9 M\Omega$ range or even in the $1 m\Omega$ e $9 m\Omega$ range?

Obviously the answer is negative because real operational amplifiers have characteristics which are different from the ideal ones and such differences may set limits on the values of external components. We will see, analyzing the internal implementation of the amplifier, why not all resistor values may be acceptable. The major differences from an ideal amplifier are:

- Non infinite gain A_d ;
- Non infinite input resistance r_d and non zero output resistance;
- Non zero input currents;
- non zero differential input voltage;
- Limited input voltage range and output voltage and current ranges.

Let us proceed step by step using models increasingly more complete with respect to the ideal one. It is important to underline that we will not analyze here the sources of the non idealities but just their effects on the overall circuit.

Limited differential gain

As a first improvement to our model let us consider the fact that the differential gain A_d is $< \infty$. Having finite A_d implies that for a given, non zero, output value we must have $v_d \neq 0$. The new model of the device will then be that of figure 2.3.

We have:

$$v_- = v_i - v_d = v_u \cdot \frac{R_1}{R_1 + R_2} = v_u \cdot \beta$$

and the input voltage is given by:

$$v_d = \frac{v_u}{A_d}$$

Combining the two equations we obtain:

$$v_i - \frac{v_u}{A_d} = v_u \cdot \beta \longrightarrow v_u \left(\beta + \frac{1}{A_d} \right) = v_i$$

and therefore:

$$v_u \cdot \frac{\beta A_d + 1}{A_d} = v_i \longrightarrow \frac{v_u}{v_i} = \frac{A_d}{1 + \beta A_d} = \frac{1}{\beta} \cdot \frac{\beta A_d}{1 + \beta A_d} = \frac{1}{\beta} \cdot \frac{T}{1 + T}$$

In feedback circuit theory, $T \triangleq \beta A_d$ is referred to as the **loop gain**.

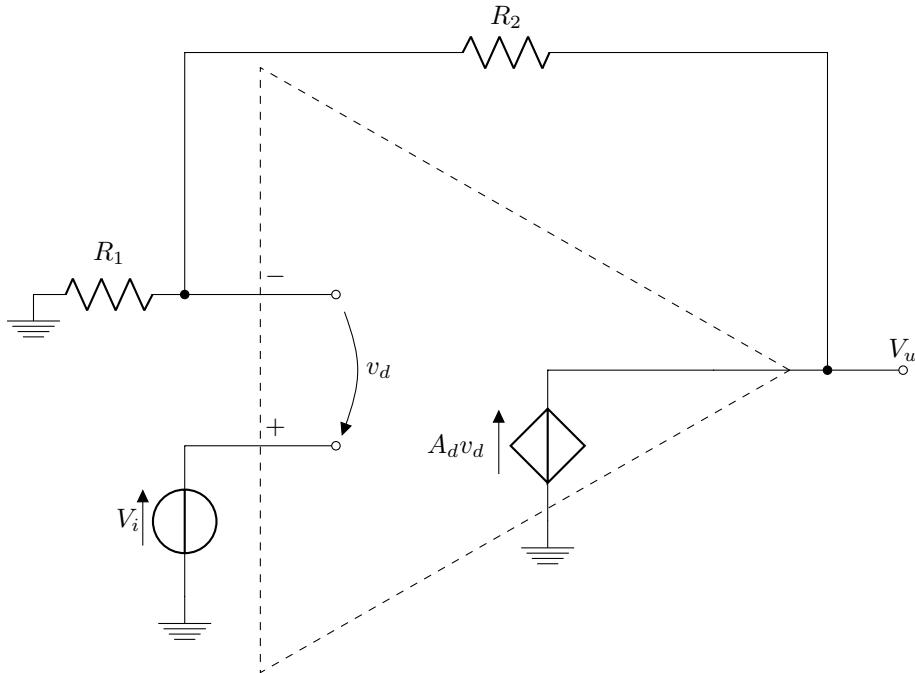


Figure 2.3: First circuit model of the non ideal operational amplifier: $A_d < \infty$.

From this model we can derive that in real cases the difference with respect to the ideal case is very small: to reach a 50 % difference the loop gain T should be 1. Even worst commercial operational amplifiers have a differential gain A_d larger than 10000; if we want to design an amplifier with a very large gain, then β could be in the range of 1/1000, but real circuits have seldom such a high gain because there would be problems related to the frequency behavior of the amplifier. Nevertheless, even in such extreme conditions we would have:

$$T \simeq \frac{10000}{1000} = 10$$

meaning a difference with the ideal case of only 10% between the real and the ideal gain of the circuit, which is normally acceptable.

Input impedance

We can now refine further our analysis by considering non ideal operational amplifier impedances. Let us start by assuming a non infinite r_d input differential resistance and a zero output impedance such as in the model of figure 2.4. We will use it to evaluate the input impedance z_i and the gain A of the amplifier.

In order to compute z_i let us evaluate the current I_x provided by a test generator V_x connected to the amplifier input.

$$V_x = I_x \cdot r_d + R_1 \cdot (I_u + I_x)$$

$$v_d = r_d \cdot I_x; \quad v_u = A_d v_d = A_d r_d I_x$$

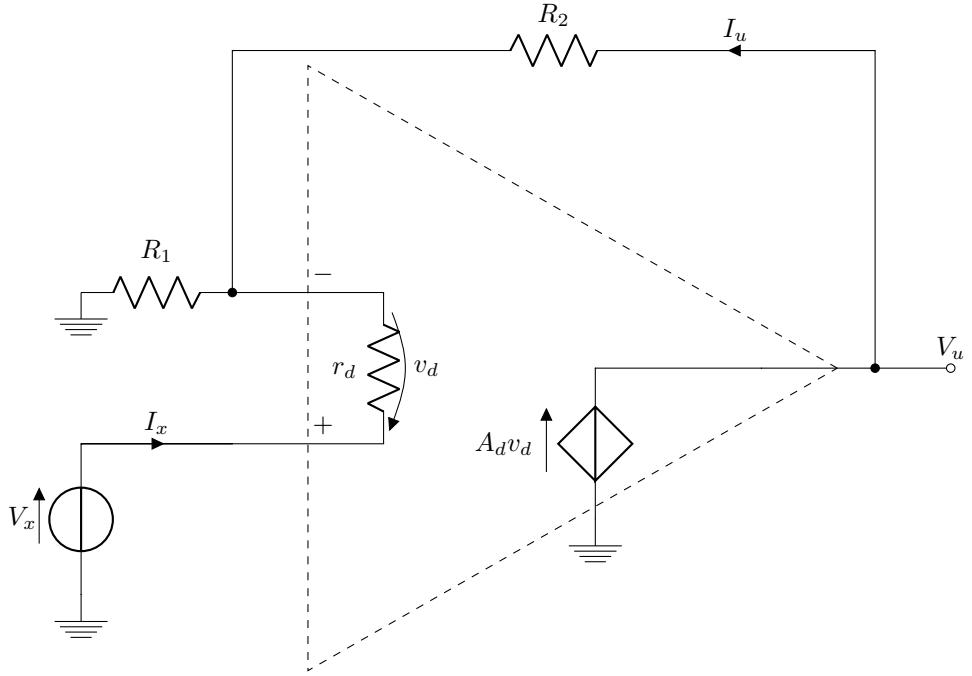


Figure 2.4: Second circuit model of the operational amplifier: finite input resistance.

Moreover we may write:

$$v_u = R_2 I_u + R_1 (I_u + I_x) \longrightarrow A_d r_d I_x = R_2 I_u + R_1 (I_u + I_x)$$

Grouping the term with I_u we get:

$$I_u (R_1 + R_2) = A_d r_d I_x - R_1 I_x \longrightarrow I_u = \frac{A_d r_d I_x - R_1 I_x}{R_1 + R_2}$$

Substituting the I_u value in the V_x expression we have:

$$V_x = I_x r_d + R_1 I_x + \left(\frac{A_d r_d I_x + R_2 I_x}{R_1 + R_2} \right) R_1$$

and separating the different terms we obtain:

$$V_x = I_x r_d + \frac{R_1}{R_1 + R_2} A_d r_d I_x + \frac{R_1 R_2}{R_1 + R_2} I_x$$

Remembering that $\beta = \frac{R_1}{R_1 + R_2}$ we finally obtain:

$$z_i = \frac{V_x}{I_x} = r_d (1 + \beta A_d) + R_1 \parallel R_2$$

Since A_d is very large the second term of the equation may often be considered negligible with respect to the first one. At any case considering only the first term is a "worst case" approximation because it leads to a value of z_i lower

than the real one. We can use these equations to re-evaluate the gain of the amplifier.

$$v_u = A_d v_d = A_d r_d I_x = A_d r_d I_x = A_d r_d \frac{V_x}{z_i}$$

from which

$$\frac{v_u}{V_x} = \frac{A_d r_d}{z_i} = \frac{A_d r_d}{r_d(1 + \beta A_d) + R_1 \parallel R_2}$$

We can immediately see that if $r_d(1 + \beta A_d) \gg R_1 \parallel R_2$ the gain of the amplifier is equal to the value computed in the previous case. Therefore this more refined model does not change the already obtained results and just gives us a more realistic value of the input impedance which with previous models was infinite.

Output impedance

Our amplifier is still ideal from the output point of view because it has a zero output impedance. In order to have a complete model we must consider the effects of a non null output impedance of the operational amplifier. Let us then consider the model of figure 2.5.

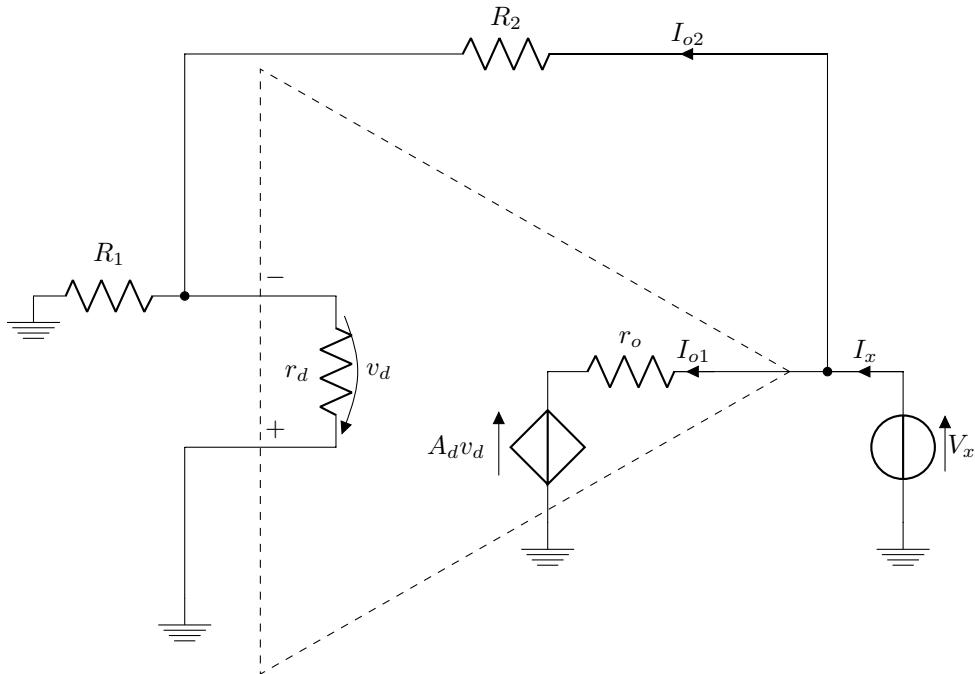


Figure 2.5: Operational amplifier model with non zero output impedance

We will use the model to evaluate the output impedance Z_o of the feedback amplifier by setting the output to voltage V_o by means of a voltage generator and by evaluating the current supplied by it. All other independent generators must be zeroed. Current I_o is composed of two terms: I_{o1} entering into the branch containing the dependent voltage source $A_d v_d$ and I_{o2} which enters R_2 and whose value is approximately given by $V_o/(R_1 + R_2)$; it is possible to simplify the analysis by assuming $I_{o2} \ll I_{o1}$. We will justify this hypothesis a posteriori but

neglecting I_{o2} will lead to an over evaluation of Z_o and is therefore a conservative assumption because we want the value of Z_o to be low. We therefore have:

$$I_o \simeq I_{o1} = \frac{V_o - A_d v_d}{r_o}$$

v_d may be expressed (considering r_d very large) as:

$$v_d = -\beta V_x = -\frac{R_1}{R_1 + R_2} V_o$$

Hence we can say:

$$I_o \simeq \frac{V_o + A_d \beta V_o}{r_o}$$

This value is very large with respect to I_{o2} for any reasonable value of r_o .

$$\frac{I_o}{V_o} \simeq \frac{1 + \beta A_d}{r_o}, \quad \beta A_d = T$$

Hence

$$Z_o = \frac{V_o}{I_o} \simeq \frac{r_o}{1 + T}$$

If our operational amplifier has an output resistance of 100Ω , higher than that of most commercial ones, and the loop gain is 1000, the feedback amplifier output impedance would be reduced by three orders of magnitude to $100 \text{ m}\Omega$.

Therefore our circuit is a good voltage amplifier with a gain value which depends only on the feedback resistances, a high input impedance and a low output impedance. In general the first, very simple model can be used for evaluating the gain while more complete models allows to evaluate the value of input and output impedances.

2.1.3 Voltage follower

A particular case of the non inverting amplifier is shown in figure 2.6.

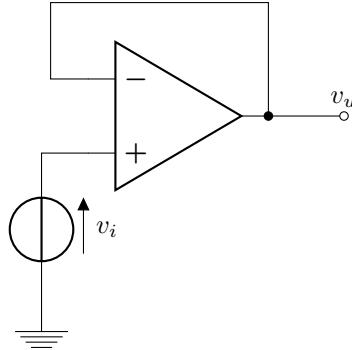


Figure 2.6: Voltage follower.

In this topology the output signal is fully fed back to the operational amplifier input with $\beta = 1$; as a consequence the amplifier gain is 1 and the loop gain

T is equal to A_d maximizing the input impedance and minimizing the output impedance. This circuit acts as an "impedance separator" with unity gain, very high input impedance and very low out impedance (ideal voltage generator). We will further analyze this circuit when considering the frequency behavior of amplifiers.

2.1.4 Transresistance Amplifier

Another circuit based on the operational amplifier is shown in figure 2.7.

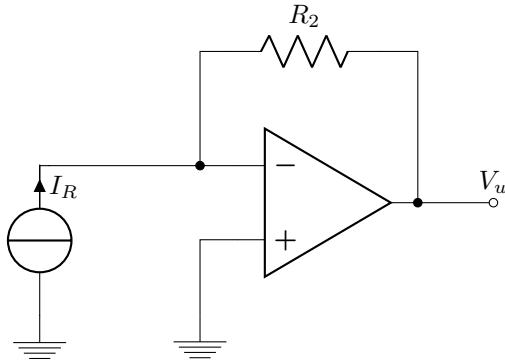


Figure 2.7: Transresistance Amplifier.

The input signal is a current while the output signal is a voltage. The dimension of the transfer function, which a ratio between a voltage and a current, is a resistance and therefore this circuit topology is referred to as *transresistance amplifier*. Since the operational amplifier input current is zero, all the signal current will enter R_2 , and the voltage V_R across the resistance will be:

$$V_R = I_R R_2$$

Since $v_- = 0$ then $V_R = -V_u$ and therefore:

$$V_u = -I_R R_2$$

In other words this circuit “transforms” the current into an output voltage proportional to R_2 and by analyzing it we may easily verify that both input and output impedance are very low.

2.1.5 Current Amplifier

Figure 2.8 shows an amplifier where the input is again a current generator I_s and whose output is the current I_L in load resistance R_L . Assuming an ideal operational amplifier the non inverting pin is at ground voltage and since $V_d \approx 0$ the inverting pin will be at virtual ground. Since no current enters the operational amplifier, current I_s will flow in resistance R_2 . The voltage of node A will be $V_A = -R_2 I_s$ and the current I_{R1} in resistance R_1 will be given by:

$$I_{R1} = \frac{V_A}{R_1} = -I_s \frac{R_2}{R_1}$$

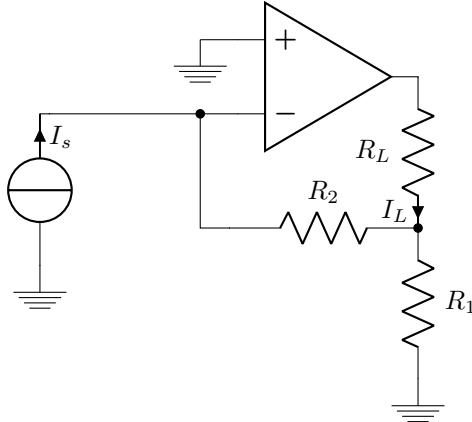


Figure 2.8: Current Amplifier

From the current equation at node A we have $I_L + I_s = I_{R1}$ and we may easily obtain:

$$I_L = -I_s \left(1 + \frac{R_2}{R_1} \right)$$

The fact that current I_L does not depend on load resistance R_L shows that the circuit behaves at its output as a current generator whose value is proportional to the input current. It is therefore a *current amplifier* with high output impedance and with gain:

$$A_I = - \left(1 + \frac{R_2}{R_1} \right)$$

Since the input voltage $v_i = -v_d$ is close to 0 the input impedance $Z_i = \frac{v_i}{I_s}$ will be very small, ideally infinitesimal. For evaluating more precisely Z_i it is necessary to use a more refined model of the O.A. taking into account the finite gain of the amplifier and this is left as an exercise for the reader. Note that this analysis of the input impedance applies as well to the transresistance amplifier of the previous section.

2.1.6 Transconductance Amplifier

The fourth type of amplifier we are going to consider is the one shown in figure 2.9. whose input is a voltage V_s and whose output is the current I_L in load resistance R_L .

The analysis is very simple. Since $v_d \approx 0$ then $v_A = v_s$ and the current I_{R1} in resistance R_1 will be given by $\frac{v_A}{R_1}$ and since no current enters the non inverting pin $I_L = I_{R1}$. We have then

$$I_L = v_s \frac{1}{R_1}$$

where the proportionality factor between the input and the output variable has the dimension of a conductance.

In this case both the input and the output impedances will be very high.

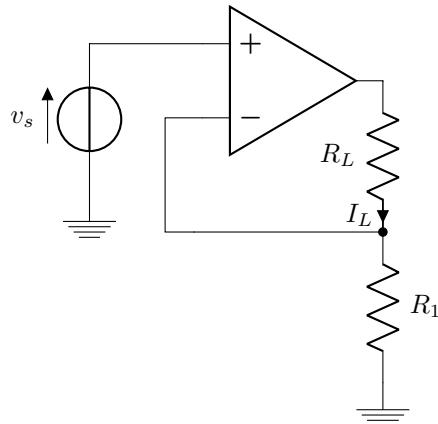


Figure 2.9: Transconductance Amplifier

2.1.7 Inverting Amplifier

This circuit is, together with the non inverting amplifier, one of the most common circuit topologies for the linear use of the operational amplifier, and, as we will see later, it is also the basis for a number of other linear circuits.

If we start from the transresistance amplifier and we substitute the current generator with a voltage generator with a series resistor R_1 , as shown in figure 2.10, we obtain a circuit in which the input voltage is first converted into a current which is then converted back to a voltage at the amplifier output.

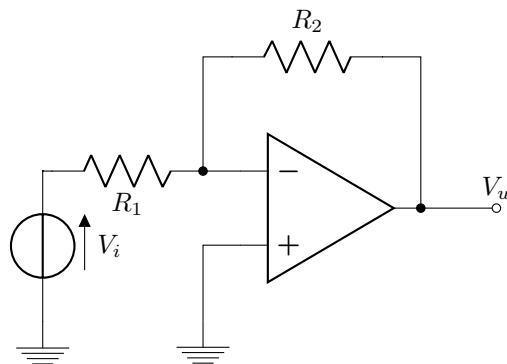


Figure 2.10: The inverting amplifier

Before computing the gain of the amplifier let us say that this circuit is not a good voltage amplifier because its input impedance is equal to R_1 and not very large as in the non-inverting amplifier. As we have already seen in the analysis of the previous circuit if we assume that the gain of the operational amplifier is very large, the differential voltage v_d will be 0 and the v^- node will be at ground potential without being physically connected to it. It is generally referred to as *virtual ground*.

The input current of the amplifier is the current flowing through R_1 and the input impedance is the ratio between the input signal and the input current.

Remembering that node v^- is at ground potential we have:

$$I_i = \frac{V_i}{R_1} \longrightarrow z_i = \frac{V_i}{I_i} = R_1$$

Since we have assumed that the input current of the operational amplifier is 0, then all current I_i will flow through R_2 generating a voltage drop across it $V_{R2} = I_i R_2$. Node v^- being at ground potential we have $V_u = -V_{R2}$ and therefore:

$$V_u = -\frac{V_i}{R_1} \cdot R_2$$

The voltage gain of the amplifier will then be:

$$\frac{V_u}{V_i} = -\frac{R_2}{R_1}$$

This circuit amplifies an input signal with a gain depending only on the ratio of the two resistors and introduces a phase inversion, that is a phase shift of 180° .

If we want to take into account a non infinite gain A_d of the operational amplifier then we can repeat the analysis assuming $v_d \neq 0$ and obtaining:

$$\frac{V_i + v_d}{R_1} = \frac{-v_d - V_u}{R_2}$$

Since $v_d = V_u/A_d$ we have:

$$V_i = -V_u \left(\frac{R_1}{R_2} + \frac{R_1 + R_2}{A_d R_2} \right)$$

Defining $\beta = R_1/(R_1 + R_2)$ and reordering the terms we finally have

$$\frac{V_u}{V_i} = -\frac{R_2}{R_1} \cdot \frac{1}{1 + \frac{1}{A_d \beta}} = \left(1 - \frac{1}{\beta} \right) \cdot \frac{1}{1 + \frac{1}{A_d \beta}} = \left(\frac{\beta - 1}{\beta} \right) \cdot \frac{A_d \beta}{1 + A_d \beta}$$

which shows that also in this case the error with respect to the ideal case is on the order of $1/A_d \beta$.

We have obtained another type of voltage amplifier which is far from ideal because its input impedance is far from infinite and therefore its characteristics depend also on the impedance of the source generator. For what concerns the output impedance if we place a test generator at the output and we set to 0 the input generator we get the same circuit of the non inverting amplifier and therefore the same value of the output impedance computed in section 2.1.2.

2.2 Differential and common mode signals and gain

Figure 2.11 shows a differential amplifier; in the general case the output is linear combination of the two inputs:

$$V_u = A_1 V_1 + A_2 V_2 \quad (2.1)$$

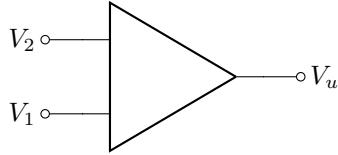


Figure 2.11: Generic differential amplifier

If we want the amplifier to be fully *differential* the output must be proportional only to the difference of the inputs, that is:

$$V_u = K(V_1 - V_2)$$

and this requires:

$$\frac{A_2}{A_1} = -1 \implies A_1 = -A_2$$

The two coefficients in equation 2.1 must therefore have the same absolute value and opposite sign.

In order to better analyze this system let us make a change of variables. Instead of expressing the output V_u as a linear combination of the input voltages, let us express it as a function of the difference v_d between the two input voltages, referred to as **differential mode** voltage and the average V_C of the two input voltages, referred to as **common mode** voltage.

Figure 2.12 shows two sinusoidal signals V_1 e V_2 together with the corresponding common mode V_C and differential mode v_d voltages. The expressions relating these new variables and the input voltages are:

$$\begin{cases} v_d = V_1 - V_2 \\ V_C = (V_1 + V_2)/2 \end{cases}$$

and conversely

$$\begin{cases} V_1 = V_C + v_d/2 \\ V_2 = V_C - v_d/2 \end{cases}$$

Substituting these values in the general V_u expression of equation 2.1 we can express it as a function of v_d e V_C :

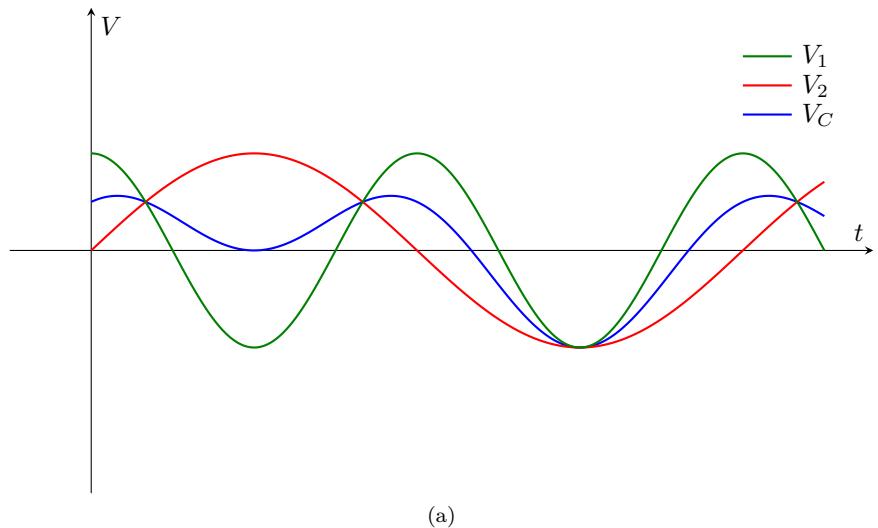
$$V_u = A_d v_d + A_C V_C$$

Where

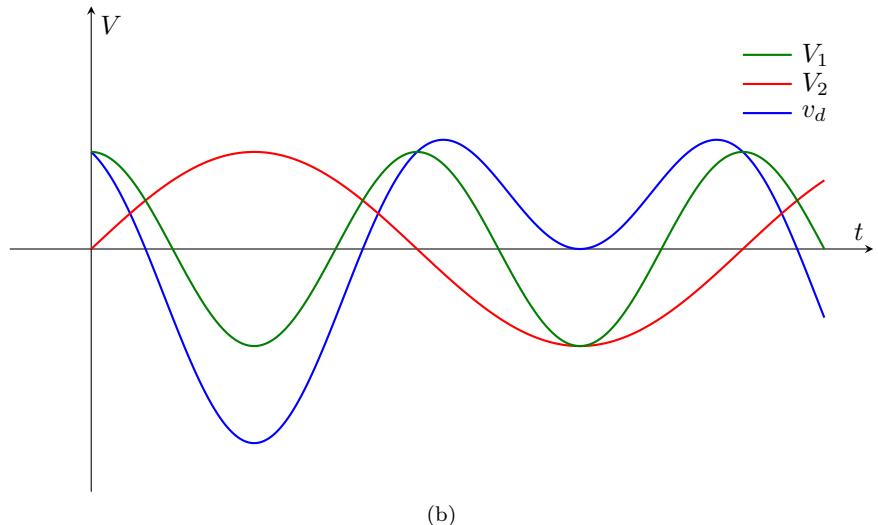
$$A_d = \frac{A_1 - A_2}{2}; \quad A_C = A_1 + A_2$$

We have separated the two *operating modes* of the amplifier, that is we consider the system as composed of two parts: the differential amplifier which amplifies only the differential signal and the common mode amplifier which amplifies (or better reduces) the average value of the two signals.

The *ideal differential amplifier* must amplify only the differential signal and have a very large A_d and $A_C = 0$ in order to cancel the common mode component of the input signals. Ideally we would like to have the output voltage V_u depending only on v_d :



(a)



(b)

Figure 2.12: Common mode (a) and differential mode (b) of a pair of sinusoidal signals V_1 e V_2 . Note the values in correspondence of the maximum, minimum and zero values of the inputs.

$$V_u = A_d v_d$$

We can rewrite the expression of V_u by introducing the ratio between the differential and the common mode gain:

$$V_u = A_d v_d \left[1 + \frac{A_C}{A_d} \frac{V_C}{v_d} \right]$$

This means that the larger is the value of A_C with respect to the "useful" gain A_d the larger will be the errors with respect to the ideal differential amplifier.

The most common figure of merit for evaluating the errors due to the common mode amplification is the CMRR (*Common Mode Rejection Ratio*), which is defined as:

$$\text{CMRR} \triangleq \frac{A_d}{A_C}$$

In general it is expressed in dB that is:

$$(\text{CMRR})_{dB} \triangleq \left(\frac{A_d}{A_C} \right)_{dB} = 20 \cdot \log_{10} \left| \frac{A_d}{A_C} \right|$$

The higher the CMRR the better are the characteristics of the differential amplifier.

As any other electronic device, the amplifier must be powered and the value of the voltage power supply determines important parameters such as the *common mode voltage input range* and the *differential mode voltage input range*.

The amplifier has a correct behavior if all its internal components operate in linearity and the above mentioned ranges define the limits for the input signals in order to insure it. We will obtain the values of those limits when analyzing the internal structure of the amplifier and now let us just state that at the inputs we must not apply:

- Common mode signals which may bring into cutoff or saturation the input transistors or the following stages due to non zero common mode voltage. The *common mode input voltage range* is the range of the common mode voltage which guarantees the linear operation of the internal components of the amplifier.
- Differential mode signals which may bring out of linearity internal active components. The allowed signal interval is called *differential mode input voltage range*

Both the above mentioned voltage ranges depend on the power supply voltage of the differential amplifier. In general there are not major problems for the differential range because the differential amplifier is in general the input stage of an operational amplifier with negative feedback which keeps the differential input voltage close to 0.

We must pay more attention to the common mode input range. Let us consider as an example the voltage follower of figure 2.6 and note that the input common mode voltage is the input voltage V_i . In fact V_i is applied to the non

inverting input and the inverting input is kept at the same voltage by the feedback network which maintains $v_d = 0$. Therefore the allowed input range of V_i is the common mode range.

This does not apply to the inverting amplifier which has no input voltage limit deriving from the common mode voltage range because for any signal both input terminals are at 0 V (either real or virtual) and voltage limits apply only to the output voltage range.

2.3 Real Operational Amplifier

With respect to ideal operational amplifier, the real device has limitations for what concerns the input and output voltage range, the pass band, the slew rate. Also parasitic input voltages and currents concur to non ideal voltage behavior.

2.3.1 Current and voltage offsets

The ideal model of the operational amplifier is valid as a first order approximation in the initial design or analysis phase of many circuits. However in the advanced design phases it is necessary to take into account the differences from the ideal behavior due to the internal structure of the amplifier and to tolerances of active and passive components.

Input currents

We have assumed that the input currents in an *ideal* operational amplifier are null by definition. However the amplifier inputs generally connect internally to the bases of bipolar transistors or to the gate of JFET or MOSFET devices. In the case of bipolar transistors, a small input bias current is necessary to maintain the device in active region, in the case of MOSFET devices, parasitic currents generally flow into input gate protection devices (diodes).

Currents I_+ e I_- are very small. Nevertheless these currents are not null and the circuit external to the amplifier must provide a DC path between each input and the reference voltage of the system. If this path is blocked, i.e. by a series capacitor, then the corresponding transistor will be off and the amplifier will not operate correctly.

It is clear that even with null differential input voltage, input currents I_+ and I_- may be very different because of variability of parameters of input transistors. If a differential voltage is applied this difference may further increase.

In order to take into account the bias currents in the analysis of operational amplifier circuits we introduce a pair of current generators, one for each input terminal IN_+ and IN_- . It is convenient to separate these currents, as we did for input voltages v_+ e v_- , into a common mode and a differential mode component. Let us then define *bias* and *offset* currents as:

$$I_{\text{bias}} := \frac{I_+ + I_-}{2} \quad I_{\text{offset}} := |I_+ - I_-|$$

$$\implies I_+ = I_{\text{bias}} + \frac{I_{\text{offset}}}{2} \quad I_- = I_{\text{bias}} - \frac{I_{\text{offset}}}{2}$$

The bias current is the average value of the input currents while the offset current is the difference between them, due to β difference of input transistors or other similar unbalance factors. We will see in the following how to reduce the effect of these currents on the circuit output and to make it negligible with respect to effect of the real input signals.

Input offset voltage

If we take a real operational amplifier with a 0 V input signal, due to internal asymmetries the output voltage will not be 0 V and it is necessary to apply a small input voltage difference to set the output to 0 V. This voltage difference is named *offset* voltage V_{off} and is normally limited to a few mV. We can take into account this non ideality by introducing a voltage generator of value V_{off} between the input pins of the ideal amplifier model.

2.3.2 Common mode input range

In section 2.2 we have noticed that the most important input voltage range is the common mode one and not the differential one which, in linear operation, is very small. In most operational amplifiers the upper and lower limits for common mode voltage are symmetrical and approximately equal to the supply voltage minus 1.5 V.

There are other input stage circuits with different input voltage ranges. Amplifiers built to be used with a single supply voltage have in general the lower limit close to 0 V. There are amplifiers whose input common voltage voltage range is equal, with a few millivolts difference, to the power supplies and they are referred to as *rail-to-rail input* amplifiers.

2.3.3 Output voltage range

The output voltage range is the range of voltage which the output terminal of the amplifier may reach. Due to voltage drops over junctions and resistances it is not possible in general to reach the power supply value and the difference may go from a few hundreds millivolts to a few volts. The output voltage range depends also on the load value because the maximum current the amplifier may provide is limited and if the load is high (low resistance) the output voltage will be automatically reduced.

Figure 2.13 shows the typical output characteristic from the datasheet of the TL081/TL082 amplifiers as a function of the load resistance. For a supply voltage of ± 15 V and a load resistance $R_L = 10 \text{ k}\Omega$ the data sheet specify also:

$$\text{Output Voltage Swing} \quad \text{min} = \pm 12 \text{ V} \quad \text{typ} = \pm 13.5 \text{ V}$$

Some amplifiers, with different final stages, may provide output voltages very close to power supplies and are referred to as *rail-to-rail output* amplifiers.

2.3.4 Input Impedance

A full model of the operational amplifier must also take into account that the input impedances are large but not infinite. Since we have separated the input signal into a common and a differential mode we will also define a *differential*

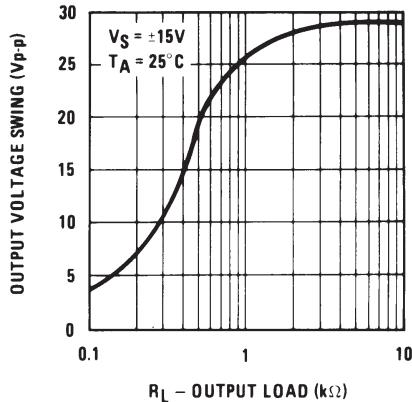


Figure 2.13: Output voltage range V_{pp} as a function of the load in $k\Omega$ for the TL081/TL082 amplifiers with symmetric power supply $V_{AL} = \pm 15$ V.

mode resistance related to the differential signal and a *common mode* resistance which applies to the common mode signal V_C only.

Differential mode resistance Given a purely differential signal v_d , the differential mode resistance is the ratio between v_d and the current i_d entering the amplifier.

$$r_{id} = \frac{v_d}{i_d}$$

Depending on the devices which create the input stage of the operational amplifier, differential mode resistance has a value ranging from $50\text{ k}\Omega$ to several $\text{M}\Omega$.

Common mode resistances The common mode input resistance is far higher than the differential one and in many circuit analyses it is possible to neglect it, but it is not affected by negative feedback, so for example in voltage follower configuration it is the lowest term limiting input resistance.

2.3.5 Differential gain

The differential gain of a real operational amplifier is generally very high, ranging from 60 dB to 80 dB or more.

2.3.6 Common mode gain

The common mode gain of an operational amplifier is ideally null, but this is impossible to achieve in real circuits, though it is normally very small. As previously discussed, one important parameter of the real operational amplifier is the ratio between differential and common mode gains, called CMRR.

2.4 Sizing the components of an amplifier

After analyzing the characteristics of a real amplifier we may now size correctly the passive components of a operational amplifier based circuit. As an example let us redesign the non inverting amplifier using a real amplifier with its non idealities.

2.4.1 Specifications

Let us use the following design specifications:

- Voltage gain: $A_V = 10$;
- LM741 Operational amplifier by Texas Instruments. Datasheets are available on the web at address www.ti.com/lit/ds/symlink/lm741.pdf
- Minimum output voltage range $V_u = \pm 10 \text{ V}$;
- Load resistance $R_L > 4 \text{ k}\Omega$;
- Power supply voltage $V_{AL} = \pm 15 \text{ V}$.

2.4.2 Design

The first design task is to verify the match between the specifications and the characteristics of the chosen components.

Maximum ratings

First we must check whether the *maximum ratings* of the chosen components operating parameters are compatible with the design specifications. In our case the maximum symmetric supply voltage for the commercial version of the component (LM741C, *Operating Temperature Range 0 °C to 70 °C*) is $V_{AL} = \pm 18 \text{ V}$ and therefore compatible with the specifications. In a real project all the maximum ratings must be verified in detail, but in our case there is no problem.

Electrical characteristics

It is then necessary to verify whether the amplifier may drive the load and to provide the requested output voltage range. We find the relevant information in the electrical characteristics section of the datasheet (page 3).

The parameter *Output Voltage Swing* is the output voltage range and, since its value decreases with the load resistance, it is necessary to verify the compatibility with the R_L of the specifications. In order to guarantee the operation in any condition it is necessary to evaluate the worst case conditions. From the datasheet we have:

- If $R_L \geq 10 \text{ k}\Omega$, $V_{min} = \pm 12 \text{ V}$, $V_{typ} = \pm 14 \text{ V}$
- If $R_L \geq 2 \text{ k}\Omega$, $V_{min} = \pm 10 \text{ V}$, $V_{typ} = \pm 13 \text{ V}$

For both $R_L = 2\text{k}\Omega$ and $R_L = 10\text{k}\Omega$ it is possible to evaluate the maximum amplifier current for the given load conditions. This value is an important parameter for the following design phases.

$$R_L = 2\text{k}\Omega \implies V_{MAX} = 10\text{V} \implies I_{MAX} = \frac{V_{MAX}}{R_L} = 5\text{mA}$$

$$R_L = 10\text{k}\Omega \implies V_{MAX} = 12\text{V} \implies I_{MAX} = \frac{V_{MAX}}{R_L} = 1.2\text{mA}$$

Let us then note that in short circuit conditions, that is for 0 output swing, the maximum current is 25 mA. Our minimum load resistance is an intermediate value for the values given and the required output voltage range is guaranteed for a lower R_L . Therefore this parameter too is compatible with our specification.

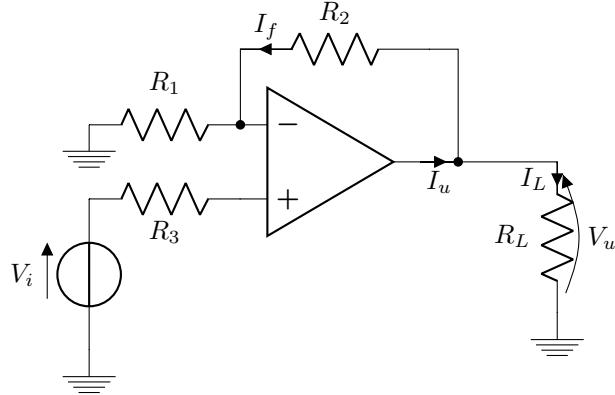


Figure 2.14: Circuit schematic of the non inverting amplifier

Circuit schematic

The amplifier schematic is shown in figure 2.14 and resistor R_3 has been added for reasons which will be soon clear. The figure explicitly shows also the load resistance R_L which normally is normally not shown because it represents the input of another circuit to be connected to the amplifier output. This fact reminds us that it is always necessary to evaluate it in order to know the current that the amplifier must provide.

As we have already seen, the ratio between R_2 and R_1 sets the transfer function of the amplifier. Since we want $A_V = 10$ then we have:

$$\frac{V_u}{V_i} = \left(1 + \frac{R_2}{R_1}\right) = 10 \implies R_2 = 9R_1$$

The analysis of the ideal amplifier only specifies the ratio between the resistances and does not give any information about their absolute value. It also does not provide any information on the value of R_3 which has no effect on the transfer function if we assume an ideal amplifier. In order to choose the optimal values of the circuit components (R_1 , R_2 and R_3) it is necessary to take into account also the output and feedback currents and the offset and bias currents.

Feedback and load currents

The current I_u provided by the operational amplifier is composed of the feedback current I_f flowing in resistance R_2 and of current I_L in the load. We have then:

$$I_u = I_f + I_L$$

$$I_f = \frac{V_u}{R_1 + R_2} = \frac{V_u}{R_2 \cdot 10/9}$$

For sizing the feedback current a possible criterion is to make it small with respect to the "useful" current, that is the one provided to the load. If we do not know the value of the load in advance then we make it small, that is one order of magnitude lower, with respect to the maximum current the amplifier may provide. In our case we have seen that the maximum amplifier current which guarantees the required output voltage range of 10 V is 5 mA. We have then:

$$I_f \ll 5 \text{ mA} \implies R_2 \cdot \frac{10}{9} \simeq R_2 \gg \frac{10 \text{ V}}{5 \times 10^{-3} \text{ A}} = 2 \text{ k}\Omega$$

$$R_2 \geq 20 \text{ k}\Omega$$

This inequality, deriving from the limit on the feedback current sets a *lower bound* to R_2 .

Offsets and bias currents

In order to find an *upper bound* to R_2 it is necessary to take into account the effect of bias currents and of offset voltages and currents. The circuit model of figure 2.15 outlines those contributions.

Offset and bias generator are independent and in the analysis of the circuit we may apply the superposition principle by considering each of them separately, all other ones been set to 0. We will also assume an amplifier with $A_d \rightarrow \infty$, therefore $v_d \rightarrow 0$ and the current in r_{id} equal to 0.

Offset voltage Using basic circuit theory it can be easily seen that V_{off} can be moved to the non inverting input of the amplifier and figure 2.16 shows the modified circuit after nulling all other generators. It is clear that the effect of the offset voltage is the same as that of an external signal and is given by:

$$V_u|_{V_{off}} = V_{off} \cdot \left(1 + \frac{R_2}{R_1}\right)$$

Therefore the voltage offset contribution to the output does not depend on the absolute value of R_2 but only on the value of the gain which is set by circuit specification. Once the gain has been defined, the only way to reduce its contribution is to use an operational amplifier with a lower value of V_{off} .

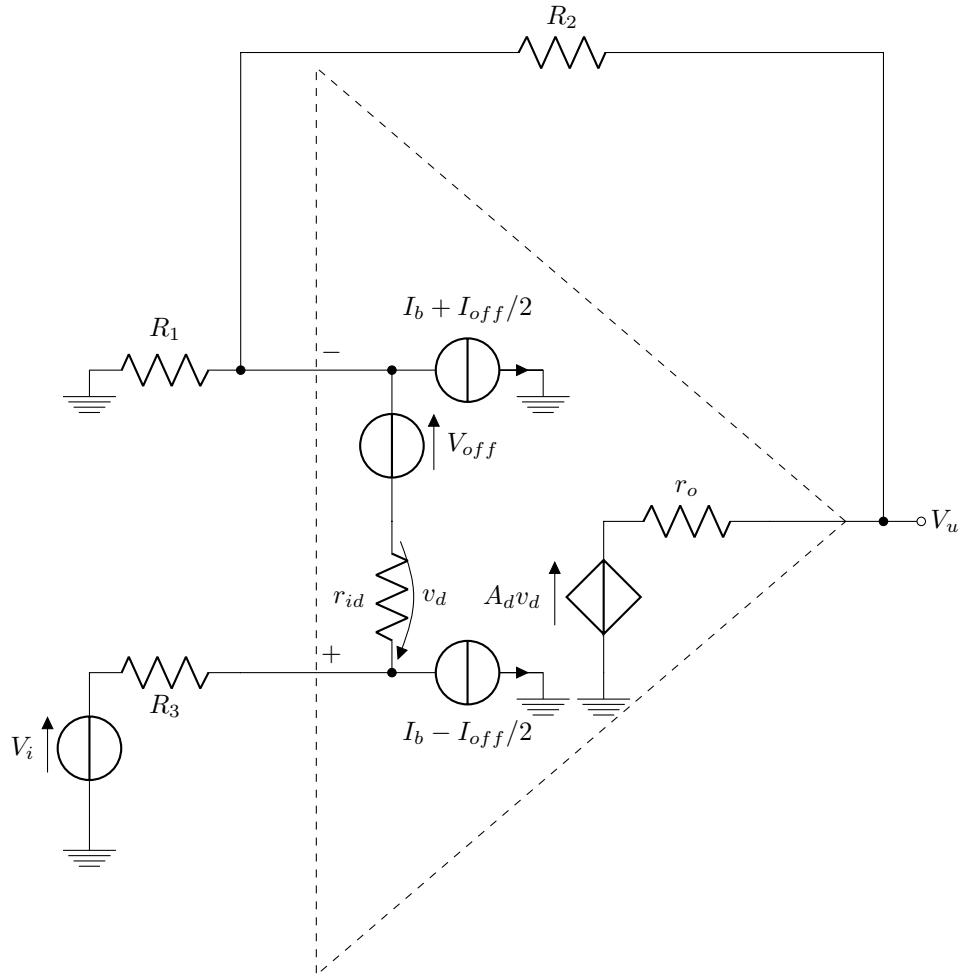


Figure 2.15: Full Circuit model of a real operational amplifier

Offset and bias currents Figure 2.17 shows the circuit when considering only the offset and bias current at the inverting input. Since there is no current flowing into R_3 and v_d being 0, then also the voltage of the inverting input node will be 0 and therefore no current will flow through R_1 .

All the current will flow through R_2 and the output voltage will be:

$$V_u|_{I_-} = R_2 \left(I_b + \frac{I_{off}}{2} \right)$$

Figure 2.18 shows the circuit when considering bias and offset currents at the non inverting input.

No current will pass in r_d and all the generator current will flow through R_3 and the voltage drop across it will be $-(I_b - I_{off}/2)R_3$. This voltage appears at the non inverting terminal of the amplifier and will be amplified by the gain A of the non inverting amplifier giving:

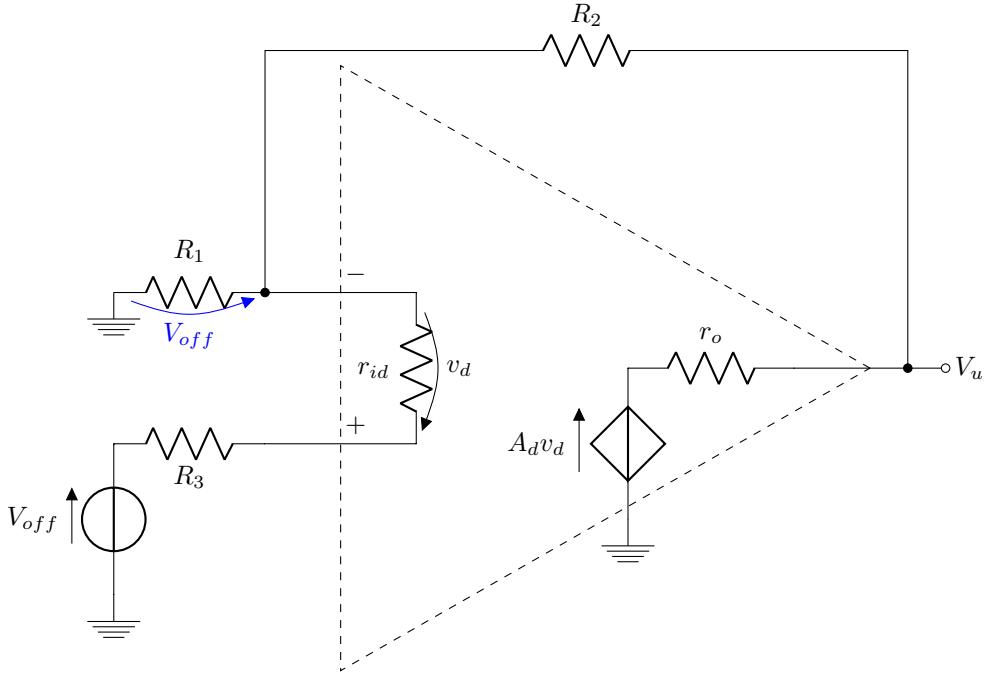


Figure 2.16: Circuit for evaluating the V_{off} contribution

$$V_u|_{I_+} = -R_3 \left(I_b - \frac{I_{off}}{2} \right) \left(1 + \frac{R_2}{R_1} \right)$$

Superposing the contributions In order to evaluate the total effect of the offset voltage and bias and offset currents we may just add the three contributions computed in the previous paragraphs obtaining:

$$V_u|_{offset} = V_{off} \cdot A_V + R_2 \left(I_b + \frac{I_{off}}{2} \right) - R_3 \left(I_b - \frac{I_{off}}{2} \right) \left(1 + \frac{R_2}{R_1} \right)$$

Let us now consider the effect of resistance R_3 which we have added to the design and whose value is an additional degree of freedom which we can exploit to minimize the effect of bias currents. In fact in the previous expression the two terms containing I_b have opposite sign and if we make their coefficients equal then they will cancel each other. Separating the terms containing I_b and I_{off} we have:

$$V_u|_{offset} = V_{off} \cdot A_V + I_b \left[R_2 - R_3 \left(1 + \frac{R_2}{R_1} \right) \right] + \frac{I_{off}}{2} \left[R_2 + R_3 \left(1 + \frac{R_2}{R_1} \right) \right]$$

And imposing the I_b coefficient to be 0 we obtain:

$$R_3 \left(1 + \frac{R_2}{R_1} \right) = R_2 \longrightarrow R_3 = \frac{R_1 R_2}{R_1 + R_2} = R_1 \parallel R_2$$

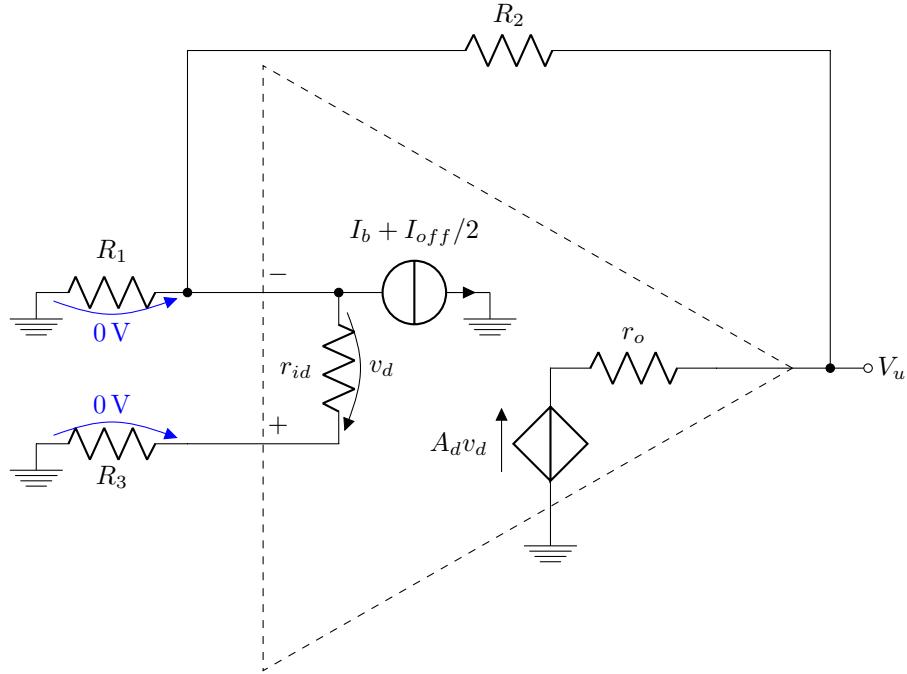


Figure 2.17: Contribution of the offset and bias current at the inverting input

That is if the resistance "seen" from the two input terminal is equal then the effect of I_b will be canceled. If we choose the R_3 value which satisfies this condition then total output voltage due to offsets will become:

$$V_u|_{offset} = V_{off} \cdot A_V + I_{off}R_2$$

Since we cannot reduce the V_{off} contribution, we can make the I_{off} contribution negligible with respect to it by appropriately sizing the value of R_2 .

$$V_{off} \cdot A_V \gg R_2 \cdot I_{off} \implies R_2 \ll \frac{V_{off} \cdot A_V}{I_{off}}$$

In our case let us take from the LM741 datasheet the worst case offset parameters and evaluate the upper limit of assuming the I_{off} contribution to be 10 times lower than the one due to V_{off} .

$$R_2 \ll 10 \cdot \frac{6 \cdot 10^{-3}}{2 \cdot 10^{-7}} = 0.3 \text{ M}\Omega$$

$$R_2 \leq 30 \text{ k}\Omega$$

Having obtained the upper and lower bounds for R_2 it is possible now to choose its absolute value among the available ones. In our case, using resistors of the normalized series E12 with a 10% tolerance, we may chose the value $R_2 = 27 \text{ k}\Omega$ and this gives for R_1 the value of $3 \text{ k}\Omega$. This value however is available in the E24 series only and in the E12 series the possible choices are

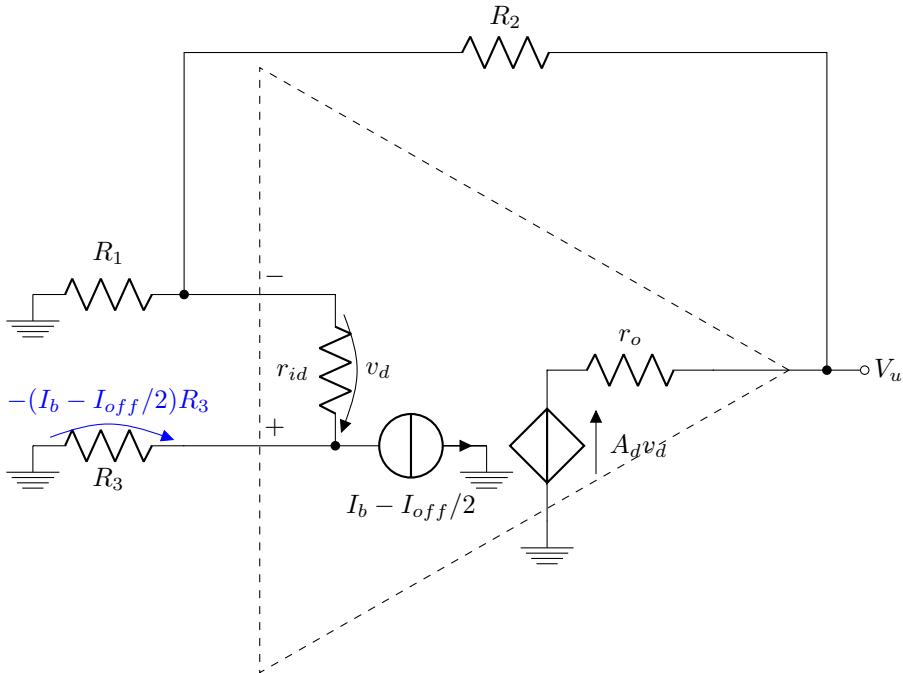


Figure 2.18: Contribution of the offset and bias current at the inverting input

values $2.7\text{ k}\Omega$ or $3.3\text{ k}\Omega$ which give a nominal value of gain higher or lower by 10% with respect to the required one. For R_3 an acceptable value is $2.2\text{ k}\Omega$.

This design methodology may be used for a large percentage of operational amplifier circuits. It should be noted that sizing the components of an inverting amplifier would lead to identical results because the effect of the offset generator or of the current in the load is independent of where the input signal is connected.

2.5 Frequency behavior

After analyzing all the DC characteristics of the amplifier let us now analyze the *frequency* behavior. The amplifier is composed of several amplification stages and each stage introduces in the open loop transfer function at least one pole (using the simple dominant pole approximation). The amplifier is always used with a negative feedback and therefore its stability is not guaranteed a priori but must be verified according to the feedback network which is used. In this analysis we will assume that the basic feedback concepts and methodologies for assessing system stability are already known and we will just recall the fundamental notions when required.

We know that a feedback system is stable, that is it behaves as an amplifier, when the feedback is negative, that is the output signal is fed back to the input with a phase shift of 180° . This may be easily obtained in DC or low frequency but may be harder to achieve at frequencies at which amplifier poles introduce an additional phase shift in the transfer function.

2.5.1 Transfer function

Let us recall that any pole in the left half-plane of the transfer function introduces a phase shift of -90° on the transfer function (t.f.). The shift begins at approximately one decade before the pole, at the angular frequency ω_p of the pole is -45° and reaches a value close to -90° one decade after the pole. The phase shifts due to several poles add up.

If we add a resistive feedback circuit as in the non inverting amplifier of figure 2.14 then the phase shift between the input and the signal which is fed back will be 180° .

Increasing the frequency, due to the phase shift of the transfer function, the phase difference between the input and the feedback signals will decrease; at the first pole frequency ω_{p1} it is $180^\circ - 45^\circ = 135^\circ$ and at a decade after the pole it becomes $180^\circ - 90^\circ = 90^\circ$.

When the second pole at ω_{p2} is reached the phase difference will be $180^\circ - 135^\circ = 45^\circ$ and after one decade $180^\circ - 180^\circ = 0^\circ$. This means that the feedback is not negative anymore but has become *positive* and, if in these conditions the loop gain is greater than 1, any signal at that frequency will be amplified infinitely and the system will become an oscillator.

Given that the operational amplifier has sometimes more than two amplification stages, this problem is even more complex, so that the designers of the amplifier normally insert in the device a compensation network (usually just a capacitor) to modify the frequency behavior of the system and provide stability in the most stringent situation, which is normally voltage follower configuration. Such situation is the standard and this kind of operational amplifiers is called *internally compensated*.

2.5.2 Loop gain

Given the number of amplifier poles, there will always be a frequency for which phase difference will be 0 and the solution is to insure that for that frequency the signal will not be amplified by the feedback loop. Referring to a generic feedback system such as the one shown in figure 2.2, in section 2.1.1, we can define the *loop gain* as the product A of the amplifier by the attenuation β of the feedback network. The value of β is the ratio value of the signal at the output of the feedback network to the amplifier output when we imagine to open the feedback network at the amplifier input. Considering that the feedback signal is added with a negative sign, then the phase shift between the amplifier input and the feedback signal is 180° plus the phase shift of the loop gain.

The frequency behavior of the loop gain depends both on those of A and of β , but if we assume the feedback network to be purely resistive, then the Bode diagram of the loop gain will be identical to that of A to which a constant term $20\log\beta$ (negative) is added.

The **phase margin** is defined as the phase difference between 180° and the phase of the loop gain at the angular frequency ω_T at which the absolute value of the loop gain is 1, that is $|A\beta| = 1$ or $20\log|A\beta| = 0 \text{ dB}$. The higher the phase margin the more stable is the system. If the phase margin is lower than 45° , the closed loop transfer function at frequencies around ω_T has an higher value with respect to the ideal one (*ringing* and *overshoot* effects).

The **gain margin** is defined as the absolute value in dB of the loop gain at

the angular frequency ω at which $\angle A\beta(\omega) = 180^\circ$. In order for the system to be stable the absolute value of $A\beta(\omega_{180})$ must be negative (in dB).

Figure 2.19 shows the definition of these two parameters.

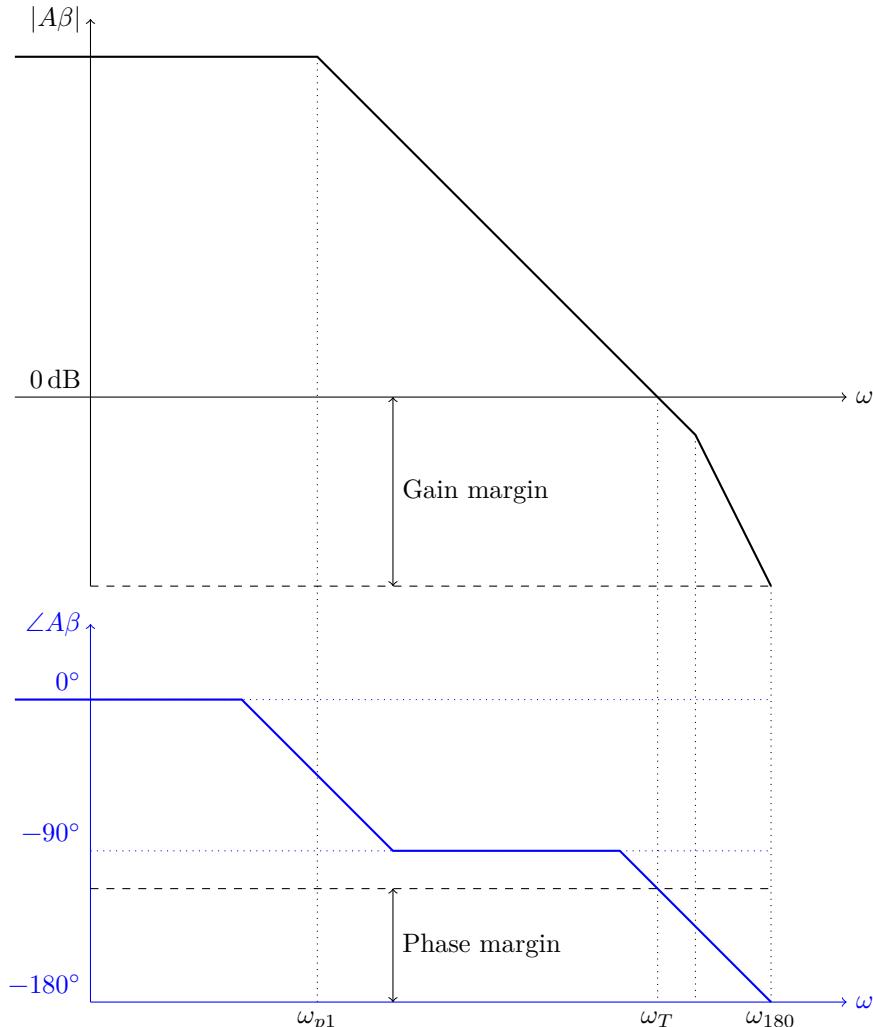


Figure 2.19: Definition of phase and gain margins.

In order to obtain a phase margin of 45° the loop gain must be 0 dB at the frequency of the second pole of the loop gain. This is what is obtained in internally compensated operational amplifiers.

2.5.3 Gain-Bandwidth product

The gain of an amplifier compensated for unity gain (voltage follower) has a Bode diagram like to one shown in figure 2.20. For frequencies lower than the second pole (at which the gain is 0 dB) the diagram may be approximated with that of a transfer function with a single dominant pole:

$$A_d(f) = \frac{A_{d0}}{1 + j \frac{f}{f_0}}$$

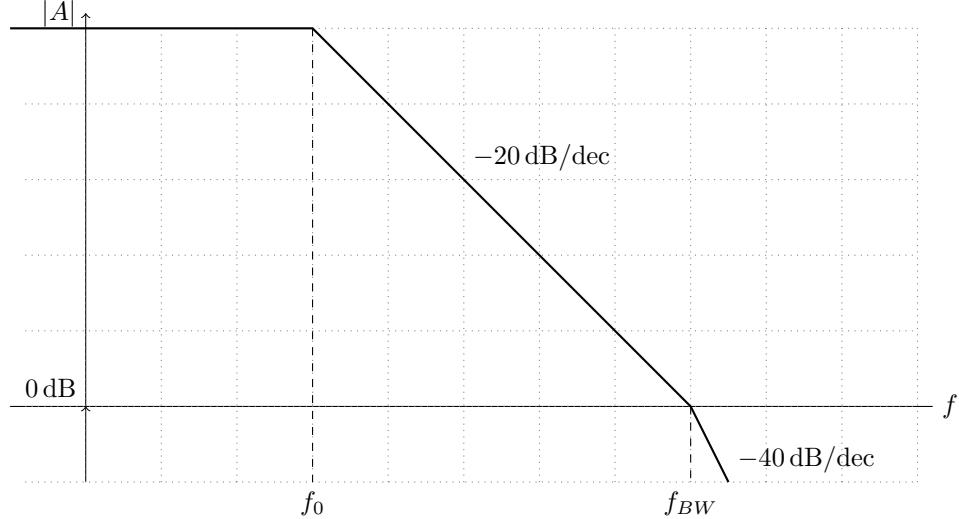


Figure 2.20: Bode diagram of an amplifier compensated for unity gain

Let us use this approximate expression to evaluate the bandwidth of a non inverting amplifier using the operational amplifier with a feedback network with value β . Let us recall that the closed loop gain has the following expression:

$$A_V = \frac{1}{\beta} \cdot \frac{1}{1 + \frac{1}{T}} \quad (2.2)$$

Where $T = A_d\beta$ is the loop gain. By substituting the value of $A_d(f)$ we get:

$$A_V = \frac{1}{\beta} \cdot \frac{1}{1 + \frac{1+jf/f_0}{A_{d0}\beta}}$$

And reordering the term we may express the closed loop gain as:

$$A_V = \frac{1}{\beta} \cdot \frac{A_{d0}\beta}{A_{d0}\beta \left(1 + \frac{1}{A_{d0}\beta} + j \frac{f}{A_{d0}\beta f_0}\right)}$$

In general we have:

$$\frac{1}{A_{d0}\beta} \ll 1 \quad (2.3)$$

from which we get:

$$A_V \approx \frac{1}{\beta} \cdot \frac{1}{1 + j \frac{f}{A_{d0}\beta f_0}} \quad (2.4)$$

from which we may evaluate the closed loop bandwidth.

Let us define now:

$$f_{BW} = A_{d0}f_0$$

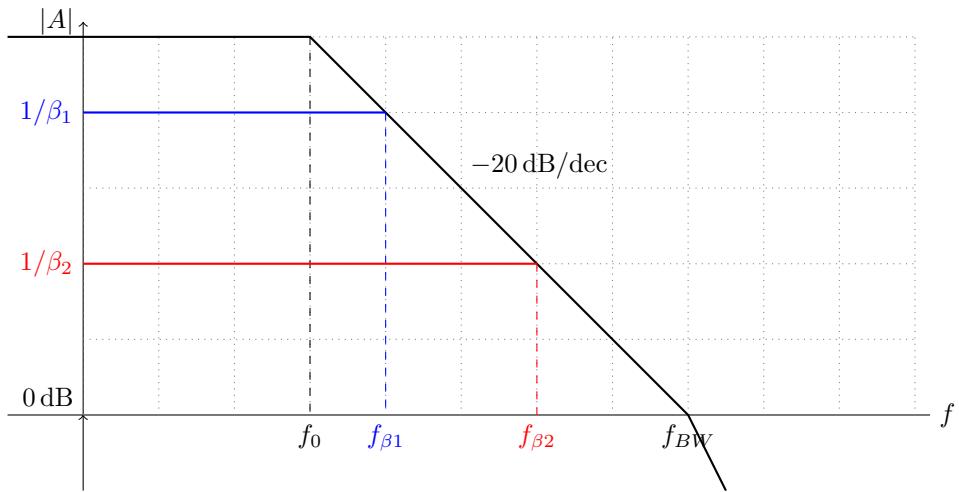


Figure 2.21: Bode diagram of the closed loop gain for different values of β

This frequency is the value for which the open loop gain is 1, that is the Bode diagram intersect the 0 dB axis. We can then write that the -3 dB cutoff frequency f_T value is $f_T = \beta f_{BW}$. Given that the low frequency closed loop gain is $A_{V0} = 1/\beta$ then the product $A_{V0}f_T$ of the gain times the bandwidth is a constant whose value is $= f_{BW}$ and which is referred to as *gain-bandwidth product*, a typical parameter of the operational amplifier. Therefore if we use an operational amplifier, which is internally compensated for unity gain, to implement a voltage amplifier then the higher the DC gain the lower will be the amplifier bandwidth. Figure 2.21 shows the Bode diagrams of the closed loop gain for different values of β .

This relation holds also for the inverting amplifier whose transfer function may be expressed as:

$$A_V = \left(1 - \frac{1}{\beta}\right) \cdot \frac{1}{1 + \frac{1}{T}}$$

It is clear that for the same feedback network β of the non inverting amplifier the amplifier bandwidth will be the same.

2.5.4 Slew Rate

The frequency limits of operational amplifiers do no come only from its bandwidth but also from a non linear parameter, the *slew rate* which is related to the time domain. Namely the slew rate is the maximum speed at which the amplifier output may vary per unit time. and is defined as:

$$SR(V_U) = \max \left(\frac{\partial V_U}{\partial t} \right)$$

Its effects are clearly visible in the amplifier response to a step input as shown in figure 2.22. Figure 2.22 shows an example of large signal output step from which its is possible to evaluate the slew rate limit. When the slew rate limit is

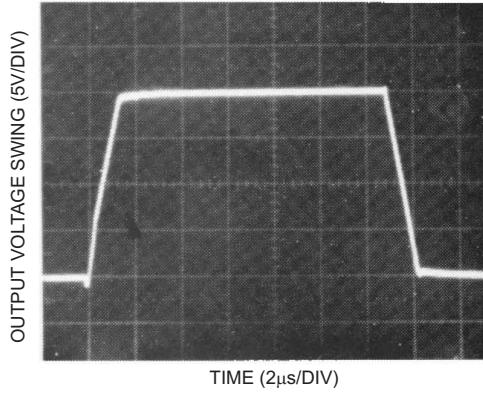


Figure 2.22: Large signal step response of operational amplifier TL082 (Texas Instruments data sheet).

reached then the output signal is affected by a large distortion and in the case of a sinusoidal signal as the frequency increases the output gradually becomes a triangular wave.

Relation between BW and SR Both the bandwidth and slew rate limit the operating frequency range of the amplifier but it is important to understand the different nature of the two mechanisms and to be able to evaluate the frequency limits introduced by each of them.

The bandwidth indicates the maximum frequency at which a signal is amplified by a given value **independently** of its amplitude, that is the limit applies in the same way to an infinitesimal signal or to a large one. The slew rate limits the maximum output signal rate of change and therefore the frequency at which it sets a limit to the signal depends also on the amplitude of the signal itself. In case of a sinusoidal signal

$$V_u(t) = V_{pk} \sin(\omega t)$$

its maximum slew rate may be computed by taking its derivative.

$$\left. \frac{dV_u}{dt} \right|_{MAX} = \omega V_{pk} \cos(\omega t)|_{MAX} = \omega V_{pk}$$

from which we have:

$$\omega V_{pk} \leq SR = \frac{I_0}{C_C} \implies V_{pk} \leq \frac{I_0}{\omega C_C} = \frac{SR}{\omega}$$

Hence the maximum undistorted amplitude decreases with frequency according to an hyperbolic relation, or, conversely, the maximum operating frequency decreases as a function of the maximum signal amplitude. Therefore it is necessary to evaluate which limit will be lower, the bandwidth or the slew rate one.

Example 2. Let us use a TL082 operational amplifier to implement an amplifier with bandwidth $f_T = 400$ kHz. Which is the maximum allowable gain and the

maximum undistorted output signal when the input sinusoidal signal has a frequency of 400 kHz?

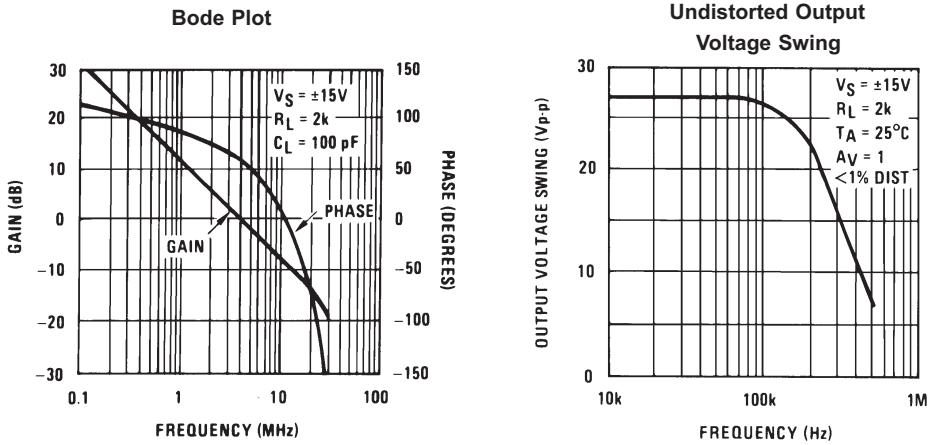


Figure 2.23: Frequency response and maximum undistorted amplitude for the TL082 amplifier Texas Instruments (data sheet).

TL82 datasheet give a slew rate value $SR = 13 \text{ V } \mu\text{s}^{-1}$ and a gain-bandwidth product $GBW = 4 \text{ MHz}$. The maximum allowable gain for the required bandwidth will be:

$$A_{V_{MAX}} = \frac{GBW}{f_T} = 10$$

The maximum undistorted amplitude at that frequency is computed by evaluating the maximum derivative of the output sinusoidal signal which must be lower than the slew rate limit:

$$V_{pk} \cdot 2\pi \cdot 4 \times 10^5 \text{ V s}^{-1} \leq 13 \text{ V } \mu\text{s}^{-1}$$

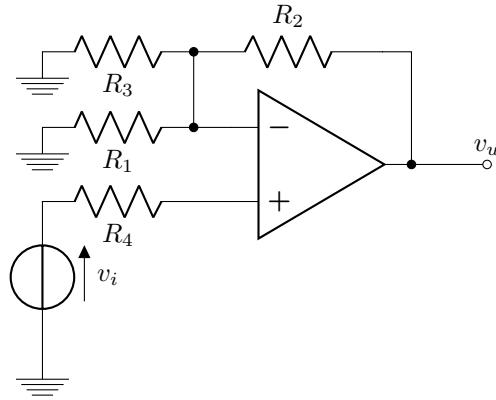
$$V_{pk_{MAX}} = \frac{13}{2\pi \cdot 0.4} \text{ V} = 5.17 \text{ V}$$

The same results may be obtained by using the diagrams in the datasheet. Figure 2.23 presents the most relevant ones.

2.6 Exercises

Exercise 2.1. Non inverting Amplifier

Given the schematic of figure 2.24



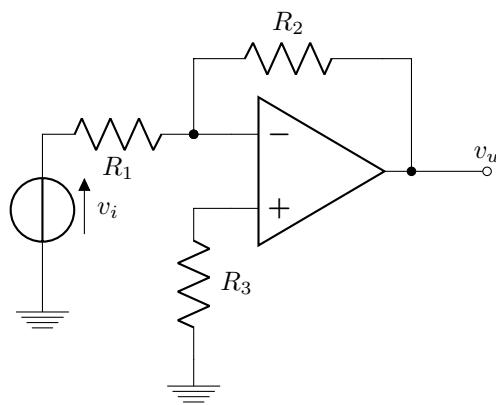
$$R_1 = 2.7 \text{ k}\Omega \quad R_2 = 27 \text{ k}\Omega \quad R_3 = 2.7 \text{ k}\Omega \quad R_4 = 10 \text{ k}\Omega$$

Figure 2.24: Operational amplifier circuit

1. Compute the gain v_u/v_i assuming that all the parameters of the operational amplifier are ideal.
2. Compute the gain v_u/v_i assuming the differential gain $A_d = 1000$ and the differential input resistance $R_{id} = 50 \text{ k}\Omega$. All other amplifier parameters are ideal.

Exercise 2.2. Inverting Amplifier

Given the inverting amplifier of figure 2.25



$$R_1 = 68 \text{ k}\Omega \quad R_2 = 220 \text{ k}\Omega \quad R_3 = 47 \text{ k}\Omega$$

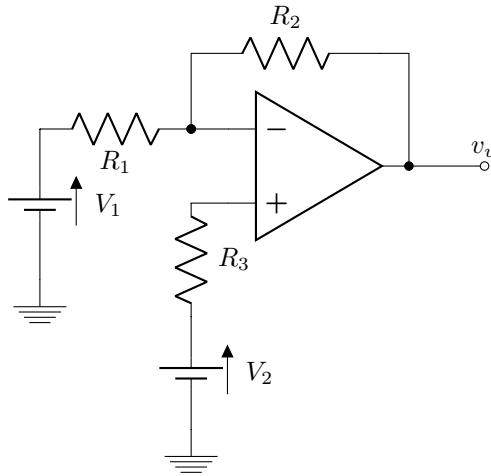
Figure 2.25: Inverting amplifier circuit

Compute the gain v_u/v_i in the two following conditions:

1. Assuming the operational amplifier to be ideal;
2. Assuming the differential gain A_d of the operational amplifier to be equal to 1000 and for all other aspects the amplifier to be ideal.

Exercise 2.3. Amplifier with DC generators and offsets

Given the operational amplifier circuit of figure 2.26



$$R_1 = 1 \text{ k}\Omega \quad R_2 = 47 \text{ k}\Omega \quad R_3 = 1 \text{ M}\Omega \quad V_1 = 0.2 \text{ V} \quad V_2 = 0.4 \text{ V}$$

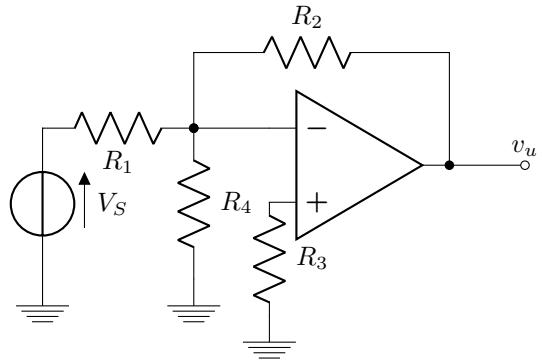
Figure 2.26: Operational amplifier circuit

1. Compute the output voltage V_u assuming that the operational amplifier is ideal.
2. Compute V_u assuming the differential gain $A_d = 1000$ and the differential input resistance $R_{id} = 50 \text{ k}\Omega$. All other amplifier parameters are ideal.
3. Compute v_u assuming that the offset voltage v_{off} and the bias current I_b of the operational amplifier are respectively equal to $\pm 3 \text{ mV}$ and to 100 nA and that all the other amplifier parameters are ideal.

Exercise 2.4. Inverting Amplifier with offsets

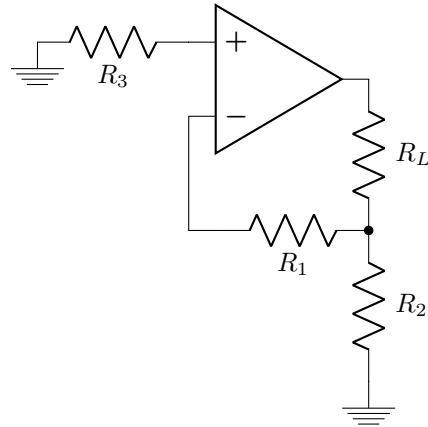
Given the operational amplifier circuit of figure 2.27

1. Compute the gain v_u/V_S assuming that the operational amplifier is ideal
2. Compute the gain v_u/V_S assuming that the differential gain $A_d = 2000$ and that all other amplifier parameters are ideal.
3. Compute the output voltage v_u for $V_S = 0$ assuming that the operational amplifier offset voltage is $\pm 10 \text{ mV}$ and that all other amplifier parameters are ideal.



$$R_1 = 6.8 \text{ k}\Omega \quad R_2 = 47 \text{ k}\Omega \quad R_3 = 3.3 \text{ k}\Omega \quad R_4 = 6.8 \text{ k}\Omega$$

Figure 2.27: Inverting amplifier circuit



$$R_1 = 1 \text{ k}\Omega \quad R_2 = 100 \Omega \quad R_3 = 1.2 \text{ k}\Omega \quad R_L = 500 \Omega$$

Figure 2.28: Circuit for offset evaluation

Exercise 2.5. Offset evaluation

Given the circuit schematic of figure 2.28

Compute the value of the current in the load resistance R_L assuming

$$v_{off} = \pm 6 \text{ mV}$$

$$I_B = 200 \text{ nA}$$

$$i_{off} = \pm 100 \text{ nA}$$

For all other aspects the operational amplifier may be considered ideal.

Exercise 2.6. Current amplifier

Using an operational amplifier, design a current amplifier using with $I_L = 10I_i$, on a load Z_L whose both terminals are available and do not require to be at a given voltage, that is, load may be floating.

The input signal range is from 0.1 mA to 1 mA. Size the feedback resistors considering that the amplifier is powered at ± 15 V, has a maximum output voltage range of 13 V and has the following offset parameters: $V_{off} = 3$ mV, $I_{off} = 50$ nA, $I_b = 200$ nA.

Which are the maximum and minimum admissible values of Z_L ?

Exercise 2.7. Current amplifier

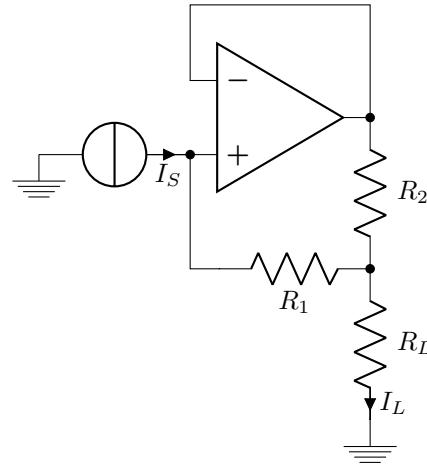


Figure 2.29: Howland Current Amplifier

Given the circuit of figure 2.29 compute the gain, that is the ratio between current I_L and current I_S .

Exercise 2.8. Voltage Follower

A voltage follower circuit is implemented using an operational amplifier with the following parameters:

- $A_d = 80.000$
- CMRR = 100 dB
- $R_{id} = 80 \text{ k}\Omega$
- $R_{out} = 470 \Omega$

If the impedance of the input generator v_s is $R_s = 470 \text{ k}\Omega$ and the load impedance is $R_L = 1 \text{ k}\Omega$ compute the real gain V_u/V_s of the amplifier.

Chapter 3

Operational Amplifier applications

Contents

3.1	Inverting configurations	3-1
3.1.1	Integrator	3-2
3.1.2	Differentiator	3-4
3.1.3	Inverting adder	3-6
3.2	Differential amplifier	3-7
3.3	Non inverting adder	3-10
3.4	Generalized adder	3-12
3.5	Instrumentation amplifiers	3-16
3.5.1	Introduction	3-16
3.5.2	Electromagnetic compatibility	3-16
3.5.3	Instrumentation amplifiers circuits	3-17
3.6	Audio amplifiers	3-21
3.7	Single supply circuits	3-24
3.8	Exercises	3-28

IN THIS CHAPTER we will analyze a number of linear applications of the operational amplifier, among which several types of adders, the differential amplifier, the instrumentation amplifier and the circuits with a single power supply.

3.1 Inverting configurations

In the introductory chapter on operational amplifiers we have analyzed two basic types of amplifiers using this component, namely the non inverting and the inverting amplifiers. An interesting feature of the inverting configuration is that the transfer function is extremely simple and is just the ratio of two resistors. It is possible to extend the analysis of this configuration including reactive elements in the feedback network obtaining, with two generic impedances Z_1 and Z_2 instead of resistors R_1 and R_2 , a more general expression:

$$\frac{V_u}{V_i} = -\frac{Z_2}{Z_1}$$

This means that by choosing the appropriate Z_1 and Z_2 it is possible to easily synthesize transfer functions which may be very different from that of a simple amplifier. It is commonly said that the inverting amplifier is the “mother” of many linear circuits based on the operational amplifier, such as active filters and others. In the next paragraphs we will present some basic circuits based on the inverting amplifier. It will be clear that the name *operational* associated to the component derives from the fact that, by using the appropriate external components, it is possible to obtain mathematical operations on input signals such as addition, integration, logarithm.

3.1.1 Integrator

Let us consider the circuit of figure 3.1.

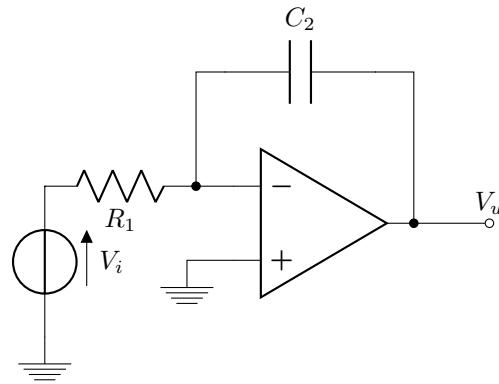


Figure 3.1: The integrator circuit.

Using the expressions of impedances Z_1 and Z_2 it is possible to derive that in the Laplace s variable domain the transfer function is:

$$\frac{V_u}{V_i} = -\frac{1}{sC_2R_1}$$

The Bode diagram is in fig. 3.2. This circuit is called *integrator*; we can obtain its behavior in the time domain by reverse transforming the transfer function and obtain:

$$v_u(t) = \mathbb{L}^{-1}\{V_u(s)\} = v_u(0) - \frac{1}{R_1C_2} \int_0^t v_i(t)dt$$

This circuit integrates the input signal in the time domain. If we were using an ideal operational amplifier it would implement that function correctly, but gain and bandwidth limits of real components introduce limits in the behavior which must be more carefully analyzed.

We will not investigate circuit stability because it is out of the scope of this course, we just note that in DC the capacitor is an open circuit and therefore the operational amplifier will be an *open loop* with gain A_d .

Another limiting factor for the integrator at low frequencies is the output voltage range of the operational amplifier because, for a constant amplitude (DC) or low frequency input signal, the output voltage is inversely proportional

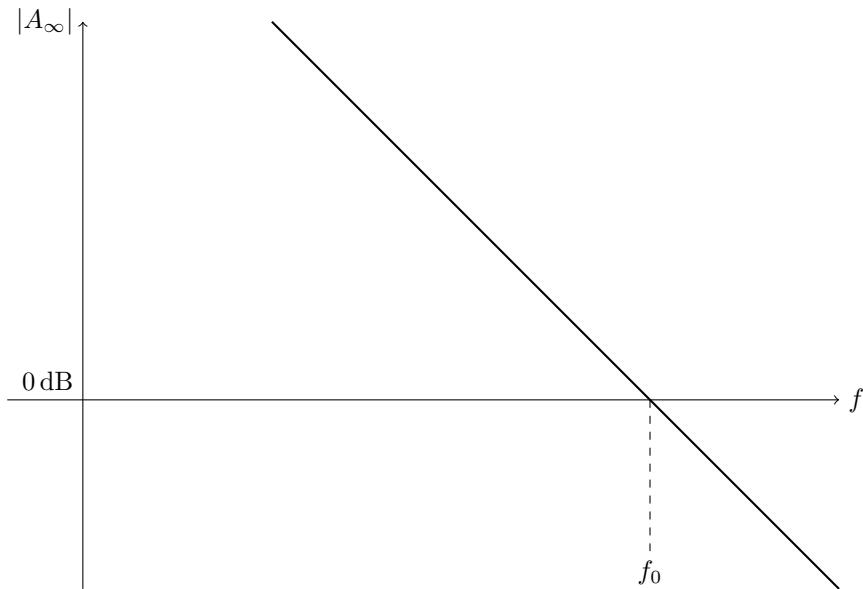


Figure 3.2: Bode diagram of the integrator.

to the frequency itself and the amplifier saturates when reaching its output voltage limit. In particular this happens when we apply DC input voltage V_{DC} . The amplifier would try to reach the value $A_d V_{DC}$ with a transient governed by the time constants of the circuit but would permanently saturate at the output voltage range limit. Even if the input signal has no DC component there are the *internal* DC components, that is the offset voltage and currents, which will cause saturation.

The circuit, as it is, does not operate correctly and there are two ways to overcome its limitation, which will be analyzed in the following paragraphs.

Periodic reset

If the circuit must operate only for a short time, then it is possible to periodically reset the circuit output. The operating time must be short enough to make negligible the contribution to the output of the offset and bias current. In practice an electronic switch is added which will short the capacitor when the integrator is not in use. At the starting of an integration cycle the switch is opened and at the end is closed again in order to reset to the same conditions for the next cycle. This technique is often employed in ramp-type voltmeters.

Integrator with losses

If the system must operate as an integrator in a limited range of frequencies, not including DC, then it is possible to use the modified circuit of figure 3.3 which has a finite DC gain, set by the additional resistance R_2 .

For low frequencies the circuit operates as an inverting amplifier and becomes an integrator at the frequencies at which the reactance of the capacitor becomes

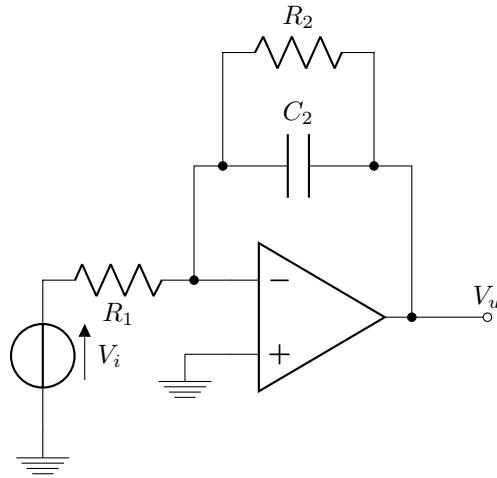


Figure 3.3: Modified integrator circuit

significantly lower than R_2 . We may study the transfer function considering the resistor in parallel to the capacitor as a single impedance Z_2 .

$$Z_2 = R_2 // \frac{1}{sC} = \frac{R_2/(sC)}{R_2 + 1/(sC)} = \frac{R_2}{1 + sR_2C}$$

$$\Rightarrow \frac{V_u}{V_i} = -\frac{Z_2}{Z_1} = -\frac{R_2}{R_1} \cdot \frac{1}{1 + sR_2C}$$

There is a pole at frequency $f_2 = 1/(2\pi R_2 C)$ and a decade above that frequency the circuit operates as an integrator.

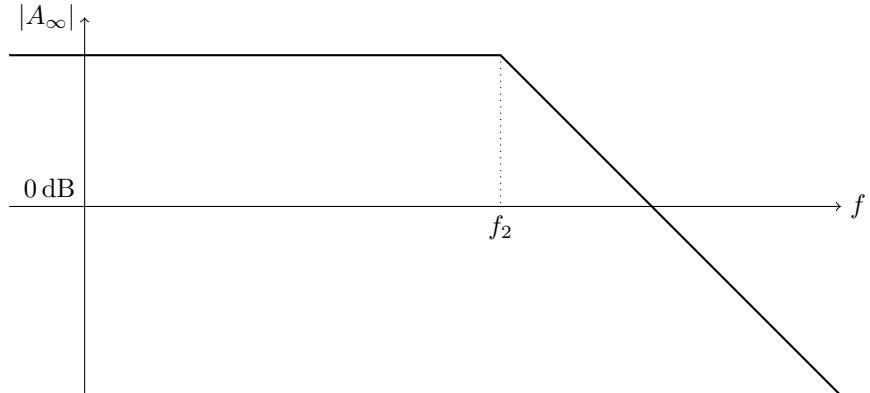


Figure 3.4: Bode diagram of the modified integrator.

The Bode diagram of the absolute value of A_∞ is shown in figure 3.4.

3.1.2 Differentiator

The circuit schematic of the differentiator, shown in figure 3.5, is dual with respect to the integrator. The differentiator transfer function is:

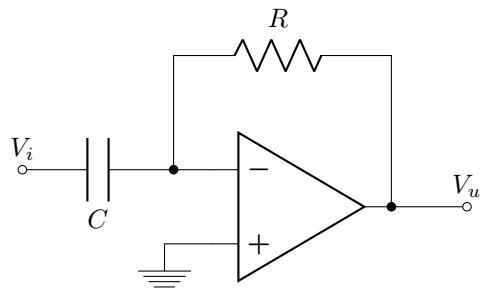


Figure 3.5: Differentiator circuit

$$\frac{V_u}{V_i} = -sRC$$

The Bode diagram of this function has a zero at zero frequency (DC) and a gain which increases proportionally to the frequency as shown in figure 3.6. The unity gain is at $f_0 = 1/(2\pi RC)$.

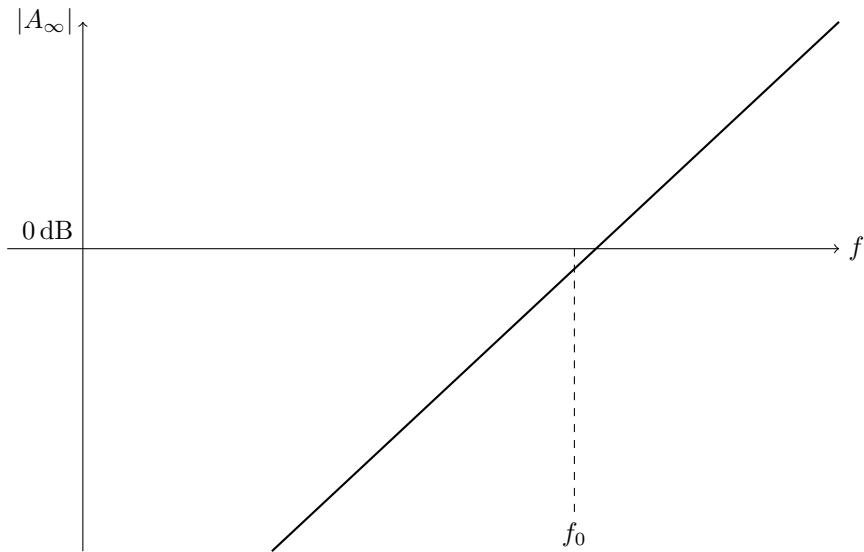


Figure 3.6: Ideal differentiator transfer function.

The operational amplifier has a limited bandwidth and will not implement correctly the transfer function beyond a certain frequency. Moreover if we analyze the circuit stability we see that the phase margin is reduced with respect to voltage follower configuration and the circuit is close to oscillating.

The stability problems may be solved by adding a resistor in series to the capacitor, as shown in figure 3.7, but this reduces the gain for high frequencies and therefore the operational frequency range of the differentiator.

The transfer function of the modified circuit may be evaluated as the ratio of the two impedances:

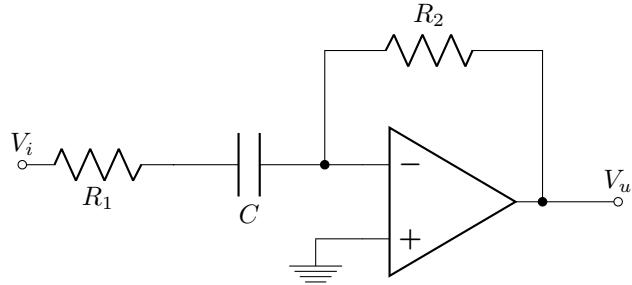


Figure 3.7: Modified differentiator circuit.

$$\frac{V_u}{V_i} = -\frac{Z_2}{Z_1} = -\frac{R_2}{1/(sC) + R_1} = -\frac{sR_2C}{1 + sR_1C}; \quad f_p = \frac{1}{2\pi R_1 C}$$

There is a zero in the origin and a pole at frequency f_p . This circuit will be analyzed again when studying filter circuits.

In this case there is a reduction of the phase shift and therefore of the stability problems.

3.1.3 Inverting adder

An adder is a circuit which provides an output which is the linear combination of a number of independent input signals. Figure 3.8 shows the schematic of an inverting adder with two input signals.

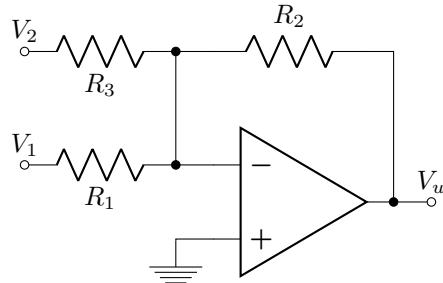


Figure 3.8: Circuit schematic of the inverting adder.

We start from the inverting amplifier and we add, for each input signal, a branch composed of the signal generator and a resistor connected to the inverting input. Since this node is at (virtual) ground in each resistor there will be a current proportional to the corresponding generator and the sum of all currents will flow into the R_2 resistor and will be converted into an output voltage. To analyze the circuit we can use the superposition principle.

Let us consider only signal V_1 , after zeroing all the other ones, as shown in figure 3.9. The non inverting input is at 0V and the inverting input will be also at (virtual) ground. The voltage across R_3 will be 0 and, by Ohm's law, also the current through it. In fact it is as if R_3 would not exist and we may consider the circuit as a simple inverting amplifier with transfer function :

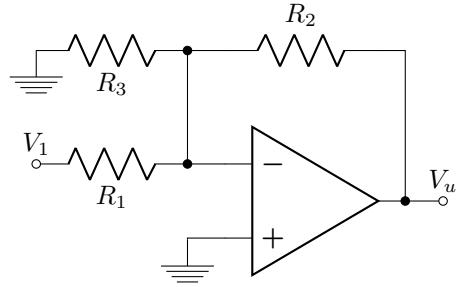


Figure 3.9: Inverting adder with only one input signal

$$\left. \frac{V_u}{V_i} \right|_{V_1} = -\frac{R_2}{R_1}$$

Repeating the same analysis for V_2 , connected to resistor R_3 , we have:

$$\left. \frac{V_u}{V_i} \right|_{V_2} = -\frac{R_2}{R_3}$$

And using the superposition principle:

$$V_u = -\frac{R_2}{R_1} \cdot V_1 - \frac{R_2}{R_3} \cdot V_2$$

This result may be extended to any number of input signals.

3.2 Differential amplifier

Having implemented the inverting adder let us see how can we design a circuit which amplifies, by a known and precise amount, the difference of two signals:

$$V_u = K(V_1 - V_2)$$

We can start by the circuit of figure 3.10, that is an inverting amplifier with an additional signal at the non inverting input.

Considering each signal independently we have the following output signals:

$$V_u|_{V_1} = \left(1 + \frac{R_2}{R_1}\right) \cdot V_1$$

$$V_u|_{V_2} = -\frac{R_2}{R_1} \cdot V_2$$

and superposing the effects we obtain:

$$V_u = \left(1 + \frac{R_2}{R_1}\right) V_1 - \frac{R_2}{R_1} V_2 = K_1 V_1 - K_2 V_2$$

We obtain a linear combination of signals V_1 and V_2 but not the desired one because $K_1 \neq K_2$ and in particular $K_1 = K_2 + 1$

We need to add other circuit elements and additional degrees of freedom in order to be able to obtain the desired gains. Since $K_1 > K_2$ it is necessary to

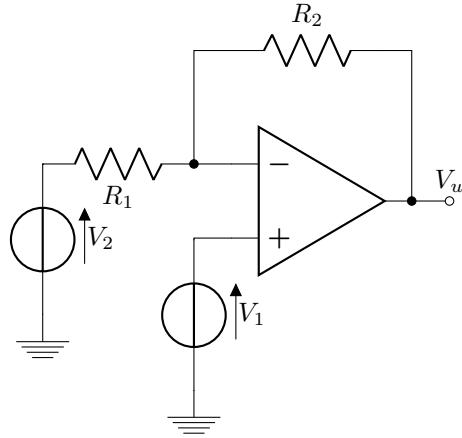


Figure 3.10: Circuit summing two signals with opposite signs

reduce the gain of V_1 and may do it by a voltage divider implemented by two resistors as shown in figure 3.11.

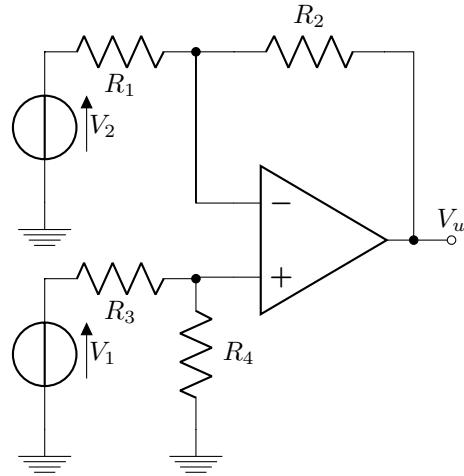


Figure 3.11: Differential Amplifier

Using again the superposition principle we will have:

$$V_+ = V_1 \cdot \frac{R_4}{R_3 + R_4}$$

$$V_u = V_+ \left(1 + \frac{R_2}{R_1} \right) - \frac{R_2}{R_1} V_2 = V_1 \cdot \frac{R_4}{R_3 + R_4} \left(1 + \frac{R_2}{R_1} \right) - \frac{R_2}{R_1} V_2$$

To obtain the same K , V_1 and V_2 coefficients must be equal and therefore::

$$\frac{R_4}{R_3 + R_4} \left(1 + \frac{R_2}{R_1} \right) = \frac{R_2}{R_1} \rightarrow \frac{R_4}{R_3 + R_4} = \frac{R_2}{R_1} \cdot \frac{R_1}{R_1 + R_2} = \frac{R_2}{R_1 + R_2}$$

If the equality is verified it will be true also for the reciprocal values:

$$\frac{R_4 + R_3}{R_4} = \frac{R_2 + R_1}{R_2} \rightarrow 1 + \frac{R_3}{R_4} = 1 + \frac{R_1}{R_2}$$

Hence the condition for which the amplifier is differential, that is amplifies only the difference of the two input signals, is:

$$\frac{R_1}{R_2} = \frac{R_3}{R_4}$$

If we want to cancel the effect of the bias currents then the above condition becomes:

$$\begin{cases} R_1 = R_3 \\ R_2 = R_4 \end{cases}$$

This simple amplifier is the base for a large family of amplifiers, named instrumentation amplifiers, which will be analyzed in the following. Its main problem is the low input impedance for both signals which makes the output dependent on the internal resistance of the input signal generators which are added respectively to R_1 and R_3 .

Common mode gain

An important parameter of the differential amplifier is the common mode rejection because if *the signal at the two inputs is the same* we would ideally expect the output of a good differential amplifier to be 0.

Let us consider the circuit of figure 3.12.

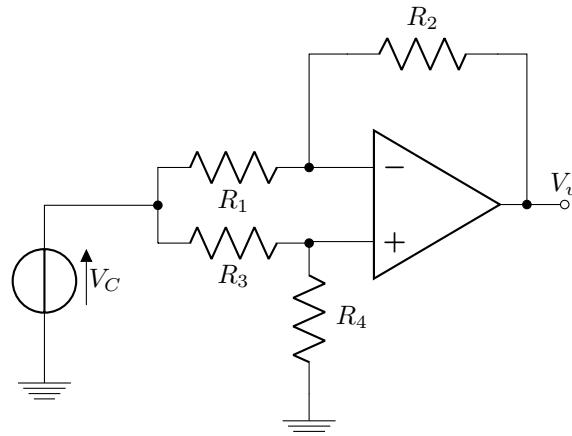


Figure 3.12: Differential amplifier with common mode input signal

If we apply an identical signal V_C to both inputs we will have a current in R_1 equal to:

$$I_1 = \frac{1}{R_1} \left[V_C - \frac{R_4}{R_4 + R_3} V_C \right] = V_C \frac{R_3}{R_4 + R_3} \cdot \frac{1}{R_1}$$

The output voltage V_u will be

$$V_u = \frac{R_4}{R_4 + R_3} V_C - I_2 R_2$$

Where I_2 is the current in R_2 and since there is no input current in the operational amplifier then $I_2 = I_1$; substituting the value we have:

$$\begin{aligned} V_u &= \frac{R_4}{R_3 + R_4} V_C - \frac{R_2}{R_1} \cdot \frac{R_3}{R_3 + R_4} V_C = \\ &= \frac{R_4}{R_3 + R_4} \left(1 - \frac{R_2}{R_1} \cdot \frac{R_3}{R_4} \right) V_C \end{aligned}$$

The common mode gain A_C is therefore:

$$A_C = \frac{V_u}{V_C} = \left(\frac{R_4}{R_3 + R_4} \right) \cdot \left(1 - \frac{R_2}{R_1} \cdot \frac{R_3}{R_4} \right) \quad (3.1)$$

If we suppose to have exactly $R_1 = R_3$ and $R_2 = R_4$, then the common mode gain would be minimized and ideally reduced to 0, but since resistor tolerance is limited, this condition will never be exactly met and A_C will be different from 0.

3.3 Non inverting adder

We may also implement the sum of several signals using a non inverting amplifier. In the case of three signals the schematic is shown in figure 3.13.

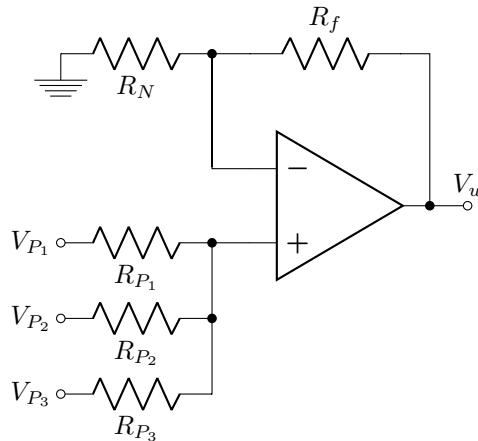


Figure 3.13: Non inverting adder

We again use the superposition principle. For signal V_{P1} with $V_{P2} = V_{P3} = 0$, we have:

$$V_u|_{V_{P1}} = \left(1 + \frac{R_f}{R_N} \right) \cdot \frac{R_{P2}/R_{P3}}{R_{P1} + R_{P2}/R_{P3}} \cdot V_{P1}$$

The contribution of the other signals is similar to this one. Defining R_P as:

$$R_P = R_{P_1} // R_{P_2} // R_{P_3}$$

and A_N as:

$$A_N = \frac{R_f}{R_N}$$

we may rewrite the gain for signal V_{P_1} as:

$$V_u|_{V_{P_1}} = (A_N + 1) \cdot \frac{R_P}{R_{P_1}} \cdot V_{P_1}$$

The full transfer function is then:

$$V_u = (A_N + 1) \cdot \left(\frac{R_P}{R_{P_1}} \cdot V_{P_1} + \frac{R_P}{R_{P_2}} \cdot V_{P_2} + \frac{R_P}{R_{P_3}} \cdot V_{P_3} \right)$$

With respect to the inverting adder, in which the gain of each signal is independent of the other ones, in the non inverting adder varying a resistance changes the gain of all the signals.

Let us now analyze the procedure to determine the relation among the resistances, given the general specification of a non inverting adder as:

$$V_u = K_{P_1} V_{P_1} + K_{P_2} V_{P_2} + K_{P_3} V_{P_3}$$

The signal coefficient are:

$$\begin{aligned} K_{P_1} &= (A_N + 1) \cdot R_P / R_{P_1} \\ K_{P_2} &= (A_N + 1) \cdot R_P / R_{P_2} \\ K_{P_3} &= (A_N + 1) \cdot R_P / R_{P_3} \end{aligned}$$

Summing the three equations we get:

$$A_N + 1 = K_{P_1} + K_{P_2} + K_{P_3}$$

And dividing them pairwise we obtain the relations which must hold between the resistances:

$$\frac{R_{P_3}}{R_{P_1}} = \frac{K_{P_1}}{K_{P_3}}$$

and

$$\frac{R_{P_2}}{R_{P_1}} = \frac{K_{P_1}}{K_{P_2}}$$

The procedure may be easily extended to any number of signals.

In the previous chapter we found that to minimize offsets contribution to the output of the amplifier it is necessary that the resistor placed in series to the non inverting input of the operational amplifier (in that case R_3) be equal to the parallel of the feedback resistor (that we called R_2) and the resistor from inverting input to ground (in that case R_1).

That result applies also in the case of our adder, because offsets are independent inputs and to compute their contribution we should short circuit to ground all other circuit inputs. In practice we should have:

$$R_P = R_f \parallel R_N$$

Applying this equation we have a very simple solution to the problem of assigning values to our resistors. We can rework this equation to obtain:

$$R_P = \frac{1}{\frac{1}{R_f} + \frac{1}{R_N}} = \frac{R_f}{1 + \frac{R_f}{R_N}} = \frac{R_f}{A_N + 1}$$

The expression of the gain of the adder can be written by collecting R_P :

$$V_u = (A_N + 1) \cdot R_P \left(\frac{1}{R_{P_1}} \cdot V_{P_1} + \frac{1}{R_{P_2}} \cdot V_{P_2} + \frac{1}{R_{P_3}} \cdot V_{P_3} \right)$$

But we can express R_P as function of R_f and finally obtain:

$$V_u = R_f \left(\frac{1}{R_{P_1}} \cdot V_{P_1} + \frac{1}{R_{P_2}} \cdot V_{P_2} + \frac{1}{R_{P_3}} \cdot V_{P_3} \right)$$

So we obtain an optimized circuit for offset contribution if we set:

$$R_{P_i} = R_f / K_{P_i}$$

and

$$R_N = \frac{R_f}{\left(\sum_{j=1}^p K_{P_j} \right) - 1}$$

Example 3. Let us design a non inverting adder with a real operational amplifier to obtain:

$$V_u = 3V_1 + 2V_2 + V_3$$

We use $R_f = 100 \text{ k}\Omega$.

The solution to this problem is very simple. The sum of the three gains is 6. We must have therefore:

$$R_N = \frac{100 \text{ k}\Omega}{6 - 1} = 20 \text{ k}\Omega$$

The resistors connected to the three inputs are:

$$\begin{aligned} R_{P_1} &= R_f / K_{P_1} = 33 \text{ k}\Omega \\ R_{P_2} &= R_f / K_{P_2} = 50 \text{ k}\Omega \\ R_{P_3} &= R_f / K_{P_3} = 100 \text{ k}\Omega \end{aligned}$$

3.4 Generalized adder

The adder circuit may be made even more general by adding input signals at the amplifier inverting input as shown in figure 3.14. Considering a system with n inputs on the inverting side and p inputs on the non inverting side we again obtain the transfer function by the superposition principle.

The gains of signals on the inverting side are:

$$V_u|_{V_{N_i}} = \frac{R_f}{R_{N_i}} \cdot V_{N_i}$$

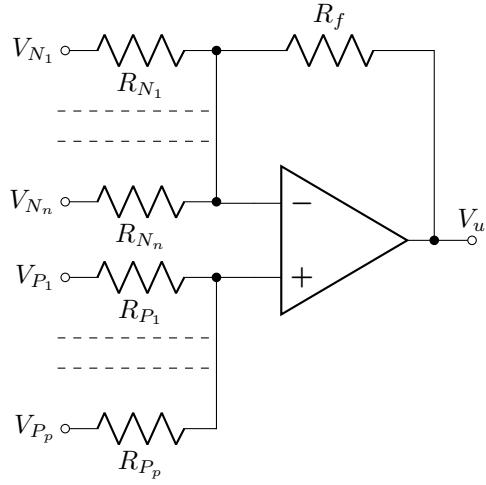


Figure 3.14: Generalized adder

and for inputs on the non inverting we can use the result derived in section 3.3 considering for R_N :

$$R_N = R_{N_1} // R_{N_2} // \dots // R_{N_n}$$

The full transfer function will be:

$$V_u = (A_N + 1) \cdot \sum_{j=1}^p \frac{R_P}{R_{P_j}} \cdot V_{P_j} - \sum_{i=1}^n \frac{R_f}{R_{N_i}} \cdot V_{N_i}$$

Let us now analyze the design methodology of a generalized adder whose expression is:

$$V_u = \sum_{j=1}^p K_{P_j} \cdot V_{P_j} - \sum_{i=1}^n K_{N_i} \cdot V_{N_i}$$

Using the schematic of figure 3.14 we may not satisfy all design conditions for positive and negative inputs because we miss a degree of freedom. In fact we have for inverting inputs:

$$\frac{R_f}{R_{N_i}} = K_{N_i}$$

and it can be easily shown that:

$$A_N = \frac{R_f}{R_N} = \sum_{i=1}^n K_{N_i}$$

but defining

$A_P = \sum_{i=1}^p K_{P_i}$, from what computed in section 3.3 we know that

$$A_N + 1 = A_P$$

This relation establishes a link between the sum of the *inverting gains* and the sum of the *non-inverting gains*. In general the specification of an adder does not satisfy the relation but it is easy to solve the problem.

Suppose we have $A_N < A_P - 1$ it means that we have to increase the sum of the inverting gains without modifying the gain of each signal to be added. We can do that by adding a resistance R_{N_x} between the inverting input and 0 V. It is like adding a 0 V signal with a R_f/R_{N_x} gain.

The relation between positive and negative gains holds if:

$$\frac{R_f}{R_{N_x}} = K_{N_x} = A_P - A_N - 1$$

If we have $A_N > A_P - 1$ we must increase the sum of the non inverting gains by adding a resistance R_{P_x} between the non inverting input and 0V. It is like adding a null input signal with positive gain. We must impose:

$$\frac{R_P}{R_{P_x}} = K_{P_x} = A_N - 1 - A_P$$

Although the generalized adder circuit is formally very simple, its major drawback is the fact that the non inverting gains are not independent of each other. An often preferred implementation is to use two cascaded inverting adder stages.

What we observed for the non inverting adder, about the optimal solution to minimize offsets contributions to the output of the circuit and the subsequent very simple method to select resistor values, can be obviously extended to the generalized adder. Focusing on the final scheme with added resistor to inverting or non inverting input (if necessary), we can easily find that we must obey the equation:

$$R_P = \frac{1}{\frac{1}{R_f} + \frac{1}{R_N}} = \frac{R_f}{1 + \frac{R_f}{R_N}} = \frac{R_f}{A_N + 1} = \frac{R_f}{A_P}$$

wherer R_N is the parallel of the resistors placed on the inverting side of the operational amplifier as defined above and including also the additional resistor R_{N_x} (if there).

Also in this case we have that:

$$R_{P_j} = R_f/K_{P_j}$$

extending this definition also to the eventual additional resistor R_{P_x} .

Example 4. Design a generalized adder with an LM741 operational amplifier with supply voltage of ± 15 V. The transfer function has to be:

$$V_u = 4V_1 + 6V_2 - 3V_3$$

Let us start by computing the inverting and non inverting gains. In our case we have $A_P = 10$ and $A_N = 3$. We must thus add a resistor R_{N_x} on the inverting input to balance the gains to obtain $A_N = 9$. The final transfer function is:

$$V_u = 4V_1 + 6V_2 - 3V_3 - 6 \cdot 0 \text{ V}$$

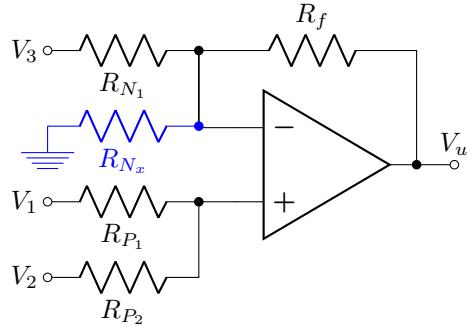


Figure 3.15: Schematic of the example circuit.

The schematic of the circuit is depicted in figure 3.15. To size the resistors we can note that, being the gain $A_P = 10$, we are in the same situation of paragraph 2.4. Therefore, we can set $R_f = 22 \text{ k}\Omega$. The value of the other resistors is:

$$\begin{aligned} R_{N_1} &= R_f/3 = 7.33 \text{ k}\Omega \\ R_{N_x} &= R_f/6 = 3.67 \text{ k}\Omega \\ R_{P_1} &= R_f/4 = 5.5 \text{ k}\Omega \\ R_{P_3} &= R_f/6 = 3.67 \text{ k}\Omega \end{aligned}$$

Depending on the required precision of the gains, we can select the normalized resistors closest to the nominal ones using the relevant *Exx* series.

Example 5. Design a generalized adder with the following transfer function:

$$V_u = 4V_1 - 6V_2 - 2V_3$$

The value of the feedback resistor is given: $R_f = 50 \text{ k}\Omega$.

Let us compute the inverting and non inverting gains. We have $A_P = 4$ and $A_N = 8$. Therefore we need to add a resistor R_{P_x} on the non inverting side to balance the gain so that we obtain $A_P = 9$. We modify the transfer function so that:

$$V_u = 4V_1 - 6V_2 - 2V_3 + 5 \cdot 0 \text{ V}$$

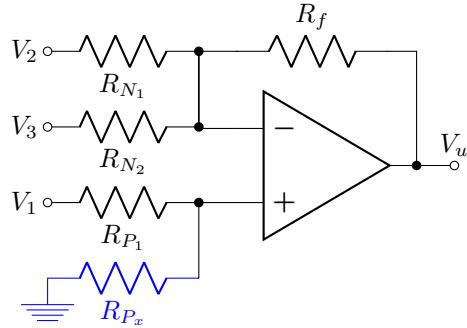


Figure 3.16: Schematic diagram of the example circuit.

Figure 3.16 shows the schematic of the circuit. Given that $R_f = 50 \text{ k}\Omega$, the value of the other resistors is:

$$\begin{aligned}R_{N_1} &= R_f/6 = 16.7\text{k}\Omega \\R_{N_2} &= R_f/2 = 25\text{k}\Omega \\R_{P_1} &= R_f/4 = 12.5\text{k}\Omega \\R_{P_3} &= R_f/5 = 10\text{k}\Omega\end{aligned}$$

Depending on the required precision, we can select the normalized resistors closest to the nominal ones using the relevant *Exx* series.

3.5 Instrumentation amplifiers

3.5.1 Introduction

We have already analyzed the circuit implementing a *differential amplifier* that is an amplifier with the transfer function of equation 3.2:

$$\frac{V_u}{V_i} = K(V_1 - V_2) \quad (3.2)$$

We want now to design a better version because these amplifiers are fundamental for a number of applications, from measurement systems to communication systems. A common example is a line receiver over a communication line. The electromagnetic noise in the environment degrades the signal and the use of a differential amplifier allows us to reject all the common mode noise components introduced by the connection.

Before analyzing the possible circuit solutions let us first briefly introduce the problems related to electromagnetic noise introduced by interconnections between system components limiting our analysis to low frequencies.

3.5.2 Electromagnetic compatibility

In an electronic system, such a telecommunication system or the embedded electronics in a car, there are several subsystems or *modules* which communicate among each other by interconnections, often very long, and signals are transmitted over them.

Each signal is a voltage measured with respect to a common *reference potential* or 0 V node which is assumed to be equipotential throughout the whole system.

Besides the reference potential there are other relevant electrical nodes, often referred to as *grounds* such as the metallic body of the equipment or other conducting bodies attached to it. Often these nodes do not have any functional purpose but are important for electromagnetic compatibility. In most cases, for security reasons they are connected to *earth*, that is the reference potential for mains power supply. They may be directly connected or not to the 0 V reference potential, but in any case since there is always a parasitic capacitive coupling between the circuit 0 V and these electrical nodes there will always a frequency at which these nodes may be considered short circuited to the 0 V.

In general grounds, wires and other interconnecting elements cannot be modeled as ideal short circuits but have a number of non ideal undesired parasitic elements such as resistances, inductances and capacitances. As an example the metallic frame of a car, when passing close to a high voltage pylon is subjected

to a low frequency electromagnetic field due to current flowing in the high voltage conductor. A current may be induced in the car frame and that current due to the frame resistance will generate voltage differences among the different parts. If the frame is used as a reference potential node than we may model the disturbance as a noise voltage generator V_N between different parts of the reference node. In the same way other currents and corresponding voltage drops may be induced on the frame by electric motors in the car or power circuits related to ignition or other functions. Mobile phones or other radio-frequency sources may inject noise on the frame.

There are methodologies which allow to reduce and cancel problems related to electromagnetic compatibility such as the study of the connections of grounds to the reference potential, the reduction of parasitic parameters and coupling and others. We will focus now on a specific simple and basic problem such as the interconnection of a remote sensor which must be read by an amplifier at some distance from it. The sensor transforms the physical variable to be measured into an equivalent voltage referred to its 0V. The sensor and the amplifier have a reference potential which is only nominally common since, as we have seen before, external electromagnetic noise may change its value. Therefore if we connect the sensor to the amplifier with a simple wire, any difference in reference potential will appear as a signal difference. Moreover the loop formed by the wire and the return through the reference potential connection is an antenna which collects further electromagnetic noise.

A solution to this problem is to perform a differential measurement using a two wire (possibly shielded) connection from the sensor to the amplifier using a twisted pair to reduce the electromagnetic coupling. In such a way the amplifier will perform a differential measurement and electromagnetic noise will appear as a common mode signal to its input. Its is therefore important that the amplifier has a high CMRR.

3.5.3 Instrumentation amplifiers circuits

For this reason a number of high performance differential amplifier circuits have been developed, which are in general referred to as *instrumentation amplifiers*. Let us start from the simple differential amplifier of figure 3.17 analyzing its defects and how to improve it.

If we compare equation 3.2 with the generalized adder relation with only two input signals we obtain that, in order to have a differential gain K , we must set the condition of equations 3.3. The first equation sets the gain K and the other two minimize the contribution of the input bias current of the operational amplifier.

$$\left\{ \begin{array}{l} R_2/R_1 = R_4/R_3 = A_D \\ R_2 = R_4 \\ R_1 = R_3 \end{array} \right. \quad (3.3)$$

The main problem of this circuit is that the input impedances for the two signals are different and not very high. For V_1 it is $R_3 + R_4$ and for V_2 is R_1 . This means that the differential gain will depend on the value of the source resistances of the two input generators. In fact if we define R_{eq1} the equivalent resistance of generator V_1 and R_{eq21} that of generator V_2 the gains of the two signals become:

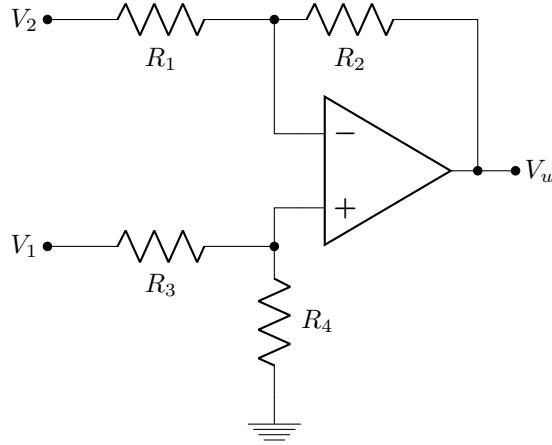


Figure 3.17: A simple differential amplifier.

$$K_2 = \frac{R_2}{R_1 + R_{eq2}}$$

$$K_1 = \left(1 + \frac{R_2}{R_1 + R_{eq2}} \right) \cdot \frac{R_4}{R_3 + R_4 + R_{eq1}}$$

and it is clear that the gains are lower than the required value and that if $R_{eq2} \neq R_{eq1}$ we have $K_1 \neq K_2$ and the common mode gain will be different from 0 and the CMRR will be degraded.

To solve the problem related to input impedances we can separate each input signal from the amplifier by means of a **buffer** stage with very high input impedance and very low output impedance such as a *voltage follower*. The circuit is shown in figure 3.18.

Gain regulation

The circuit of figure 3.18 is a good differential amplifier but with *fixed* gain due to the fact that resistances are integrated in order to obtain a high CMRR. If we want a high CMRR amplifier in which the gain may be modified by changing a single resistance value we may use the circuit of figure 3.19 where the voltage followers are substituted by two non inverting amplifiers which have a common resistance R in their feedback network.

The analysis of the circuit is very simple and will be carried on separately for the common and the differential mode voltages. If we apply a common mode signal V_C to both inputs and we assume that both operational amplifiers operate in the linear region with $v_d = 0$, then both inverting inputs will also be at V_C and the voltage across R will be 0. Therefore the current in R will also be 0 by Ohm's law as well as currents in R_a and R_b because no current enters the operational amplifier.

Since there is no voltage drop on R_a and R_b the output of both input amplifiers will be V_C and they behave as voltage followers. The common mode gain will then depend on the final differential amplifier which has the same common mode input voltage of the system.

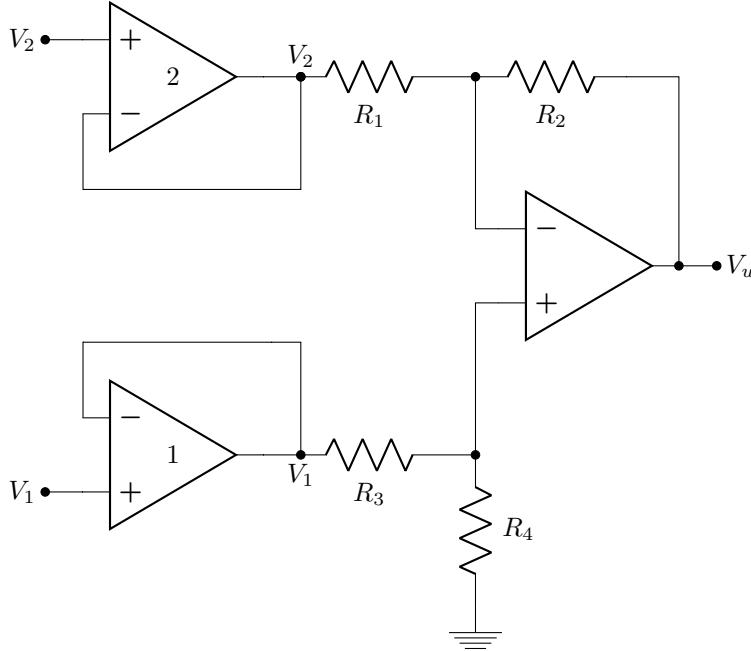


Figure 3.18: Differential amplifier with *voltage follower buffers*.

A differential input signal $v_d = V_1 - V_2$ will appear also across resistance R and the current I in it will be

$$I = \frac{V_1 - V_2}{R}$$

since there is no input current in the operational amplifiers the same current flows through R_a and R_b and the difference of the output voltages of the two amplifiers is the sum of the voltage drop across the three resistances:

$$V_{u1} - V_{u2} = \frac{V_1 - V_2}{R} \cdot (R_a + R + R_b) \quad (3.4)$$

This is also the differential input voltage to the final stage and combining equation 3.4 with the transfer function of the differential amplifier we obtain the full transfer function between input and output voltages. Keeping the relations 3.3 which optimize offsets and CMRR we have:

$$\frac{R_2}{R_1} = \frac{R_4}{R_3}$$

$$V_u = (V_{u1} - V_{u2}) \frac{R_2}{R_1} = \frac{R_a + R + R_b}{R} \cdot \frac{R_2}{R_1} \cdot (V_1 - V_2)$$

$$A_D = \frac{R_a + R + R_b}{R} \cdot \frac{R_2}{R_1} = \left(1 + \frac{R_a + R_b}{R}\right) \cdot \frac{R_2}{R_1} \quad (3.5)$$

The differential gain $A_D(R)$ is a function of R and may be easily modified.

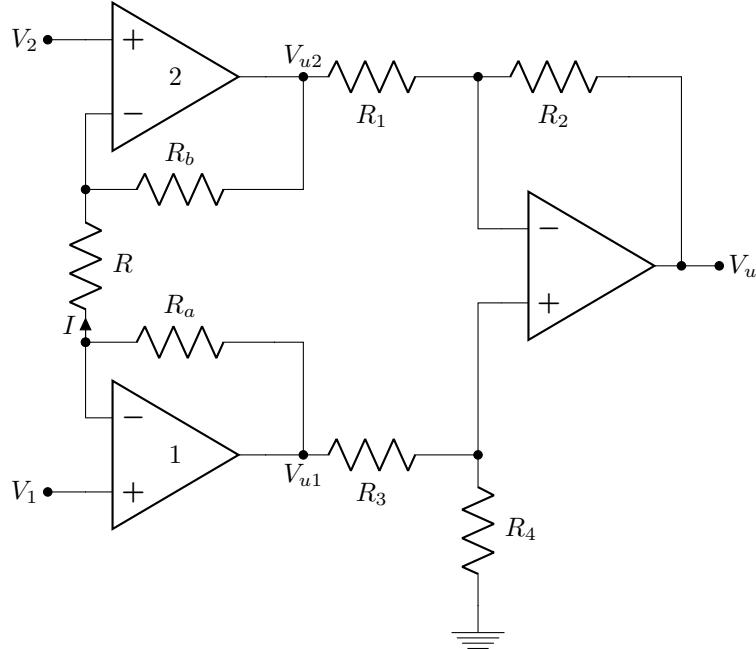


Figure 3.19: A real instrumentation amplifier

Offsets

Since the operational amplifiers are not ideal when the differential input is 0 the output will be different from 0 due to offset parameters. In general offset voltage are dominant over offset and bias currents and in this short analysis we will consider only them. The offset voltages of input operational amplifiers 1 and 2 are in the same position as the input signals and therefore their contribution to the output will be:

$$v'_{uoff} = A_D(|v_{off1}| + |v_{off2}|)$$

The contribution of the offset voltage of the third amplifier is:

$$v''_{uoff} = |v_{off3}| \frac{R_1 + R_2}{R_1}$$

In instrumentation amplifiers data sheet the term $|v_{off1}| + |v_{off2}|$ is often referred as *input offset voltage* or v_{ioff} and the term v''_{uoff} as *output offset voltage*. The total output offset is therefore given by:

$$v_{uoff} = A_D v_{ioff} + v''_{uoff}$$

Alternately the datasheet may specify the total offset at the *input* of the amplifier as:

$$v_{off} = v_{ioff} + \frac{v''_{uoff}}{A_D}$$

Final remarks

Commercial instrumentation amplifiers of this type are implemented by integrating the operational amplifiers and the R_1 , R_2 , R_3 , R_4 , R_a and R_b feedback resistances in the same *package* and having as external pins the power supplies, the inputs, the output, the R gain resistor terminals and the sense and reference terminals. Their name derives from the fact that it is largely used in all measurement and amplification systems processing signals from transducers where it is important to have a high rejection of the common mode noise. An example of such amplifier is the INA114 components, a general purpose instrumentation amplifier which may provide a gain $1 < G < 10^4$, has a $CMRR = 115\text{dB}$ @ $G = 1000$. The gain may be set with a single resistor according the the following equation which corresponds to the one (eq.3.5) which has been previously derived.

$$G = 1 + \frac{50\text{k}\Omega}{R}$$

3.6 Audio amplifiers

In all the application circuits of the operational amplifier analyzed until now (except the differentiator) the pass-band starts from 0, that is DC. However there are a number of applications in which this is not necessary, or in which it is required to block the DC component of the signal which could limit the output voltage range of the signal. In general we refer to as *audio amplifiers* those circuits which operate over a frequency band which does not include DC but is not as wide to be classified as radiofrequency.

The nominal audio band spans from 20 Hz to 20 kHz for high quality music and from 300 Hz to 3200 Hz for telephone signals but the techniques and the the circuits for these kind of amplifiers may be used also at lower and higher frequency, obviously modifying accordingly the circuit time constants.

The DC signal components may be canceled using the techniques used also in single transistor amplifier stages, but it is important to play attention to some points which will be outlined in the following analysis of a non inverting amplifier for the audio band.

The circuit schematic is shown in figure 3.20. and at the input pin there is capacitor C_3 which blocks any signal DC component and resistor R_3 whose presence is *necessary* because a DC path must always be provided from the power supplies or the reference voltage node to the amplifier input in order to provide the input bias currents. These two components introduce a zero at $f = 0$ and a pole at low frequency.

Capacitor C_1 decouples the in band behaviour of the amplifier from the DC one. In DC C_1 is an open circuit and therefore the amplifier gain is unitary. It is important to reduce the DC gain of the amplifier, even if no DC input signal may be present, in order to reduce the effects of *internal* DC components. In this case V_{off} component has unity gain and if we set $R_3 = R_2$ then the contribution of the input bias currents will be minimized. C_1 , R_1 and R_2 introduce a zero-pole pair in the transfer function and the pole frequency should be set at a decade lower than lower limit of the pass band, set by R_3 and C_3 in order not to modify its value.

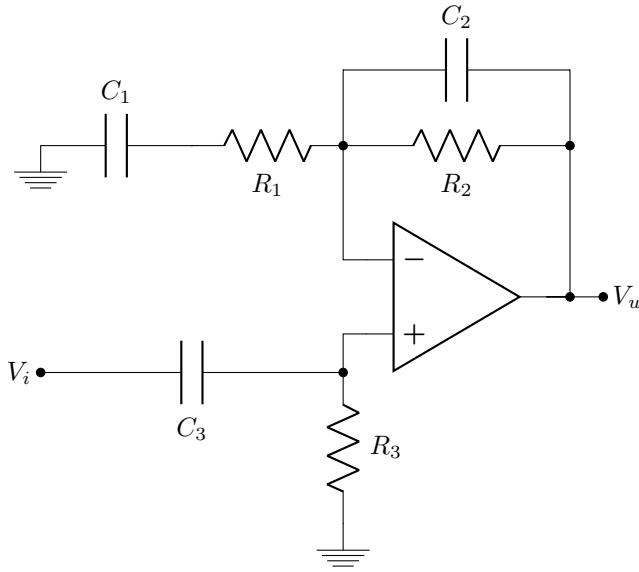


Figure 3.20: Circuit schematic of the audio amplifier.

Capacitor C_2 operates at high frequency introducing a pole which defines the upper limit of the pass band. This is useful to reduce the noise outside the signal band which is not amplified and transmitted to following stages. C_2 introduces also a zero and after the zero frequency the capacitor may be considered a short circuit and the amplifier will have unity gain.

Example 6. Design a non inverting audio amplifier according to the following specifications:

- -3 dB pass band: from 20 Hz to 20 kHz ;
- in band gain: $10 \pm 10\%$;
- use of normalized components from the E12 series;
- operational amplifier: LM741;
- input impedance: $100 \pm 10\text{ k}\Omega$;
- power supplies: $\pm 15\text{ V}$.

The in band design specifications are that of a standard non inverting amplifier and this sets the ratio between R_1 and R_2 . The specification on the input impedance sets the value of R_3 to $100\text{ k}\Omega$ and the minimization of the effects of the bias current will set $R_2 = R_3$. At this point also the value of R_1 will be set and the resistors will have the following values:

$$\begin{cases} R_2 = 100\text{ k}\Omega \\ R_1 = 10\text{ k}\Omega \\ R_3 = 100\text{ k}\Omega \end{cases}$$

The in band nominal value will be 11 which is within the specification limits. Let us now define the value of the capacitors. Since C_2 operates at frequencies

far higher than C_3 and C_1 , we may define its value considering the others two as short circuits. The high frequency transfer function is:

$$\frac{V_u}{V_i} = \frac{R_1 + R_2}{R_1} \cdot \frac{s(R_1/R_2)C_2 + 1}{sR_2C_2 + 1}$$

We may evaluate the pole frequency as:

$$f_{p2} = \frac{1}{2\pi R_2 C_2} \implies C_2 = \frac{1}{2\pi \cdot 20 \text{ kHz} \cdot 100 \text{ k}\Omega} = 80 \text{ pF}$$

Let us set the normalized value $C_2 = 68 \text{ pF}$.

It is now necessary to set the value of the two capacitors which are responsible for the low frequency behaviour. Their behaviour may be evaluated independently and then added to the final transfer function of the amplifier. Let us first consider C_1 , evaluating the transfer function from the non inverting input of the amplifier to the output as:

$$\frac{V_u}{V_+} = \frac{s(R_1 + R_2)C_1 + 1}{sR_1C_1 + 1}$$

Hence the frequency of the pole introduced by C_1 is:

$$f_{p1} = \frac{1}{2\pi R_1 C_1}$$

Regarding C_3 if we assume that the input impedance of the operational amplifier is very high the voltage on the non inverting input will be:

$$\frac{V_+}{V_i} = \frac{sR_3C_3}{sR_3C_3 + 1}$$

and the frequency of the pole introduced by C_3 will be:

$$f_{p3} = \frac{1}{2\pi R_3 C_3}$$

We must decide which of the two poles will be the low frequency limit of the amplifier, the other one been placed at a frequency one decade lower in order not to modify the value of the -3 dB frequency. Since R_3 is larger than R_1 it is convenient to assign it to the lower frequency in order to reduce the value of the associated capacitor. We may then write:

$$C_1 = \frac{1}{2\pi \cdot 20 \text{ Hz} \cdot 10 \text{ k}\Omega} = 800 \text{ nF}$$

$$C_3 = \frac{1}{2\pi \cdot 2 \text{ Hz} \cdot 100 \text{ k}\Omega} = 800 \text{ nF}$$

Then both C_1 and C_3 will be have the normalized 820 nF value.

Finally we may recompute the pole and zero frequencies, using the normalized values of the capacitors and draw the complete Bode diagram of the transfer function of the amplifier which is shown in figure 3.21.

$$f_{z1} = 1.8 \text{ Hz}; \quad f_{p3} = 1.9 \text{ Hz}; \quad f_{p1} = 19 \text{ Hz}; \quad f_{p2} = 23 \text{ kHz}; \quad f_{z2} = 260 \text{ kHz}.$$

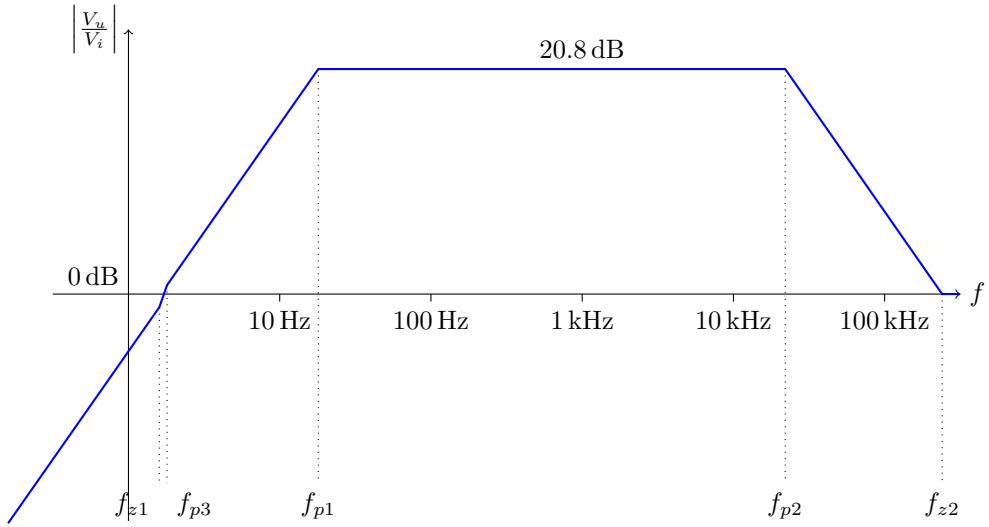


Figure 3.21: Bode diagram of the absolute value of the transfer function of the audio amplifier.

3.7 Single supply circuits

Until now we have assumed that the power supply of the operational amplifier is symmetric, that is with a positive voltage V_{AL} with respect to the reference voltage and a negative one $-V_{AL}$, but in many applications the system has a single positive power supply and the negative supply pin is connected to the reference voltage.

All the circuits previously described may operate with a single supply at the cost of reducing both the input and the output voltage range: as an example using the $V_{AL} = 15$ V supply instead of the pair $V_{AL} = \pm 15$ V the amplifier will operate correctly but the output range will be halved (no negative output voltages) and the input range will also be at least halved.

A circuit with a single positive supply may have negative input voltage in two cases:

- if the input is AC coupled and the input is properly biased;
- in case of an inverting amplifier it is possible to have negative input values, preferably adding protection circuits operating in case the power supply is missing.

In the following we will examine two circuits with single power supply.

Example 7. Let us design an amplifier according to the following specifications:

- Input voltage range: 1 V–2 V;
- Output voltage $V_u = 0.5$ V when $V_i = 1$ V;
- Output voltage $V_u = 4.5$ V when $V_i = 2$ V;
- $V_{AL} = 5$ V, $-V_{AL} = 0$ V;

- Rail-to-rail operational amplifier LMH6645;
- For resistors: normalized values of the E96 series.

The first step is to determine graphically the transfer function of the circuit, which is a linear segment as shown in figure 3.22.

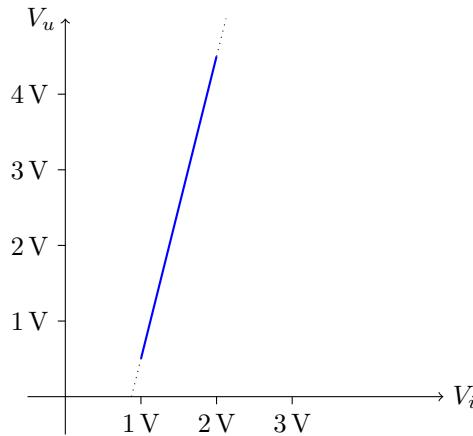


Figure 3.22: Transfer function according to the specification of example 7.

The analytical expression of the straight line comprising the segment may be derived by the coordinates of the two points belonging to it:

$$V_u = 4(V_i - 0.875 \text{ V}) = 4V_i - 3.5 \text{ V}$$

In order to implement the -3.5 V constant value we connect to the inverting input as a signal through resistor R_4 the power supply V_{AL} .

The transfer function may be determined using the superposition principle and is given by:

$$V_u = V_i \left(1 + \frac{R_2}{R_1//R_4} \right) - V_{AL} \cdot \frac{R_2}{R_4}$$

The R_2/R_4 ratio is set considering:

$$V_{AL} \cdot \frac{R_2}{R_4} = 3.5 \text{ V} \implies \frac{R_2}{R_4} = \frac{3.5 \text{ V}}{5 \text{ V}} = 0.7$$

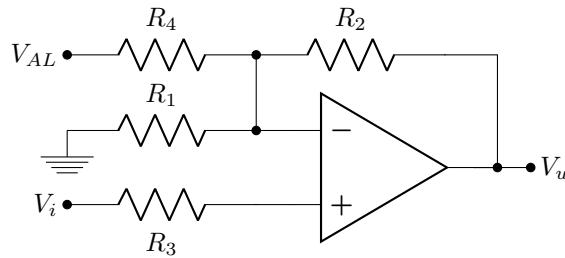


Figure 3.23: Amplifier circuit of example 7.

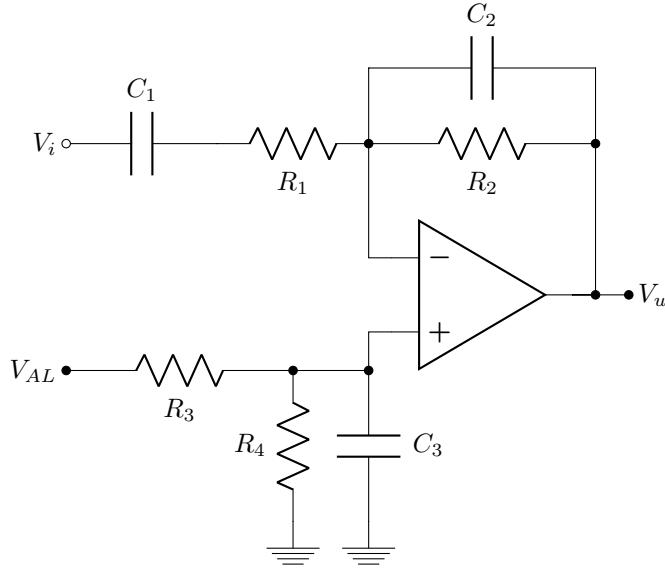


Figure 3.24: Audio amplifier of example 8.

The other condition is:

$$1 + \frac{R_2}{R_1//R_4} = 4 \quad (3.6)$$

To determine the resistor values let us note that the typical offset voltage of the operational amplifier is 1 mV and offset current is 1 nA. Therefore:

$$R_2 \ll \frac{A_V \cdot V_{off}}{I_{off}} = \frac{4 \cdot 1 \text{ mV}}{1 \text{ nA}} = 4 \text{ M}\Omega$$

Let us then chose $R_2 = 100 \text{ k}\Omega$. This gives $R_4 = 143 \text{ k}\Omega$ (a value existing in the E96 series). Substituting these values in the expression 3.6 we obtain $R_1 = 43.5 \text{ k}\Omega$. The closest normalized value in the E96 series is $R_1 = 43.2 \text{ k}\Omega$. R_3 value is the parallel of the other three resistances and corresponds to $24.9 \text{ k}\Omega$, a value which exists in the E96 series.

Example 8. Let us implement an inverting amplifier for a audio band signal in the frequency range from 20 Hz to 20 kHz with a gain $A_V = 10 \pm 10\%$ and single supply $V_{AL} = 5 \text{ V}$. We will use an LM324 operational amplifier and resistors of the normalized series E12.

The reference schematic is given in figure 3.24. Capacitors C_1 and C_2 limit the pass band respectively at low and high frequencies. Resistors R_3 and R_4 provide the input biasing which maximizes the output voltage range and capacitor C_3 reduces the noise which may be injected from the power supply into the non inverting input.

Let us first determine the output voltage range of the amplifier. The output voltage range of the LM324 amplifier with 5 V supply is approximately 0 V–3.5 V. The bias network composed of R_3 and R_4 determines the DC value of the output signal and the DC gain of the amplifier from the non inverting input is $A_{DC} = 1$. If we set the DC output voltage to half of the full output voltage

range, that is 1.75 V, then the voltage at the non inverting input is the same and therefore the ratio of the two resistors must satisfy the equation:

$$V_{AL} \cdot \frac{R_4}{R_3 + R_4} = 1.75 \text{ V}$$

Hence $R_3/R_4 = 1.86$. In order to minimize the input bias currents effect the parallel of these two resistors must be equal to the DC resistance seen from the inverting input, that is R_2 and therefore we have $R_3//R_4 = R_2$.

The passband gain is $A_V = -R_2/R_1 = 10$ and to set the R_2 value we may, as usual, consider the offset voltage and current which, from the data sheet are respectively equal to 2 mV and 5 nA. We have therefore:

$$R_2 << \frac{A_{DC} \cdot V_{off}}{I_{off}} = \frac{1 \cdot 2 \text{ mV}}{5 \text{ nA}} = 400 \text{ k}\Omega$$

We may then chose $R_2 = 39 \text{ k}\Omega$. This in turn gives for R_4 a value of $60 \text{ k}\Omega$ and the closest normalized value in the E12 series is $R_4 = 56 \text{ k}\Omega$. The value of R_3 will be $56 \text{ k}\Omega \cdot 1.86 = 104 \text{ k}\Omega$ which may be normalized to $R_3 = 100 \text{ k}\Omega$. At last the value of R_1 is set by the passband gain to $R_1 = 3.9 \text{ k}\Omega$.

We can now evaluate the value of the capacitors which may be analyzed separately because the two time constants are separated by more than one decade. Their contribution has been already discussed in section 3.1.1 for C_2 and in section 3.1.2 for C_1 . The frequency of the two poles is therefore given by:

$$\begin{aligned} f_{p2} &= \frac{1}{2\pi R_2 C_2} \implies C_2 = \frac{1}{2\pi 39 \text{ k}\Omega \cdot 20 \text{ kHz}} = 204 \text{ pF} \\ f_{p1} &= \frac{1}{2\pi R_1 C_1} \implies C_1 = \frac{1}{2\pi 3.9 \text{ k}\Omega \cdot 20 \text{ Hz}} = 2.04 \mu\text{F} \end{aligned}$$

The closest normalized values are $C_2 = 180 \text{ pF}$ and $C_1 = 2.2 \mu\text{F}$. (Choosing the normalized value in such a way that the passband is never smaller than the specifications).

The last component to be chosen is C_3 which is not strictly necessary but is useful to reduce the noise injected from the power supply into the non inverting input which is amplified by a factor

$$A_{V_{AL}} = \frac{1.75}{5} (1 + A_V) = 3.85$$

The type of noise depends on the power supply but the most important for an audio amplifier is the ripple at twice the frequency of the mains supply which may reach an amplitude of few tens of millivolt and therefore be a significant source of noise. The frequency of the pole due to C_3 is:

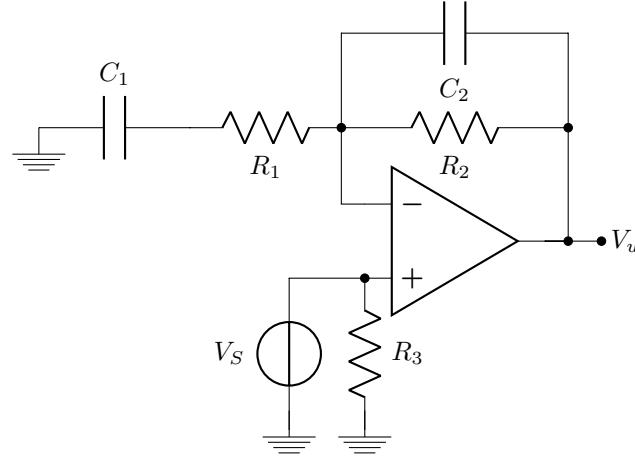
$$f_{p3} = \frac{1}{2\pi(R_3//R_4)C_3}$$

Imposing $f_{p3} = 2.5 \text{ Hz}$ a noise at the frequency of a 100 Hz will be reduced by 20 dB taking also into account the above mentioned amplification. This corresponds to $C_3 = 51 \mu\text{F}$ and the normalized closest value is $C_3 = 47 \mu\text{F}$. We may choose a electrolytic capacitor because the voltage across it has always a positive value.

3.8 Exercises

Exercise 3.1. Wide band bandpass amplifier

Given the circuit schematic of the bandpass amplifier of figure 3.25



$$R_1 = 1 \text{ k}\Omega \quad R_2 = 100 \text{ k}\Omega \quad C_1 = 1.8 \text{ }\mu\text{F} \quad C_2 = 180 \text{ pF} \quad R_3 = 100 \text{ k}\Omega$$

$$v_{off} = \pm 3 \text{ mV} \quad i_b = 50 \text{ nA} \quad i_{off} = 0$$

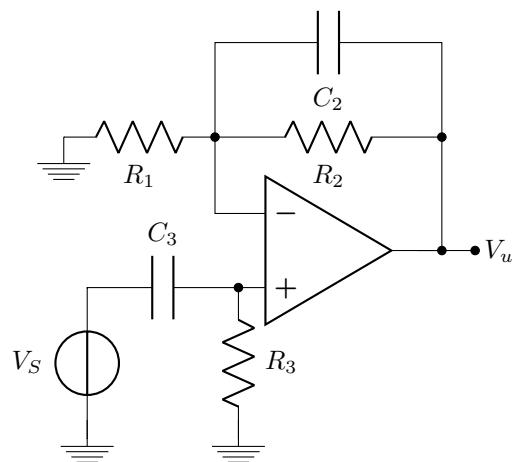
Figure 3.25: Bandpass amplifier circuit

1. Compute the ratio V_u/V_S as a function of the frequency
2. Compute the value of V_u for $V_S = 0$.

Exercise 3.2. Bandpass amplifier

Given the circuit schematic the bandpass amplifier of figure 3.26

1. Compute the ratio V_u/V_S as a function of the frequency
2. Compute the value of V_u for $V_S = 0$.



$$R_1 = 8.2 \text{ k}\Omega \quad R_2 = 120 \text{ k}\Omega \quad C_2 = 220 \text{ pF} \quad C_3 = 27 \text{ nF} \quad R_3 = 330 \text{ k}\Omega \\ v_{off} = \pm 3 \text{ mV} \quad I_b = 50 \text{ nA} \quad i_{off} = 0$$

Figure 3.26: Band pass amplifier circuit

Exercise 3.3. Differential amplifier

With reference to the circuit schematic of the differential amplifier shown in figure 3.11:

1. Compute A_d and A_C assuming $R_1 = 12 \text{ k}\Omega$, $R_2 = 330 \text{ k}\Omega$, $R_3 = 12 \text{ k}\Omega$, $R_4 = 330 \text{ k}\Omega$.
2. What is the worst case change in A_d and A_C if we assume that all resistors have 1% tolerance?

Exercise 3.4. Differential amplifier

A differential amplifier is used to read the output of a resistive bridge in which we have three equal resistors R_A and one strain gauge R_B . The schematic is in figure 3.27.

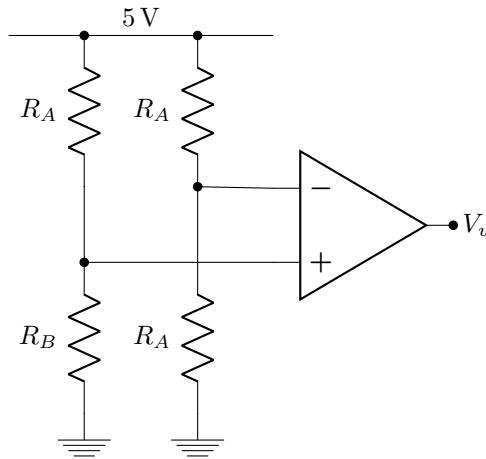


Figure 3.27: Bridge amplifier circuit

We have $R_A = 1 \text{ k}\Omega$. The strain gauge has nominal resistance $R_B = 1 \text{ k}\Omega$ and the resistance will change with deformation with a rate of $\Delta R_B/\text{mm} = 10 \Omega \text{ mm}^{-1}$.

1. Compute A_d of the differential amplifier to obtain $V_u = 5 \text{ V}$ when stretching the strain gauge by 0.5 mm.

Exercise 3.5. Adder circuit

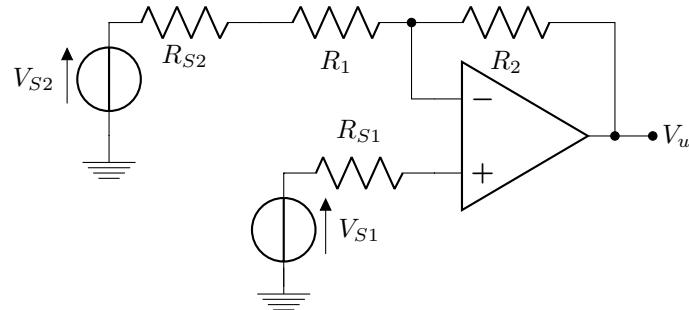
Given the circuit schematic of figure 3.28

Assuming that the operational amplifier has ideal parameters compute the function $v_u = f(v_1, v_2)$ in the following conditions:

1. $R_{S1} = R_{S2} = 0$
2. $R_{S1} = 10 \text{ k}\Omega$, $R_{S2} = 2 \text{ k}\Omega$

Exercise 3.6. Adder circuit Using an operational amplifier design a circuit whose output is:

$$v_u = 3v_1 - 2v_2 - v_3$$



$$R_1 = 6.8 \text{ k}\Omega \quad R_2 = 120 \text{ k}\Omega$$

Figure 3.28: Adder circuit

where v_1 , v_2 and v_3 are three independent voltage sources. If the maximum input offset voltage of the operational amplifier is $\pm 4 \text{ mV}$ determine the possible output voltage range of the circuit output if $v_1 = v_2 = v_3 = 0$

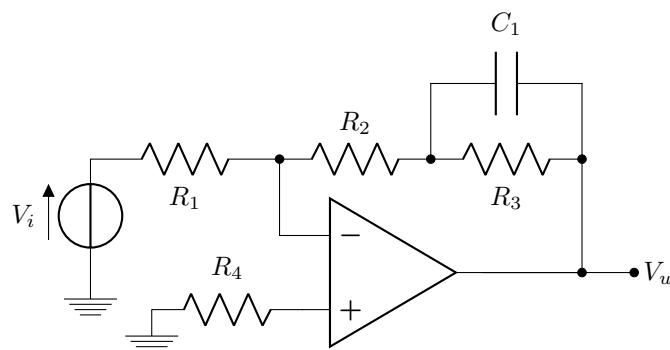
Assume that apart from the offset voltage all other parameters of the operational amplifier are ideal.

Exercise 3.7. Instrumentation amplifier Size the resistors of an instrumentation amplifier, assuming $R_2 = R = 50 \text{ k}\Omega$. Refer to figure 3.19 for the circuit's schematic and resistor names.

1. Set the remaining resistors to have $A_d = 5$.
2. Add an additional resistor R_p in parallel to R to change the gain to 10.
3. Modify the value of R_p to obtain $A_d = 50$.

Exercise 3.8. AC amplifier

Given the circuit schematic of figure 3.29



$$\begin{aligned} R_1 &= 12 \text{ k}\Omega & R_2 &= 39 \text{ k}\Omega & C_1 &= 2.2 \text{ nF} \\ R_3 &= 120 \text{ k}\Omega & R_4 &= 10 \text{ k}\Omega \end{aligned}$$

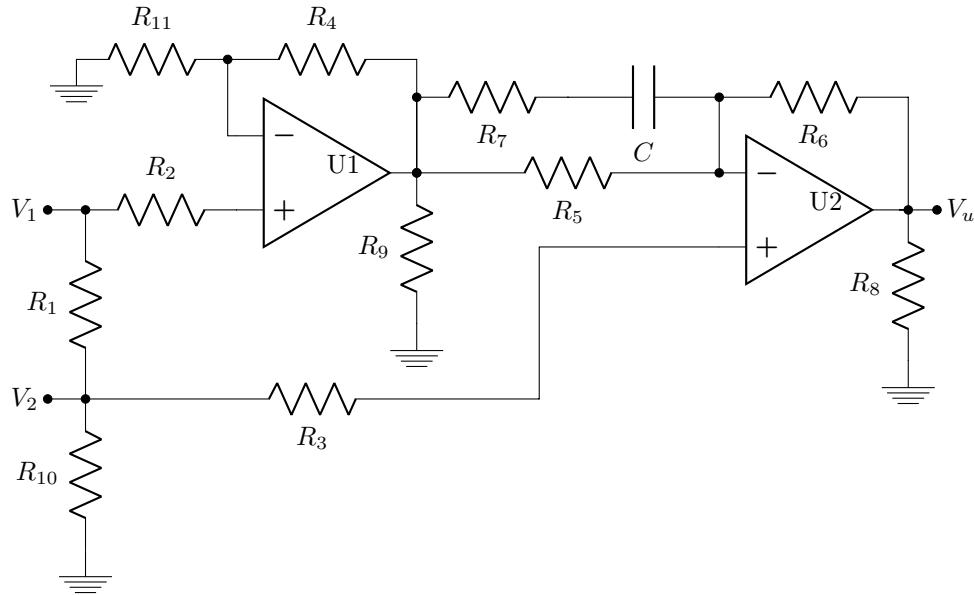
Figure 3.29: Circuit schematic diagram

1. Compute the asymptotic gains in DC and high frequency.

2. write the equation of the complete transfer function (V_u/V_i).
3. Draw its Bode plot.

Exercise 3.9. AC amplifier

Given the circuit schematic of figure 3.30



$$\begin{array}{llll}
 R_1 = 270 \Omega & R_2 = 11.2 \text{ k}\Omega & R_3 = 60 \text{ k}\Omega & R_4 = 12 \text{ k}\Omega \\
 R_5 = 120 \text{ k}\Omega & R_6 = 120 \text{ k}\Omega & R_7 = 10 \text{ k}\Omega & R_8 = 12 \text{ k}\Omega \\
 R_9 = 27 \text{ k}\Omega & R_{10} = 10 \text{ k}\Omega & R_{11} = 12 \text{ k}\Omega & C = 4.7 \text{ nF}
 \end{array}$$

Figure 3.30: Circuit schematic diagram

1. Identify and remove from the schematic diagram all the resistors which do not contribute to the V_u/V_1 or V_u/V_2 transfer functions.
2. Write the equation of the V_u/V_1 and V_u/V_2 DC transfer functions and compute their value (linear and in dB).
3. Write the equation of the V_u/V_1 and V_u/V_2 high frequency transfer functions and compute their value (linear and in dB).
4. Write the equation of the complete V_u/V_1 and V_u/V_2 frequency dependent functions.
5. Draw their Bode plot.

Chapter 4

Non linear Amplifiers

Contents

4.1	Precision half wave rectifier	4-1
4.1.1	Shift of the transfer characteristic	4-3
4.2	Full wave rectifier	4-4
4.2.1	Extended full wave rectifier	4-7
4.2.2	Design example	4-11
4.2.3	Frequency behaviour	4-12
4.2.4	Changing the diode polarity	4-13
4.3	Exercises	4-14

THE circuits analyzed in this chapter use **non linear components** (diodes and transistors) in the feedback network, but the operational amplifier continues to operate in linearity and therefore the basic equations of the ideal O.A. ($v_d = 0$ e $i_+ = i_- = 0$) are still valid. The transfer function is determined by the feedback network and by introducing into it non linear components, we may obtain a non linear V_u/V_i behavior.

4.1 Precision half wave rectifier

The ideal half wave rectifier, also often referred to as *ideal diode* is a circuit which amplifies (positively or negatively) the positive part of a signal and has a zero output when $V_i < 0$ V. Figure 4.1 shows the input and output voltage of an inverting version of the circuit with a sinusoidal input waveform.

This circuit must not be confused with rectifiers used in power systems with high voltages and currents. It is not intended to transfer power from the input to the output but its purpose is to obtain a rectified signal which is not affected by errors due to diode threshold voltage for *any* amplitude of the input signal, even far smaller than the diode threshold itself.

Let us analyze the active circuit of figure 4.2 which is derived from the basic inverting amplifier. When V_i is positive the current in R_2 flows in the direction shown in diode D_1 and its value is V_i/R_1 . Since the inverting input is at virtual ground the output voltages V_u and V'_u will be:

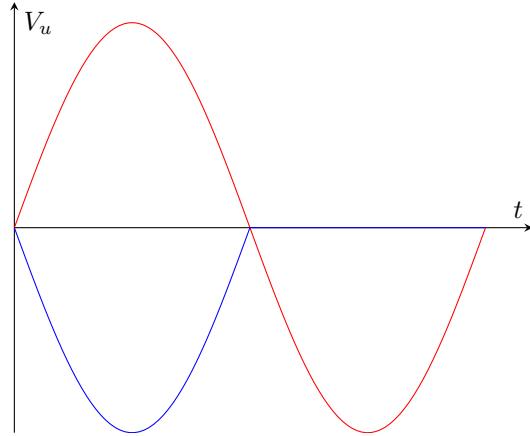


Figure 4.1: Output of an *inverting half wave rectifier* (in blue) with an input sinusoidal waveform (in red).

$$\begin{aligned} V_i > 0 \implies V_u &= -\frac{R_2}{R_1}V_i - V_\gamma \\ \implies V'_u &= -\frac{R_2}{R_1}V_i \end{aligned}$$

where V_γ is the diode threshold voltage, that is the voltage drop across it when it is in direct conduction.

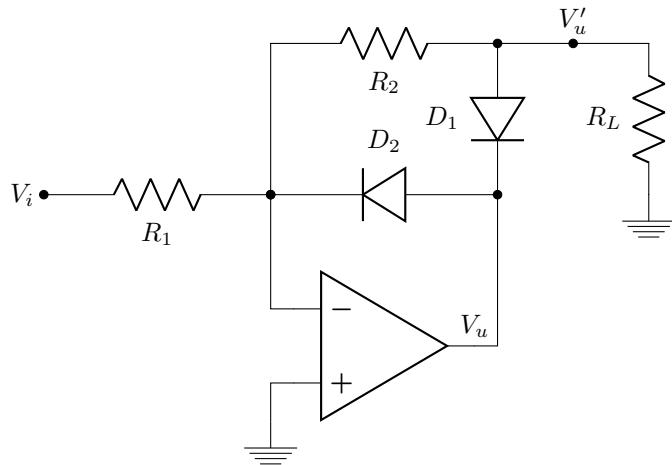


Figure 4.2: Half wave rectifier circuit.

For $V_i < 0$ the current in R_1 flows in the opposite direction and cannot pass in D_1 . It flows in D_2 which is now forward biased and the operational amplifier output voltage V_u will be V_γ higher than the virtual ground of the inverting pin. Since there is no current in R_2 the drop across it will be 0 and therefore the V'_u will be at virtual ground too. Figure 4.3 shows the transfer characteristics for both outputs and we may immediately notice that the V'_u output corresponds to the required one.

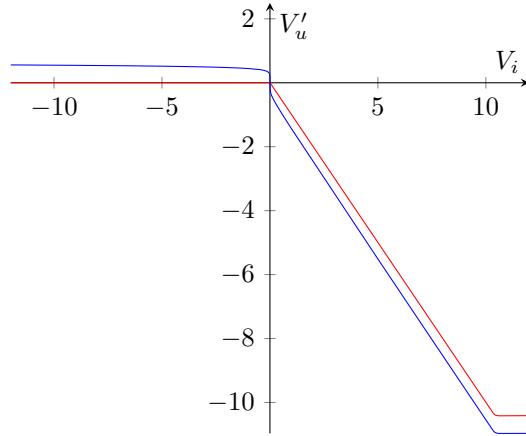


Figure 4.3: Transfer characteristic of the circuit of figure fig. 4.2: V'_u is in red and V_u in blue.

Output impedance

There is still a problem with this circuit and to understand it let us evaluate the output impedance as seen from output V'_u . When the input signal is positive D_1 conducts and V'_u is inside the feedback loop of the operational amplifier. As a consequence if we were computing the output impedance with a test generator, as we did for the non inverting amplifier, we would obtain the same result substituting to the output impedance r_o of the operational amplifier the series of it and of the differential resistance of the diode and dividing it by the loop gain. In summary the output impedance would be very low.

If we recompute the impedance when the input signal is negative the result is completely different. In these conditions diode D_1 is off and R_2 is not anymore in the amplifier feedback loop. The negative input pin remains at virtual ground because D_2 closes the feedback loop and therefore the impedance seen from the V'_u output is R_2 .

If we want to connect a load to that output then it must be necessarily referred to 0 V, as shown in figure 4.2 because when the output is 0 (negative input voltage) no current should flow in R_2 . If the output were referred to any other voltage, a current would flow in R_2 modifying the transfer characteristic of the circuit which would become dependent on the load resistance value. This is true also if the load is another amplifier stage.

4.1.1 Shift of the transfer characteristic

It is possible to modify the circuit in order to obtain a shift of the angular point of the characteristic from the origin of the axes, that is to have a null output voltage V'_u for input voltage V_i lower than a threshold voltage different from 0 V and that in the following we will refer to as V_{PA} . Analytically this function may be described as a piece-wise linear function by the equations 4.1 corresponding to the graph of figure 4.5.

$$V_u = \begin{cases} -(R_2/R_1) \cdot (V_i - V_{PA}), & V_i > V_{PA} \\ 0, & V_i < V_{PA} \end{cases} \quad (4.1)$$

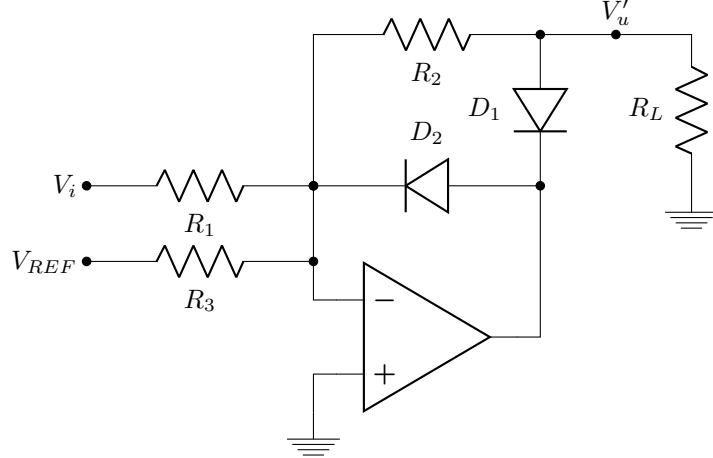


Figure 4.4: Single wave rectifier with an additional reference voltage

To obtain the shift of the characteristic we introduce a reference voltage V_{REF} as an additional input through resistor R_3 as shown in figure fig.4.4. Considering currents in R_1 and R_3 positive when entering the circuit, the output V'_u is 0 when the total current coming from the inputs flows in D_2 and is therefore negative. Conversely V'_u will be negative when the sum of the currents is positive, that is when:

$$I_1 + I_3 > 0 \implies \frac{V_i}{R_1} + \frac{V_{REF}}{R_3} > 0 \quad (4.2)$$

The transition from D_2 conducting to D_1 conducting takes place when the total input current is 0 that is when:

$$\frac{V_i}{R_1} = -\frac{V_{REF}}{R_3}$$

and the input voltage V_i at which this takes place is the x coordinate of the corner point V_{PA} of the transfer characteristic shown in figure 4.5. Its value is given by:

$$V_{PA} = -V_{REF} \frac{R_1}{R_3}$$

This circuit has a behavior close to that of an ideal diode and is sometimes referred to as *superdiode*.

4.2 Full wave rectifier

Often it is necessary to rectify a signal that is, in the time domain, to change sign to its negative component leaving unmodified the positive one. The circuit which performs the task is the *double or full wave rectifier* and its transfer characteristic is shown in figure 4.6. This circuit is not intended as a

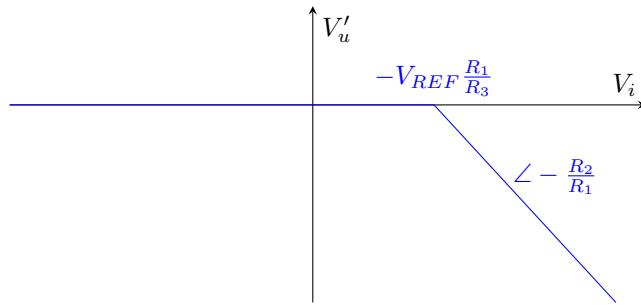


Figure 4.5: Transfer characteristic of the single wave rectifier with modified threshold

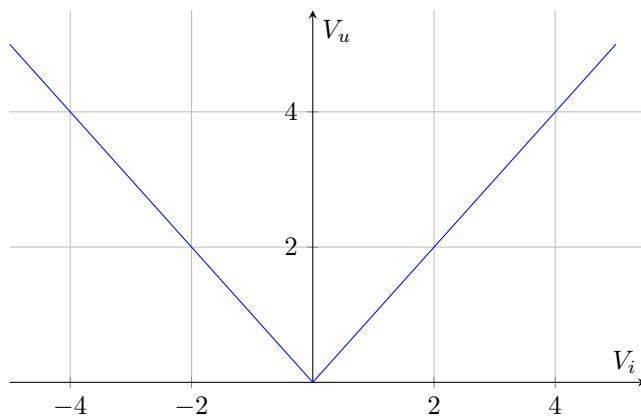


Figure 4.6: Transfer characteristic corresponding to the function $v_u = |V_i|$.

power rectifier to transform ac into dc power but as a signal processor, in general for measurement purposes.

From the mathematical point of view the circuit performs the *absolute value* function. The circuit is composed of two stages : an half wave rectifier such as the one previously analyzed and a second stage which adds to it the input signal as shown in figure 4.7.

The half wave rectifier inverts the positive part of the V_i waveform which enters the inverting amplifier giving as output a positive output. The second stage adds to the output the non inverted original input waveform and by setting correctly the gains of the different stages we may obtain the desired result.

Quantitative analysis

Let us define as V_1 the output of the first stage. The total output of the circuit will then be:

$$V_u = -V_1 \frac{R_4}{R_3} - V_i \frac{R_4}{R_5}$$

If $V_i > 0$ then diode D_1 is on and the output of the first stage is that of an inverting amplifier:

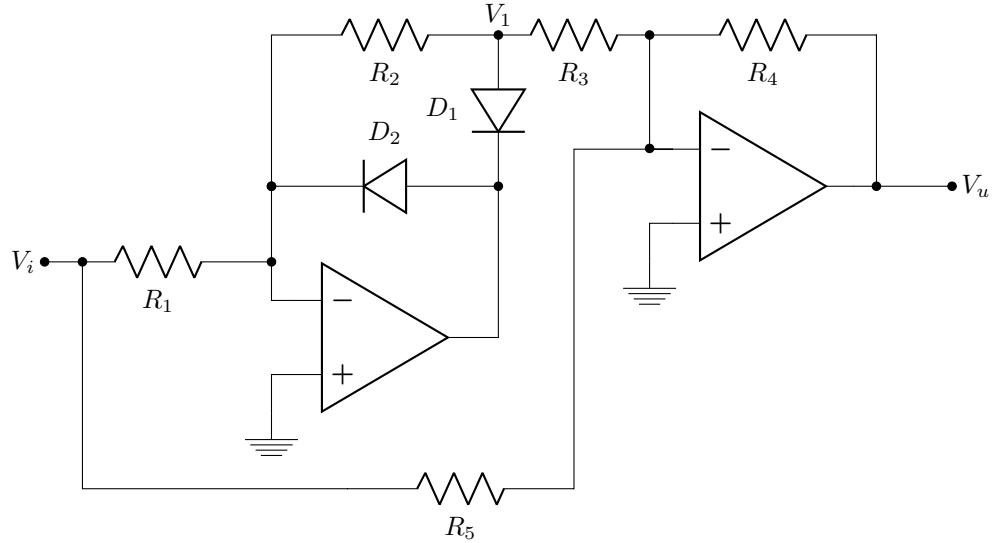


Figure 4.7: Circuit schematic of the full wave rectifier.

$$V_i > 0 \implies V_1 = -V_i \frac{R_2}{R_1} \implies V_u = V_i \left(\frac{R_2}{R_1} \cdot \frac{R_4}{R_3} - \frac{R_4}{R_5} \right)$$

When $V_i < 0$, D_1 does not conduct and $V_1 = 0$ V. Therefore V_u depend only on the input V_i connected to R_5 :

$$V_i < 0 \implies V_1 = 0 \implies V_u = -V_i \frac{R_4}{R_5} > 0$$

It is possible to obtain any value of the slope of the two segments of figure 4.6 by setting first the ratio between R_4 and R_5 for $V_i < 0$ and then the ratio between R_2 and R_1 for $V_i > 0$. The following example allows to verify numerically the design methodology.

Example 9. Let us design a full wave rectifier with both unit slopes as those of figure 4.6.

As we have previously derived the output voltage is given by:

$$V_u = \begin{cases} -V_i \cdot (R_4/R_5), & V_i < 0 \\ V_i [(R_4/R_3) \cdot (R_2/R_1) - (R_4/R_5)] & V_i > 0 \end{cases}$$

When the input voltage is negative the output depends only on the ratio between R_4 and R_5 and if we want it to unitary we must have:

$$\frac{R_4}{R_5} = 1 \implies R_4 = R_5$$

When the input is positive we have:

$$\frac{R_4}{R_3} \cdot \frac{R_2}{R_1} - \frac{R_4}{R_5} = 1 \implies \frac{R_4}{R_3} \cdot \frac{R_2}{R_1} = 2 \quad (4.3)$$

We may satisfy condition 4.3 in infinite ways, but in order to simplify the choice and increase precision it is common to have one of the two ratios unitary that is either $R_1 = R_2$ or $R_3 = R_4$. The two possible choices are then:

$$\frac{R_2}{R_1} = 2; \quad \frac{R_4}{R_3} = 1$$

$$\frac{R_2}{R_1} = 1; \quad \frac{R_4}{R_3} = 2$$

and which one we take depends on which parameter we want to optimize.

- The first one, seldom used, optimizes the effect of offset on the output because it amplifies less the offset components of the second operational amplifier. It is used when we want a **precise system**, insensitive to offset voltages and currents. However, since the gain of the first amplifier is 2 the input voltage range of the whole system is limited to 1/2 of the output voltage range of the first operational amplifier.
- the second choice **maximizes the input voltage range** and is rather common.

A simple example clearly shows the difference. Using the first solution with the following parameters:

$$V_{AL} = \pm 15 \text{ V} \quad V_{u\max} = \pm 10 \text{ V} \quad \frac{R_2}{R_1} = 2$$

The input voltage range is limited to $V_i = \pm 5 \text{ V}$ because for larger voltages the first operational amplifier would go out of linearity. The second solution does not have the same problems because the first amplifier has a unity gain and the second one at the same time adds two signals with different signs and the output range is limited by their sum and not by the value of one of them.

The example shows that in the design it is important to take care of the input and output ranges of each of the stages which compose the system.

4.2.1 Extended full wave rectifier

Starting from the circuit analyzed in the previous section, it is possible to obtain a system capable of providing a more elaborate characteristic, introducing reference voltages that translate the coordinates of the angular point and at the same time programming slopes of the two segments that are different from each other.

The resulting system could find applications in processing signals that need to be “selectively” amplified with respect to voltage level. For example, one might want to obtain a trans-characteristic like the one in figure 4.8.

The starting point is the full wave rectifier to which two reference voltages V_{R1} and V_{R2} are added (fig. 4.9) which represent new degrees of freedom of the system. The operating principle is similar to that of precision half wave rectifier with the addition of V_{REF} .

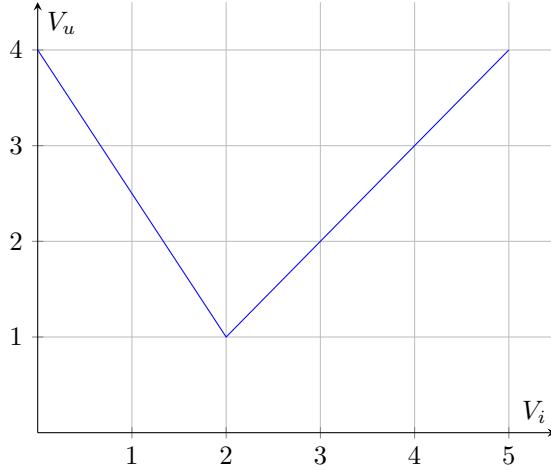


Figure 4.8: Piecewise linear trans-characteristic.

Effect of V_{R1} voltage

Let us first consider a system with only the reference voltage V_{R1} . The purpose of V_{R1} is to modify the V_i corresponding to the corner point. Also in this case it is necessary to write the equation to the node to which the inverting terminal of the first opamp is connected when there is D_2 in reverse bias and therefore all the current coming from R_1 and R_6 circulates in D_1 .

$$I_{R2} = \frac{V_i}{R_1} + \frac{V_{R1}}{R_6} > 0 \implies V_1 < 0$$

$$V_i > -V_{R1} \cdot \frac{R_1}{R_6} = V_{PA} \implies V_1 < 0$$

Therefore when $V_i > V_{PA}$ the first stage (precision half wave rectifier) inverts and eventually amplifies V_i and V_{R1} producing $V_1 < 0$ depending on these two inputs.

$$V_i > V_{PA} \implies V_1 = -\left(\frac{V_i}{R_1} + \frac{V_{R1}}{R_6}\right) R_2$$

At the same time V_{R1} has an effect on V_u when not blocked by the rectifier (that is when $V_i > V_{PA}$), so the net effect of this reference voltage is to shift the original characteristic both with respect to V_i and to V_u (oblique transformation in geometrical terms). In fig. 4.10 we can see the translation effect of the sole V_{R1} on a full wave rectifier with gain of $-3/2$ for $V_i < V_{PA}$ and 1 for $V_i > V_{PA}$, with $V_{PA} = 2$ V.

$$V_i < V_{PA} \implies V_u = -V_i \frac{R_4}{R_5} \quad (4.4)$$

$$V_i > V_{PA} \implies V_u = V_i \left(\frac{R_2 R_4}{R_1 R_3} - \frac{R_4}{R_5} \right) + V_{R1} \frac{R_2 R_4}{R_6 R_3} \quad (4.5)$$

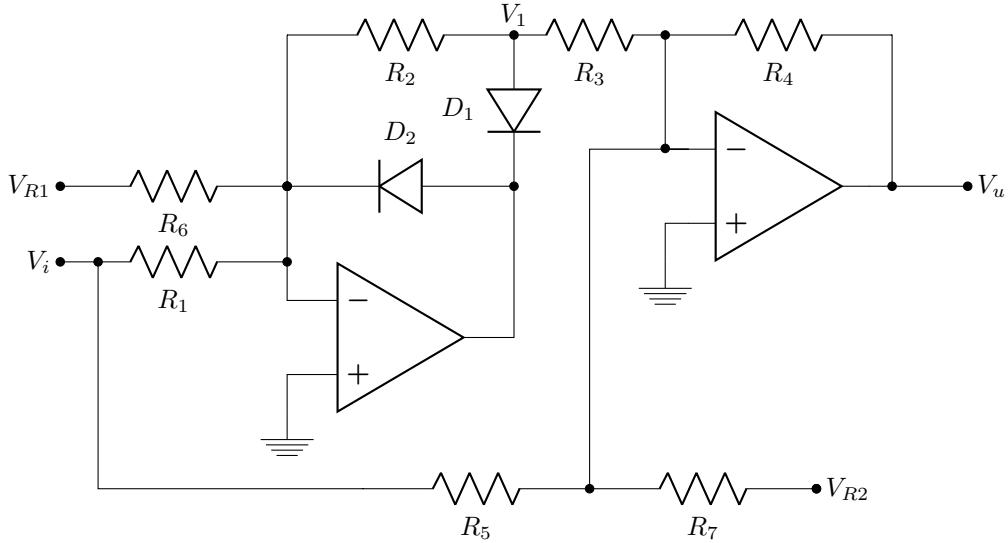


Figure 4.9: Schematic circuit of extended full wave rectifier.

We can determine the electrical coordinates of the corner point in the V_u/V_i plane by imposing $V_i = V_{PA}(V_{R1})$ and calculating the corresponding output. The corner point belongs to both segments, but the simpler expression 4.4 allows for easier substitutions.

$$V_{PA} = -V_{R1} \frac{R_1}{R_6} \quad V_u(V_i = V_{PA}) = V_{R1} \frac{R_1 R_4}{R_6 R_5}$$

As V_{R1} varies, both V_i and V_u undergo variations as stated qualitatively speaking of the oblique translation of the characteristic. It is also very important to note that the displacement of the angular point P.A. occurs along a line with a constant slope that can be determined analytically.

$$\frac{V_u(V_{PA})}{V_i} = -\frac{R_4}{R_5}$$

In fact, the P.A. corresponds to an input voltage such as to bring the precision half-wave rectifier into the limit situation in which its output V_1 vanishes. When $V_1 = 0$ then the output of the system V_u depends only on V_i as stated above. Therefore V_u varies with V_{PA} with the characteristic of the second stage (inverting amplifier).

Effect of V_{R2} voltage

At the end of the previous paragraph we have shown that V_{R1} translates *obliquely* the V_u/V_i curve; to be able to translate the angular point in any position of the Cartesian plane it is necessary to introduce a further reference voltage V_{R2} .

V_{R2} is introduced into the system in parallel to the half-wave rectifier so that its effect on the output does not depend on the value of V_i unlike what happens for V_{R1} . The contribution on V_u due to V_{R2} is therefore always present and so

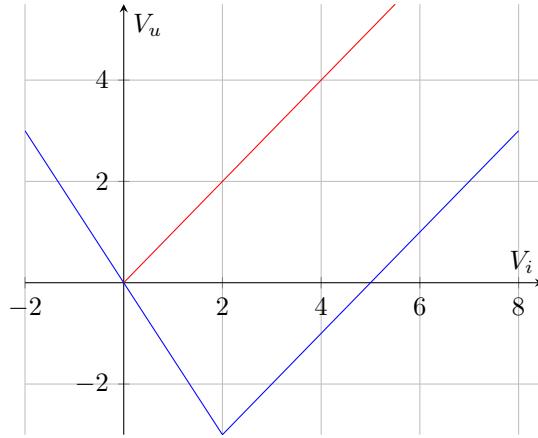


Figure 4.10: Shift due to V_{R1} at the output of a full wave rectifier. In red: positive side of the characteristic without V_{R1}

allows the graph of the characteristic to be translated vertically by a desired value.

In other words, the first reference voltage translates the characteristic obliquely and the second translates it vertically. The effect of V_{R2} on V_u is that typical of a bias signal introduced into an inverting amplifier.

$$V_u|_{V_{R2}} = -V_{R2} \cdot \frac{R_4}{R_7}$$

Overall characteristic

The overall characteristic which takes into account the contributions of the various design parameters is depicted in the plot of figure 4.11.

There are far more degrees of freedom than in previous circuits, with the possibility of obtaining more elaborate waveforms. The purpose of this paragraph is to summarize the waveform parameters obtainable with this circuit. In particular, the value of the output with null input has also been added.

- the coordinates of the **corner point**, or **angular point**, the limit between the region of operation as a rectifier and a non-inverting amplifier of the input signal V_i , depend on the reference voltages.

$$V_{i,PA} = -V_{R1} \frac{R_1}{R_6} \quad V_{u,PA} = V_{R1} \frac{R_1 R_4}{R_6 R_5} - V_{R2} \frac{R_4}{R_7}$$

- the amplification of the signal to the right of the corner point depends on the values of the resistors R_1 , R_2 , R_3 , R_4 and R_5 ; it can be positive or negative depending on their values:

$$A_V^+ = \left. \frac{\partial V_u}{\partial V_i} \right|_{V_i > V_{PA}} = \frac{R_2 R_4}{R_1 R_3} - \frac{R_4}{R_5}$$

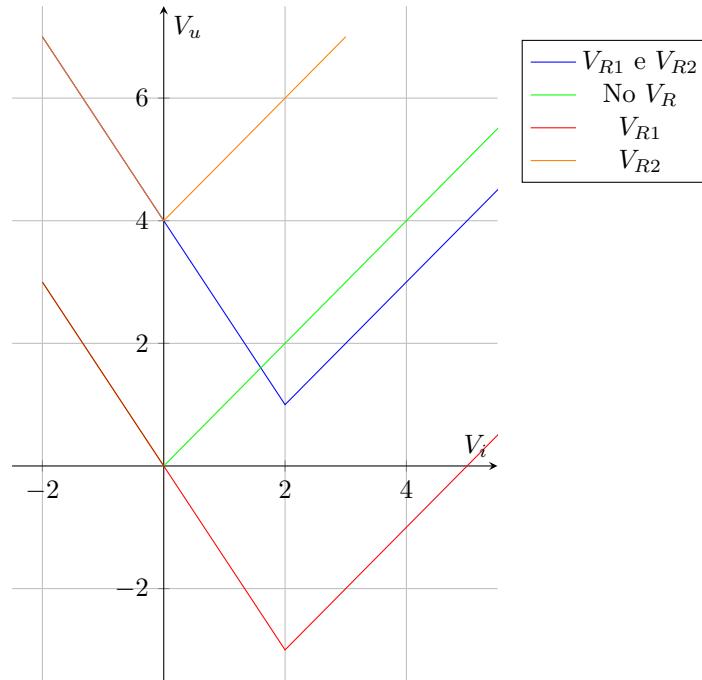


Figure 4.11: Single and combined effects of the shifts imposed by V_{R1} and V_{R2} on the characteristic of the circuit of fig. 4.9.

- the amplification of the signal to the left of the corner point depends on the values of the resistors R_4 and R_5 and it is negative because only the second stage is involved and it is an inverting amplifier.

$$A_V^- = \left. \frac{\partial V_u}{\partial V_i} \right|_{V_i < V_{PA}} = -\frac{R_4}{R_5}$$

- the value of V_u when the input is null depends on the reference voltages.

$$V_{PA} < 0 \implies V_u(V_i = 0) = V_{R1} \frac{R_2 R_4}{R_6 R_3} - V_{R2} \frac{R_4}{R_7}$$

$$V_{PA} > 0 \implies V_u(V_i = 0) = -V_{R2} \frac{R_4}{R_7}$$

4.2.2 Design example

Size the full-wave rectifier circuit that provides the trans-characteristic shown above (figure 4.8), using $V_{AL} = \pm 15$ V.

From the graph it is possible to easily obtain the slopes of the two segments corresponding to the amplifications of the stages to which the resistance ratios are linked.

$$A_V^- = -\frac{R_4}{R_5} = \frac{1\text{ V} - 4\text{ V}}{2\text{ V} - 0\text{ V}} = -\frac{3}{2}$$

$$A_V^+ = \frac{4V - 1V}{5V - 2V} = 1$$

From the intersection of the graph line with the ordinate axis we obtain the value of the output in correspondence with the null input, to which only the reference voltage V_{R2} is linked since the angular point has a positive abscissa.

$$V_u(V_i = 0) = -V_{R2} \cdot \frac{R_4}{R_7} = 4V \implies V_{R2} < 0$$

Since a negative V_{R2} is needed, it can be set equal to the supply voltage $-15V$ which is available from specifications, dimensioning R_4 and R_7 so that, instead of amplifying, it attenuates the $-15V$ bringing a contribution of $4V$ requested on the output.

$$V_u|_{V_{R2} = -V_{AL}} = -15V \frac{R_4}{R_7} = 4V \implies \frac{R_4}{R_7} = \frac{4}{15}$$

The amplification to the right of the corner point is unitary, therefore we find ourselves in a situation similar to that encountered during the design of the precision half-wave rectifier carried out previously. Assuming we want to maximize *input dynamic range* we impose unity gain for the first amplification stage.

$$\frac{R_4}{R_3} \cdot \frac{R_2}{R_1} - \frac{3}{2} = 1 \implies \frac{R_4}{R_3} \cdot \frac{R_2}{R_1} = \frac{5}{2}$$

$$\frac{R_2}{R_1} = 1 \implies \frac{R_4}{R_3} = \frac{5}{2}$$

The last parameter to be established is V_{R1} and to do this we consider the abscissa of the angular point. Also in this case the reference voltage is set equal to V_{AL} and the ratio between R_1 and R_6 is determined which fixes the desired $V_{i,PA}$.

$$V_{i,PA} = -V_{R1} \frac{R_1}{R_6} = -V_{AL} \frac{R_1}{R_6} = 2V \implies \frac{R_1}{R_6} = \frac{2}{15}$$

It is necessary to fix the absolute value of at least two resistors and obtain the others, considering the current and offset characteristics of the amplifier operational used. Assuming a typical amplifier and wanting to use only E12 series components, we can for example fix $R_2 = 27\text{k}\Omega$ and $R_4 = 68\text{k}\Omega$. With those values we find: $R_5 = 47\text{k}\Omega$, $R_3 = 27\text{k}\Omega$, $R_1 = 27\text{k}\Omega$, $R_6 = 220\text{k}\Omega$ and $R_7 = 270\text{k}\Omega$.

4.2.3 Frequency behaviour

The frequency behaviour of the full wave rectifier is not very good and the circuit may be used only for signals whose bandwidth is far below the gain bandwidth product of the operational amplifiers. This is because the first amplifier at high frequencies introduces a phase shift and when this signal is added to the input at the second stage the two signals are not any more in phase and the result is a high distortion at the output.

4.2.4 Changing the diode polarity

In all the previous analysis we have assume the polarity of diodes D_1 and D_2 to be that shown in figure 4.2 because this is the most common case. However we may build a single and a full wave rectifier exchanging the anode of each diode with the cathode. In this case the negative part of the waveform is rectified and amplified and the student is encouraged to derive by himself, as an exercise, the transfer characteristic of the circuit in this case.

4.3 Exercises

Exercise 4.1. Half-wave rectifier

Design a precision half wave rectifier with $V_{PA} = -3\text{ V}$ and gain $A_V = -4$. Use one of the supply voltages $V_{AL} = \pm 12\text{ V}$ as reference voltage.

Exercise 4.2. Full wave rectifier

Draw the schematic of a full wave rectifier and size the components in order to obtain

$$Vu = 2|Vi|$$

Exercise 4.3. Full wave rectifier Design an extended full wave rectifier whose characteristic conforms with the graph of figure 4.12. Use $V_{AL} = \pm 5\text{ V}$ and a rail-to-rail operational amplifier.

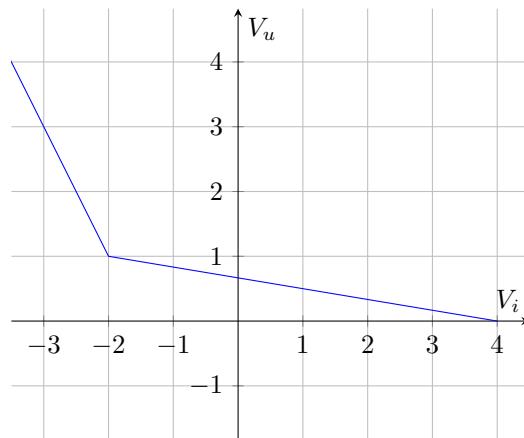


Figure 4.12: Circuit specifications.

Chapter 5

Use of the Operational Amplifier out of linearity

Contents

5.1	Voltage comparators	5-1
5.1.1	Inverting comparator	5-2
5.1.2	Non inverting voltage comparator	5-3
5.1.3	Comparators and binary decision	5-3
5.1.4	Noise sensitivity	5-3
5.1.5	Comparators with hysteresis	5-4
5.1.6	Circuit implementation of voltage comparators	5-7
5.2	Astable Multivibrator	5-8
5.2.1	Circuit analysis	5-8
5.2.2	Quantitative analysis	5-10
5.3	Triangular wave generator	5-11
5.3.1	Quantitative analysis	5-11
5.3.2	Design example	5-13
5.4	Exercises	5-21

The operational amplifier may be used out of linearity, that is without a negative feedback loop, in order to obtain circuits very different from the amplifiers analyzed until now. When the operational amplifier does not operate in linearity, then the equations used in the previous chapters to study its behavior are no more valid as well as the assumption of zero differential input voltage and current. Without a negative feedback the large gain of the amplifier makes it to behave differently from an amplifier. We will first analyze circuits operating in open loop and later circuits using positive feedback.

5.1 Voltage comparators

Voltage comparators are the first step in the transition from the analog world to the digital one.

The simplest circuit is the open loop operational amplifier which operates as a voltage comparator and whose transfer characteristic is shown in figure 5.1.

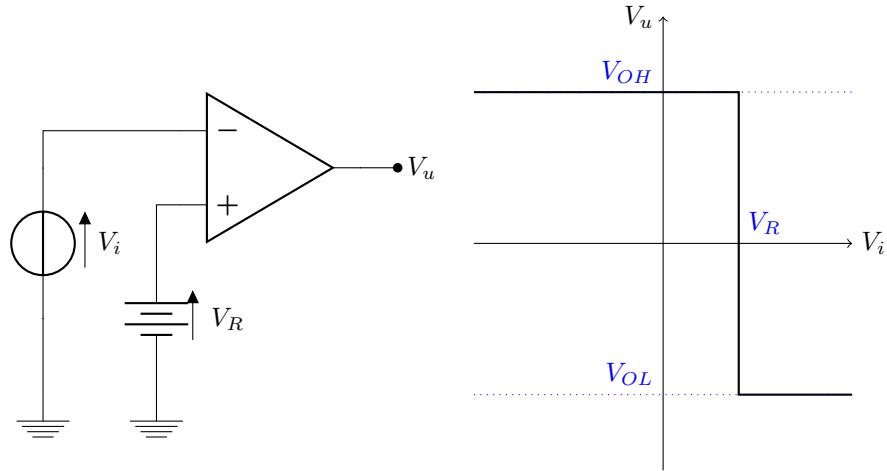


Figure 5.1: Circuit schematic of the inverting voltage comparator and its transfer characteristic

The circuit inputs are two *continuous amplitude signals*, which may have any value in a given input range and the output is a *binary signal* with only two possible states.

5.1.1 Inverting comparator

The circuit shown in figure 5.1 is called *inverting voltage comparator* and the constant voltage V_R on the non inverting input pin is in general referred to as *reference voltage* or *threshold*. The input signal V_i is connected to the inverting pin.

The differential gain of the amplifier is very high (up to 10^6) and therefore, when operating in linearity, a very small input signal would produce a very high output signal. However the value of the output is limited to the output voltage range bound by the power supply range. When the output voltage range limits are reached the amplifier saturates and the output voltage does not increase or decrease further. We will refer to the upper limit as V_{OH} and to the lower limit, reached when the differential voltage is negative, as V_{OL} .

$$V_u = \begin{cases} V_{OH} & v_d > 0 \implies V_R - V_i > 0 \implies V_i < V_R \\ V_{OL} & v_d < 0 \implies V_R - V_i < 0 \implies V_i > V_R \end{cases}$$

When V_i reaches a value very close to V_R , we may consider the transfer characteristic $V_u(V_i)$ a *vertical* line since the real slope is the amplifier differential gain $A_d \rightarrow \infty$.

The term *inverting* given to this circuit derives from the fact that when the input voltage is *higher* than the reference one then the output value is $V_u = V_{OL}$; when $V_i < V_R$ then $V_u = V_{OH}$. The input signal is connected to the inverting pin.

5.1.2 Non inverting voltage comparator

If we connect the input signal to the non inverting pin and the reference voltage to the inverting, one we obtain the *non inverting voltage comparator* (fig. 5.2).

The differential input voltage v_d is equal to $= V_i - V_R$ and therefore the output is at the high voltage V_{OH} when V_i is higher than V_R .

$$V_u = \begin{cases} V_{OH} & v_d > 0 \implies V_i - V_R > 0 \implies V_i > V_R \\ V_{OL} & v_d < 0 \implies V_i - V_R < 0 \implies V_i < V_R \end{cases}$$

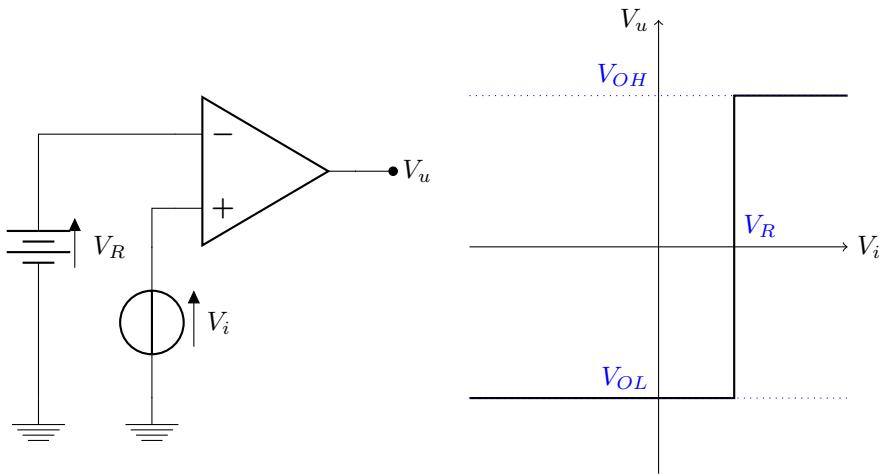


Figure 5.2: Circuit schematic and transfer characteristic of the non inverting voltage comparator.

Information about output voltage ranges of a given operational amplifier, and therefore on V_{OH} and V_{OL} may be found in the component datasheet. Component implemented with different technologies may have very different values of output range, although in general close to the power supplies $\pm V_{AL}$ which are in any case the upper and lower limit to the output voltage.

5.1.3 Comparators and binary decision

The voltage comparator is the first example of an *interface between the analog world and the digital one* because when the input varies over a continuous range the output may take only two possible values V_{OL} and V_{OH} . It compares the value of the input with the reference one, takes a binary decision and outputs a binary value.

To obtain this kind of behavior we need a strongly non linear component because the comparator associates a *single* output value to all the values of one of two sets in which the range of the input voltage V_i is divided..

5.1.4 Noise sensitivity

The comparators analyzed in the previous paragraphs are extremely sensitive to noise when the signal level is close to the reference one. All signals have a certain amount of noise either of thermal origin in active and passive components

or deriving from the power supplies or by coupling with other electromagnetic sources. In most cases it may be modeled as a stochastic signal with null average value. Normally the amplitude of the noise is small with respect to the signal and therefore when the signal is far larger or smaller than the reference V_R the probability that the noise added to V_i make the comparator change its output is very low. However if the input signal is close to the reference value, then a small noise voltage is sufficient to bring the input of the comparator over or under V_R generating an output toggling which is spurious because not due to V_i (upper graphs of figure 5.3). As V_i approaches V_R the number of output transitions increases and, at the limit, when $V_i = V_R$ all switching is due to noise and V_u becomes a function of the noise, instead of the input signal.

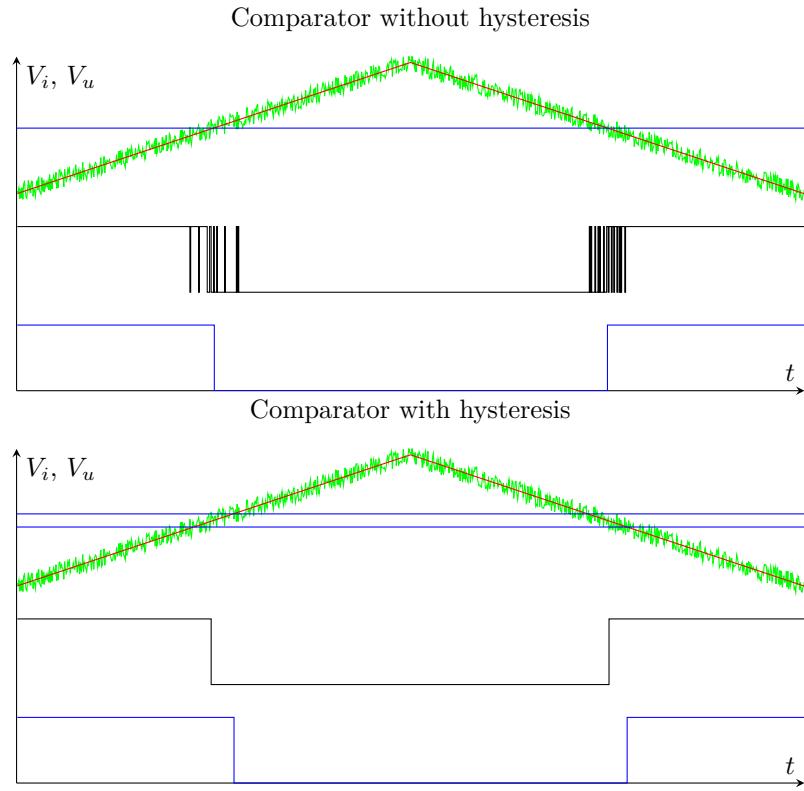


Figure 5.3: RED ideal input signal, GREEN added with noise; BLUE output when no noise is present; BLACK output when noise is present; BLUE horizontal lines: thresholds

5.1.5 Comparators with hysteresis

In order to improve the behavior of the circuit we introduce a *double* threshold with two reference voltages and the concept of hysteresis. Switching from one output state to the other one takes place only when the input signal crosses both thresholds in the same direction and this increases the amplitude of the noise signal which is necessary to produce a spurious toggle (the difference between

the two reference voltages) reducing its probability as shown in the lower timing diagrams of figure 5.3.

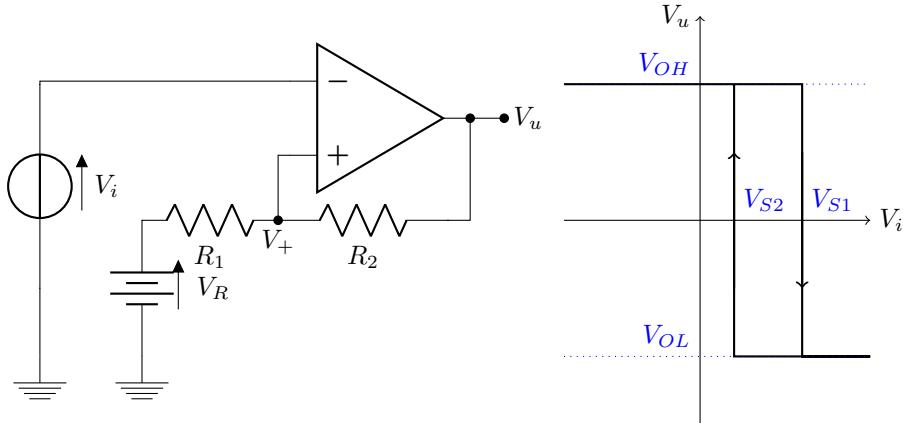


Figure 5.4: Inverting comparator with hysteresis, schematic and transfer characteristic

Inverting comparator

In order to have an hysteresis in the input output transfer function switching must take place at different values of input voltage V_i , according to the current output state. This may be achieved by a **positive** feedback as shown in the circuit of figure 5.4. The output voltage is fed back to the positive input and therefore added to V_i . It is easy to show that in these conditions the output may have only two states, V_{OL} and V_{OH} , and that the operational amplifier operates out of linearity. The linearity assumptions are no more valid and therefore in general we will have $V_+ \neq V_-$. However we will still assume that the input current will remain small and in most cases negligible, being bounded by the input bias currents.

By analyzing the circuit of figure 5.4 we may evaluate the voltage at the non inverting input $V_+(V_u, V_R)$ and compute which are the values of V_i , dependent of the output voltage V_u , at which the comparator will switch states.

$$V_+ = V_R \cdot \frac{R_2}{R_1 + R_2} + V_u \cdot \frac{R_1}{R_1 + R_2}$$

Let us first assume that the input is at a very low voltage and since it is connected to the inverting input that the output will therefore be at the maximum voltage V_{OH} . When V_i rises the switching will take place when it will be at the same voltage of the non inverting pin, that is at voltage V_{S1} :

$$V_{S1} = V_R \cdot \frac{R_2}{R_1 + R_2} + V_{OH} \cdot \frac{R_1}{R_1 + R_2}$$

The output will then switch to V_{OL} , V_+ will decrease, the differential input voltage will become more negative. Let us now assume that the output is at V_{OL} , and the input is higher than V_+ which now has the value V_{S2} given by:

$$V_{S2} = V_R \cdot \frac{R_2}{R_1 + R_2} + V_{OL} \cdot \frac{R_1}{R_1 + R_2}$$

When V_i decreases and reaches this value the output will switch back to V_{OH} . V_{S1} and V_{S2} are the thresholds of the comparator and note that $V_{S1} > V_{S2}$ and that V_R is not the average value of V_{S1} and V_{S2} .

Two important parameters of the system are the amplitude $HY = V_{S1} - V_{S2}$ of the hysteresis and the average value of the thresholds:

$$HY = V_{S1} - V_{S2} = (V_{OH} - V_{OL}) \frac{R_1}{R_1 + R_2}$$

$$\frac{V_{S1} + V_{S2}}{2} = V_R \cdot \frac{R_2}{R_1 + R_2} + \frac{V_{OH} + V_{OL}}{2} \cdot \frac{R_1}{R_1 + R_2}$$

If the output voltage range is symmetric, that is $V_{OL} = -V_{OH}$ then the hysteresis and the average value of the thresholds become:

$$V_{OH} = -V_{OL} \implies V_{S1} - V_{S2} = 2V_{OH} \cdot \frac{R_1}{R_1 + R_2}$$

$$V_{OH} = -V_{OL} \implies \frac{V_{S1} + V_{S2}}{2} = V_R \cdot \frac{R_2}{R_1 + R_2}$$

The average value is always lower than V_R , if $|V_R| > 0$.

Non inverting comparator

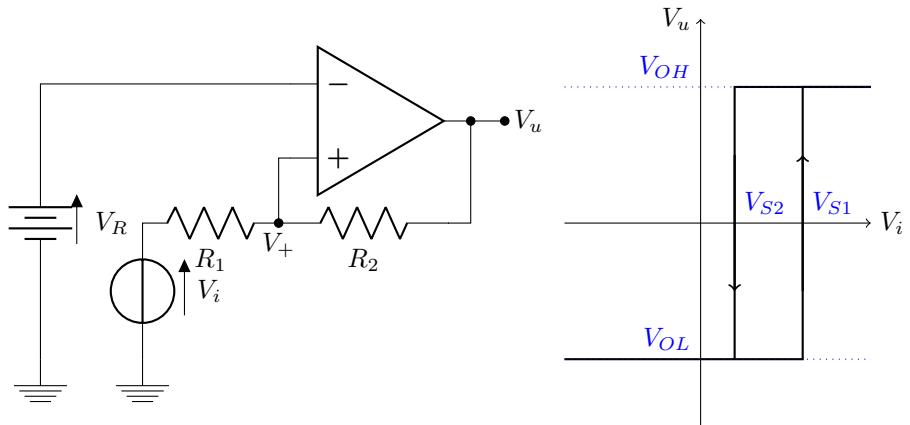


Figure 5.5: Non inverting comparator: schematic and transfer functions

The circuit schematic of the non inverting comparator is shown in figure 5.5; in this case the value of the voltage at the non inverting pin may be computed by applying the superposition principle to V_u and V_i and is given by:

$$V_+ = V_+(V_i, V_u) = V_i \cdot \frac{R_2}{R_1 + R_2} + V_u \cdot \frac{R_1}{R_1 + R_2}$$

Switching takes place when $V_+ = V_R$; if we assume a low output voltage V_{OL} the input value V_{S1} which determines the transition from low to high state may be computed solving equation:

$$V_R = V_{S1} \cdot \frac{R_2}{R_1 + R_2} + V_{OL} \cdot \frac{R_1}{R_1 + R_2}$$

and obtaining:

$$V_{S1} = V_R \cdot \frac{R_1 + R_2}{R_2} - V_{OL} \cdot \frac{R_1}{R_2}$$

In the same way assuming a V_{OH} output state we may evaluate the input voltage V_{S2} corresponding to the high to low transition of the output voltage:

$$V_{S2} = V_R \cdot \frac{R_1 + R_2}{R_2} - V_{OH} \cdot \frac{R_1}{R_2}$$

The values of the hysteresis HY and of the average threshold value are:

$$V_{S1} - V_{S2} = \frac{R_1}{R_2} (V_{OH} - V_{OL})$$

$$\frac{V_{S1} + V_{S2}}{2} = V_R \cdot \frac{R_1 + R_2}{R_2} - (V_{OH} + V_{OL}) \cdot \frac{R_1}{2R_2}$$

If the output voltage range is symmetric, that is when $V_{OL} = -V_{OH}$, the comparator parameters become:

$$V_{S1} - V_{S2} = 2 \frac{R_1}{R_2} V_{OH}$$

$$\frac{V_{S1} + V_{S2}}{2} = V_R \cdot \frac{R_1 + R_2}{R_2}$$

Also for the non inverting comparator the average of the two thresholds is different from V_R and is always higher than it, except for $V_R = 0$.

5.1.6 Circuit implementation of voltage comparators

Any operational amplifier may be used as a voltage comparator but there are circuits which are optimized for that use. Important features are a high slew rate, low offset, a large differential input voltage range, high input impedance and low bias currents, also outside the linear operation range. Many amplifiers designed to operate in linearity have at their input protection diodes which are off when the differential voltage is small and are conducting when the differential voltage is high as in the case of positive feedback. In this case the input currents may become very large. Moreover the internal compensation capacitor has no function when the circuit operates out of linearity but heavily reduces the amplifier slew rate. In addition, operational amplifiers designed for linearity are not optimized for a fast exit from saturation, when switching from one state to the other one.

As already noted, the comparator output is a digital signal which have only two possible values and in many cases the comparator output is connected to the input of a logic gate. A common schematic of the output stage of a comparator is shown in figure 5.6; the power stage is substituted by a single BJT whose

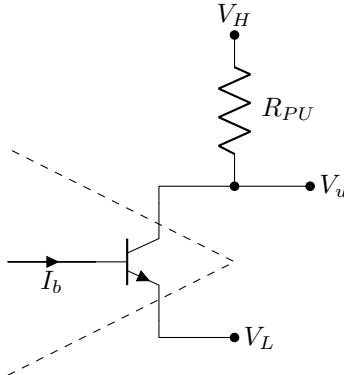


Figure 5.6: Circuit model of the output stage of an open collector voltage comparator.

emitter is connected to a V_{OL} reference voltage. The collector may be connected externally to a V_{OH} voltage through a *pull-up* resistor R_{PU} .

This transistor operates as a switch: if I_b is large enough the BJT saturates, $V_C \sim V_E$ (apart a small drop of approximately 0.2 V), and the output voltage is close to V_{OL} , if $I_b = 0$, the BJT is off and the output voltage is the external voltage V_{OH} . This stage is referred to as **open collector** and we will analyze it better when studying transistors in switching mode and logic circuits.

5.2 Astable Multivibrator

Comparators are a basic component of the *astable multivibrator* which is a square wave generator. It takes the *astable* name from the fact that it has two operating states, each of them being stable only for a finite time, and it switches continuously from one to the other changing periodically its output.

The core element is an inverting comparator with hysteresis with an input signal which rises when the output is high until the switching occurs and then decreases when the output is low until the comparator switches again.

In order to obtain a varying input voltage we may use a capacitor charged by the output voltage. When the voltage across it reaches the threshold voltage, the output switches to the other state and the capacitor discharges until the second threshold is reached. The *charge-discharge cycle* repeats continuously and a periodic waveform is obtained.

5.2.1 Circuit analysis

The astable multivibrator schematic is shown in figure 5.7 and we will assume that the comparator input currents are negligible. Therefore R and C are connected in series.

The inverting comparator has two possible output voltages V_{OH} or V_{OL} and in this case the comparator reference voltage is 0V. This makes the circuit analysis simple but it is possible only if the circuit has a dual power supply. If it is powered by a single power supply then a reference voltage approximately at half of the power supply is used. The analysis of this case is left as an exercise.

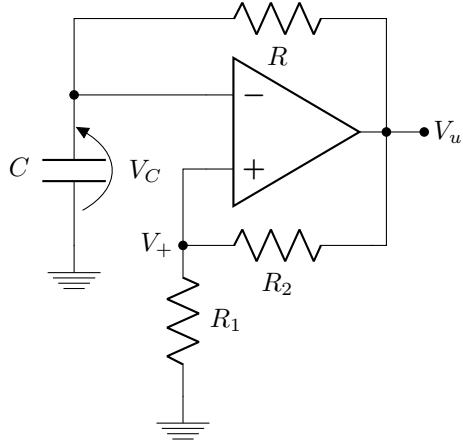


Figure 5.7: Circuit schematic of the astable multivibrator.

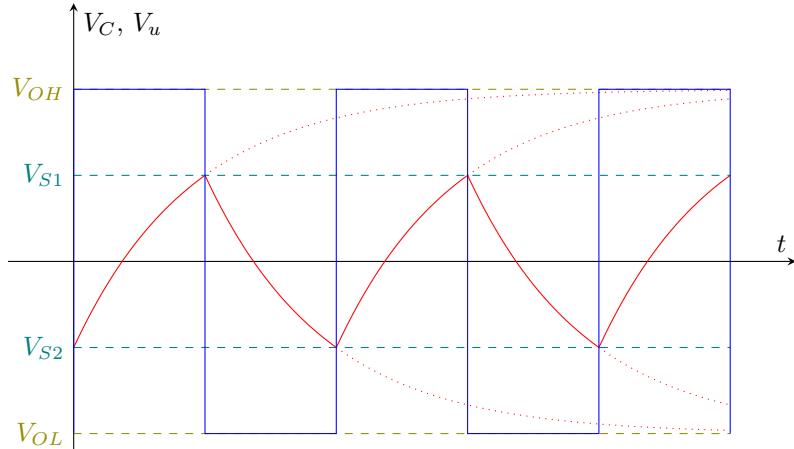


Figure 5.8: V_u e V_C waveform respectively in blue and red.

When the circuit is powered up the output will be either at V_{OL} or V_{OH} . Let us suppose that the initial state will be $V_u = V_{OH}$ and the initial capacitor voltage is 0 V; V_u being constant, the capacitor will charge with an exponential transient with V_{OH} as target final value V_∞ .

$$V_C \propto \left(1 - e^{-t/\tau}\right) \quad \tau = RC$$

When the capacitor voltage reaches the upper comparator threshold, that is when $V_C = V_+ = V_{S1}$ the comparator will switch and the output will become $V_u = V_{OL}$.

At this point the capacitor voltage will start a discharge transient with final target value $V_\infty = V_{OL}$ and when it reaches the V_{S2} threshold, the comparator output will switch to V_{OH} and the cycle will start again.

In summary the capacitor voltage waveform will be a sort of sawtooth, as shown in figure 5.8, with alternating charges and discharges, while the output waveform will be a square wave.

5.2.2 Quantitative analysis

To compute the parameters of the waveform let us assume as $t = 0$ the time at which the output switches from the low to the high level. Moreover considering to output range symmetric, that is $V_{OL} = -V_{OH}$, we may say that the waveform full period is twice the value of one of the two *half periods*.

After the switching we have that $V_\infty = V_u = V_{OH}$, $V_{0+} = V_{S2}$ and that the impedance seen by the capacitance C is just the resistance R and therefore the time constant of the transient is RC . The capacitor voltage is then:

$$v_C(t) = (V_{0+} - V_\infty) e^{-t/\tau} + V_\infty$$

$$\begin{cases} V_{0+} = V_{S2} \\ \tau = R_{eq}C = RC \\ V_\infty = V_{OH} \end{cases} \implies v_C(t) = (V_{S2} - V_{OH}) e^{-t/RC} + V_{OH}$$

After an halfperiod $t = T/2$, v_C has reached V_{S1} , and the comparator switches again. Substituting $T/2$ to T and V_{S1} to v_C we obtain:

$$v_C(T/2) = V_{S1} = (V_{S2} - V_{OH}) e^{-T/(2RC)} + V_{OH}$$

from which:

$$e^{-T/(2RC)} = \frac{V_{S1} - V_{OH}}{V_{S2} - V_{OH}} \implies T = -2RC \ln \left(\frac{V_{S1} - V_{OH}}{V_{S2} - V_{OH}} \right)$$

Using logarithm properties and changing sign both to numerator and denominator we obtain the value of T :

$$T = 2RC \ln \left(\frac{V_{OH} - V_{S2}}{V_{OH} - V_{S1}} \right)$$

Substituting the values of the comparator thresholds we have:

$$\begin{aligned} V_{S1} &= V_{OH} \frac{R_1}{R_1 + R_2}; \quad V_{S2} = V_{OL} \frac{R_1}{R_1 + R_2}; \quad V_{OH} = -V_{OL} \\ V_{OH} - V_{S1} &= V_{OH} \left(1 - \frac{R_1}{R_1 + R_2} \right) = V_{OH} \frac{R_2}{R_1 + R_2} \\ V_{OH} - V_{S2} &= V_{OH} \left(1 + \frac{R_1}{R_1 + R_2} \right) = V_{OH} \frac{2R_1 + R_2}{R_1 + R_2} \\ \implies T &= 2RC \ln \left(\frac{\frac{V_{OH}}{R_1 + R_2} \frac{2R_1 + R_2}{R_2}}{\frac{V_{OH}}{R_1 + R_2} \frac{R_2}{R_1 + R_2}} \right) = 2RC \ln \left(\frac{2R_1 + R_2}{R_2} \right) \end{aligned}$$

This expression shows that if the output range is symmetric, then the frequency of the square wave does not depend on the value of V_{OH} e V_{OL} and therefore is not sensitive to voltage supply variations.

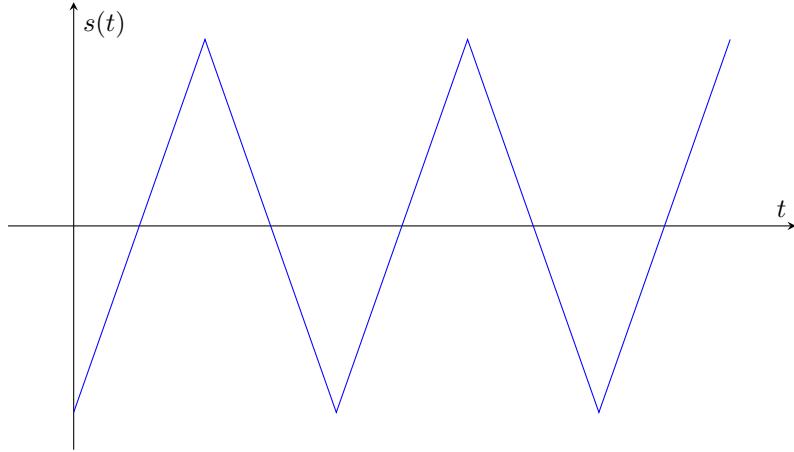


Figure 5.9: Example of triangular waveform.

5.3 Triangular wave generator

The triangular waveform shown in figure 5.9, composed of two linear ramps, is used in a number of circuits ranging from switching power supply controllers to the deflection subsystems of analog oscilloscopes.

This signal may be considered the integral of a square wave and in fact if we analyze the astable multivibrator we may note that the $R-C$ group output is an approximate integral of the output square wave. The approximation comes from the fact that the $R-C$ transfer function is

$$\frac{V_C}{V_u} = \frac{1}{1 + sRC}$$

instead of the real integrator function $(1/(sRC))$. Therefore to obtain a triangular waveform we may use two elementary blocks that is a voltage comparator which provides the constant output at two alternating levels and an integrator providing the triangular waveform. Since we are able to realize only an inverting integrator with transfer function $-1/(sRC)$ we may use a non inverting comparator to obtain a correct operation. The basic circuit is shown in figure 5.10.

5.3.1 Quantitative analysis

Figure 5.11 shows the waveforms at the output of the voltage comparator and of the integrator as a function of time.

Let us assume the output voltage to be V_{OL} . As the time increases the capacitor charges and when the integrator output reaches the upper threshold V_{S1} the comparator (non inverting) switches from V_{OL} to V_{OH} . At this point the output of the inverting integrator will begin to decrease until it reaches the lower threshold V_{S2} of the comparator which will switch to V_{OL} , starting a new cycle. The output voltage of the integrator is:

$$v_T(t) = v_T(0) - \frac{1}{RC} \int_0^t v_Q(t) dt$$

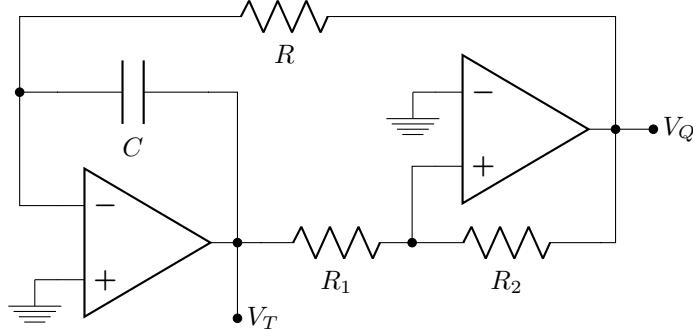


Figure 5.10: Circuit schematic of the triangular waveform generator.

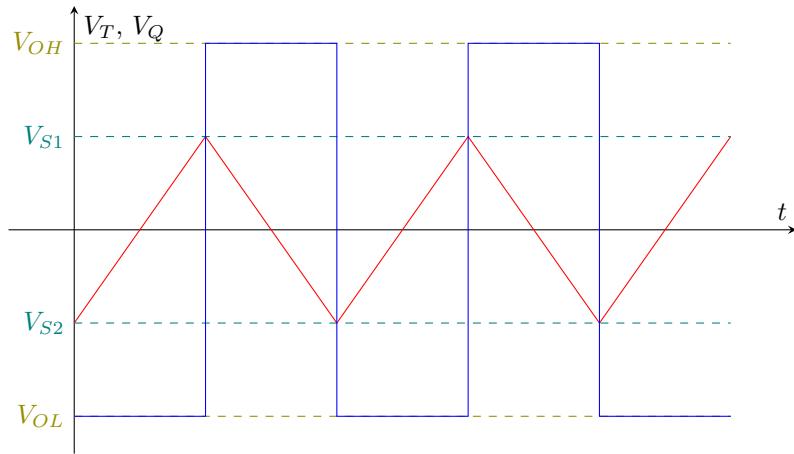


Figure 5.11: Triangular generator waveforms.

To compute the waveform period let us assume as reference time $t = 0$ the switching of the comparator output V_Q from V_{OL} to V_{OH} . Substituting the values in the previous expression we have:

$$v_T(t) = V_{S1} - \frac{V_{OH}}{RC} \cdot t$$

If we also assume a symmetric output voltage range, that is $V_{OL} = -V_{OH}$ the two half periods will be equal and the waveform duty cycle will be 50%. Then it takes a time equal to $T/2$ for the integrator output to reach V_{S2} . Substituting these values in the previous expression and solving it for T we obtain:

$$V_{S2} = V_{S1} - \frac{V_{OH}}{RC} \cdot \frac{T}{2} \implies T = 2RC \frac{V_{S1} - V_{S2}}{V_{OH}}$$

Substituting the values of the comparator thresholds and for symmetric output range the value of the waveform period becomes:

$$V_{S1} - V_{S2} = 2V_{OH} \cdot \frac{R_1}{R_2} \implies T = 4RC \frac{R_1}{R_2}$$

From this basic circuit, with few additional components, we may obtain some additional features as shown in the following design example.

5.3.2 Design example

Basic circuit

Design a triangular waveform generator according to the following specifications:

- Output frequency $f = 500 \text{ Hz}$;
- triangular waveform peak-to-peak output $V_{Tpp} = 8 \text{ V}$;
- Operational amplifier power supply $V_{AL} = \pm 15 \text{ V}$;
- LM741 operational amplifier. Assume for the output voltages: $V_{OH} = 12 \text{ V}$ and $V_{OL} = -12 \text{ V}$.

The triangular waveform amplitude is the comparator hysteresis:

$$V_{S1} - V_{S2} = 8 \text{ V} = 2V_{OH} \cdot \frac{R_1}{R_2}$$

$$\implies 8 \text{ V} = 24 \text{ V} \cdot \frac{R_1}{R_2} \implies R_2 = 3R_1$$

R_1 draws current from the integrator output while R_2 draws it from the comparator and they must be high enough to limit these currents but not too much to reduce the effect of the bias and offset currents which otherwise change the thresholds of the comparator. A reasonable value is $R_2 = 150 \text{ k}\Omega$ giving $R_1 = 47 \text{ k}\Omega$ (values normalized according to the E12 series).

At this point we may determine the values of capacitor C and resistor R by considering the value of the waveform period:

$$T = \frac{1}{f} = 2 \text{ ms} = 4RC \frac{R_1}{R_2}$$

$$\implies 2 \text{ ms} = \frac{4}{3}RC \implies RC = 1.5 \text{ ms}$$

For the choice of the absolute value we should consider that R must be large with respect to the minimum load resistance of the amplifier but not too high. A reasonable choice is $R = 100 \text{ k}\Omega$. Lower values are also acceptable but this choice will be useful for the next modified versions of the circuit.

$$R = 100 \text{ k}\Omega \implies C = 1.5 \times 10^{-3} \cdot 1 \times 10^{-5} \text{ F} = 15 \text{ nF}$$

Variable frequency triangular waveform generator

The triangular waveform generator may be modified in order to make some of its parameters variable. Let us first consider the regulation of the output frequency and let us modify one of the original specifications into the following one:

- Triangular waveform frequency regulation from $f_{min} = 50 \text{ Hz}$ to $f_{MAX} = 500 \text{ Hz}$.

keeping unchanged all the other ones.

From the equation giving the waveform period we note that the period is inversely proportional to V_{OH} , that is the frequency varies *linearly* with V_{OH} which is the output of the comparator and the *input* of the integrator.

$$T = 2RC \frac{V_{S1} - V_{S2}}{V_{OH}} \implies f = \frac{V_{OH}}{2RC(V_{S1} - V_{S2})}$$

We will try to change this voltage considering that the change in other circuit parameters, such as R or C leads to a non linear change in frequency.

The simplest technique for changing the integrator input voltage (which in the following we will refer to as V_3) is to connect a potentiometer as shown in figure 5.12. The potentiometer is equivalent to two resistors with a common node, the cursor, with the property that the sum of the two resistances is constant while the value of each one of them depends on the cursor position.

There are linear and logarithmic potentiometers and in linear ones the value of the resistances is linearly proportional to the potentiometer angular position. Let us refer to the potentiometer position as $x \in [0; 1]$ and $x = 0$ when the cursor is in the bottom position and $x = 1$ when it is in the topmost one. Then x will be the fraction of P_1 between the cursor of P_1 and R_3 . Assuming the current in R to be negligible we have:

$$V_3 = V_{OH} \cdot \frac{xP_1 + R_3}{xP_1 + R_3 + (1-x)P_1} = V_{OH} \cdot \frac{xP_1 + R_3}{P_1 + R_3};$$

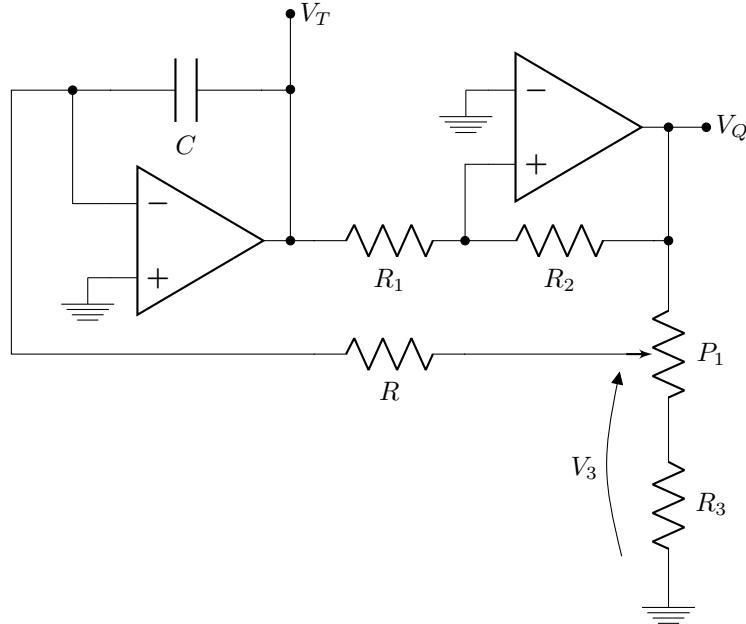


Figure 5.12: Triangular waveform generator with frequency variable by potentiometer P_1 .

Moving the cursor and changing x the integrator input voltage is changed and therefore also the triangular signal slope and frequency. If $x = 1$ then $V_3 = V_{OH}$

and the circuit operates as in the previous case without the potentiometer; if $x = 0$, the input voltage to the integrator is minimum and the frequency the lowest one, corresponding to the f_{min} of the specifications and resistance R_3 determines this frequency.

Remembering that $f_{MAX} = 500 \text{ Hz}$ as in the previous design and that $f_{min} = 50 \text{ Hz}$, we may determine the relative values of P_1 and R_3 by setting for x values 1 and 0 and obtaining:

$$\frac{f_{min}}{f_{MAX}} = \frac{\min V_3}{\max V_3} = \frac{V_3(x=0)}{V_3(x=1)} = \frac{R_3}{P_1 + R_3}$$

$$\frac{50 \text{ Hz}}{500 \text{ Hz}} = \frac{1}{10} = \frac{R_3}{P_1 + R_3} \Rightarrow R_3 = \frac{P_1}{9} \simeq \frac{1}{10} P_1$$

The final approximation guarantees a frequency range of one decade even considering resistor tolerances.

Setting the design parameters The previous analysis is approximate because V_3 has been evaluated without taking into account the equivalent resistance R of the voltage divider composed of $(1-x)P_1$ and $xP_1 + R_3$. We may consider the Thevenin equivalent circuit at the integrator input, as shown in figure 5.13.

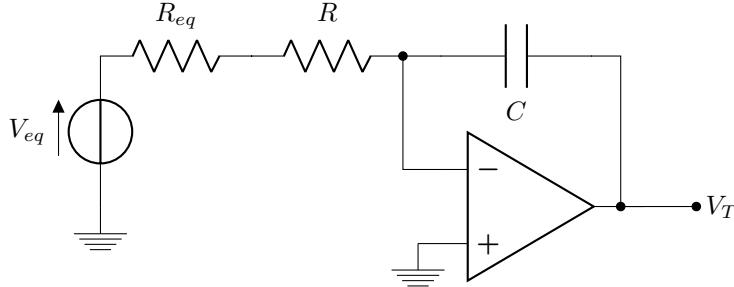


Figure 5.13: Equivalent circuit of the P_1 potentiometer.

Resistance R_{eq} is in series to R and therefore modifies the input frequency. In addition, since R_{eq} is not constant, the frequency variation is not any more linear with the potentiometer position. One solution is to insert a voltage follower after the potentiometer but the circuit would use a larger area and become more expensive, another one is to make the influence of R_{eq} negligible with respect to R . Let us compute the value of ∂R_{eq} .

$$R_{eq} = R_{eq}(x) = \frac{(xP_1 + R_3)(1-x)P_1}{P_1 + R_3}$$

To minimize its contribution it is necessary that $R \gg \max R_{eq}$. If we employ 5% resistors and 10% capacitors then, imposing the maximum $R_{eq} = 5\%R$, its error contribution is comparable to that due to passive components tolerance.

$R_{eq}(x)$ has a parabolic dependence on x and it is possible to compute its maximum value by setting to 0 the value of the derivative obtaining:

$$\max R_{eq} = \begin{cases} \frac{P_1 + R_3}{4} & \text{if } P_1 > R_3 \\ \frac{P_1 \cdot R_3}{P_1 + R_3} & \text{if } P_1 \leq R_3 \end{cases}$$

using the previously derived relation between R_3 and P_1 , we have

$$\begin{aligned} \max R_{eq} &= \frac{11}{10} P_1 \cdot \frac{1}{4} \simeq \frac{1}{4} P_1 \\ \implies \frac{1}{4} P_1 &\leq 5\%R \implies P_1 \leq 4 \cdot 5 \text{ k}\Omega = 20 \text{ k}\Omega \end{aligned}$$

there are commercially available potentiometers of the E3 series with values 1; 2.2; 4.7 per decade. We may chose $P_1 = 10 \text{ k}\Omega$ and this value satisfies the previous condition and at the same time does not load too much the output of the operational amplifier

Variable average value triangular waveform generator

An additional feature to the waveform generator is to make the average value of the triangular waveform variable, that is to add to it a constant value which may be varied independently of other parameters. In the basic circuit this value is 0 because the two thresholds are symmetric. Let us therefore add a further specification to our design, leaving unchanged the other ones:

- Average value of the triangular waveform variable from $V_{offset,T_{min}} = -4 \text{ V}$ e $V_{offset,T_{MAX}} = 4 \text{ V}$.

As noted before we must change the average value of the thresholds and it may be done by adding a reference voltage to the inverting pin of the voltage comparator. Since the voltage must be variable under user control we add to the circuit a second potentiometer P_2 as shown in figure 5.14.

$$V_{offset,T} = \frac{V_{S1} + V_{S2}}{2} = V_{REF} \cdot \frac{R_1 + R_2}{R_2} = V_{REF} \left(1 + \frac{R_1}{R_2} \right) = \frac{4}{3} V_{REF}$$

For maximum $V_{offset,T}$ we must have maximum reference voltage V_{REF} and the same holds for the minimum value, which is symmetrical.

$$\max V_{offset,T} = +4 \text{ V} \implies \frac{4}{3} \max V_{REF} = 4 \text{ V} \implies \max V_{REF} = 3 \text{ V}$$

$$\min V_{offset,T} = -4 \text{ V} \implies \frac{4}{3} \min V_{REF} = -4 \text{ V} \implies \min V_{REF} = -3 \text{ V}$$

V_{REF} is obtained from the two power supplies using the voltage divider composed of P_2 , R_4 and R_5 and the reference voltage is dependent on the position of the potentiometer's cursor. Since the minimum and maximum reference voltages must be symmetric we will have $R_4 = R_5$. The maximum value corresponds to the cursor in the rightmost position in figure 5.14. The circuit and the voltages

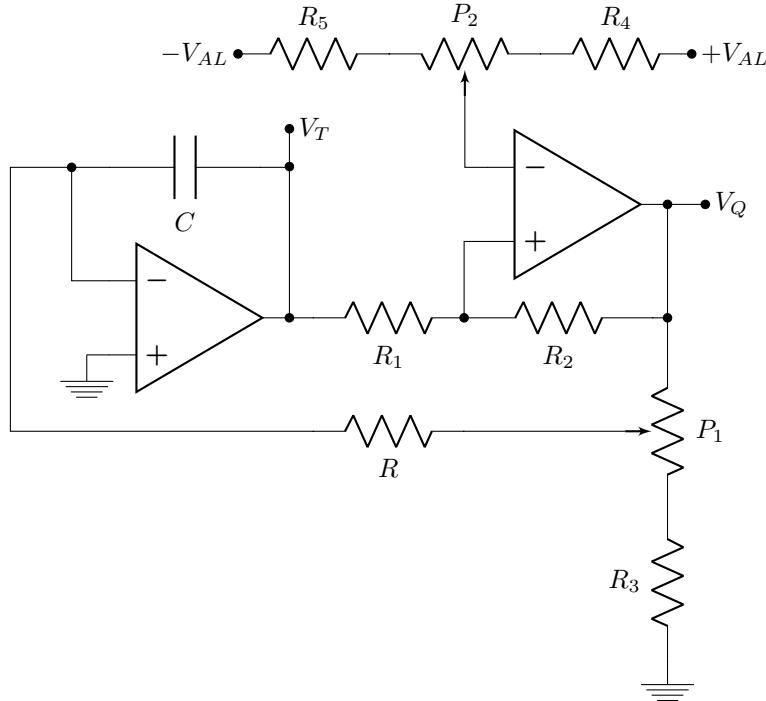


Figure 5.14: Variable triangular waveform offset by using P_2

being symmetric, the central position of P_2 will be at 0 V and we may therefore write:

$$\max V_{REF} = \frac{P_2/2}{P_2/2 + R_4} \cdot V_{AL}$$

$$\max V_{REF} = 3 \text{ V} = \frac{P_2/2}{P_2/2 + R_4} \cdot (+15 \text{ V}) \implies \frac{P_2}{2} = \frac{1}{5} \left(\frac{P_2}{2} + R_4 \right)$$

$$\frac{1}{5}R_4 = \frac{P_2}{2} - \frac{1}{5}P_2 \cdot \frac{1}{2} \implies R_4 = 2P_2$$

We may choose from the E12 series the values $R_4 \simeq 18 \text{ k}\Omega$, and $P_2 = 10 \text{ k}\Omega$ using a potentiometer equal to P_1 .

Triangular waveform generator with variable duty-cycle

With a further slight circuit change it is possible to obtain a waveform generator with independently variable duty cycle. The rising and falling times of the triangular waveform are changed, without changing their sum, that is the period and therefore keeping the frequency constant. Let us then add this last specification keeping the other ones unchanged:

- Square waveform duty-cycle variable from $DC_{min} = 25\%$ to $DC_{MAX} = 75\%$.

Starting from the definition of Duty-cycle DC we see that it depends on the ratio between the triangular wave falling time T_1 and the total period T . T_1 depends on a number of parameters

$$DC = \frac{T_1}{T}; \quad T_1 = \frac{V_{S1} - V_{S2}}{V_{OH}} RC; \quad T_2 = T - T_1$$

Considering that the threshold amplitude is already set and that the capacitance may not be easily varied the only parameter free for duty cycle variation is the integrator resistance R . The basic idea is to employ a non linear network, based on diodes and shown in figure 5.15 to change the value of the resistance during the charge and the discharge of the capacitor. The diodes operate as switches which are directly controlled by the polarity of the output voltage of the comparator. Diode D_1 is on when the output voltage is high and D_2 when the output is low.

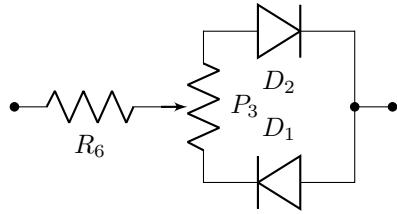


Figure 5.15: The diode resistor network substituting resistor R for obtaining a variable duty cycle.

In each phase of the waveform the current will flow in only one of the two variable resistances composing the potentiometer and, according to the cursor position of P_3 , the capacitor will charge/discharge through a different resistance giving as a result different duration of the phases. However the sum of the two resistances remains constant and therefore also the waveform period will be constant. The complete circuit is shown in figure 5.16.

If we define $R_a = yP_3$ and $R_b = (1 - y)P_3$ and we substitute the value of these resistances in the expressions of T_1 and T_2 we obtain:

$$T = T_1 + T_2 = \frac{V_{S1} - V_{S2}}{V_{OH}} \cdot C \cdot (R_a + R_b + 2 \cdot R_6)$$

$$DC = \frac{T_1}{T_1 + T_2} = \frac{R_a + R_6}{R_a + R_b + 2 \cdot R_6}$$

We have also that $R_a + R_b = P_3$ and by varying the position of the cursor, that is y , it is possible to vary the waveform duty-cycle. Resistor R_6 sets the minimum value of the duty cycle and avoids short circuiting the comparator output to the integrator virtual ground. Substituting the values of R_a and R_b we have :

$$T = \frac{V_{S1} - V_{S2}}{V_{OH}} \cdot C \cdot (P_3 + 2 \cdot R_6)$$

$$DC = \frac{yP_3 + R_6}{P_3 + 2 \cdot R_6}$$

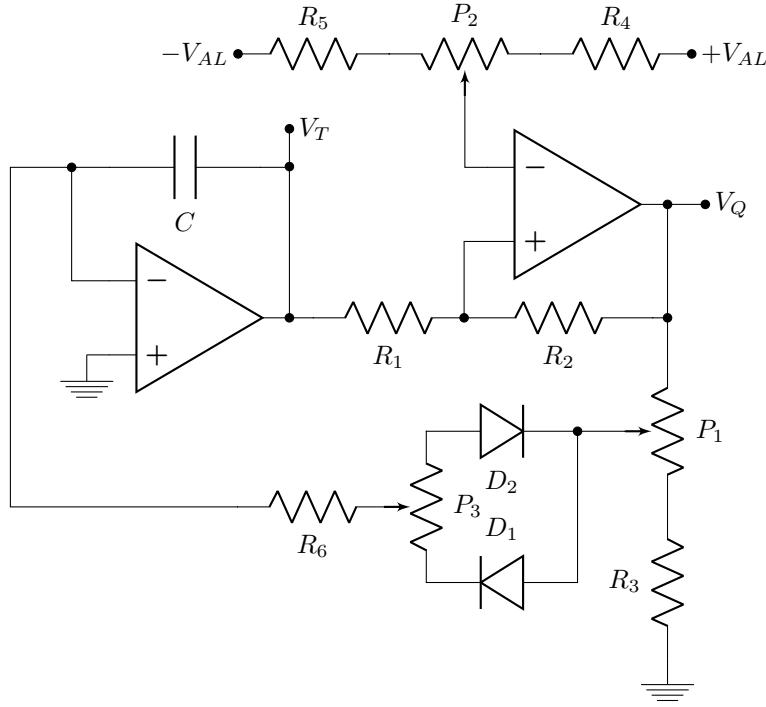


Figure 5.16: Triangular waveform generator with variable *dutycycle*.

Component values In order to add the duty cycle variation by keeping unchanged the other circuit parameters, let us observe that in the previous design with fixed $R = 100 \text{ k}\Omega$ the minimum period was given by:

$$T = \frac{V_{S1} - V_{S2}}{V_{OH}} \cdot C \cdot 2R$$

and in new circuit by:

$$T = \frac{V_{S1} - V_{S2}}{V_{OH}} \cdot C \cdot (P_3 + 2R_6)$$

Comparing the two expression we have $2R = P_3 + 2R_6$. The duty cycle as a function of resistances is given by:

$$DC = DC(y) = \frac{yP_3 + R_6}{P_3 + 2R_6}$$

The minimum duty cycle, that is 25% , corresponds to the lowest potentiometer position, that is $y = 0$,. Putting this value in the previous equation we may obtain the values of P_3 and R_6 :

$$\min DC = DC(y = 0) = \frac{R_6}{P_3 + 2R_6} = \frac{1}{4} \implies R_6 = \frac{200 \text{ k}\Omega}{4} = 50 \text{ k}\Omega \simeq 47 \text{ k}\Omega$$

$$2R_6 \simeq 100 \text{ k}\Omega \implies P_3 + 100 \text{ k}\Omega \simeq 200 \text{ k}\Omega \implies P_3 \simeq 100 \text{ k}\Omega$$

If we set the cursor in the opposite position, that is $y = 1$ we obtain the complementary value of DC which is:

$$\left. \frac{R_6 + yP_3}{P_3 + 2R_6} \right|_{y=1} \cdot 100 = \frac{147 \text{ k}\Omega}{200 \text{ k}\Omega} \simeq 75\%$$

and we note that the duty-cycle range is symmetric, as required by the specifications. Since R_6 is common both to capacitor charge and discharge the duty-cycle range is symmetric by construction and the value of this resistor is sufficiently high to make negligible the influence of P_1 and R_3 on the waveform period.

Asymmetric duty-cycle range If we need an **asymmetric** duty-cycle range then it is necessary to insert two different fixed resistors in series to diodes D_1 and D_2 as shown in figure 5.17. These resistors set two different limits to minimum duration of the two phases. The analysis and the design may be carried on with the same methodology used for the symmetric range.

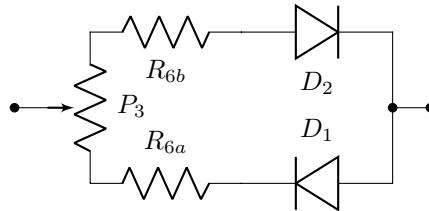


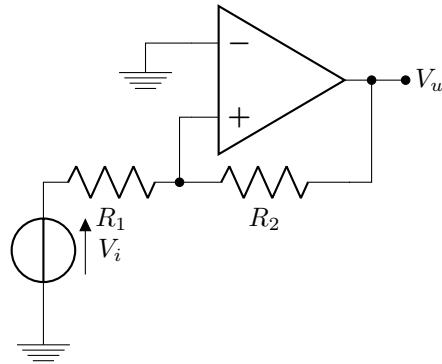
Figure 5.17: The diode resistor network for variable asymmetric duty cycle.

The previous analysis is approximate because we have assumed the diodes to be ideal. If necessary, we could perform a more exact analysis and design taking into account the voltage drop V_γ across them when conducting. The influence of these voltages is negligible when the cursor is in the top position (high frequency) and increases for lower V_3 values (lower frequencies) giving a value of frequency lower than that for ideal diodes.

5.4 Exercises

Exercise 5.1. Non inverting comparator

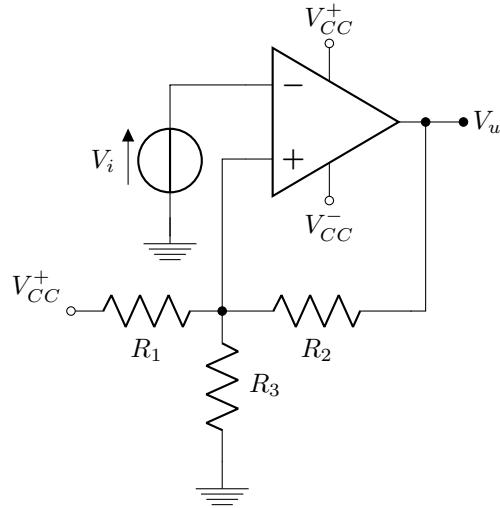
Given the non inverting comparator circuit shown in figure 5.18 and assuming $V_{OH} = 10V$ and $V_{OL} = -8V$, draw the input-output transfer characteristic of the circuit.



$$R_1 = 18 \text{ k}\Omega \quad R_2 = 47 \text{ k}\Omega$$

Figure 5.18: Non inverting comparator

Exercise 5.2. Inverting comparator Given the circuit of figure 5.19



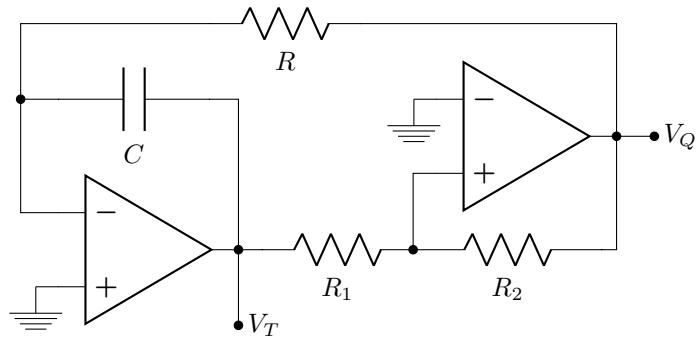
$$R_1 = 180 \text{ k}\Omega \quad R_2 = 68 \text{ k}\Omega \quad R_3 = 18 \text{ k}\Omega \\ V_{CC}^+ = 12 \text{ V} \quad V_{CC}^- = 0 \text{ V} \quad V_{OH} = V_{CC}^+ - 2 \text{ V} \quad V_{OL} = V_{CC}^-$$

Figure 5.19: Inverting comparator

compute and draw its input-output transfer characteristic.

Exercise 5.3. Triangle wave generator

Given the circuit of figure 5.20



$$R_1 = 18 \text{ k}\Omega \quad R_2 = 82 \text{ k}\Omega \quad R = 10 \text{ k}\Omega \quad C = 33 \text{ nF} \quad V_{umin} = -12 \text{ V} \quad V_{umax} = 12 \text{ V}$$

Figure 5.20: Triangle wave oscillator

determine:

- the frequency at which the circuit oscillates.
- the peak-to-peak amplitude of the triangular waveform

Chapter 6

Silicon diodes, BJTs and MOSFETs

Contents

6.1 Basic concepts on semiconductors	6-2
6.1.1 Band model	6-3
6.1.2 Generation and recombination	6-4
6.1.3 Doped semiconductors	6-4
6.1.4 Semiconductor currents	6-5
6.2 Diode	6-7
6.2.1 pn junction	6-7
6.2.2 The diode	6-8
6.2.3 Diode static characteristic	6-10
6.2.4 Using diodes	6-12
6.2.5 Zener diodes	6-18
6.2.6 Diode applications	6-21
6.3 Bipolar transistor	6-25
6.3.1 The BJT as amplifier	6-29
6.4 MOSFET devices	6-42
6.4.1 MOS capacitor	6-42
6.4.2 MOSFET transistor	6-43
6.4.3 p channel MOSFET	6-46
6.4.4 Breakdown of MOSFET transistors	6-48
6.4.5 MOSFET symbols	6-49
6.4.6 The MOSFET as amplifier	6-49
6.5 Exercises	6-54

In this chapter we will introduce the structure and physical behavior of semiconductors. We will then use silicon to build basic devices such as diodes, bipolar junction transistors (BJT) and metal oxide semiconductor field effect transistors (MOSFET).

We will examine the behavior of BJTs and MOSFETs as amplifiers, which are the base of analog circuits, while in subsequent chapters we will study how to use them as switches, to build digital circuits.

6.1 Basic concepts on semiconductors

One of the most important properties that characterize the electrical behavior of a material is its electrical conductivity σ , [S cm^{-1}], or the resistivity $\rho = 1/\sigma$, [$\Omega \text{ cm}$].

The conductivity is used to classify materials as:

- Insulators, if σ is very low ($\sigma < 10^{-10}$);
- Conductors, if σ is very high ($\sigma > 2 \cdot 10^2$);
- Semiconductors if σ is in the intermediate range.

Semiconductor based electronic devices are generally built using silicon, germanium or gallium arsenide. Nowadays silicon is the most used material to create integrated circuits, so we will focus on silicon in the following, warning that, while gallium arsenide or other new semiconductors are at the moment used only in very specific applications, in future they will probably extend to other areas.

Silicon is a semiconductor that belongs to the group IV of the periodic table¹. It has therefore four valence electrons in the outer shell.

Pure silicon (also named *intrinsic* silicon) has a crystalline structure characterized by an ordered spatial arrangement:

- Atoms are arranged in a regular and periodic *crystal lattice*;
- Atoms share the four electrons in the outermost shell with four adjacent atoms creating *covalent bonds*.

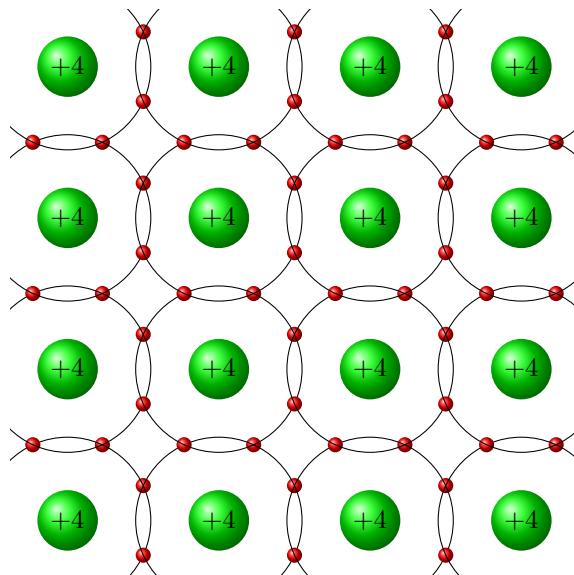


Figure 6.1: 2D representation of silicon crystal showing covalent bonds.

Figure 6.1 shows a 2D representation of silicon crystal structure.

¹In modern physics it is named group 14, or carbon group, but in semiconductor physics the old notation of group IV is still in use.

6.1.1 Band model

In a crystal formed by N silicon atoms there are $4N$ valence electrons, two in the s level and 2 in the p level for each atom.

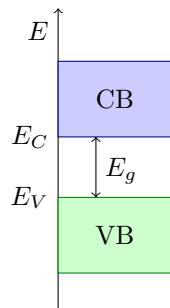


Figure 6.2: Graphic representation of band model.

This results in a typical energy band structure in which there are three bands, as depicted in figure 6.2:

- a high energy *conduction band* CB with $4N$ states;
- a low energy *valence band* VB with $4N$ states;
- an empty *forbidden band* FB which separates CB from VB. The difference between the highest energy level of the valence band and the lowest energy level of the conduction band is called *energy gap* E_g .

At very low temperature (close to 0K) all of the $4N$ valence electrons occupy the VB, increasing the temperature a few of them can occupy energy levels in the CB. When all electrons are in the VB there cannot be charge motion even in presence of an electric field.

At room temperature some electrons move to the CB by thermal ionization. When an electron moves to the CB, it creates a *hole* in the VB, that is in the valence band there is a positive charge $+q$ due to the missing charge of the electron. In this case there can be electric conduction in the presence of an applied electric field for two reasons:

- electrons in the CB are free to move in the crystal lattice;
- holes in the VB can move from one atom to the other, that is the broken covalent bond can move between adjacent atoms.

Given that the number of atoms is very large, it is convenient to consider the average concentration of free charged particles (called *carriers*) per unit volume. We define as n the density of free electrons in the CB and as p the density of holes in the VB. Each electron has an associated charge of $-q$, each hole has associated charge $+q$.

In order to jump from VB to CB, an electron has to absorb an amount of energy at least as great as the energy gap E_g . The energy source can be the temperature. Therefore the smaller the E_g the more probable is the transition to VB to CB. Table 6.1 compares band gap and free electron density data at room temperature of a conductor, a semiconductor and an insulator.

	Conductor (Sn)	Semiconductor (Si)	Insulator (Diamond)
Lattice constant a / Å	6.49	5.43	3.57
Band gap E_g / eV	0.1	1.124	5.47
Free electrons density n / cm ⁻³	10 ²²	10 ¹⁰	10

Table 6.1: Comparison between different materials at 300 K.

6.1.2 Generation and recombination

In an intrinsic semiconductor at thermal equilibrium, the number of free electrons and holes results from the equilibrium of two mechanisms:

- *generation* of one free electron and one hole from a valence electron. The electron absorbs energy greater or equal to E_g and thus moves to CB.
- *recombination* of electron-hole pairs. One of the free electrons is captured by a broken covalent bond and returns to the VB.

The consequence is that the concentration of free electrons n and the concentration of free holes p must be the same:

$$n = p = n_i$$

where n_i is the *intrinsic concentration* of the semiconductor and depends only on the material. The concentration increases exponentially with the temperature.

6.1.3 Doped semiconductors

The importance of semiconductors in electronics is due to the possibility of changing the electrical conductivity of the material by several orders of magnitude by altering the crystal structure. This is done by substituting some silicon atoms with atoms with different outer shell properties. These atoms are called *dopant* atoms.

Dopant atoms are usually from different groups of the periodic table different from the host semiconductor group, such that, in silicon, one has:

- more than 4 electrons in the outer shell: they are called *donors* and the resulting semiconductor is called n doped;
- less than 4 electrons in the outer shell: they are called *acceptors* and the resulting semiconductor is p doped.

Figure 6.3 shows a 2D representation of a n doped silicon crystal.

n doped silicon is normally obtained by using arsenic as dopant. Arsenic has 5 electrons in the outer shell. Four of these electrons form a covalent bond with neighboring silicon atoms, while the fifth electron is only loosely coupled to the arsenic nucleus.

At ambient temperature, the thermal energy is sufficient to set free the fifth electron, thus we have one more free electron in the CB, while the arsenic atom gets ionized so that we have a fixed positive charge in its position. If we call N_D the concentration of donor atoms per cubic centimeter, at 300 K that concentration is equal to that of ionized donor atoms.

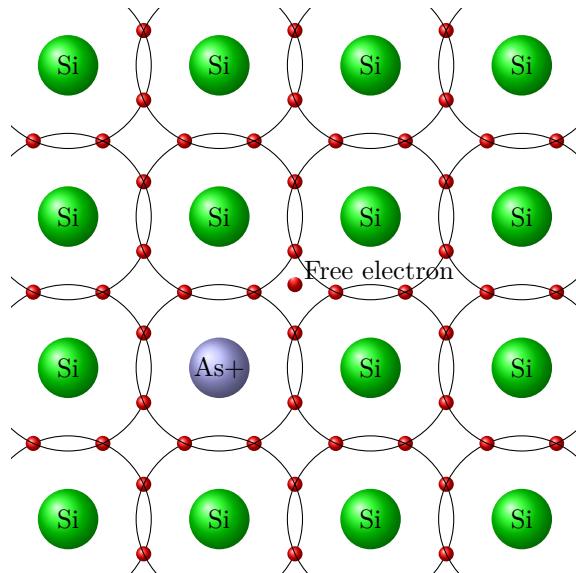


Figure 6.3: 2D representation of n doped silicon crystal.

Figure 6.4 shows a 2D representation of a p doped silicon crystal.

p doped silicon is normally obtained by using boron as dopant. Boron has 3 electrons in the outer shell. All the boron electrons form a covalent bond with neighboring silicon atoms. A small amount of energy is sufficient to capture one more electron from the valence band to complete the fourth bond. At room temperature the thermal energy is enough to capture the electron from VB thus creating a free hole and ionizing the dopant boron atom which becomes a fixed negative charge.

If we call N_A the concentration of acceptor atoms per cubic centimeter, at 300 K that concentration is equal to that of ionized acceptor atoms.

The two types of doped silicon have complementary structure. In the n -type silicon we have free electrons, that is mobile negative charges, and positive ions, that is fixed positive charges, whereas in p -type silicon we have free holes, that is mobile positive charges, and fixed negative ions, that is bound negative charges. Electrons in n -type and holes in p -type are called *majority carriers*, the others *minority carriers*.

In figure 6.5 we can see a graphic representation of the two types of doped silicon, where fixed charges are depicted in circles.

6.1.4 Semiconductor currents

When a semiconductor is in thermal equilibrium there is no exchange of energy with the external world. In this case there is no net current in the device. When we apply electric fields and include our silicon device in a circuit, we move it out of equilibrium and, under certain conditions, there will be a current in the device. In general, carrier density in the device will be different from the equilibrium value and, in general, it will change with position and time.

The variations depend on three different mechanisms:

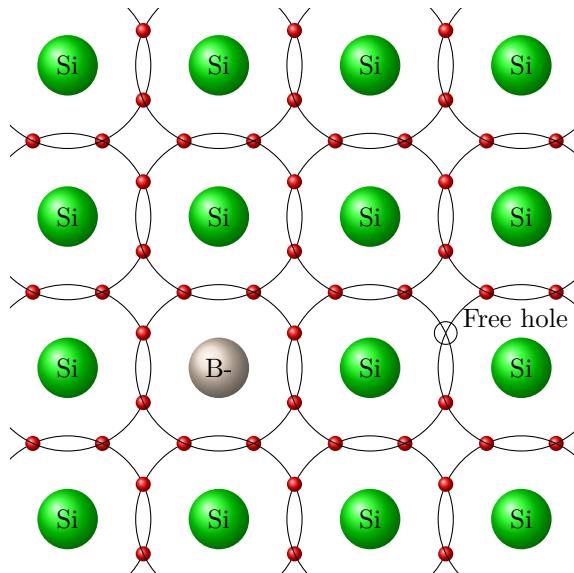


Figure 6.4: 2D representation of *p* doped silicon crystal.

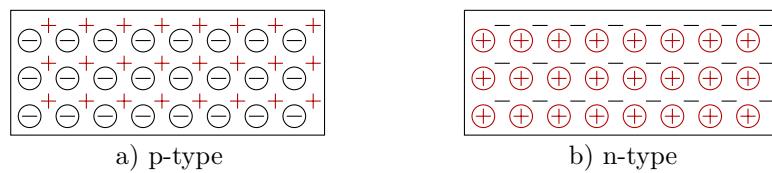


Figure 6.5: Majority and minority carriers in doped silicon

- carrier motion by drift (induced by applied electric field E): *drift current*, defined by the microscopic Ohm law;
- carrier motion by diffusion: *diffusion current*;
- generation and recombination mechanisms.

We will briefly overview the two currents, but the detailed study of the mechanisms are out of the scope of this course.

Drift current

When we apply an external electric field E to a semiconductor sample, the applied field induces a force that superimposes to the instantaneous random motion of the single charges due to thermal energy. This force creates therefore an electric current. The carriers will move with a certain velocity which is proportional to the applied electric field through a coefficient which is called *mobility* and indicated with the greek letter μ . It is possible then to express the speed of electrons, v_n as:

$$v_n = -\mu_n E$$

and the speed of holes v_p as:

$$v_p = \mu_p E$$

Mobility μ_n and μ_p are positive quantities and are expressed in square centimeters per Volt per second. Speed of electrons is negative because they are negative charges so they move in the opposite direction with respect to E .

The conductivity of a doped silicon sample is proportional to the concentration of carriers multiplied by their mobility. The mobility depends on the amount of doping but as general rule we have that mobility of electrons is about three times the mobility of holes.

We can simply find the conductivity of a sample having n free electrons per unit volume and p free holes per unit volume, by summing their contributions:

$$\sigma = qn\mu_n + qp\mu_p$$

When we apply an electric field to a doped silicon sample, we will therefore observe a current depending on the conductivity and the geometric properties of the sample. This current is called *drift current*.

Diffusion current

Free charge carriers in a semiconductor behave like gas particles in a closed volume: they tend to distribute evenly in all available space. If for any reason the charge density is non uniform, then the particles will *diffuse* from high density areas to low density ones with a diffusion intensity proportional to the *gradient* of the charge density, that is the first order spacial derivative of the density. The carriers direction is opposite to the gradient.

Given that the particles are charged, the diffusion process causes an electric current proportional to the diffusion intensity.

It is possible to quantify the diffusion current density J_{diff} with the following equations:

$$J_{n,\text{diff}} = qD_n \frac{\delta n}{\delta x} \quad J_{p,\text{diff}} = -qD_p \frac{\delta p}{\delta x}$$

Where we considered a variation in density in the x direction. D_n and D_p are called diffusion coefficients respectively of electrons and holes. The sign of the two currents are opposite due to negative charge of electrons and positive charge of the holes.

The total current in a semiconductor is the sum of diffusion current and drift current. If no external electric field is applied to the device, the two currents compensate, giving a null total current.

6.2 Diode

The first semiconductor device that we will introduce is the silicon diode. To understand the working principles of the diode we have to study the main physical properties of the so-called pn junction.

6.2.1 pn junction

A pn junction is made by two adjacent regions of a monocrystal semiconductor, one being p doped, the other n doped.

The most simple pn junction is the *abrupt* junction where the two sides of the junction are uniformly doped, with dopant densities respectively N_A for the *p* side, N_D for the *n* side. It is not possible to simply weld two pieces of silicon with different doping. We will see in the technology section how a junction is created in practice.

Let's consider what happens if we ideally join two pieces of silicon with opposite doping. Before joining the two pieces, the *p* region and the *n* region were in equilibrium, like in the previous figure 6.5. The *p*-side was rich of free holes and the *n* side was rich of free electrons. In both regions, the free carriers were compensated respectively by ionized acceptor and donor atoms. When we connect the two regions, we have a transient where the majority carriers of one region diffuse thru the other region because of the gradient in the carrier density. This movement creates a region around the junction where there are no free carriers and only the fixed ionized dopant atoms remain. This region is called *depletion region* or *space charge region*. The situation is depicted in figure 6.6.

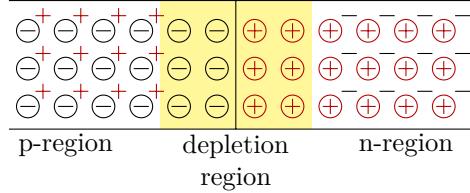


Figure 6.6: pn junction.

In the depletion region, the diffusion of majority carriers to the opposite region exposes the fixed charge of acceptor/donor ions. Having a non null total charge density, means that a potential barrier grows up in the depletion region. This barrier opposes to further diffusion of majority chargers but accelerates the drift current due to thermal energy of minority carriers, which is proportional to the electric field. The system reaches an equilibrium state in which a small drift current due to minority carriers is perfectly balanced by a small diffusion current due to majority carriers.

If the junction is isolated, that is if no external potential is applied, the net current is 0 and a certain potential V_{bi} (builtin potential) develops across the junction.

Figure 6.7 shows the charge density and the potential barrier in the depletion region, with no external voltage applied to the junction.

6.2.2 The diode

Now we want to understand what happens if we build a bipole by connecting two electrodes at the far ends of a pn junction and apply a voltage difference. The device we build is called *diode*. The terminal connected to the *p* side of the junction is called *anode*, the other *cathode*. Figure 6.8 shows a positive potential applied to a diode, with positive current entering the diode, together with the conventional symbol for the diode.

If we apply a voltage difference V to the junction, we alter the equilibrium, so the junction is brought out of balance and we can measure a current I which

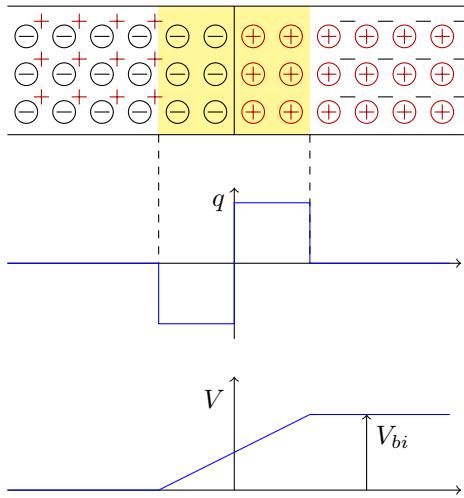


Figure 6.7: Depletion region, charge and built-in potential.

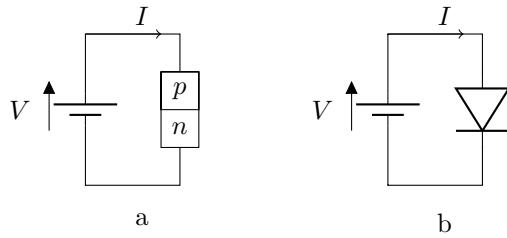


Figure 6.8: a) voltage and current in a pn junction. b) using conventional symbol for diode.

crosses the junction.

We want to evaluate the static characteristics of the diode, that is an equation linking the values of applied voltage and current crossing the junction, assuming that the device works under static conditions (the voltage was applied long time before the measurement of the current and any transient is already finished).

If we apply an external voltage to the device, we alter the total potential drop across it, and this results in a change of the width of the depletion region. If the current crossing the device is small, the ohmic contribution of the part of the device outside the depletion region is negligible, so that all the potential is applied to the depletion region itself. This will result in a change in the amplitude of the depletion region.

Reverse bias

If we apply a negative voltage to the junction (always refer to figure 6.8 for polarity), we increase the voltage drop across the depletion region, like depicted in figure 6.9. In this case we say that the junction is *reverse biased*.

The potential energy barrier increases its amplitude to $V_{bi} - V$ (which is larger than V_{bi} because V is negative), so the diffusion of majority carriers is

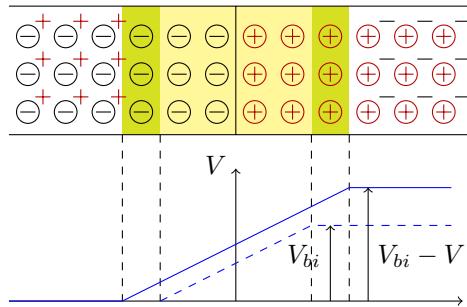


Figure 6.9: Reverse biased junction.

further prevented, therefore they give no contribution to the current. The only current flowing in the device is due to the drift current of minority carriers, electrons drifting to the n side, holes towards the p side. The resulting current is negative and, given that the available minority carriers are few, its magnitude is very small.

It can be shown that the amplitude of the current is almost independent of the applied negative voltage. This current is called *reverse saturation current* and will be named I_s in the following.

Forward bias

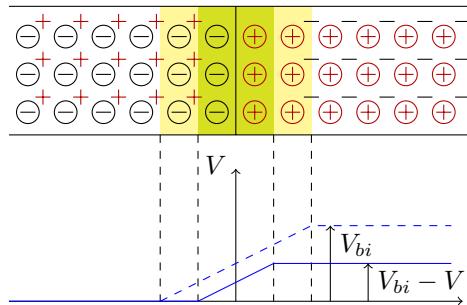


Figure 6.10: Forward biased junction.

If we apply a positive voltage to the junction, we decrease the voltage drop across the depletion region to the value $V_{bi} - V < V_{bi}$. In this case we say that the junction is *forward biased*. The amplitude of the potential energy barrier decreases. This is represented in figure 6.10.

The decrease of the amplitude of the potential energy barrier creates an exponential increase of the diffusion current of majority carriers towards the opposite side. We then have positive current and its amplitude can be very large, due to the large amount of available majority carriers.

6.2.3 Diode static characteristic

The processes outlined above can be expressed quantitatively and it is possible to demonstrate that the current in the diode obeys an exponential function. We

will not go thru the demonstration because this is out of the scope of this course, but we will study the resulting equation (Shockley ideal diode equation):

$$I = I_s \left(e^{V/(\eta V_T)} - 1 \right)$$

The parameters of the above equation are:

- I_s , previously introduced, is the reverse saturation current of the junction. It is very small but it increases exponentially with the temperature. It doubles every 10°C .
- η is a coefficient ranging approximately from 1 to 2. Normally it is close to 1.
- V_T is the voltage equivalent of the temperature. Its can be computed as:

$$V_T = \frac{KT}{q}$$

where

- K is the Boltzmann constant, $K = 1.38 \times 10^{-23} \text{ J K}^{-1}$
- T is the absolute temperature in Kelvin
- q is the charge of an electron, $q = 1.6 \times 10^{-19} \text{ C}$.

So $V_T = (T/11600)V$, and, at $T = 300 \text{ K}$, $V_T \simeq 26 \text{ mV}$.

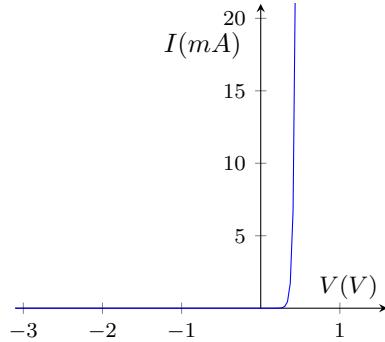


Figure 6.11: Diode V-I characteristic.

If we plot the diode V-I characteristic, setting $I_s = 1 \mu\text{A}$, $\eta = 1$, $V_T = 26 \text{ mV}$, we obtain the drawing of figure 6.11.

We note that if the applied voltage is negative, the current in the diode is very low (in fact it can be approximated by I_s). If the applied voltage is positive, then the current is low up to a certain voltage, in this case about 0.5 V. Above this voltage, the current increases very rapidly, assuming a nearly vertical slope, and above 0.6 V it is over 10 A. The equation we are considering does not take into account the ohmic resistance of the silicon outside the depletion area, so it is valid only for currents not exceeding 10 or 20 mA. When we consider larger currents we should add to our model a resistor in series to the junction that

takes ohmic voltage drop into account, so the slope will not be close to vertical and the current will continue to increase linearly and not exponentially with the applied voltage starting from some fraction of ampere.

The complete equation for large currents becomes:

$$V = I \cdot R + \eta V_T \cdot \ln \left(\frac{I}{I_s} + 1 \right)$$

where R is the ohmic resistance of the diode.

6.2.4 Using diodes

The Shockley equation is inherently non linear and the equation for large currents is even more complex. How can we analyze or design circuits with diodes? Every time we have to work with non linear devices we can use different level of approximations in order to ease our work.

Ideal diode

The simplest model we can derive is called ideal diode. This is a switch model:

- if the voltage applied to the diode is negative, that is if $V < 0$, then the diode is an open circuit and no current flows into the diode
- if the circuits imposes a positive current into the diode, the diode is modeled as a short circuit and $V = 0$

The ideal diode has the characteristic shown in figure 6.12. This very simple

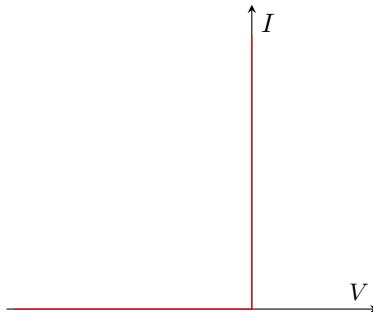


Figure 6.12: Ideal diode V-I characteristic.

approximation can be used when we need to analyze a circuit in which we can neglect the small voltage drop across the diode when the diode is conducting.

Let's clarify the concept with a simple example. Figure 6.13 shows a very simple diode circuit. To analyze the circuit with the ideal diode model, we need to find which are the conditions that allow current flow from anode to cathode in the diode. We consider the diode as a short circuit if the current can flow in that direction, an open circuit otherwise. If we substitute the short circuit to the diode, we can easily see that the current flows in the right direction whenever V is positive. This means that only positive values of V will make current flow in R , while if $V < 0$ there will be no current in R .

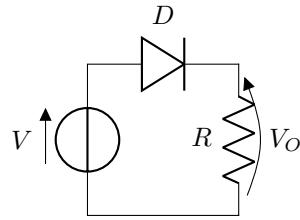


Figure 6.13: Simple diode circuit.

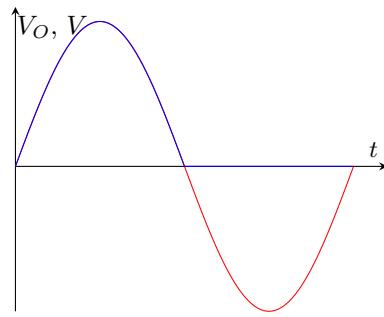


Figure 6.14: V_O (blue), V (red) for the simple diode circuit.

Now let us impose $V = V_p \sin(\omega t)$. What will be the waveform on V_O ? From the above discussion, we have

$$\begin{cases} V_p \sin(\omega t) \geq 0 \implies V_O = V_p \sin(\omega t) \\ V_p \sin(\omega t) < 0 \implies V_O = 0 \end{cases}$$

Figure 6.14 graphically shows input and output waveforms. Is this analysis enough accurate? It depends on the characteristics of the signal. If for example the peak value of the sine input is $V_p = 100$ V and R is big with respect to the internal resistance of the diode, then this analysis is probably enough, but if V_p is much smaller, then the results probably are not accurate enough.

Threshold voltage

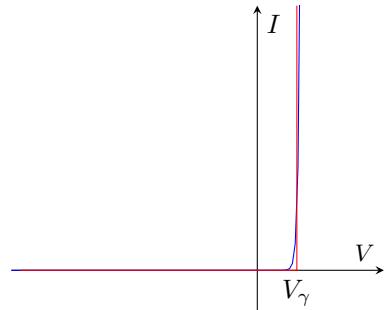


Figure 6.15: Diode simple model with threshold voltage.

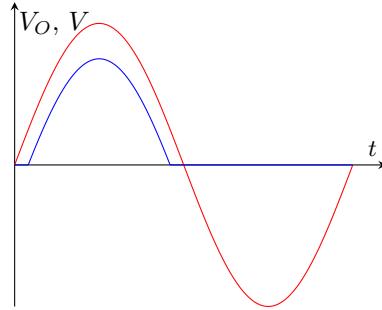


Figure 6.16: V_O (blue), V (red) for the simple diode circuit taking V_γ into account.

By examining figure 6.11, we see that a better approximation of the exponential transfer function can be obtained by shifting the characteristic of the ideal diode by an amount that we will call diode *threshold voltage*, or V_γ . This new characteristic is superimposed to that of the diode in figure 6.15.

For a silicon diode the value of V_γ can range from 0.5 V to 0.7 V depending on the current flowing in the diode. The idea behind this model is that the exponential model presents a *knee*. If the voltage applied to the diode is less than V_γ , the current is very low, almost negligible. In a neighborhood of V_γ the current increases rapidly.

What is the circuit equivalent of this new model? It is simply an ideal diode in series with a constant voltage generator. The new conditions for conduction are:

- if the voltage applied to the diode is less than the threshold voltage, that is if $V < V_\gamma$, then the diode is an open circuit and no current flows into the diode
- if the circuit imposes a positive current into the diode, the diode is modeled as a constant voltage generator and $V = V_\gamma$.

Let us again refer to the simple diode circuit introduced in figure 6.13. Now let us assume that $V_p = 2$ V and $V_\gamma = 0.5$ V.

With our more accurate model, the conditions for diode conduction are:

$$\begin{cases} V_p \sin(\omega t) \geq V_\gamma \implies V_O = V_p \sin(\omega t) - V_\gamma \\ V_p \sin(\omega t) < V_\gamma \implies V_O = 0 \end{cases}$$

The output waveform is displayed in figure 6.16.

Example

Let us further clarify how to work with diodes using the simple models, by analyzing a circuit slightly more complex. Figure 6.17 shows the circuit's schematic.

In the circuit, let us assume that:

- $R_1 = 1\text{ k}\Omega$
- $R_2 = 3\text{ k}\Omega$

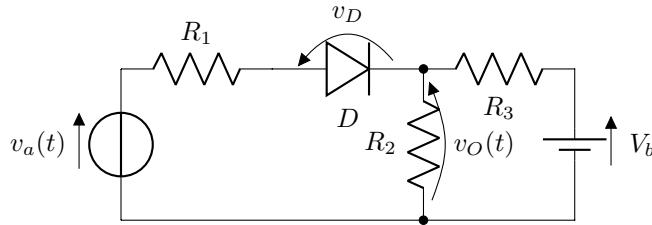


Figure 6.17: Example circuit.

- $R_3 = 3 \text{ k}\Omega$
- $v_a(t) = 2 \text{ V} \cdot \sin(\omega t)$
- $V_b = 1 \text{ V}$

We want to compute and plot $v_O(t)$, assuming ideal diode. We will then repeat the analysis assuming $V_\gamma = 0.5 \text{ V}$.

Ideal diode In the first case, we can first identify the voltage v_a that makes the diode switch from off to on. It is convenient to substitute the diode with an open circuit and to evaluate the voltage difference across it.

$$v_O = V_b \frac{R_2}{R_2 + R_3} = 0.5 \text{ V}$$

$$v_D = v_a - v_O < 0$$

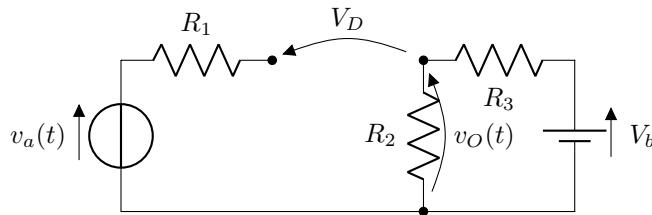


Figure 6.18: Example circuit when the diode is an open circuit.

Figure 6.18 displays the new circuit. It is evident that, to have reverse biasing, we need to have $v_a < V_O = 0.5 \text{ V}$. In this case we can substitute an open circuit to the diode, so we will have $v_O = 0.5 \text{ V}$.

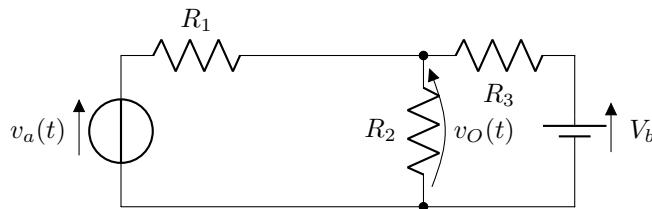


Figure 6.19: Example circuit when diode is conducting.

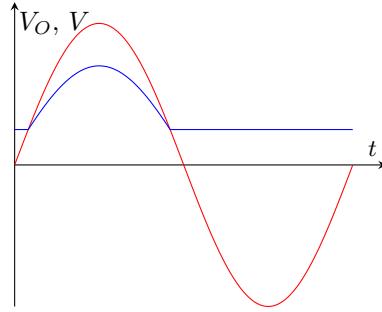


Figure 6.20: v_O (blue), v_a (red) for the example circuit solved with ideal diode model.

When $v_a > 0.5$ V, the diode becomes a short circuit. The equivalent circuit is displayed in figure 6.19.

We can use the superposition principle to find v_O .

$$V_O = 2 \text{ V} \frac{R_2 \parallel R_3}{R_1 + R_2 \parallel R_3} \sin(\omega t) + V_b \frac{R_1 \parallel R_2}{R_1 \parallel R_2 + R_3}$$

We can then write the equations for all working conditions:

$$\begin{cases} 2 \text{ V} \sin(\omega t) \geq 0.5 \text{ V} \Rightarrow V_O = 1.2 \text{ V} \sin(\omega t) + 0.2 \text{ V} \\ 2 \text{ V} \sin(\omega t) < 0.5 \text{ V} \Rightarrow V_O = 0.5 \text{ V} \end{cases}$$

Figure 6.20 displays the results graphically.

Threshold voltage. If we use the more accurate model that takes a threshold voltage into account, we can proceed analogously by first finding the voltage v_a that makes the diode switch from off to on. In this case, the voltage drop across the diode must be $v_D > V_\gamma$. We can visualize this condition by substituting a constant voltage generator V_γ to the diode, like in figure 6.21.

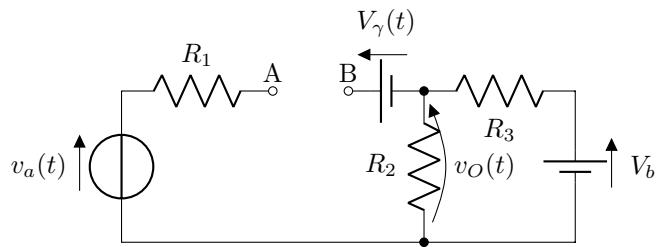


Figure 6.21: Example circuit when diode is not conducting.

As far as voltage at node A is lower than that at node B, the diode is not conducting. Again, we have:

$$v_O = V_b \frac{R_2}{R_2 + R_3} = 0.5 \text{ V}$$

The voltage on A is lower than the voltage on B if:

$$v_a = 2 \text{ V} \sin(\omega t) < 0.5 \text{ V} + V_\gamma = 1 \text{ V}$$

When the voltage generated by v_a is higher than 1 V, we substitute the diode with a short circuit in series with a V_γ voltage generator. Figure 6.22 displays the equivalent circuit. As in the previous case, we can use superposition principle

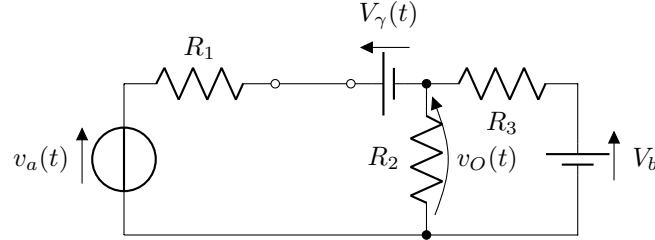


Figure 6.22: Example circuit when diode is conducting.

to find v_O when the diode is conducting:

$$V_O = [2 \text{ V} \sin(\omega t) - V_\gamma] \frac{R_2 \parallel R_3}{R_1 + R_2 \parallel R_3} + V_b \frac{R_1 \parallel R_2}{R_1 \parallel R_2 + R_3}$$

We can synthesize the expression of v_O as:

$$\begin{cases} 2 \text{ V} \sin(\omega t) \geq 1 \text{ V} \implies V_O = 1.2 \text{ V} \sin(\omega t) - 0.1 \text{ V} \\ 2 \text{ V} \sin(\omega t) < 1 \text{ V} \implies V_O = 0.5 \text{ V} \end{cases}$$

If we plot these expressions we obtain the graph of figure 6.23.

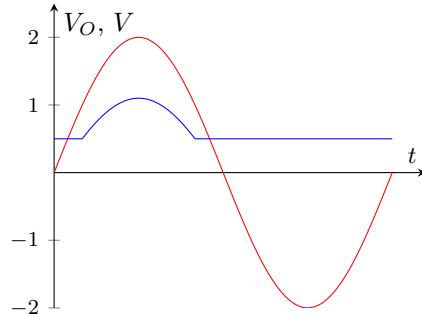


Figure 6.23: v_O (blue), v_a (red) for the example circuit solved with $V_\gamma = 0.5 \text{ V}$.

Dynamic behavior

We will investigate more in details in the following chapters the behavior of a diode during transients. For now we will only note that the junction will not respond instantaneously to voltage or current changes since there are accumulated charges, both in the depletion region and in the two neutral regions away from the depletion region.

The depletion region is equivalent to a parallel plate capacitor, charged with the fixed charge of the dopant ions. When the diode is conducting, more charge is accumulated due to the excess free carriers due to diffusion current. In reverse

bias, the width of the depletion region changes with the applied voltage, modifying the value of the capacitor equivalent to the diode, but there is no charge associated to diffusion current.

In practice, a parasitic capacitance is always associated with a diode, and its value depends on the applied voltage. This is used in some applications (varicap diodes).

Breakdown

If we apply a negative voltage to a diode, only a very small current will flow through the diode. This is true only if the magnitude of the negative applied voltage is not too high. For every diode there is a certain voltage, called breakdown voltage, beyond which we observe a steep increase in the current, so the junction exhibits a new quasi vertical V-I characteristic. The breakdown is due to two different mechanisms, avalanche multiplication and Zener effect. We will not investigate the physical phenomena involved in the two mechanisms.

6.2.5 Zener diodes

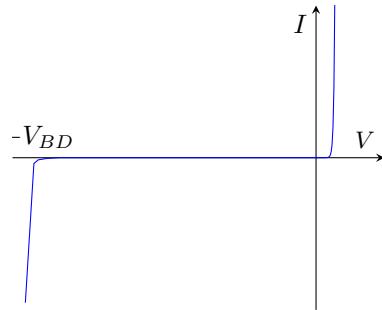


Figure 6.24: Diode V-I characteristic including breakdown

The complete diode V-I characteristic including breakdown is depicted in figure 6.24. The breakdown condition can be catastrophic in a diode because the increase in current also creates an increase in the power that the diode has to dissipate. Every diode can sustain a maximum power P_D . If there are means of controlling the maximum current, then breakdown effect can be used in a circuit. Manufacturers build diodes in which the breakdown voltage is guaranteed to have a nominal value and a certain tolerance. Such diodes are called Zener diodes (even if the breakdown in some diodes comes by avalanche multiplication and not Zener effect). They are used as voltage reference in power supplies or to limit the voltage in some node (for example, protection of circuit inputs connected via cables from voltage spikes due to electromagnetic noise or other events like thunderbolts captured by the electrical connection).

A Zener diode has a special symbol, shown in figure 6.25. Voltage and current in such device normally are measured in the opposite direction with respect to normal diodes, because these devices are mainly used in reverse bias region to exploit breakdown effect.

As Zener diodes are frequently used as voltage reference, ideally beyond breakdown the characteristic should be vertical. This does not happen in prac-

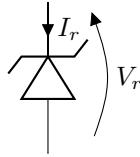


Figure 6.25: Zener diode.

tice, so, to be able to compute the voltage drop of the diode in breakdown, one should take into account the slope of the characteristic in that region. A non vertical characteristic means that we cannot substitute a Zener diode in breakdown with an ideal voltage source, but instead we need to add a series resistance, like in figure 6.26.

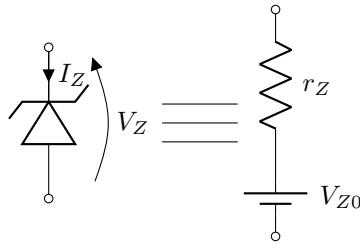


Figure 6.26: Zener diode equivalent circuit in breakdown region.

To find the value of r_Z , called differential Zener resistance, and V_{Z0} , we can refer to figure 6.24. The breakdown region is characterized by a knee, for low currents, followed by a nearly linear region. In general we don't want to use the Zener in the knee region, so we define the minimum Zener current $I_{Z\min}$ as the current above which the characteristic is linear. We can derive also a maximum Zener current $I_{Z\max}$ as the current that causes the maximum allowable power dissipation in the device.

$$I_{Z\max} = \frac{P_{d\max}}{V_Z}$$

If we call $V_{Z\max}$ the Zener voltage corresponding to $I_{Z\max}$ and $V_{Z\min}$ the voltage at $I_{Z\min}$, we can define r_Z as

$$r_Z = \frac{V_{Z\max} - V_{Z\min}}{I_{Z\max} - I_{Z\min}}$$

We can define V_{Z0} as the intercept of the straight line passing through the two points $(V_{Z\max}, I_{Z\max})$ and $(V_{Z\min}, I_{Z\min})$:

$$V_{Z0} = \frac{V_{Z\max} + V_{Z\min}}{2} - r_Z \frac{I_{Z\max} + I_{Z\min}}{2}$$

Example circuit

To understand how to work with Zener diodes, let us introduce a very simple circuit using such device. Figure 6.27 shows the circuit's schematic diagram.

Assume that the Zener diode has the following parameters:

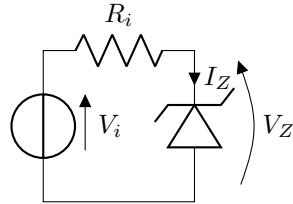


Figure 6.27: Example circuit.

- $V_{Z0} = 6.7 \text{ V}$;
- $r_Z = 4 \Omega$;
- $I_{Z\min} = 1 \text{ mA}$;
- $I_{Z\max} = 50 \text{ mA}$;

If $R_i = 560 \Omega$, compute the range of acceptable input voltage V_i and the corresponding output voltage V_Z .

Solution We want to find the range of V_i that maintains the diode in the breakdown region with current ranging from $I_{Z\min}$ to $I_{Z\max}$. We can therefore substitute to the diode the model for breakdown region. The circuit becomes the one in figure 6.28.

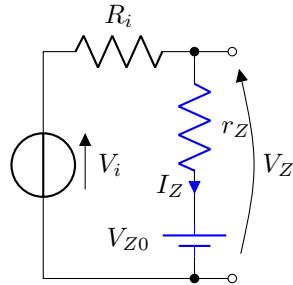


Figure 6.28: Example circuit with breakdown region model of the diode

We can easily compute the Zener current as there is no other current in the circuit.

$$I_Z = \frac{V_i - V_{Z0}}{R_i + r_Z}$$

Solving for V_i we can find:

$$V_{i\min} = V_{Z0} + (R_i + r_Z)I_{Z\min} = 6.7 \text{ V} + 0.564 \text{ k}\Omega \cdot 1 \text{ mA} = 7.264 \text{ V}$$

$$V_{i\max} = V_{Z0} + (R_i + r_Z)I_{Z\max} = 6.7 \text{ V} + 0.564 \text{ k}\Omega \cdot 50 \text{ mA} = 34.9 \text{ V}$$

The corresponding minimum and maximum output voltages are therefore:

$$V_{Z\min} = V_{Z0} + r_Z \cdot I_{Z\min} = 6.7 \text{ V} + 0.004 \text{ k}\Omega \cdot 1 \text{ mA} = 6.704 \text{ V}$$

$$V_{Z\max} = V_{Z0} + r_Z \cdot I_{Z\max} = 6.7 \text{ V} + 0.004 \text{ k}\Omega \cdot 50 \text{ mA} = 6.9 \text{ V}$$

The effect of stabilization of the voltage due to the Zener is clear : a variation of more than 27 V in the input voltage results in less than 0.2 V of variation of the output voltage.

6.2.6 Diode applications

Diodes have many fields of application: mainly in power electronics, in signal processing or as Electromagnetic Compatibility (*EMC*) protection devices. We will introduce a few examples that will be useful for the rest of the course.

Half wave rectifier

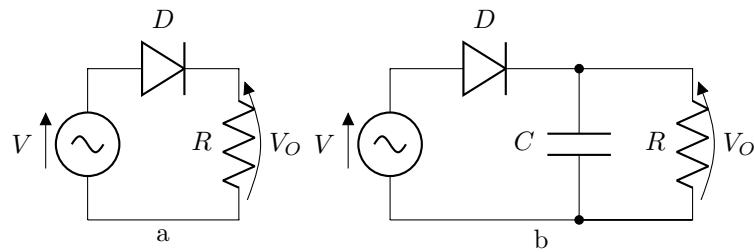


Figure 6.29: Half wave rectifier: a) basic circuit; b) with capacitive filter.

We already introduced the basic version of a half wave rectifier, redrawn in figure 6.29a. This circuit is used in power electronics to convert a sine waveform in DC, in AC-DC converters. If we insert the diode alone, the positive half period of the incoming sine wave is passed to the output while the negative is stopped, this is the reason of the name of the circuit. The output waveform is the green line of figure 6.30.

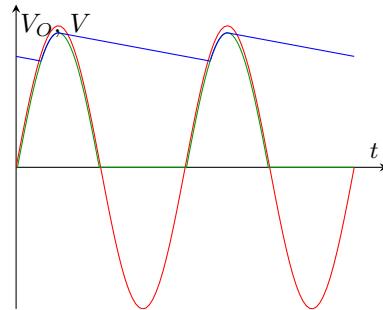


Figure 6.30: Waveforms of the circuits of figure 6.29: V (red), V_O of circuit a (green), V_O of circuit b (blue), .

For a better conversion to DC, we have to add a large capacitor, whose function is to store charge from input waveform when it is close to the positive peak value and release it to the load when the input waveform is low voltage or negative. The smoothing function of C can be appreciated by looking at figure 6.30, blue line.

Full wave rectifier

The half wave rectifier can absorb energy from the incoming sine wave only during the positive half cycle. In power electronics this is not optimal, it would be much better to have a AC-DC conversion using both half period of the incoming waveform.

This is easily done by inserting four diodes instead of one, in what is called Graetz bridge configuration, shown in picture 6.31.

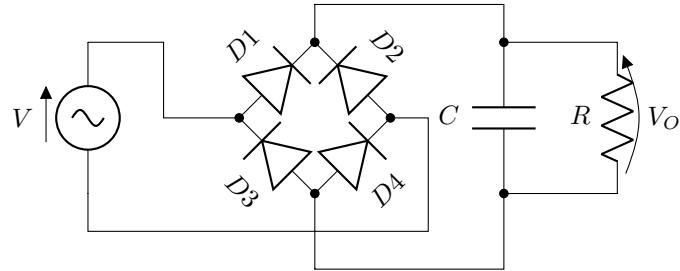


Figure 6.31: Full wave rectifier.

In this circuit when V is positive, current flows through diodes D_1 and D_4 while, when V is negative, diodes D_3 and D_2 conduct. Using a filter capacitor C , the output voltage is smoother than in the half wave rectifier with the same capacitor. Waveforms with and without capacitor are drawn in figure 6.32.

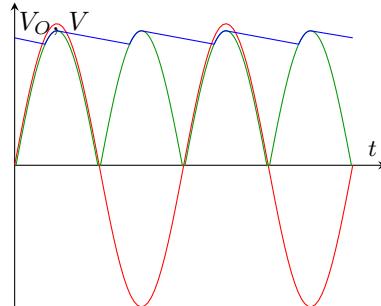


Figure 6.32: Waveforms of the full wave rectifier: V (red), V_O without filter capacitor C (green), V_O with C (blue).

The full wave rectifier is a very popular circuit, so it is possible to find it as a four pin device called bridge rectifier, with four diodes integrated and internally connected.

AM demodulator

Diodes can be used to process signals. One historical use of diodes is in the demodulation of amplitude modulated signal (AM). Amplitude modulation is a very simple way of transmitting a signal at radio frequency. The low frequency signal (typically audio) changes the amplitude of an radio frequency sine wave (RF) which is then amplified and transmitted. At the receiver input we have a

signal similar to the one drawn in figure 6.33a, where the low frequency signal is a sine wave for simplicity.

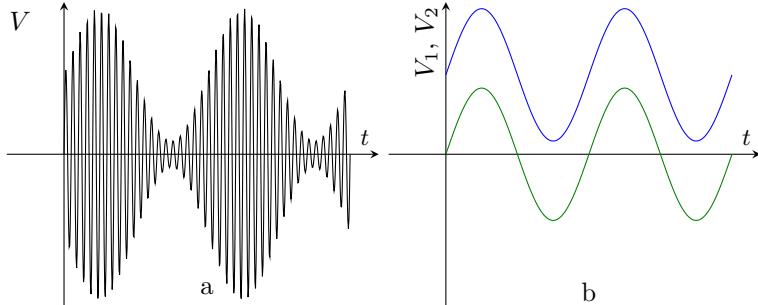


Figure 6.33: AM modulated RF carrier (a); demodulated output signal V_2 (b, green); intermediate signal V_1 (b, blue).

Using a diode and a low pass filter, we can reconstruct the low frequency modulating signal, which will be affected by a DC bias. We can remove the DC bias by adding an high pass filter at the output of the previous one.

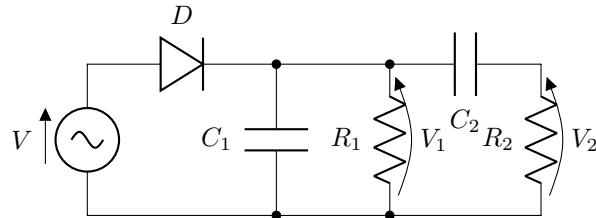


Figure 6.34: AM demodulator.

Figure 6.34 shows a possible implementation of a AM demodulator. Signal V_2 is the output of the final high pass filter, V_1 is the intermediate signal with DC offset. Both signals are shown in graph b of figure 6.33, when the signal V applied at the input of the demodulator is the one of graph a of the same figure. The circuit is also called *envelope detector*.

Signal clamping

One of the applications of diodes is in the protection of circuit inputs from signals out of normal range. This is a simple form of signal conditioning and is called *signal clamping*.

Figure 6.35 show three example of clamping circuits. The first circuits is used when the input signal V should be always positive. In this case, if the signal V becomes more negative than $-V_\gamma$, the diode starts conducting and avoids transferring a negative voltage to subsequent circuits.

The second circuit uses a Zener diode and provides control not only on the polarity of the input signal, but also on the magnitude. If the signal is positive but higher than V_Z , the Zener diode conducts, effectively avoiding propagating voltages higher than V_Z to subsequent circuits. Protection against negative voltages is provided by the same diode: in case of negative input voltages, the Zener

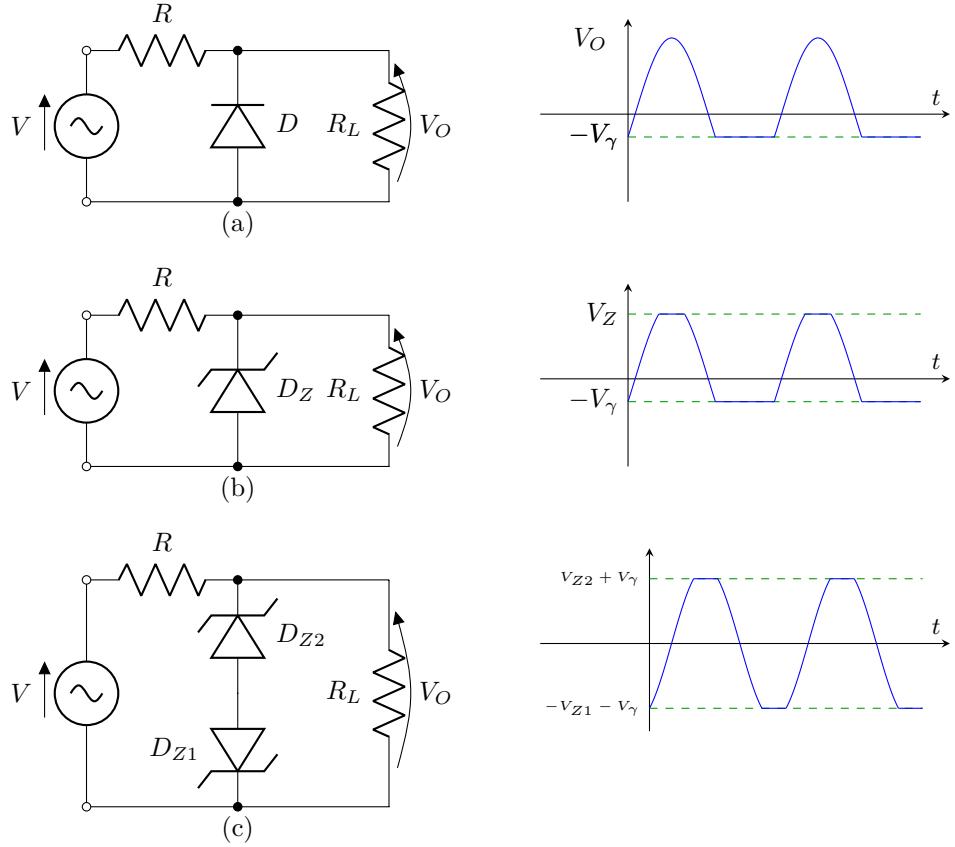


Figure 6.35: Example of clamping circuits.

diode is forward biased and starts conducting like any normal diodes when the input signal is more negative than $-V_\gamma$.

The third circuit is used to protect inputs that can accept bipolar signals up to a certain maximum amplitude. In case of positive signal, when the input amplitude is higher than the sum of the Zener voltage of D_{Z2} and V_γ of D_{Z1} , both diodes start conducting (D_{Z2} in breakdown region and D_{Z1} in forward region) and clamp the input voltage to $V_{Z2} + V_\gamma$. The same happens for wide negative input signal, where the limiting voltage is $-V_{Z1} - V_\gamma$.

Zener diodes specially built for clamping unidirectional signals and couples of Zener diodes for bidirectional signals, are known as Transient Voltage Suppressors (TVS) and sold under commercial name of Transzorb, Transil or others.

Diode as sensor

The diode can be used as temperature or light sensor. In the first case, we use the strong dependence of I_s from temperature. In the second, if we expose a pn junction to light, photons hitting the junction transfer energy to the carriers, creating free hole-electron pairs (photoelectric effect), creating a photocurrent proportional to the light intensity that can be read by a circuit.

Let us see in detail the first case. We know from section 6.2.3 that the fundamental diode equation is:

$$I = I_s \left(e^{V/(\eta V_T)} - 1 \right)$$

We also know that I_s increases exponentially with the temperature, it doubles every 10°C . If we drive a diode with a constant current I_0 , like in figure 6.36, we can easily compute the resulting V_D .

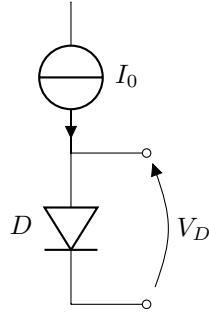


Figure 6.36: Temperature sensor.

Solving the diode equation for V_D , we obtain:

$$V_D = \eta V_T \ln \left(\frac{I_0}{I_s} \right)$$

The exponential variation of I_s with the temperature gives a linear variation of V_D thanks to the logarithmic function. V_D decreases by approximately $-2.5 \text{ mV } ^\circ\text{C}^{-1}$.

A single diode used as temperature sensor will not be very precise, but a more complex circuit based on the same principle is widely used in integrated temperature sensors like the LM35, which achieve $\pm 0.5^\circ\text{C}$ over the -55°C - 125°C temperature range.

6.3 Bipolar transistor

A transistor (from TRANSfer resISTOR) is a three terminals device, one of which is common to input and output ports. The output port's current depends not only from the output voltage but also from one of the input port's electrical variables (voltage, current). Figure 6.37 shows a black-box representation of this kind of devices. A common abbreviation of bipolar transistor is *BJT* which stands for Bipolar Junction Transistor.

A bipolar transistor is in fact made of two pn junctions, with one doped region in common (the central one) and very thin. We can then have two types of transistors, depending on the kind of doping of the central region: NPN transistor and PNP transistor. The central region is called *base* and identified by *B*, while the other two regions are called *emitter*, *E*, and *collector*, *C*.

Collector and emitter are not equivalent, because emitter is heavily doped while collector is lightly doped. The base is more doped than collector, dopant

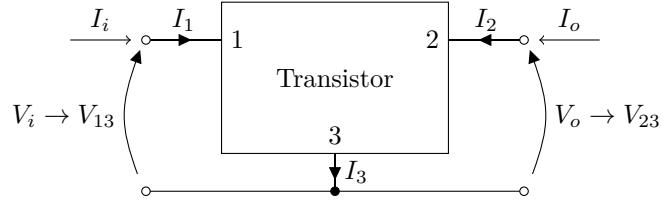


Figure 6.37: Black box representation of a bipolar transistor.

concentration is about ten times more, but far less than the emitter. In normal use, the base-emitter junction is forward biased and the collector-base junction is reverse biased, that is $V_{BE} > 0$ and $V_{CE} > V_{BE}$. We will define the region identified by these conditions as the active region. We normally consider as input the current entering the base, and the emitter as the terminal common to the input and output ports, although different configurations are possible. Figure 6.38 shows a NPN transistor with applied input and output voltages.

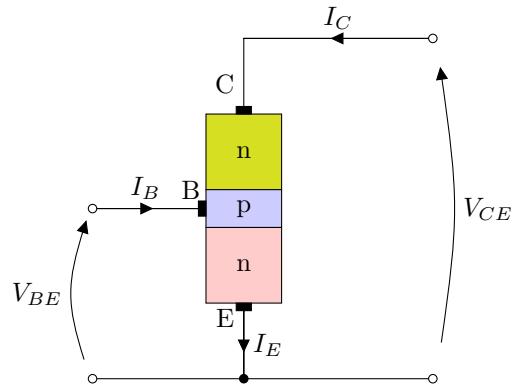


Figure 6.38: Junctions of a NPN bipolar transistor.

As said before, the base region is very thin. When we impose a small current from base to emitter, by forward biasing the base-emitter junction ($V_{BE} > 0$), a much larger current will flow from collector to emitter. As with diodes, we will not investigate the physics that allows this phenomenon, we only note that the key role is played by the difference in dopant concentration, that creates large gradients and therefore large diffusion currents, and the thin base region that allows electrons injected by the emitter to reach the collector without recombining in the base.

The following equations define the operation of a BJT in its active region:

$$\begin{cases} I_C = \alpha I_E \\ I_B = (1 - \alpha) I_E \end{cases}$$

Where the coefficient α is very close to 1 and depends on transistor technology (essentially doping levels and base length). As we noted above, we have two kinds of BJTs, NPN and PNP. Their schematic symbols are shown in figure 6.39.

Looking at the symbols we can note that:

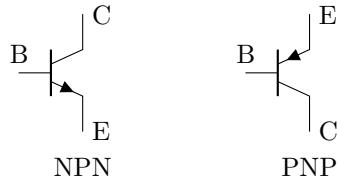


Figure 6.39: Schematic symbols of NPN and PNP bipolar transistors.

- The emitter terminal is identified by an arrow. The direction of the arrow is the direction of the current when the emitter-base junction is forward biased.
- For the PNP transistor to have emitter-base junction forward biased, the voltage at the emitter has to be higher than the voltage at the base, so normally the symbol is drawn with emitter on top and collector on bottom, reversed with respect to NPN devices.
- the vertical bar symbolizes the semiconductor bar with the base connected in the middle.

NPN transistor

Let us first examine input and output characteristics of a NPN transistor.

The input characteristic of a BJT in the active region is similar to that of a forward biased silicon diode, because of the base-emitter junction being forward biased. As the transistor comprises a diode between base and emitter, we can show that emitter current and base-emitter voltage obey an equation very similar to that of the silicon diode:

$$I_E = I_s \left(e^{V_{BE}/V_T} - 1 \right)$$

Comparing with the diode equation, we notice that the only difference is the η parameter, which is not present because for a BJT transistor it is always equal to 1.

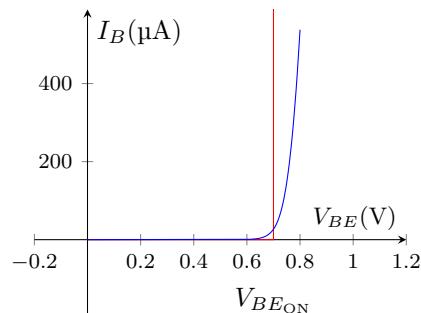


Figure 6.40: BJT input characteristic.

Figure 6.40 plots the typical I_B versus V_{BE} curve (here for transistor model 2N2222). Sometimes we can approximate the curve with an on-off characteristic like we did for the diode. Usually $V_{BE_{ON}} = 0.6\text{ V}$ to 0.7 V at room temperature.

If we use, as it is normal, the base current as the input variable and the collector current as the output variable, we can find a very simple relation between the two, which holds in active region:

$$\begin{cases} I_C = \alpha I_E \\ I_B = (1 - \alpha)I_E \\ I_C = \frac{\alpha}{1-\alpha} I_B = \beta I_B \end{cases}$$

The β parameter indicates the current gain of the device, when used as a current amplifier. Given that α is less than one but very close to this value (typically 0.98 or more), β can be a large number. For small transistors, the typical value of β ranges from 50 to 300.

One of the problems in using BJTs in circuits is that the manufacturer cannot control accurately the β parameter, that can change by a factor of three between transistors of the same type in the same working conditions.

From the above discussion, the collector of an ideal transistor should perform like an ideal current source, with the current determined only by the base current and not by the V_{CE} voltage. This is valid only as a first approximation and only when the base-collector junction is reverse biased, that is in the active zone of the transistor.

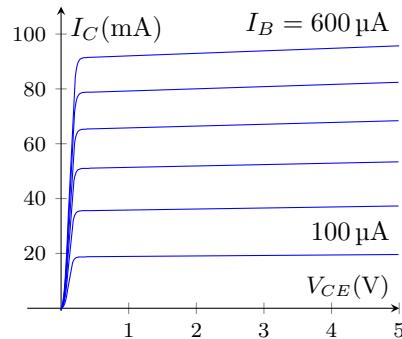


Figure 6.41: BJT output characteristic.

In figure 6.41 we show the output characteristic of a small power transistor, in our case the 2N2222. It can be seen that the collector current is not constant when V_{CE} is close to 0, because in this region the base-collector junction is not reverse biased. This is called *saturation* region. We will study the behavior of BJTs in saturation region in the following.

When V_{CE} rises above 0.5 V the transistor enters the active region and the collector current flattens. However, specially for high currents, we can notice that the current is not perfectly independent of V_{CE} and the collector of the transistor does not behave exactly as a current source.

If we increase the collector-emitter voltage above a certain limit, the reverse biased collector-base junction experiences a breakdown condition, similar to that of a diode. In case of breakdown, the collector current suddenly increases to very high values and it is possible to damage the device.

PNP transistor

A PNP transistor is the dual of the NPN transistor. As already noted, active region is obtained when the emitter is at a more positive voltage than the base and collector is more negative than the base.

We will not show graphs of input and output characteristics because they are easily obtained by reversing the polarity of V_{BE} , V_{CE} , I_B and I_C with respect to the NPN transistor. We only note that the performances of PNP transistors are reduced compared to NPN devices because the majority carriers in emitter and collector are holes instead of electrons. Holes have reduced mobility and that means that PNP transistors are slower and exhibit higher ohmic resistance.

6.3.1 The BJT as amplifier

The main uses of a bipolar transistor are as an amplifier or as a switch. In this chapter we will focus on the amplifier applications. As seen in the previous section, the BJT works as a nearly linear current to current amplifier, but this operation is limited to the active zone. Also, even in the active zone, the characteristic is not linear. How can we set a transistor to work in the active zone and how can we obtain a satisfactory linear behavior from this device?

As seen previously, an amplifier is a two ports circuit. We connect a signal (voltage or current) to the input of the circuit, and we obtain a signal at the output of the circuit that is proportional to the input but with increased voltage and/or current levels.

The most simple BJT based amplifiers cannot work in DC, they can amplify signals starting at a certain frequency. They are referred as AC amplifiers. They cannot work in DC because the transistor is surrounded by a set of components that are used to bring the transistor to work in the active region. To understand why we need to work in this way, let us consider a very simple circuit with a DC input voltage source.

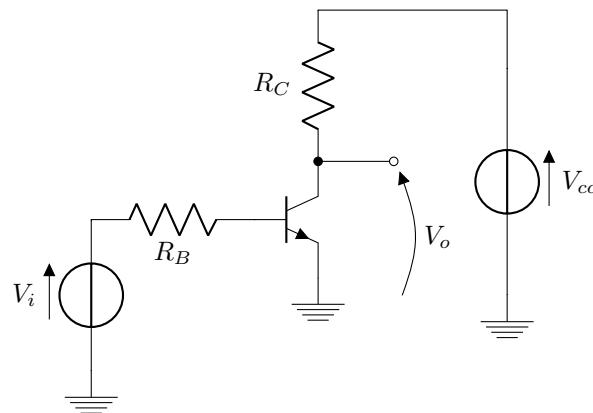


Figure 6.42: Simple amplifier.

Figure 6.42 shows the schematic. We apply a constant V_{cc} voltage to the upper terminal of R_C and a signal V_i at the left terminal of R_B .

We can refer to the input characteristic of the transistor, figure 6.40, to understand what is the base current I_B induced by our input voltage V_i . We

can easily understand that no current will flow in the base of the transistor if $V_i < 0.6$ V. If no current is flowing in the base, there will be no current in the collector either, and the voltage V_o will be V_{cc} . Starting from that voltage, an increase in V_i will produce a current in R_B , but the increase in current will not at all be linear. Nevertheless, a base current will produce a collector current, so the voltage on V_o will start to decrease. As we keep increasing V_i , the base current will increase and the voltage on the base will reach a region in which a further increase of the current will produce a small increase of the voltage and it will be almost linear. The voltage V_o will continue to decrease and at a certain point it will be close to 0.

When this happens, no further decrease in the output voltage is possible. We are out of the active region and into the saturation region.

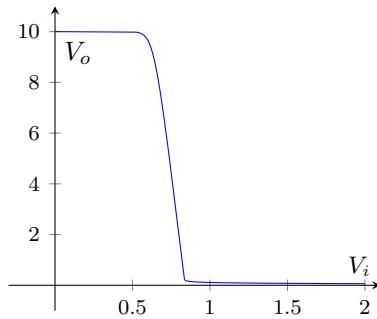


Figure 6.43: Transfer function.

Figure 6.43 shows a spice simulation of the circuit with the parameters: $R_B = 10\text{ k}\Omega$, $R_C = 3.3\text{ k}\Omega$, $V_{cc} = 10$ V, transistor: 2N2222. The behavior is nearly linear for the input voltage from 0.65 V to 0.83 V. For lower voltages the transistor is off, for higher voltages it is in saturation. Looking more in details at simulation data, we see that for $V_i = 0.67$ V we have $V_o = 8$ V and for $V_i = 0.82$ V, $V_o = 1$ V. So we can compute the local gain in this region as $A_v = \Delta V_o / \Delta V_i \simeq 47$.

The result of this analysis is that we can use the circuit as an amplifier but only if the input signal varies in the above defined range.

There are other problems with this circuit, because the simulation refers to a transistor at room temperature and with typical parameters. We already noted that in the active region the ratio between collector and base current, β , has big variations from device to device. Furthermore, β is not constant within the same device but depends on: I_C , V_{CE} and the junction temperature. Also the V_{BEON} varies with the temperature, so the transfer function of figure 6.43 will shift left or right and the slope will change from device to device and due to temperature variations.

Let us now think about practical applications of this circuit as an amplifier. It is very difficult to find an application where the input signal range is 0.65 V to 0.83 V. It is much easier to have to amplify for example the output of a microphone, which is a signal with bandwidth ranging from 20 Hz to 20 kHz (for a very good microphone) and an amplitude of a few mV. We can think of this signal as being composed of several sine waves at different frequencies. Let us therefore focus on an input signal in the form $V_i = V_p \sin(\omega t)$ with $f = \omega/(2\pi)$

in the above mentioned range.

Quiescent operating point

Can we modify the circuit so that this signal gets amplified almost linearly? For the time being we do not bother with the other problems due to parameters variations.

The easiest way to obtain an amplifier suitable for our input signal is to add a constant source that will bring the output to about half the output range when our input signal is at 0 V. The input signal will then add a small positive or negative variation to the base voltage that in turn will induce a much larger variation to the output voltage. Given that the input frequency range does not include DC, we can use a capacitor as a means of superimposing the required DC voltage to our AC signal. We name the DC circuit necessary to bring the output voltage to mid range as *bias circuit*, and the input and output levels of the electrical variables with zero applied AC signal as *quiescent operating point* (q.o.p.) of the circuit.

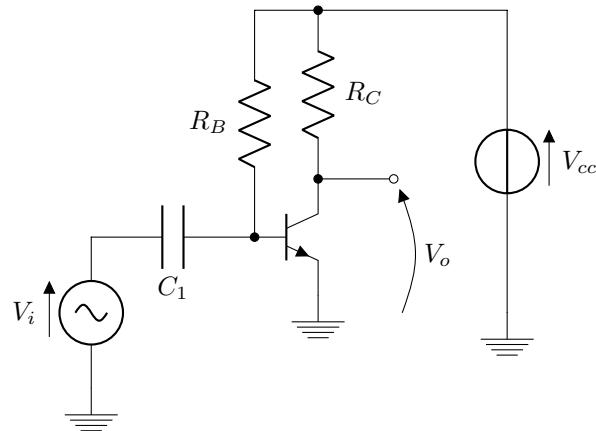


Figure 6.44: Simple amplifier.

The circuit is shown in figure 6.44. We use R_B to bias the transistor so that the output voltage is at half the available dynamic range when there is no voltage at the input. In our example, we want $V_o = 5$ V. The circuit has the same $V_{cc} = 10$ V and the same $R_C = 3.3$ k Ω . We can note that this specification means that

$$I_C = \frac{V_{cc} - V_o}{R_C} = 1.52 \text{ mA}$$

We also have that $I_B = (V_{cc} - V_{BE})/R_B$ but there is a dependency between V_{BE} and I_B , as shown in figure 6.40. Can we simplify the problem? The answer is yes, because we note that the variation of V_{BE} is very small compared with V_{cc} . We can therefore think that V_{BE} assumes a fixed value, that we call $V_{BE_{ON}}$. Looking at the graph, we can set $V_{BE_{ON}} = 0.7$ V.

We know that in the active region the collector current is proportional to the base current: $I_C = \beta I_B$. How do we estimate β ? The manufacturers list possible ranges of β as function of I_C and V_{CE} in the datasheets. The parameter name

is usually *DC current gain* and its acronym is normally h_{FE} , but other names are possible. We will usually refer to datasheet, but for our first example we look at the spice simulation and see that the spice model of 2N2222 in our case predicts $\beta = 200$.

Now we have all what we need to find the value of R_B :

$$\begin{cases} I_B = \frac{I_C}{\beta} = \frac{1.52 \text{ mA}}{200} = 7.6 \mu\text{A} \\ R_B = \frac{V_{cc} - V_{BEON}}{I_B} = \frac{9.3 \text{ V}}{7.6 \mu\text{A}} = 1.28 \text{ M}\Omega \end{cases}$$

We obtained this result by effectively substituting to our transistor a very simple model, which is very useful to find the quiescent operating point of most circuits.

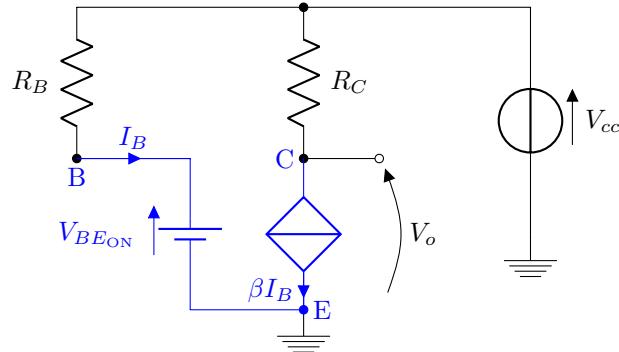


Figure 6.45: Equivalent circuit for q.o.p. (in blue the BJT model)

The circuit is shown in figure 6.45. Note that we did not show the capacitor, because we are only interested in DC analysis and in DC the capacitor is an open circuit. The model of the BJT is drawn in blue. It consists of a constant voltage source, from base to emitter, and in a current controlled current source, from collector to emitter. We will use this circuit every time we will need to perform a q.o.p. computation.

Small signal analysis

We can now apply our signal to the input of the circuit, by selecting a sufficiently big capacitor that can be considered short circuit at signal frequency. If we simulate the circuit of figure 6.44, using a 1mV peak, 1kHz sine input and the computed R_B , we obtain the graph shown in figure 6.46. As expected, the output DC component is 5V.

If we retrieve the voltage values from the simulation graph, we find that the minimum voltage is 4.82V, while the maximum is 5.18V, while for 0V at the input we have, correctly, 5V at the output. Let us compute the voltage gain, dividing the peak to peak amplitude of the output by the peak to peak amplitude of the input. We obtain $A_v \simeq 180$. This is much larger than what we computed from the circuit of figure 6.42. We can explain this because the input schematic of our circuit is different from that one, even if we use the same R_C and the same I_C . Is it possible to compute the gain without using a simulator?

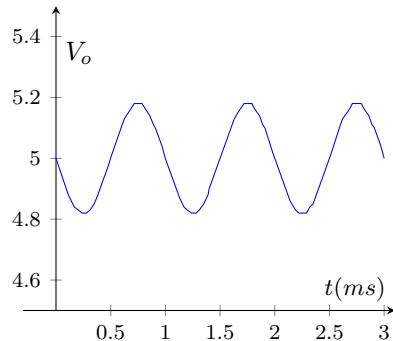


Figure 6.46: Time domain simulation.

The problem is that our transistor has a very complex non linear transfer function. LTspice simulator uses a model called Gummel-Poon, but even the simpler and less accurate Ebers-Moll model, being a combination of exponential functions, is very difficult to apply. It is however possible to obtain a simple linear model of the transistor if we limit our analysis to small variations around the q.o.p. We will perform what is called a *small signal analysis*. The idea is to have a linear differential model of our non linear device. In practice, we will compute the *derivative* of our circuit to obtain the *variations* of the output signal with respect to variations of the input signal.

The process is simple. Passive linear components, such as resistors and capacitors, are transferred without modification to our differential model. Constant voltage sources become short circuits, because their derivative is zero. Constant current sources become open circuits for the same reason. But how do we deal with bipolar transistors? When we look at the input port of the transistor, from base to emitter we see a diode, so there is an exponential link between voltage and current. Given that a variation of the input current is linked to a variation of the base-emitter voltage and we want a linear device emulating this behavior, we need to insert a resistor from base to emitter of our differential model. We call this resistor r_π . We name V_{BE0} and V_{CE0} the computed base-emitter and collector-emitter voltages at q.o.p. We use small letters for r and the other parameters, indicating this equivalence is only for variations around q.o.p. Therefore:

$$r_\pi = \left. \frac{\partial v_{be}}{\partial i_b} \right|_{V_{BE0}, V_{CE0}}$$

When we consider the output port, a variation in the base current causes a proportional variation in the collector current. That means that we have to model the output port with a current generator dependent on the variation of the input current. If we look at figure 6.41, we see that the output characteristic is not perfectly horizontal. To model the slope of the characteristic we have to add a resistor in parallel to the current source. This resistor is most of the times omitted, because its influence is normally negligible. In summary we have:

$$\left\{ \begin{array}{l} i_c = h_{fe} i_b \\ h_{fe} = \frac{\partial i_c}{\partial i_b} \Big|_{V_{BE0}, V_{CE0}} \end{array} \right.$$

$$r_o = \frac{\partial v_{ce}}{\partial i_c} \Big|_{V_{BE0}, V_{CE0}}$$

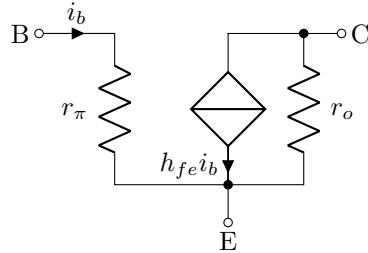


Figure 6.47: Equivalent circuit of BJT for small signal analysis.

Graphically, the model of the transistor is shown in figure 6.47

Given that the variation of the base current causes a proportional variation of the base-emitter voltage, sometimes it is preferable to calculate i_c from v_{be} instead of i_b . This brings us to an equivalent model in which the current source does explicitly depend on v_{be} . The coefficient is called small signal transconductance, indicated with g_m . The circuit is shown in figure 6.48.

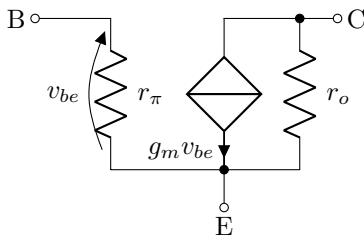


Figure 6.48: Equivalent circuit with small signal transconductance.

The value of g_m is obviously:

$$g_m = \frac{\partial i_c}{\partial v_{be}} = \frac{h_{fe}}{r_\pi}$$

Now the assumptions made to obtain the small signal model should be clear. We still need to compute the values of the parameters from the q.o.p. We can do this easily if we assume that β is sufficiently high, $\beta \gg 1$. In this case, $I_C \approx I_E$ and we can use the equation:

$$I_E = I_s \left(e^{V_{BE}/V_T} - 1 \right)$$

We first observe that the small signal current gain h_{fe} can be approximated by the DC gain with the same level of collector current and collector-emitter

voltage. That is:

$$h_{fe} = \left. \frac{\partial i_c}{\partial i_b} \right|_{V_{BE0}, V_{CE0}} \approx \beta$$

Naming I_{C0} the value of I_C in the q.o.p., we can evaluate g_m :

$$g_m = \left. \frac{\partial i_c}{\partial v_{be}} \right|_{V_{BE0}, V_{CE0}} \approx \left. \frac{\partial i_e}{\partial v_{be}} \right|_{V_{BE0}, V_{CE0}} = \left. \frac{\partial}{\partial v_{be}} \left[I_s e^{V_{be}/V_T} \right] \right|_{V_{BE0}, V_{CE0}} = \frac{I_{C0}}{V_T}$$

$$r_\pi = \frac{h_{fe}}{g_m} = \frac{\beta V_T}{I_{C0}}$$

We can apply the small signal model to the circuit of figure 6.44. We obtain the circuit of figure 6.49.

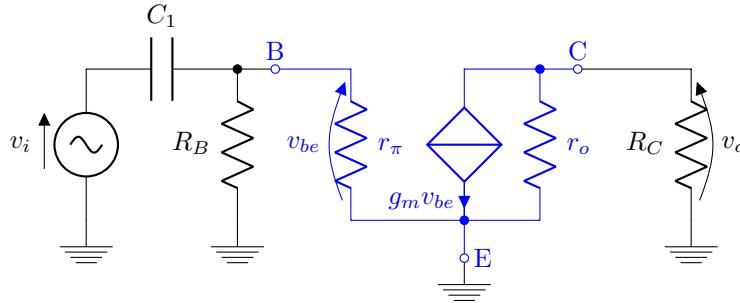


Figure 6.49: Equivalent circuit for small signal analysis.

Let us first solve the circuit considering C_1 as a short circuit. In this case, $v_i = v_{be}$ and the output voltage is:

$$v_o = -g_m v_i \cdot (r_o \parallel R_C)$$

If we neglect r_o , which is typically $200\text{ k}\Omega$, we obtain a gain of

$$A_v = -g_m R_C = \frac{I_{C0} R_C}{V_T} \simeq 192$$

This result is comparable with the simulation, where the full Gummel-Poon model was used. This formula gives us the variation of the output voltage with respect to the q.o.p., so the complete output voltage will be:

$$V_o = V_{o_{DC}} - g_m R_C v_i$$

We can now take into account the effect of C_1 . C_1 introduces a voltage divider between v_i and v_{be} :

$$v_{be} = \frac{R_B \parallel r_\pi}{R_B \parallel r_\pi + \frac{1}{sC_1}} v_i = \frac{sC_1(R_B \parallel r_\pi)}{1 + sC_1(R_B \parallel r_\pi)} v_i$$

C_1 introduces a zero in DC and a pole at the frequency:

$$f_p = \frac{1}{2\pi C_1 (R_B \parallel r_\pi)}$$

We need to compute r_π to be able to set the value of the capacitor so that the pole frequency is lower than the minimum signal frequency.

$$r_\pi = \frac{\beta V_T}{I_{C0}} = 3.42 \text{ k}\Omega$$

If we want to position the pole at 100 Hz, we need a 470 nF capacitor.

We managed to design an amplifier using a BJT, but we obtained a circuit with several problems. One of these problems is that the output of the amplifier has an offset with respect to the input equal to the q.o.p. value of V_o . We can easily modify the circuit by adding a series capacitor at the output. We obtain the circuit of figure 6.50.

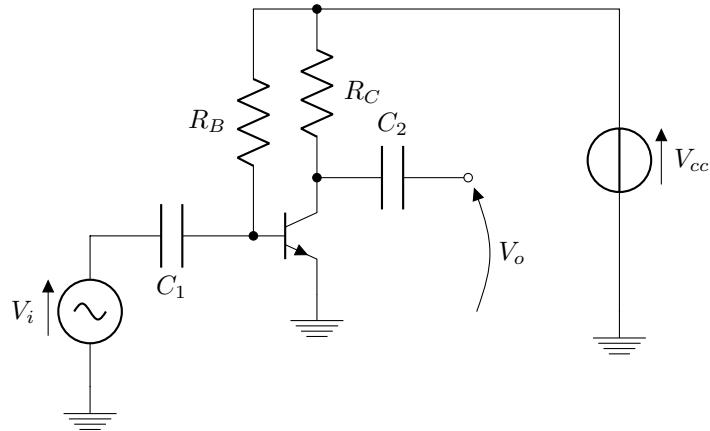


Figure 6.50: Modification to remove output offset.

C_2 introduces a zero in DC and a pole. The pole frequency depends on R_C , C_2 and the load impedance. The calculations are left as exercise.

To understand the other problems, we can compute how gain and q.o.p. vary if β changes from the design value of 200 to the minimum value of 100 or the maximum value of 300. For the q.o.p. we obtain that I_C changes from a minimum of 0.76 mA to a maximum of 2.28 mA, while V_o ranges from 2.48 V to 7.49 V. The gain A_v is comprised between 96 and 289.

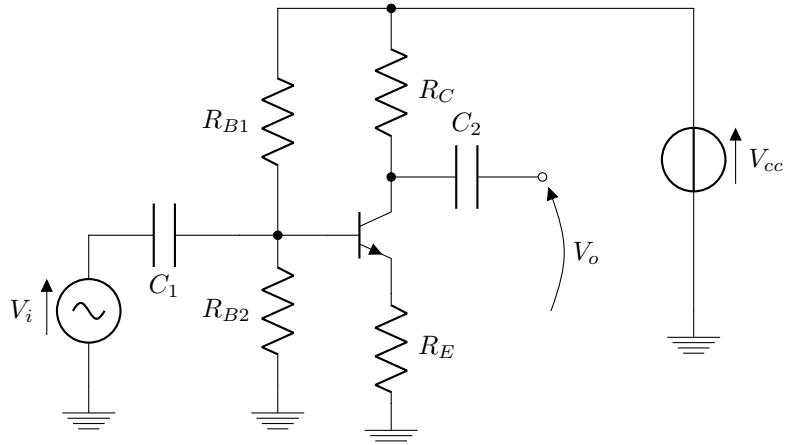
This variability of all the parameters of the amplifier is not acceptable for many applications. That is why, normally, single transistor amplifiers use some more complex scheme.

Example circuit

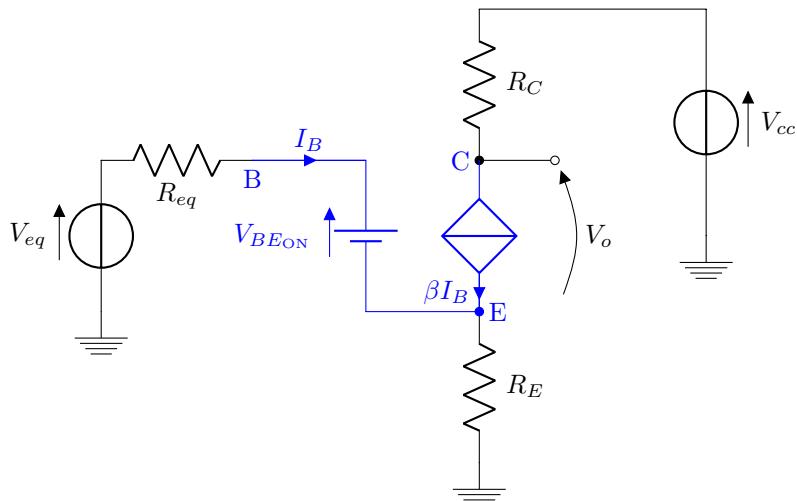
A better amplifier can be obtained just adding two resistors. We show the schematic in figure 6.51a.

We will analyze the circuit assuming the following values: $R_{B1} = 82 \text{ k}\Omega$, $R_{B2} = 10 \text{ k}\Omega$, $R_C = 4.7 \text{ k}\Omega$, $R_E = 330 \text{ }\Omega$, $V_{cc} = 10 \text{ V}$, $C_1 = C_2 = 470 \text{ nF}$.

Quiescent operating point. As we did before, we have to start calculating the quiescent operating point. That means that we have to use the q.o.p. model



a) Schematic diagram



b) Equivalent circuit for q.o.p.

Figure 6.51: Example schematic.

for the transistor, and the resulting schematic is in figure 6.51b. In the schematic we substituted the circuit at left of the transistor with its Thevenin equivalent:

$$V_{eq} = \frac{R_{B2}}{R_{B1} + R_{B2}} V_{cc} = 1.09 \text{ V}$$

$$R_{eq} = R_{B1} \parallel R_{B2} = 8.9 \text{ k}\Omega$$

By KVL at the input loop, we can compute I_B :

$$V_{eq} = R_{eq}I_B + V_{BE_{ON}} + (1 + \beta)I_B R_E$$

$$I_B = \frac{V_{eq} - V_{BE_{ON}}}{R_{eq} + (1 + \beta)R_E}$$

Using the typical value of 200 for β and $V_{BE_{ON}} = 0.7\text{ V}$, we obtain $I_B = 5.2\mu\text{A}$. We can then compute $I_C = \beta I_B = 1.03\text{ mA}$, and the voltage level on collector and emitter: $V_C = V_{cc} - I_C R_C = 5.14\text{ V}$, $V_E = (\beta + 1)I_B R_E = 0.34\text{ V}$.

To be sure that the transistor is in active zone, we need to check that the voltage level of the collector is at least 0.5 V higher than that on the emitter, otherwise the BC junction is not reverse biased and we are in saturation region. In our case $V_{CE} = 4.80\text{ V}$, so we are in the active region.

Small signal analysis. Once the q.o.p. is known, we have to perform a small signal analysis to find the gain of the circuit. We use the transistor equivalent circuit for small signal and we obtain the schematic shown in figure 6.52.

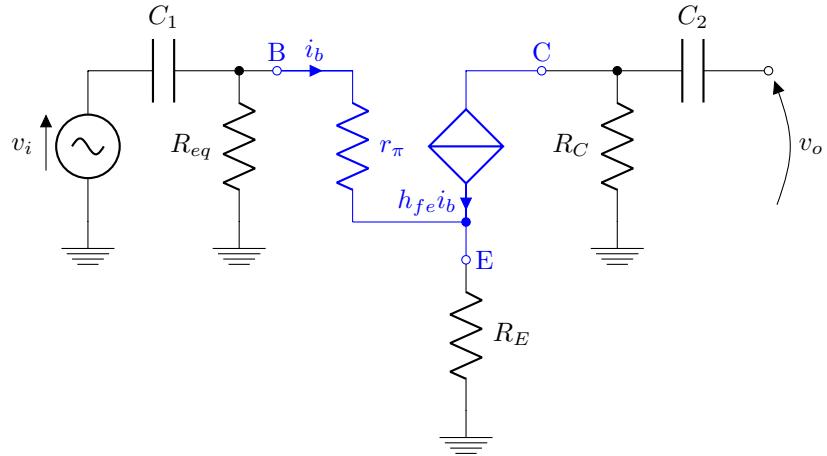


Figure 6.52: Equivalent circuit for small signal analysis.

We omitted r_o because it is much higher than R_C , so its effects are negligible. At first, we consider C_1 and C_2 as short circuit.

We can find i_b by KVL at the input loop. Then we find v_o by considering that, with C_2 short circuit, v_o is the voltage drop across r_o . We finally obtain:

$$\frac{v_o}{v_i} = -\frac{h_{fe}R_C}{r_\pi + (1 + h_{fe})R_E}$$

We know that

$$r_\pi = \frac{\beta V_T}{I_{C0}} = 5\text{ k}\Omega$$

Finally we have: $A_v = -13.2$.

We can compare the behavior of this circuit with the one of figure 6.44. We already noted that by changing the β parameter in the first circuit, we had big changes both in q.o.p. and gain of the amplifier. This is not true for the new circuit. If we compute I_C with $\beta = 100$, we obtain $I_C = 0.92\text{ mA}$. If we use $\beta = 300$, $I_C = 1.07\text{ mA}$. Even better stability is insured for the gain: A_v ranges from -13.0 to -13.23 with the above mentioned variations in β .

We still need to analyze the effects of the two capacitors. Let us start from C_1 . It is convenient to compute the input resistance of our amplifier by calculating the resistance seen from the right terminal of C_1 to ground. We note that we

have the parallel of $R_{eq} = R_{B1} \parallel R_{B2}$ and the resistance seen looking inside the base of the transistor. This is easily computed noting that the current through r_π is i_b and the current through R_E is $(1 + h_{fe})i_b$. The resistance is the total voltage drop on the two resistors, divided by i_b . In summary, if we call r_i the resistance seen by the capacitor, we have:

$$r_i = [r_\pi + (1 + h_{fe})R_E] \parallel R_{eq} = 7.92 \text{ k}\Omega$$

If we name v'_i the voltage drop across R_{eq} , we have:

$$v'_i = \frac{sC_1 r_i}{sC_1 r_i + 1} v_i$$

But the gain from v'_i to v_o is the same we computed above, considering C_1 as a short circuit. The total gain then is

$$\frac{v_o}{v_i} = -\frac{sC_1 r_i}{sC_1 r_i + 1} \cdot \frac{h_{fe} R_C}{r_\pi + (1 + h_{fe})R_E}$$

Capacitor C_1 introduces a zero in DC and a pole at frequency:

$$f_{p1} = \frac{1}{2\pi C_1 r_i} = 74.4 \text{ Hz}$$

We have to note that the frequency of the pole we just computed is valid only if the input signal comes from an ideal voltage generator. If the Thevenin equivalent of the input generator has a series resistance, that resistance will alter the pole frequency, lowering it.

It is convenient to study the frequency response due to C_2 taking into account also the load resistance, which is not shown in the schematic. If we assume that the load we connect to our amplifier does not exhibit an infinite resistance, that will alter the frequency response of the amplifier. So we will compute the output frequency dependency due to C_2 by inserting a resistor of value R_L from v_o to ground. Figure 6.53 shows the new circuit for small signal analysis.

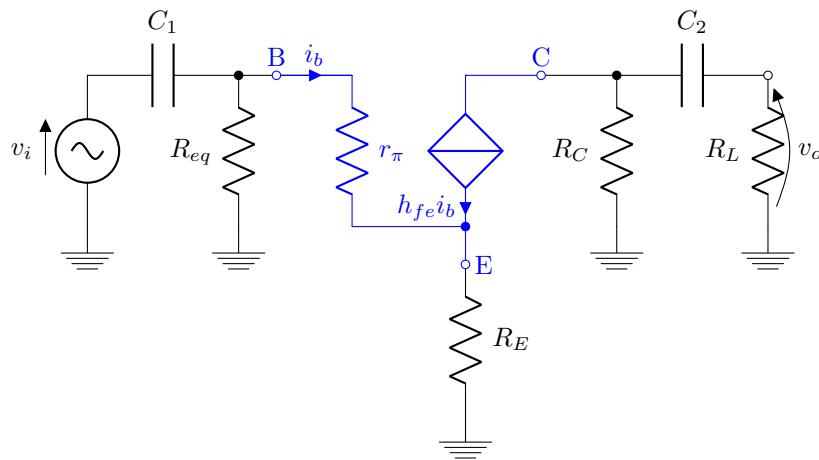


Figure 6.53: Equivalent circuit for small signal analysis with load resistor.

To find the voltage on the load resistor, we can note that not all the collector current flows through the load, but there is a current divider between R_C and the series of C_2 and R_L . We can write the current in R_L and then multiply it by R_L to find v_o .

$$v_o = -h_{fe}i_b \frac{sC_2R_C R_L}{sC_2(R_C + R_L) + 1}$$

If, for example, $R_L = 10\text{ k}\Omega$, the pole frequency is 40 Hz. The final transfer function is:

$$\frac{v_o}{v_i} = -\frac{sC_1r_i}{sC_1r_i + 1} \cdot \frac{h_{fe}}{r_\pi + (1 + h_{fe})R_E} \cdot \frac{sC_2R_C R_L}{sC_2(R_C + R_L) + 1}$$

The addition of R_L has the side effect of reducing the AC gain of the circuit. If we compute A_v at high frequency, we obtain $A_v = 8.96$.

Other amplifier circuits

Many other configurations can be used to design bipolar transistor based amplifiers. The examples we made so far use the configuration called *common emitter* and a single device. Common emitter refers to the fact that the input signal is applied to the base of the transistor, the output is on the collector and the emitter is common to input and output ports. It is also possible to obtain amplifiers by using the emitter as output. In this case the configuration is called *common collector* and the behavior of the device is very different (voltage gain less than one, but large current gain). We will not study this configuration or others more complex.

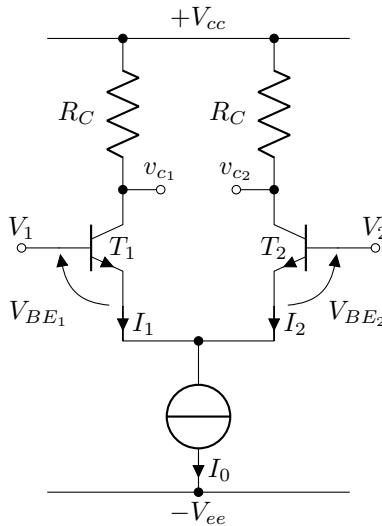


Figure 6.54: BJT based differential amplifier.

Differential amplifier Figure 6.54 shows a special case of amplifier which is used when we want to amplify also the DC component of the input signal. It is called differential amplifier and it is normally used as the input stage of a much

more complex circuit called operational amplifier. In the following we will use operational amplifiers but we will not go into details of their internal circuits. We can explain the behavior of the differential amplifier by noting that the two emitter currents depend on the respective V_{BE} , but the difference of the two voltages V_1 and V_2 is exactly equal to the difference of the two V_{BE} . We name this difference v_d :

$$v_d = V_1 - V_2 = V_{BE_1} - V_{BE_2}$$

If we consider as the q.o.p. the situation where $v_d = 0$ V, at the q.o.p. we have $I_1 = I_2 = I_0/2$ because I_1 and I_2 are emitter currents of a BJT and depend on the value of the base-emitter voltage of the respective transistors. If $v_d = 0$, the two base-emitter voltages are the same and so are the currents.

At the q.o.p., if β is high, we have $V_{c_1} = V_{c_2} = V_{cc} - R_C I_0/2$ (collector current nearly equal to emitter current). Note that the currents do not depend on the values of V_1 and V_2 but only on their difference. This is the reason for calling this amplifier a differential amplifier.

We can calculate the small signal gain when we apply a small differential signal v_d to the circuit. We can think of splitting this signal in $v_d/2$ applied to T_1 and $-v_d/2$ applied to T_2 . If we use the small signal model of the transistors, we can say immediately that the variation of collector current is g_m times the variation of the input voltage, where $g_m = I_{C0}/V_T$. But, when considering T_1 , the input voltage is $v_d/2$ and $I_{C0} = I_1 = I_0/2$, so:

$$v_{c_1} = -\frac{g_m v_d}{2} R_C = -\frac{I_0}{4V_T} v_d R_C$$

We can also write

$$v_{c_2} = \frac{g_m v_d}{2} R_C = \frac{I_0}{4V_T} v_d R_C$$

If we consider as output the difference between v_{c_1} and v_{c_2} , the gain will double and finally

$$v_o = v_{c_1} - v_{c_2} = -\frac{I_0}{2V_T} R_C v_d$$

Darlington pair Sometimes we need a very high gain between the base and the collector of a BJT. We already stated that normally the β of a transistor is in the range 50-300. If we need values in the ten thousands range, we can use a couple of transistors in what is known as Darlington pair.

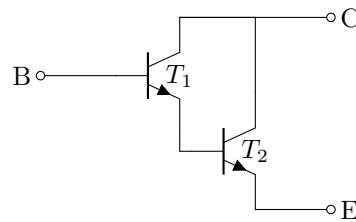


Figure 6.55: Darlington pair.

Figure 6.55 shows a Darlington pair. In this circuit the base current of T_2 is the emitter current of T_1 , and the collector current of T_2 can be expressed as function of the base current of T_1 :

$$I_{C2} = \beta_2 I_{B2} = \beta_2 \cdot (1 + \beta_1) I_{B1}$$

The current gain of the Darlington Pair is the product of the gains of the two transistors.

If we think of the Darlington pair as a single device, we can consider it as a super-transistor with a very large β (and also a larger $V_{BE_{ON}}$ to take into account). It is possible to buy Darlington pairs as integrated devices.

6.4 MOSFET devices

MOSFET stands for Metal Oxide Semiconductor Field Effect Transistor. It is inherently different from a bipolar transistor because a MOSFET is not based on the junction theory we used so far. The basis to understand how a MOSFET transistor is build and works, is to study a special kind of capacitor, the MOS capacitor.

6.4.1 MOS capacitor

A Metal Oxide Semiconductor capacitor is a very simple structure created by interposing a very thin oxide layer between a doped silicon sample and a metal layer.

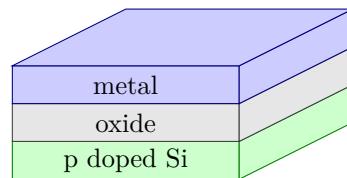


Figure 6.56: MOS capacitor.

Figure 6.56 shows the principle of a MOS capacitor created on a p-doped silicon sample. The thickness of the layers in the figure is not in scale. In standard MOS processes the MOS capacitors use silicon dioxide (SiO_2) as oxide and polycrystalline highly n-doped (n+) silicon as metal layer (poly), while in advanced processes the (SiO_2) is substituted by some high dielectric constant materials and poly is replaced by metals like copper. We name the metal layer *gate* of our device and the p doped silicon *substrate*.

If we apply a voltage V_G between gate and substrate, we can have different situations, as depicted in figure 6.57, where we see a vertical section of our structure.

- If V_G is negative, we obtain an increase of free holes in the interface between substrate and dielectric.
- If V_G is positive, free holes are removed from the region close to the interface between substrate and dielectric, so that a depleted region is created in the area

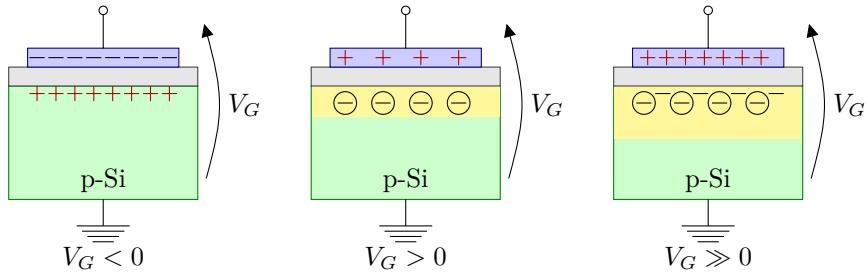


Figure 6.57: Gate voltage effects.

- If we further increase V_G , free electrons are accumulated in the interface between substrate and dielectric: this phenomenon is called *inversion*, because the p substrate behaves like n doped silicon in the area.

In inversion, we name *n-channel* the region where inversion happens. The density of free charge in the channel, Q_n , is nearly proportional to V_G after the inversion is obtained. Although some charge is present also for lower voltages, we define a *threshold voltage*, V_{th} as the voltage where the free electrons start being present consistently in the channel. The situation is described in figure 6.58. The charge

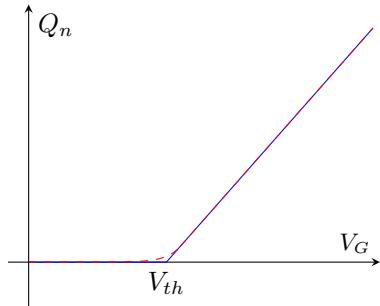


Figure 6.58: Free charge density in channel with respect to gate voltage.

control law is:

$$\begin{cases} V_G < V_{th} \implies Q_G = 0 \\ V_G > V_{th} \implies Q_G = C_{ox}(V_G - V_{th}) \end{cases}$$

Where C_{ox} is the capacitance per unit area of the MOS capacitor.

The oxide layer is not perfect and under some conditions can hold some trapped charge. Variations in the manufacturing process and trapped charge can change the threshold voltage, so the V_{th} parameter has a large tolerance, is not perfectly known.

6.4.2 MOSFET transistor

The MOSFET transistor is based on the MOS capacitor that we just introduced. We obtain a n-channel MOS field effect transistor (a.k.a. nMOS) by adding two heavily n-doped regions at two opposite sides of a MOS capacitor.

The section of the device is shown in figure 6.59. The two n+ regions are called Source, S, and Drain, D, while the substrate contact is normally called

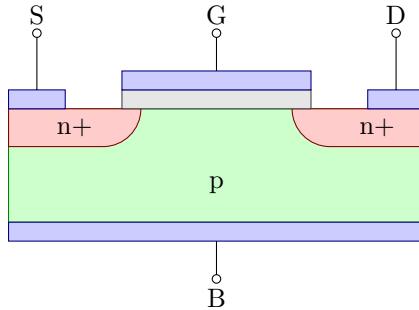


Figure 6.59: Section view of n-channel MOSFET.

Bulk, B. The distance between source and drain is the length of the device, L, while the other dimension is called width, W.

In most devices the source and drain regions are symmetrical. In this case, we define as source the terminal at lower potential. All voltages are referred to source, so we have the gate-source voltage, V_{GS} , drain-source voltage, V_{DS} , bulk-source voltage, V_{BS} . Normally $V_{BS} = 0$, and we will always assume this condition to write MOSFET equations.

Linear region

Studying the MOS capacitor, we found that if we apply a positive voltage to the gate, above the threshold voltage, we obtain a channel with free electrons. In the MOSFET, the channel creates a low resistance path between the drain and source terminals.

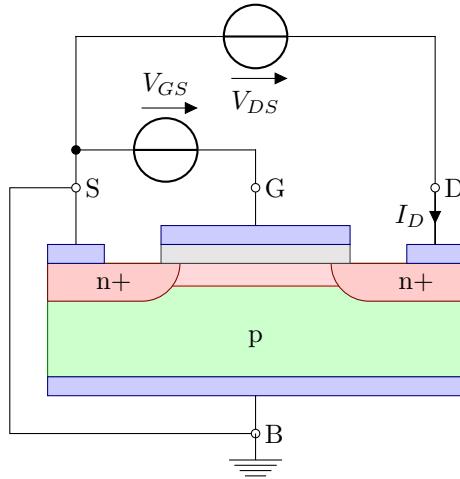


Figure 6.60: Biasing the nMOS transistor.

Figure 6.60 shows a nMOS transistor when we apply $V_{GS} > V_{th}$ and a low positive voltage V_{DS} . The transistor behaves like a resistor and a drain current flows into the drain terminal and through the channel. In the figure, the substrate is kept at the same potential as the source. In general, the substrate

must be at a potential equal or lower than source and drain, because otherwise the diodes formed by the p substrate and the source or drain regions will be forward biased and will start conducting.

If V_{DS} is close to 0, the drain current of the nMOS transistor obeys the following equation:

$$I_D = \mu_n C_{ox} \frac{W}{L} (V_{GS} - V_{th}) V_{DS}$$

where μ_n is the mobility of the electrons, C_{ox} is the above introduced capacity per unit surface of the oxide, W and L are respectively the width and length of the channel. The first two parameters are process dependent, while the others are the physical dimensions of the transistor, and can be set by the designer. Often the four parameters together are indicated as $k_n = \mu_n C_{ox} W/L$.

This is called linear or ohmic region. The MOSFET behaves like a resistor whose value is determined by the gate voltage. We can define a resistance value, called R_{ON} :

$$R_{ON} = \frac{V_{DS}}{I_D} = \frac{1}{k_n(V_{GS} - V_{th})}$$

Quadratic region

If we increase the drain voltage, the voltage drop across the channel is no more negligible. In this case, the electric field between gate and substrate is not constant but reduces when moving from source to drain.

The consequence is that the height of the channel reduces with the distance from the source. Figure 6.63 shows the channel in this region.

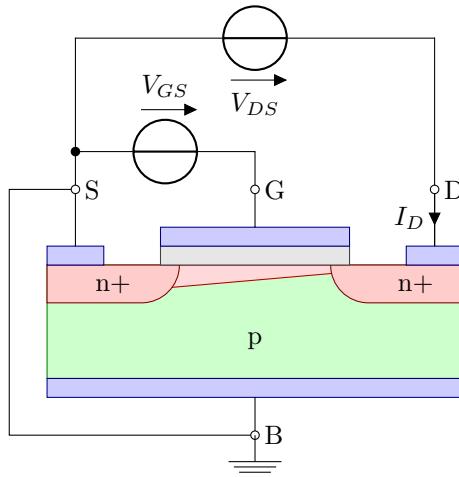


Figure 6.61: nMOS transistor in quadratic region.

The equation of the drain current is the following:

$$I_D = \mu_n C_{ox} \frac{W}{L} \left[(V_{GS} - V_{th}) V_{DS} - \frac{V_{DS}^2}{2} \right]$$

This region is called quadratic region or triode region. The latter name depends on the fact that the equation is similar to that of the old triode vacuum

tubes. The equation is a parabola and has a maximum in $V_{DS} = (V_{GS} - V_{th}) = V_{DSS}$. The drain voltage V_{DSS} is called *pinch-off* voltage. When the drain voltage is the pinch-off voltage, at the drain the channel is reduced to almost zero height. Above pinch-off the equation is no more valid (the drain current does not reduce when the drain voltage is above pinch-off voltage), but the device moves to a new region.

Saturation region

If the drain voltage rises above V_{DSS} , the drain current becomes almost independent of the drain voltage, so that the device's output port can be approximated to a current source. Note that the saturation region of a MOSFET device corresponds to the active or linear region of a bipolar transistor, while the linear region of a MOSFET transistor is the equivalent of the saturation region of a bipolar transistor. To avoid confusion, we will call the saturation region of the MOSFET *current generator* region.

The expression of I_D in this region becomes:

$$I_D = \frac{1}{2} \mu_n C_{ox} \frac{W}{L} (V_{GS} - V_{th})^2$$

This is an approximated expression, because a small linear increase of I_D with V_{DS} still holds. This is called Early effect or channel modulation. The complete equation is:

$$I_D = \frac{1}{2} \mu_n C_{ox} \frac{W}{L} (V_{GS} - V_{th})^2 (1 + \lambda V_{DS})$$

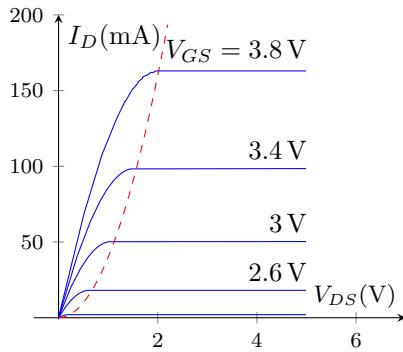


Figure 6.62: nMOS output characteristic (BS170).

Figure 6.62 shows the output characteristic of a small power nMOS BS170 for different values of V_{GS} . The left part of the graph is the linear region, then the slope of the curves reduces in the quadratic region, and finally the curves are nearly horizontal in the saturation region. The approximate border between quadratic and saturation region is indicated by the red dashed line.

6.4.3 p channel MOSFET

We obtain a pMOS, or p channel MOSFET, if we build the device starting from an n-doped substrate. The drain and source regions are p+ doped. In this case

the channel is of p type and that means that the majority carriers are holes instead of electrons.

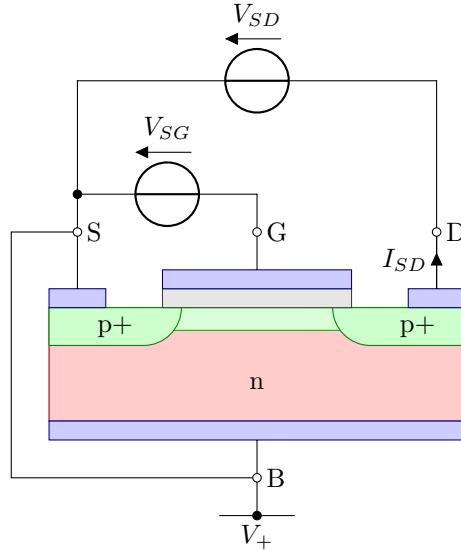


Figure 6.63: A pMOS transistor in linear region.

Figure 6.63 shows the section of a pMOS transistor together with the polarization to work in linear region.

Due to the different polarity of the charges, the equations defining the working conditions of this kind of transistor are reversed with respect to nMOS:

- the threshold voltage V_{th} is negative
- the gate-source voltage V_{GS} must be more negative than V_{th} to enable channel formation
- the drain current is negative
- the drain-source voltage V_{DS} is negative
- the substrate must be kept at a voltage larger than source and drain voltages (usually the substrate is connected to positive supply).

With these differences, the working regions and equations are the same as for the nMOS. To avoid working with negative numbers, we can rewrite the equations using V_{SG} instead of V_{GS} and V_{SD} instead of V_{DS} . We will also use I_{SD} (the current flowing from source to drain) instead of I_D , so that also the current is positive.

Linear region

When $V_{SG} > -V_{th}$ and $V_{SD} < V_{SG} + V_{th}$, with V_{SD} close to 0, we are in linear region.

The drain current obeys the following equation:

$$I_{SD} = \mu_p C_{ox} \frac{W}{L} (V_{SG} + V_{th}) V_{SD}$$

Like we did for nMOS, we can define $k_p = \mu_p C_{ox} W / L$

$$R_{ON} = \frac{V_{SD}}{I_{SD}} = \frac{1}{k_p (V_{SG} + V_{th})}$$

Quadratic region

When $V_{SG} > -V_{th}$ and $V_{SD} < V_{SG} + V_{th}$, with V_{SD} not close to 0, we are in quadratic region.

The equation for the quadratic region is:

$$I_{SD} = \mu_p C_{ox} \frac{W}{L} \left[(V_{SG} + V_{th}) V_{SD} - \frac{V_{SD}^2}{2} \right]$$

This equation is valid until we reach the pinch-off voltage $V_{SD} = (V_{SG} + V_{th}) = V_{SDSS}$.

Saturation region

If $V_{SD} > V_{SDSS}$ and $V_{SG} > -V_{th}$, we are in channel saturation region.

The equation of I_{SD} in this region is:

$$I_{SD} = \frac{1}{2} \mu_p C_{ox} \frac{W}{L} (V_{SG} + V_{th})^2 (1 + \lambda V_{SD})$$

6.4.4 Breakdown of MOSFET transistors

While the MOSFET is not based on the junction concept, there are two junctions respectively between substrate and drain and between substrate and source. These are normal diodes that are reverse biased, so they do not conduct in normal working regions. If the voltage between drain and substrate exceeds the breakdown voltage of the underlying junction, we have a sudden increase in the drain current, because the diode works as a zener diode. In this case we have also an increase of power dissipation. This situation can happen frequently in a MOSFET device used as a switch, if the load has inductive components. The breakdown is non destructive if the energy that has to be dissipated by the MOSFET is below a level specified on the transistor's datasheet. There is a second phenomenon creating breakdown, and it is due to avalanche multiplication of the carriers in the channel when the electric field between source and drain is too high and the channel is in pinch-off (saturation).

Another breakdown phenomenon, this time destructive, is related to V_{GS} . If the electric field between gate and channel is too high, the thin oxide layer can break, allowing current flow and perforating the oxide, that will be permanently damaged. To avoid this, we should guarantee that the gate voltage will never exceed a maximum positive (and negative) value specified in the datasheet.

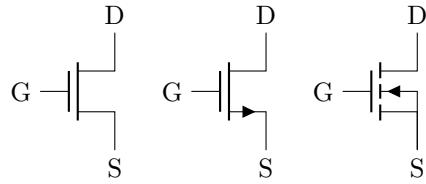


Figure 6.64: Different symbols used for nMOS transistor.

6.4.5 MOSFET symbols

Figure 6.64 shows different nMOS symbols usually encountered in schematic diagrams. The one at left is used when dealing with devices inside an integrated circuit, where the substrate is connected to the IC ground and the device is symmetric (source and drain can be interchanged, source is the terminal at lower voltage). It is used mostly in digital circuits.

The symbol in the middle is used mostly in analog circuits and the arrow denotes the source terminal and the direction of the source current.

The symbol at right is used for discrete devices: in this case the device should have four terminals. To reduce the number of pins, the substrate is connected to the source terminal, so that the device is no more symmetric: if we connect the source terminal at a voltage higher than the drain terminal, the substrate diode starts conducting. This is normally undesirable.

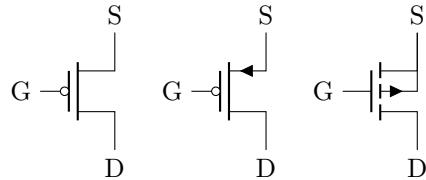


Figure 6.65: Different symbols used for pMOS transistor.

Figure 6.65 shows the corresponding symbols for pMOS devices. In the symbol at left, the different type of device is evidenced by the circle on the gate terminal. The meaning is that a low level voltage is required to switch-on the device, so the circle has the same function as in digital circuits. The symbol in the middle uses as addition the arrow on the source terminal, like in PNP BJTs. For the symbol at right, we can repeat the same considerations we made about the corresponding nMOS symbol.

For all three symbols we can also note that the source terminal is normally drawn as the upper one for pMOS transistors and as the lower one in nMOS transistors. This is due to the fact that the source must be at the lowest voltage for nMOS devices, at the highest for pMOS ones.

6.4.6 The MOSFET as amplifier

The MOS transistors can be used as amplifiers in a manner similar to BJT transistors. The main difference is that no DC current can flow in the gate terminal, because it is essentially a capacitor's plate.

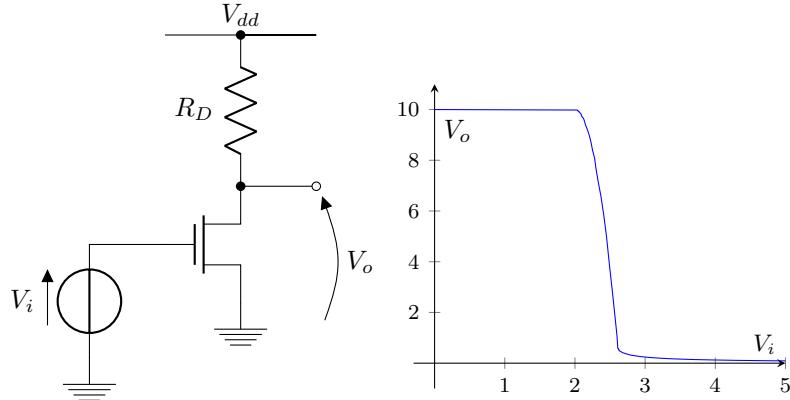


Figure 6.66: Simple nMOS amplifier.

We can start analyzing a circuit similar to the first one we introduced for BJTs. This circuit is shown in figure 6.66. We can simulate the DC transfer function with Spice, using a small power nMOS BS170 and $R_D = 500\Omega$. From the simulation results, we note that the threshold voltage of this transistor is 2 V. Above this voltage, the device enters channel saturation region. A further increase of V_i will move the device to quadratic and then ohmic regions. The circuit can be used as a nearly linear amplifier in the input voltage range from about 2.1 V to 2.6 V.

As for BJTs, the simulation does not take into account that the threshold voltage, having a typical value of 2 V, can range from 0.8 V to 3 V. This makes the practical use of this circuit as an amplifier very difficult. Nevertheless, we will use this circuit for digital applications.

We can add some components to build a better amplifier for AC signals. We need to bias the transistor so that it works in channel saturation region. As for BJTs, we build a model to find the quiescent operating point of the device, then we study small variations around the q.o.p. with a second model for small signal analysis.

Quiescent operating point model

The quiescent operating point model of a MOSFET transistor in channel saturation region is very simple. In DC, no current flows from gate to source, so the gate terminal is an open circuit. The voltage V_{GS} drives a voltage controlled current source which connects drain to source. The control equation is quadratic. Normally, the λV_{DS} term is neglected.

Let us introduce the model by applying it to a simple AC amplifier. The schematic is in figure 6.67. Since we are interested in DC operating point, we discard the capacitor and V_i . We substitute its Thevenin equivalent to the bias network set by R_1 , R_2 and their connection to V_{dd} , and use for the MOSFET the q.o.p. model we just introduced. The equivalent schematic is therefore the one displayed in figure 6.68.

Let us compute the q.o.p. by assuming the following data: $V_{dd} = 10\text{ V}$, $R_1 = 560\text{ k}\Omega$, $R_2 = 390\text{ k}\Omega$, $R_D = 2.2\text{ k}\Omega$, $k_n = \mu_n C_{ox} W/L = 1\text{ mA V}^{-2}$,

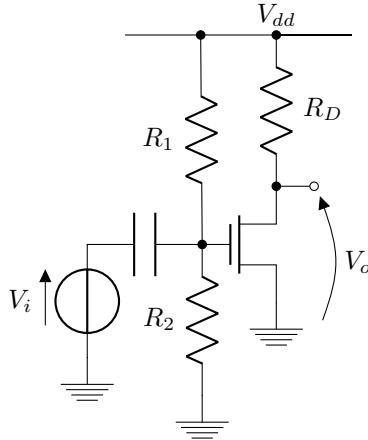


Figure 6.67: AC NMOS amplifier with simple bias circuit.

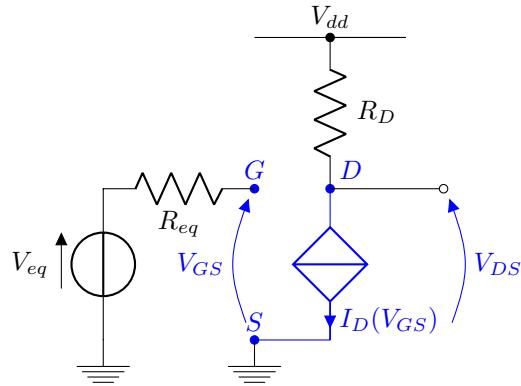


Figure 6.68: Quiescent operating mode circuit.

$V_{th} = 2 \text{ V}$, $\lambda = 0.01 \text{ V}^{-1}$. We have:

$$V_{eq} = V_{dd} \frac{R_2}{R_1 + R_2} = 4.1 \text{ V}$$

$$I_D = \frac{1}{2} k_n (V_{GS} - V_{th})^2 = 2.22 \text{ mA}$$

$$V_{DS} = V_{dd} - R_D I_D = 5.12 \text{ V}$$

We did not compute the value of R_{eq} because no current flows into the gate of the transistor. When performing the calculations one should always check two points:

- $V_{GS} > V_{th}$. If this is not true, the MOSFET is not conducting and there is no current in the drain.
- $V_{DS} > V_{GS} - V_{th}$. If this is not true, the MOSFET is not working in channel saturation zone. In this case, we have to repeat the computations using the model and the equations for quadratic or linear region. We will cover the model for linear region in the chapters focused on digital electronics.

Small signal model

As for bipolar transistors, once we know the quiescent operating point of the device, we apply the signal and we are interested in the *variations* of the output signal with respect to the q.o.p. If the signal is sufficiently small so that also output variations are small, it is possible to use a linear differential model for the device. We recall that, in channel saturation region, I_D has the following expression:

$$I_D = \frac{1}{2}k_n(V_{GS} - V_{th})^2(1 + \lambda V_{DS})$$

where we see a dependence on V_{GS} and a (small) dependence on V_{DS} . Under small signal conditions, we want to find a linear expression:

$$i_{D,ss} = \left. \frac{\partial I_D}{\partial V_{GS}} \right|_{q.o.p.} \cdot v_{GS,ss} + \left. \frac{\partial I_D}{\partial V_{DS}} \right|_{q.o.p.} \cdot v_{DS,ss}$$

We name the two partial derivatives respectively g_m and g_o , mutual and output transconductances, so we can write:

$$i_{D,ss} = g_m v_{GS,ss} + g_o v_{DS,ss}$$

To find g_m we neglect the λ term:

$$g_m = \left. \frac{\partial I_D}{\partial V_{GS}} \right|_{V_{GS0}, I_{D0}} = k_n(V_{GS0} - V_{th}) = \sqrt{2k_n I_{D0}}$$

where I_{D0} and V_{GS0} are the values of I_D and V_{GS} computed in the quiescent operating point. The expression of g_o is:

$$g_o = \left. \frac{\partial I_D}{\partial V_{DS}} \right|_{V_{GS0}, I_{D0}} = \frac{1}{2}k_n(V_{GS0} - V_{th})^2\lambda = \lambda I_{D0}$$

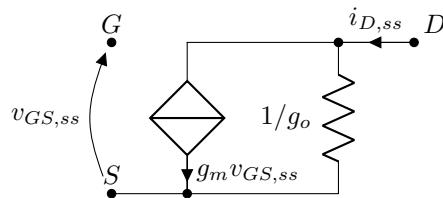


Figure 6.69: Small signal model of MOSFET.

Figure 6.69 shows the schematic of the small signal model of the MOSFET.

Let us apply our model to find the AC gain of the amplifier of figure 6.67. Let us use the same parameters we used for q.o.p. computation.

When we substitute our model to the transistor, we obtain the small signal equivalent circuit of figure 6.70. In the figure we omitted the *ss* subscript on all voltages and currents.

Our goal is to find the small signal gain of the amplifier. We note that the capacitor C introduces a high pass filter, with a zero in DC and a pole at frequency $f_p = 1/(2\pi R_{eq}C)$. Above this frequency, we are in the pass band of

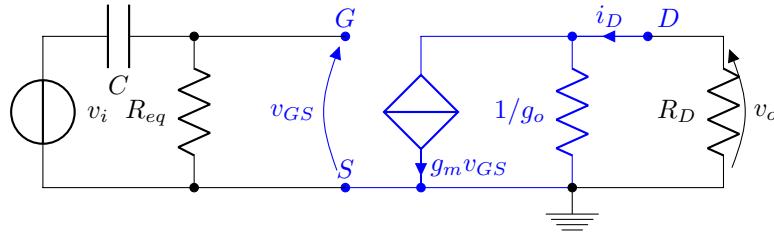


Figure 6.70: Small signal equivalent of the amplifier.

the amplifier. We will only compute the gain within the amplifier's bandwidth, so we neglect the effects of the capacitor. The results will be valid for $f > f_p$.

In this frequency range, $v_{gs} = v_i$. the output voltage v_o is:

$$v_o = -g_m v_{GS} \frac{R_D}{1 + R_D g_o} = -g_m v_i \frac{R_D}{1 + R_D g_o}$$

The gain of the amplifier is therefore:

$$A_V = -g_m R_D \parallel 1/g_o$$

Let us compute g_m and g_o :

$$g_m = \sqrt{2k_n I_{D0}} = \sqrt{2 \cdot 1 \text{ mA V}^{-2} \cdot 2.22 \text{ mA}} = 2.11 \text{ mA V}^{-1}$$

$$g_o = \lambda I_{D0} = 0.01 \text{ V}^{-1} \cdot 2.22 \text{ mA} = 0.022 \text{ mA V}^{-1}$$

The gain is:

$$A_V = -2.11 \text{ mA V}^{-1} \cdot 2.01 \text{ k}\Omega = 4.43$$

Obviously, MOSFET devices can be used in more complex configurations. For example, the differential amplifier we introduced with bipolar transistors can be implemented also with MOS devices.

6.5 Exercises

Exercise 6.1. Diode as rectifier

Referring to the circuit of figure 6.71, assume the following values for the passive components: $R_1 = 1 \text{ k}\Omega$, $R_2 = 2 \text{ k}\Omega$.

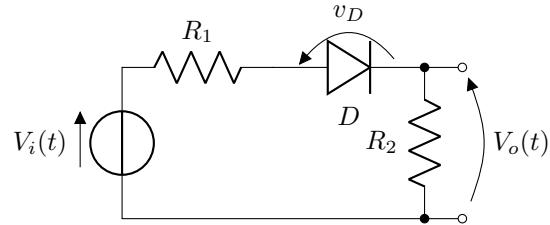


Figure 6.71: Diode circuit.

- If V_i is a square wave, with $V_H = 3 \text{ V}$ and $V_L = -6 \text{ V}$ what is the output waveform if the diode is ideal?
- In the same situation, what is the output waveform if the diode has $V_\gamma = 0.7 \text{ V}$?

Exercise 6.2. Diode as rectifier.

Referring to the circuit of figure 6.71, assume the following values for the passive components: $R_1 = 1 \text{ k}\Omega$, $R_2 = 2 \text{ k}\Omega$.

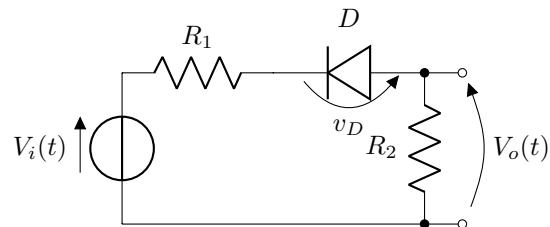


Figure 6.72: Diode circuit.

1. If V_i is a square wave, with $V_H = 3 \text{ V}$ and $V_L = -6 \text{ V}$ what is the output waveform if the diode is ideal?
2. In the same situation, what is the output waveform if the diode has $V_\gamma = 0.7 \text{ V}$?

Exercise 6.3. Zener diode.

Referring to circuit of figure 6.73, the diode has $V_{Z0} = 5 \text{ V}$, $V_\gamma = 0.7 \text{ V}$ and $r_Z = 10 \Omega$. $R_i = 100 \Omega$

1. draw the DC transfer characteristic $V_Z(V_i)$.
2. If V_i is a triangular wave, with peak high value $V_H = 10 \text{ V}$ and peak low value $V_L = -10 \text{ V}$ what is the waveform on V_Z ? Draw a time diagram.

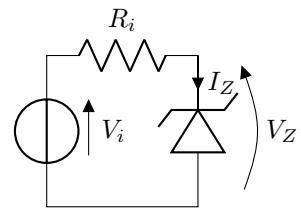


Figure 6.73: Circuit with zener diode.

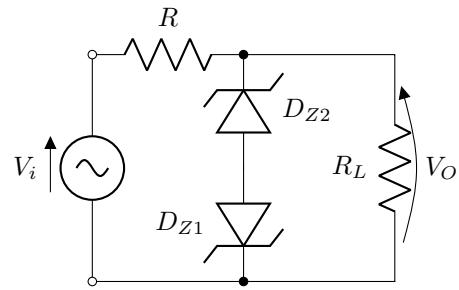


Figure 6.74: Clamping circuit.

Exercise 6.4. Clamping circuit.

In the circuit of figure 6.74 we have: $R_L = 9\text{ k}\Omega$ and $R = 1\text{ k}\Omega$. The two zener diodes are identical and their parameters are: $V_{Z0} = 4\text{ V}$, $r_Z = 0\Omega$ and $V_\gamma = 0.7\text{ V}$.

1. draw the DC transfer characteristic $V_O(V_i)$.
2. If V_i is a triangular wave, with peak high value $V_H = 10\text{ V}$ and peak low value $V_L = -10\text{ V}$ what is the waveform on V_O ? Draw a time diagram.

Exercise 6.5. Simple BJT amplifier.

In the circuit of figure 6.75 we have: $R_C = 2.2 \text{ k}\Omega$, $R_B = 470 \text{ k}\Omega$, $C_1 = 10 \mu\text{F}$, $V_{cc} = 15 \text{ V}$. The BJT has $\beta = 100$, $V_{BEON} = 0.7 \text{ V}$.

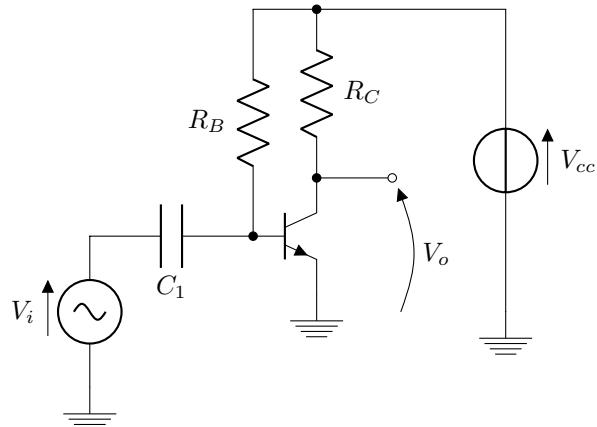


Figure 6.75: Simple amplifier.

1. Compute the quiescent operating point of the circuit: I_B , I_C , V_{CE} .
2. Compute the small signal high frequency gain of the circuit (C_1 short circuit).
3. Compute the frequency of the pole introduced by C_1 .

Exercise 6.6. BJT amplifier.

In the circuit of figure 6.76 we have: $R_C = 2.2 \text{ k}\Omega$, $R_B = 470 \text{ k}\Omega$, $R_E = 220 \Omega$, $C_1 = 10 \mu\text{F}$, $V_{cc} = 15 \text{ V}$. The BJT has $\beta = 100$, $V_{BEON} = 0.7 \text{ V}$.

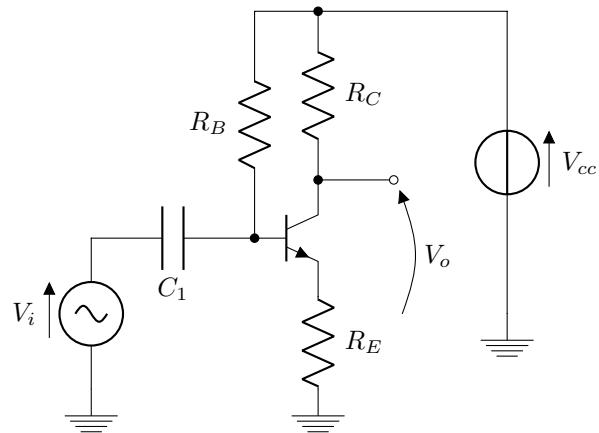


Figure 6.76: BJT amplifier.

1. Compute the quiescent operating point of the circuit: I_B , I_C , V_{CE} .
2. Compute the small signal high frequency gain of the circuit (C_1 short circuit).
3. Compute the frequency of the pole introduced by C_1 .

Exercise 6.7. Simple MOSFET amplifier.

The circuit of figure 6.77 has the following parameters:

R_1	100 k Ω	R_2	50 k Ω	R_D	2 k Ω
V_{dd}	12 V	C	100 nF		
V_{th}	2.5 V	λ	0.01 V $^{-1}$	k_n	2 mA V $^{-2}$

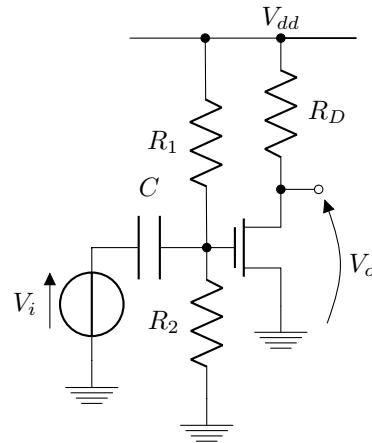


Figure 6.77: AC NMOS amplifier with simple bias circuit.

1. Compute the quiescent operating point of the circuit: V_{GS} , I_D , V_{DS} .
2. Compute the small signal high frequency gain of the circuit (C short circuit).
3. Compute the frequency of the pole introduced by C .

Chapter 7

Transistors in switching Mode

Contents

7.1	BJT switches	7-2
7.1.1	Operation in Saturation	7-3
7.1.2	Design example	7-4
7.2	MOSFET switches	7-5
7.2.1	Open and close switch	7-7
7.2.2	Design example	7-7
7.3	Dynamic behaviour	7-9
7.3.1	Dynamic behaviour of diodes	7-9
7.3.2	Dynamic behaviour of BJTs	7-11
7.3.3	MOSFET dynamic behaviour	7-12
7.4	Exercises	7-16

A VERY important use of bipolar (BJT) and field effect (MOS) transistors is in the implementation of *electronic switches*.

Such switches are used in a variety of electronic circuits, ranging from logic gates to switched capacitor filters, to switching power supplies, to power loads driving and many others. In all these applications we need a device which may **close** or **open** an electrical connection between two parts of a circuit, according to the status of a control signal.

In most cases one of the nodes of the switch is either the power supply or the reference voltage and electronic switches may be classified according to their positioning with respect to the power supply and the load as shown in figure 7.1.

1. **Low-side** switch: the load is directly connected to the power supply and the switch is between the load and the reference node (0 V);
2. **High-side** switch: the switch is between the power supply and the load and the load is directly connected to the reference node at 0 V.

In general it is easier to implement low-side switches than high-side ones. It should also be noted that many switches, especially those using BJTs, are unidirectional because the current may flow in only one direction.

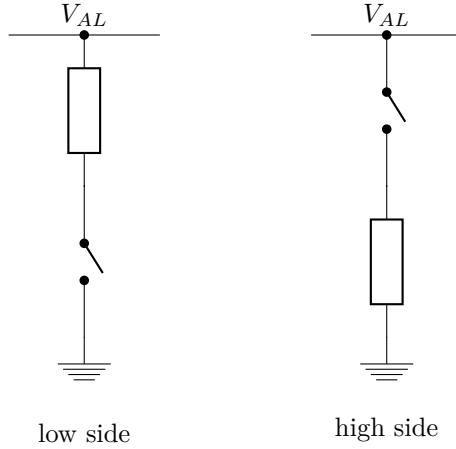


Figure 7.1: *Lowside* and *highside* switches.

7.1 BJT switches

When operating in the linear area the BJT is a device which provides an output signal approximately proportional to the input one, that is it behaves as an amplifier. In order to obtain a switch we would like to have only two output voltage values associated to the ON/OFF states or, in logic circuits to TRUE/FALSE, 0/1, HIGH/LOW signals.

In order to understand the differences between linear and switching applications let us analyze a circuit, shown in figure 7.2, which may operate both in linear and switching mode. It is composed of a BJT whose collector is connected to a load R_C and a power supply V_{AL} and whose input is a control voltage V_i transformed into a current through a resistance R_B .

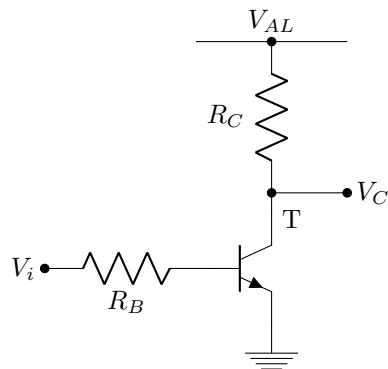


Figure 7.2: A simple BJT circuit

When V_i is 0 the BJT is not biased and is in the *cutoff* state in which the two junctions, base-collector and base-emitter may be considered as open circuits. No current is flowing in the load R_C and therefore the output voltage is equal to the power supply.

When V_i increases over the threshold voltage V_{BEON} , the base-emitter junc-

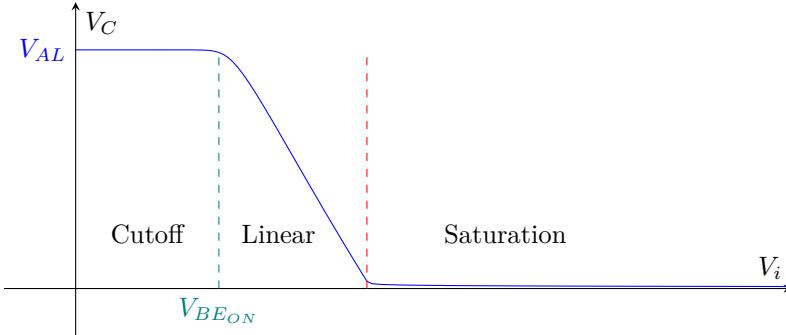


Figure 7.3: Transfer characteristic of the circuit of figure 7.2 showing the different BJT operating regions.

tion is forward biased and the BJT is in the direct active region. The voltage drop across the base-emitter junction is approximately constant and the relation relating the current I_B to voltage V_i is:

$$V_i > V_{BE_{ON}} \implies I_B = \frac{V_i - V_{BE_{ON}}}{R_B}$$

The collector current is proportional to the base current through the $\beta_{DC} = h_{FE}$ coefficient and the voltage drop on the load R_C will be proportional to it. We will then have:

$$I_C = h_{FE} I_B; \quad V_C = V_{AL} - R_C I_C$$

$$V_C = V_{AL} - R_C \cdot h_{FE} \cdot I_B = V_{AL} - R_C \cdot h_{FE} \cdot \frac{V_i - V_{BE_{ON}}}{R_B}$$

The output voltage will be linearly decreasing as a function of V_i . This is an approximation because the h_{FE} parameter depends on I_C and is not constant as it is clearly shown in the first part of the linear region in figure 7.3. The relation is valid as long as the transistor is in the direct active area: when the collector voltage decreases to a value which directly biases the base-collector junction, the transistor enters the *saturation* region.

7.1.1 Operation in Saturation

When the transistor enters the saturation region, the collector-emitter voltage continues to decrease as the base current increases but the asymptotically decrease to zero of V_C is very slow and for most practical applications it is usual to consider the output voltage to be constant at a value a $V_{CE_{SAT}} = 0.2 - 0.4V$.

The saturation region may be modeled in detail using the complete Ebers-Moll model.

$$I_E = I_S \left(e^{\frac{V_{BE}}{V_T}} - e^{\frac{V_{BC}}{V_T}} \right) + \frac{I_S}{\beta_F} \left(e^{\frac{V_{BE}}{V_T}} - 1 \right)$$

However it is far more practical to use other design tools to design a circuit in which the saturation is guaranteed under certain conditions.

Let us observe first that, when in saturation, the I_C and I_B currents in the BJT may have any value, provided that their ratio is lower than that in linear operation in the active region. Let us further observe that in saturation the collector current is determined by the load and not by the transistor because the voltage drop on the load is constant and equal to the power supply. The base current is determined by the base circuit and the two currents may be independently set. Let us then introduce a parameter referred to as β_{forced} and defined as:

$$\beta_{forced} = \frac{I_{Csat}}{I_{Bsat}} < h_{FE}$$

The term β_{forced} derives from the fact that the designer may choose any desired value, within the limits of the previous inequality, by setting independently the value of the base and the collector current.

To implement a good switch it is necessary to design the circuit in such a way that, for the set of values of the input voltage V_i , the transistor is never operating in the linear region, where the power dissipation is high. It is necessary to associate the states of *open* and *closed* switch respectively to cutoff and saturation.

The BJT is operating in cutoff when $V_i < V_{BE}$; the base emitter junction is not forward biased, the collector current is null and the transistor is OFF. For what concerns operation in saturation we must know the **minimum** β of the transistor.

By observing the characteristic of figure 7.3 one may note that the width of the linear region depends on the slope of the linear segment ΔV_i and is given by the ratio $\Delta V_C / \Delta V_i$. This ratio depends on the value of the resistances in the circuit and on the β parameter (also referred to as h_{FE}) as obtained in the previous paragraph.

A high value of β corresponds to a fast transition from high to low voltage. Therefore it is necessary to know the minimum value of β in order to compute the maximum amplitude of the range of V_i with a linear operation of the BJT. Knowing this range we may then compute the minimum value of V_i which guarantees transistor saturation and therefore the condition of closed switch.

7.1.2 Design example

With reference to fig. 7.2, given $V_{AL} = 10\text{ V}$, using as load R_C a lamp with a power of 0.5 W when operating at 10 V (approximated as a linear load), size the circuit parameters in such a way that the lamp is on when $V_i > 5\text{ V}$ and is off when $V_i < 0.5\text{ V}$.

The design parameters are the transistor and resistance R_B . The former must be able to withstand a collector current larger than the one which is flowing in the lamp when it is on, which may be computed as:

$$P_L = I_C V_C = 0.5\text{ W} \implies 10\text{ V} \cdot I_C = 0.5\text{ W} \implies I_C = 50\text{ mA}$$

The BJT maximum collector current must be greater than 50 mA and the maximum voltage V_{CEO} greater than 10 V. Transistor 2N2222A satisfies these specifications.

We must now determine the value of resistance R_B . In saturation we have determined that :

$$I_C = \beta_{forced} \cdot I_B$$

Where β_{forced} is a parameter set by the designer, lower than the minimum value of the transistor β , as specified from the datasheet which must be analyzed accurately.

The value of $h_{FE_{min}}$ is specified under very precise conditions of I_C , V_{CE} and temperature, for instance $I_C = 10$ mA, $V_{CE} = 10$ V, $T = 25$ °C. To be sure to maintain the saturation condition for any temperature and for the value of I_C defined by the circuit, it is often necessary to consider a value of h_{FE} 3 or 4 times lower than one in full linearity (for instance $V_{CE} = 10$ V). In our case let us divide by 3 the value read in the datasheet:

$$\beta_{forced} = \frac{h_{FE_{min}}}{3} = \frac{75}{3} = 25$$

$$\Rightarrow I_B = \frac{I_C}{\beta_{forced}} = \frac{50 \text{ mA}}{25} = 2 \text{ mA}$$

At this point it is necessary to evaluate the base resistance R_B and to do this it is necessary to known the base-emitter voltage in saturation, which is slightly higher than that in linearity. From the datasheet we get: $V_{BE_{SAT}} = 1$ V.

$$I_B = \frac{V_i - V_{BE_{SAT}}}{R_B} = 2 \text{ mA} \Rightarrow R_B = \frac{V_i - V_{BE_{SAT}}}{2 \text{ mA}} = 2 \text{ k}\Omega$$

Taking components from the E12 series, we may choose $R_B = 1.8 \text{ k}\Omega$.

7.2 MOSFET switches

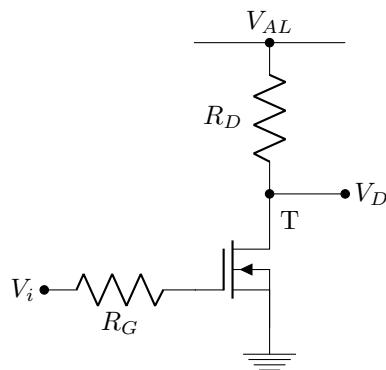


Figure 7.4: Simple switch circuit with MOSFET.

The circuit of a switch implemented with a MOSFET is shown in figure 7.4.

The transfer characteristic V_i/V_D is shown in the graph 7.5. The switch input voltage corresponds to the gate to source voltage V_{GS} of the MOSFET: when it

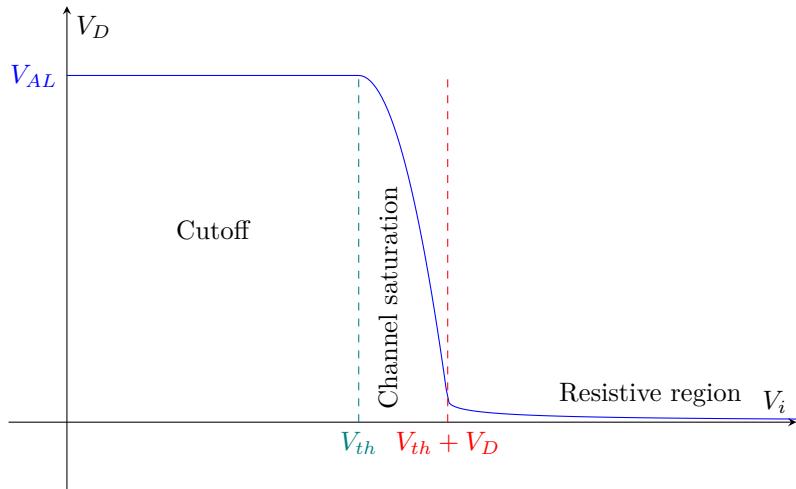


Figure 7.5: Transfer characteristic of the circuit of figure 7.4 showing the different operating *regions*.

is lower than the threshold, there is no channel between drain and source , they are disconnected and therefore the switch is an open circuit for $V_i < V_{th}$.

When the input is greater than the threshold voltage, a channel between source and drain is formed and the current I_D increases quadratically with V_{GS} . In this *channel saturation region* the output behaves as a *current generator*):

$$I_D = \frac{1}{2} \mu C_{ox} \frac{W}{L} (V_{GS} - V_{th})^2 (1 + \lambda V_{DS})$$

As we did for the BJT switch we may obtain the expression of $V_D(V_i)$:

$$V_D = V_{AL} - \frac{1}{2} R_D \mu C_{ox} \frac{W}{L} (V_i - V_{th})^2$$

in which we have neglected the term λV_D . Also in this case V_D is a decreasing function of V_i because when the current is increasing, the voltage drop across R_D increases too.

When the drain voltage becomes low enough ($V_{DS} < V_{GS} - V_{th}$), the device enter the operating region named *triode* or *quadratic* in which:

$$I_D = \mu C_{ox} \frac{W}{L} \left[(V_{GS} - V_{th}) V_{DS} - \frac{V_{DS}^2}{2} \right]$$

Reducing further V_{GS} the term $V_{DS}^2/2$ in the previous equation becomes negligible and the device enters the *resistive region*, which is the equivalent of the saturation region of the BJT.

If voltage V_i is high enough to bring the MOSFET in the resistive region, then load resistance R_D may be considered connected to 0 V through an equivalent resistance R_{ON} corresponding to the MOSFET and whose value is given by:

$$R_{ON} = \frac{1}{\mu C_{ox} \frac{W}{L} (V_i - V_{th})}$$

Let us observe that, although in static conditions no current flows through resistance R_G , it has however an important role as far as the reliability of the circuit is concerned. The input gate capacitance associated to the inductance of the interconnecting wires behaves as a resonant circuit and if the rising or falling edges of the control signal are sharp enough, they may trigger oscillations whose amplitude may lead to the gate oxide breakdown.

The purpose of R_G is to reduce the Q of the resonant circuit and therefore the value of the voltage overshoot. Its value is in general in the order of 10Ω .

7.2.1 Open and close switch

The two operating regions of the MOSFET which are associated to the open and closed switch conditions are respectively:

- $V_i \rightarrow 0$: MOSFET in cutoff and open circuit between source and drain. There is no current through R_D and no voltage drop across it and therefore $V_D \simeq V_{AL}$;
- $V_i \rightarrow V_{AL}$: MOSFET in resistive region and modeled from drain to source as a small resistance with a value far lower than R_D . Therefore $V_D \simeq 0\text{ V}$.

As we did for the BJT let us now analyze a design example in order to understand the design differences between BJT and MOSFETS used as electronic switches.

7.2.2 Design example

Design the switch of the example 7.1.2 previously analyzed for a BJT, using a MOSFET.

With respect to the BJT design the main difference is that it is not necessary to size R_G whose choice of value has been discussed in section 7.2. The only choice we have to make is that of the device which we are going to use as a switch.

The main specifications which must be satisfied are:

- Threshold voltage: the MOSFET must be ON for V_i ON (5 V) and OFF for V_i OFF (0.5 V);
- Maximum drain current: must be larger than the maximum load current;
- R_{ON} : must be low with respect to the equivalent load resistance in such a way that in the ON state most of the power supply is on the load;
- maximum V_{DS} voltage when in OFF state: must be larger than the power supply of the circuit.

We know that the drain current I_D required to switch on the lamp is:

$$I_D = \frac{500\text{ mW}}{10\text{ V}} = 50\text{ mA}$$

Hence the lamp equivalent resistance R_{eq} is:

$$R_{eq} = \frac{10\text{ V}}{50\text{ mA}} = 200\Omega$$

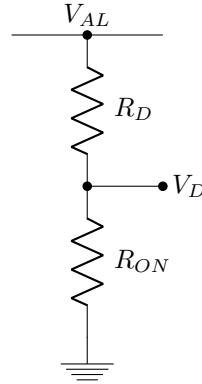


Figure 7.6: MOSFET in the resistive region in series to the lamp modeled as load resistance R_D .

When the MOSFET is ON it operates in the resistive region and can be modeled as a resistance R_{ON} as shown in fig. 7.6.

Since the MOSFET behaves like a resistance if we impose that, when it is ON, the maximum voltage drop on R_{ON} is 0.5 V, then we may determine the maximum value of R_{ON} from the voltage partition across it:

$$10 \text{ V} \cdot \frac{R_{ON}}{R_{ON} + R_D} \leq 0.5 \text{ V} \implies R_{ON} \leq 10 \Omega$$

Let us verify whether the BS170 MOSFET if it is suited to this application. From the datasheet, available on web site <https://www.onsemi.com/pub/Collateral/BS170-D.PDF>, we may obtain the parameters shown in figure 7.1.

ABSOLUTE MAXIMUM RATINGS $T_A = 25^\circ\text{C}$ unless otherwise noted

Symbol	Parameter	Conditions	Min	Typ	Max	Units
V_{DSS}	Drain-Source Voltage				60	V
V_{GSS}	Gate-Source Voltage				± 20	V
I_D	Drain Current	Continuous			500	mA
		Pulsed			1200	mA
P_D	Maximum Power Dissipation Derate above 25°C				800	mW
					2.4	$\text{mW } ^\circ\text{C}^{-1}$

ELECTRICAL CHARACTERISTICS $T_A = 25^\circ\text{C}$ unless otherwise noted

Symbol	Parameter	Conditions	Min	Typ	Max	Units
OFF CHARACTERISTICS						
BV_{DSS}	Drain-Source Breakdown Voltage	$V_{GS} = 0 \text{ V}$, $I_D = 100 \mu\text{A}$	60			V
ON CHARACTERISTICS						
$V_{GS(th)}$	Gate Threshold Voltage	$V_{DS} = V_{GS}$, $I_D = 1 \text{ mA}$	0.8	2.1	3	V
$R_{DS(ON)}$	Static Drain-Source On-Resistance	$V_{GS} = 10 \text{ V}$, $I_D = 200 \text{ mA}$		1.2	5	Ω

Table 7.1: Some parameters of the BS170 MOSFET

The parameters relative to the supply voltage (V_{DSS} e BV_{DSS}) and to the maximum drain current are far larger than the ones expected in the circuit and do not cause any problem. It is however necessary to analyze further the value of

$R_{DS(ON)}$ because it is specified for a V_{GS} far higher than that of the application.

We must therefore verify the dependence of the resistance from V_{GS} . From the graphs of figures 2, 3 and 4 of the datasheet (not shown in the table) we see that it is reasonable to expect at maximum a doubling of R_{ON} with respect to maximum value shown in the table. We are in any case within the limits of the design. A further parameter to be verified is the maximum power which may be dissipated by the device. Considering a maximum drain current of 50 mA and a maximum R_{ON} di 10 Ω , then the maximum power dissipation is:

$$P_D = I_D^2 \cdot R_{ON} = 25 \text{ mW}$$

which is compatible with the maximum value shown in the datasheet.

7.3 Dynamic behaviour

Until now we have analyzed the switching behaviour of a transistor in static conditions and we will now look at the behaviour during the transients from close to open and open to close conditions.

7.3.1 Dynamic behaviour of diodes

In order to better understand the characteristics of a bipolar transistor, let us start by analyzing a simpler circuit, composed of a square wave generator, a diode and a load resistor, shown in figure 7.7.

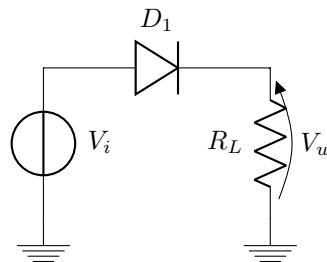


Figure 7.7: Simple circuit for the analysis of the dynamic input-output behaviour.

If we consider the simple diode model used for the dc analysis we would expect at the output a voltage with the same waveform as in the input minus the constant voltage drop across the diode. Using a 1N4007 diode, a resistance R_L of 10 Ω , a generator with an output square waveform with a 10 V amplitude and assuming a 1 V drop across the diode we would expect an output as that shown in figure 7.8.

A more realistic model of the diode would consider a slight delay of the output with respect to the input. However simulating the circuit with Spice, as shown in figure 7.9(a), we may clearly see that after the falling edge of the input signal on the output there is a negative voltage of about -1 V for a short time (approximately 26 μs in the example).

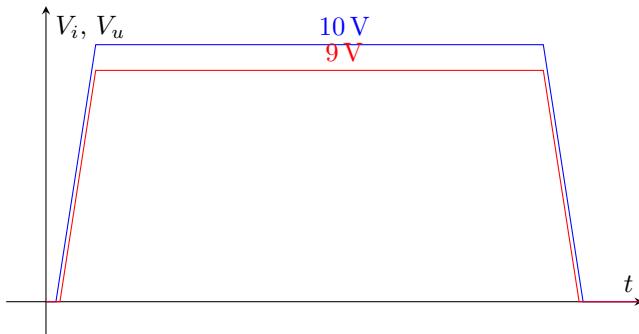


Figure 7.8: Ideal signals in the circuit of fig. 7.7

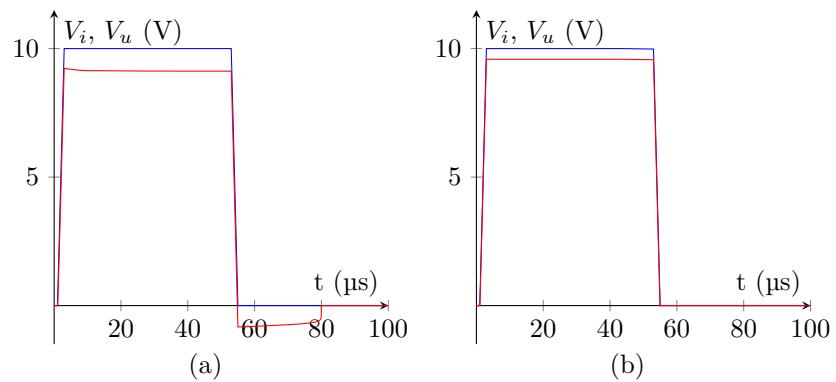


Figure 7.9: Spice simulation of the circuit of fig. 7.7 (a) with silicon power diode 1N4007, (b) with 1N5819 Schottky diode).

This voltage is stored across the diode by the *junction capacitance* which remains charged to V_γ (V_{BE} in the transistor case) when the diode (or transistor) is brought from conduction to cutoff. To bring the diode back into cutoff it is necessary to discharge this capacity making the charge to flow into the opposite direction and the time required depends on the device: a characteristic parameter is t_{RR} (*Reverse Recovery Time*) defined as the time required for the diode to return to the cutoff state in a specific circuit given in the datasheet. An alternative parameter is the Q_{RR} , that is the charge to be removed to obtain the cutoff. For power diodes devices are classified as *standard* recovery, *fast* recovery, *super-fast* recovery ones.

A particular class is that of Schottky diodes made by a metal to semiconductor junction instead of a semiconductor to semiconductor one. They have a low V_γ and a very low junction capacitance but they may stand low maximum reverse voltages (not beyond 100 V, a low limit for a power diode). Using a 1N5819 Shottky diode at the place of the 1N4007 in the Spice simulation of the circuit of figure 7.7 the timing diagram of figure 7.9(b) is obtained.

7.3.2 Dynamic behaviour of BJTs

Bipolar transistors in switching mode have a behaviour similar to diodes and in the transition from saturation to cutoff it is necessary to remove from the base region the charges stored by the direct biasing of the base-emitter and base collector junctions. The more the transistor is in saturation, the larger is the charge to be removed.

There will be a short delay in the transition from cutoff to saturation because a large voltage is available to charge the junction capacitance and much larger delay in the opposite transition because the available voltage is lower. Let us verify the switching behaviour of the test circuit of figure 7.2 by Spice simulation. The transistor used is the 2N2222 device, the power supply is 10 V, the resistances are $R_B = 330 \Omega$ and $R_C = 4.7 \text{ k}\Omega$ and the input signal is a square wave with rise and fall times equal to 100 ns. The simulation results are shown in figure 7.10.

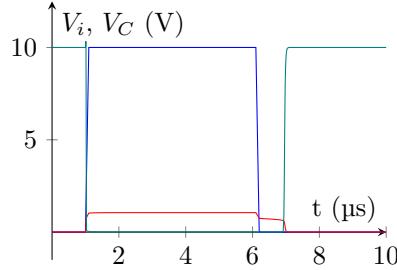


Figure 7.10: Spice simulation of the circuit of figure 7.2. In blue the input, in red V_{BE} , in green V_C .

It is possible to improve the dynamic performances of the switch by adding a capacitor referred to as *speed-up capacitor* in parallel to the base resistance. Figure 7.11 shows both the circuit and a Spice simulation of it with the same parameters of the previous one and a 1 nF capacitor.

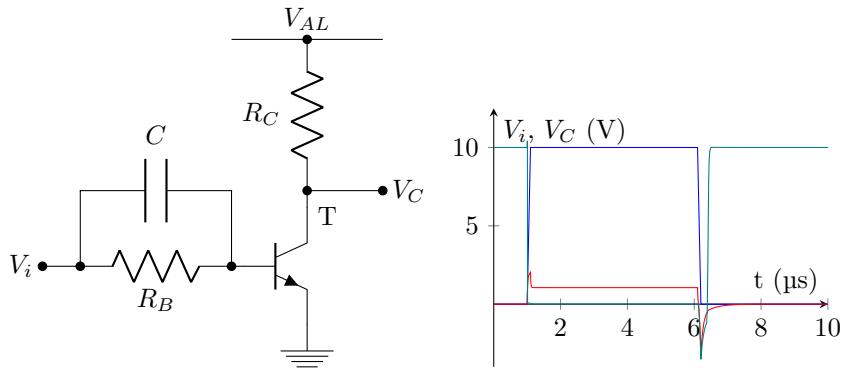


Figure 7.11: Circuit schematic and simulation with a speed-up capacitor

In static conditions the capacitor is charged to a voltage equal to the difference between V_i and the transistor base voltage. During transitions the capacitor keeps instantaneously its charge and then transfers it to the transistor base.

During the rising transition (transistor switch-on) the capacitor transfers the V_i voltage to the transistor base providing a charge pulse which speeds up the transition. In the falling transition (transistor switch-off) the capacitor brings the base voltage to a negative value close to $-V_i$. This forces the removal of the excess charge stored in the base during saturation. It is necessary to properly size the capacitor according to the charge stored into the base because too high capacitance values or too high ΔV_i may damage the transistor.

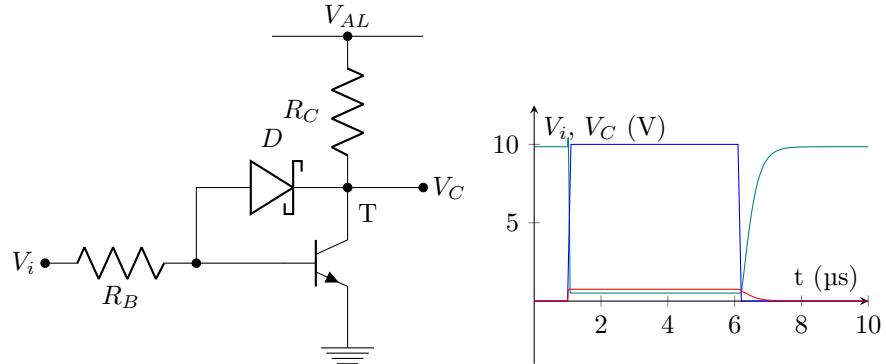


Figure 7.12: Circuit using a Schottky diode between base and collector to control the junction behaviour.

Another way to speedup the switch-off transition uses a Schottky diode, as shown in figure 7.12, in parallel to the base collector junction of the transistor. It avoids the transistor saturation because, having a lower threshold value, it enters into conduction before the base collector junction. Therefore the excess current beyond linearity does not flow through the transistor and there is not a charge accumulation in the base region. This technique is used in bipolar logic circuits, but may not be employed for high supply voltages because Schottky diodes do not allow high reverse voltages.

7.3.3 MOSFET dynamic behaviour

The transient behaviour of field effect transistors is different from that of bipolar ones because there are no junctions which change state and switching delays are basically related to the charge which is necessary to transfer to change the voltages across the parasitic capacitances of the device.

The limiting factor is therefore the capability of the driving circuit to provide current to the transistor gate. As a consequence the switch-on and switch-off transients are substantially symmetric.

MOSFETs have several parasitic capacitances, but the ones to be mainly considered for switching are two: the gate-source capacitance C_{GS} and the gate-drain capacitance C_{GD} outlined in figure 7.13. Datasheet often present other values of capacitances, referred to as C_{iss} , C_{rss} e C_{oss} which are related to the previous ones by the expressions:

$$\begin{aligned} C_{iss} &= C_{GS} + C_{GD} \\ C_{rss} &= C_{GD} \\ C_{oss} &= C_{DS} + C_{GD} \end{aligned}$$

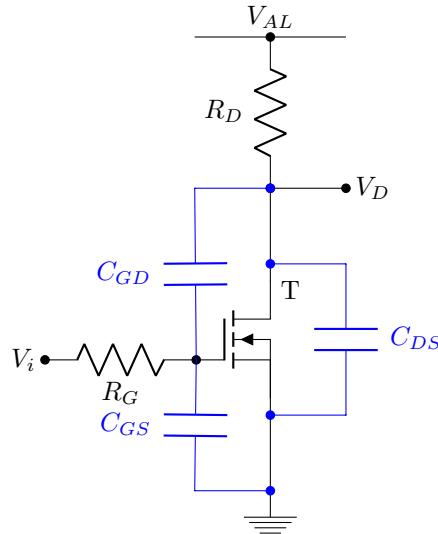


Figure 7.13: MOSFET switch showing the capacitances between each pair of terminals.

To switch the device we must change the gate voltage, that is charge and discharge C_{GS} e C_{GD} capacitors which are not linear, because their value varies with the applied voltage. The charge must be provided by the driver circuit which has in general a limit on the maximum current it may provide. To understand the behaviour during the transition let us analyze the circuit of figure 7.14.

Let us suppose that at the beginning we have $V_{GS} = 0$ V and let us apply to the gate a constant current I_G . Capacitors C_{GS} and C_{GD} may be considered in parallel because drain voltage V_D remains constant and equal to V_{AL} until V_{GS} does not reach the threshold value V_{th} . V_{GS} does increase linearly with slope inversely proportional to $C_{GS} + C_{GD}$.

Once the threshold has been reached any further charge on the gate will increase the device current and decrease the V_D voltage until the switching has been completed. To a small increase of the gate voltage corresponds a large decrease of the output voltage and the device behaves like an inverting amplifier with a rather high gain. Capacitor C_{GD} is between the input and the output and by the Miller effect the input capacitance will become $C_{GS} + C_{GD} \cdot (1 - k)$ where k (negative) is the gain of stage. Since the capacitance has increased, V_{GS} will increase more slowly than before.

Once the transition has been completed and the output is constant, the input capacitance will become again the sum of C_{GS} and C_{GD} , slightly increased due to non linearity effects and V_{GS} will rise faster.

The time behaviour of the gate voltage in the circuit of figure 7.14 depends on the Miller effect (difficult to evaluate) and it is not enough to know the value of the parasitic capacitances to evaluate the switching time. It is more simple to know the amount of charge which is necessary to transfer on the gate to complete the transition. Then knowing the voltage/current characteristics of the driver circuit it is possible to evaluate the switching time.

Figure 7.15 shows the results of a Spice simulation using a IRF510 power

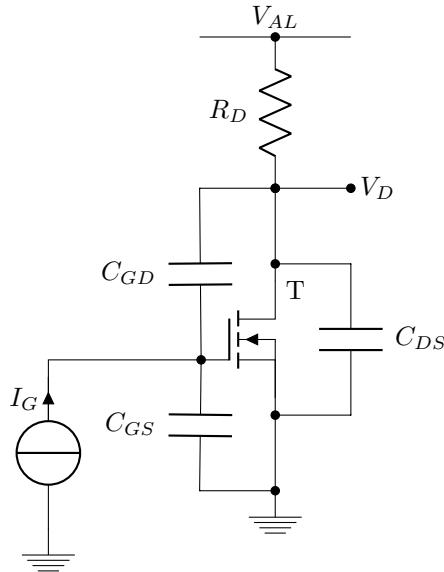


Figure 7.14: Reference circuit for analyzing the MOSFET switching.

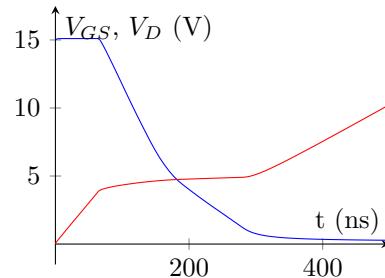


Figure 7.15: Spice simulation of the circuit of figure 7.14, using a IRF510 transistor, $I_G = 11 \text{ mA}$, $R_D = 20 \Omega$, $V_{AL} = 15 \text{ V}$.

MOSFET. It is necessary to provide a charge of 5.5 nC to bring the voltage on the gate from 0 V to 10 V , which is the nominal ON voltage for this device. If this charge is provided by a 11 mA source, as in the simulation, then the transition will take approximately 500 ns . If the driver could provide 1 A current the switching could take place in 5.5 ns .

The manufacturer reports in the datasheet the amount of charge required for the switching, sometimes specifying the amount required to reach the threshold, the amount necessary to cross the nearly horizontal region (Miller effect) and finally the amount required to reach the nominal gate value. In the design phase, knowing the total charge required and having as specification the device switching time it is possible to size the parameters of the driving circuit. As an example if $Q_{tot} = 100 \text{ nC}$, then if we want to turn on the device in a maximum time of 100 ns charging the MOSFET gate at constant current we have:

$$I = \frac{Q}{\Delta t} = \frac{100 \text{ nC}}{100 \text{ ns}} = 1 \text{ A}$$

Consider that this current must be provided by the driver for a very short time. This current pulse must be provided (through the driver) by the power supply subsystem. The connections, by wire or by printed circuit lines from the supply to the driver circuit are not ideal and have parasitic inductances. The voltage drop across these inductance in the presence of current pulses with large $(\delta i / \delta t)$ reduces the value of the voltage at the driver and in other circuit sections sometimes provoking malfunctions. To reduce these problems we may place ceramic capacitors close to the driver circuits which will act as charge tanks, reducing the amplitude of the current peaks in the wires and therefore voltage fluctuations.

7.4 Exercises

Exercise 7.1. Bipolar Switch

The output of voltage comparator which has an output voltage range of ± 9 V and a maximum output current of ± 5 mA is used to drive ON and OFF a motor which has a resistance R_M of $35\ \Omega$ is powered by a 12 V power supply and requires a minimum current I_{Mmin} of 320 mA. Determine if a bipolar transistor with the following parameters may be used to interface the comparator to the load.

Transistor parameters:

$$i_{Cmax} = 1\text{ A} \quad P_{Dmax} = 1.3\text{ W} \quad \beta_{min} = 70$$

Exercise 7.2. Bipolar Switch

Use a bipolar transistor to switch ON and OFF a 10 W lamp with a supply voltage of 40 V. Assume the following parameters for the transistor:

$$\beta_{typ} = 100 \quad \beta_{min} = 60 \quad T = 25^\circ\text{C} \quad V_{CEsat} = 0.3\text{ V} \quad V_{BE} = 0.73\text{ V}$$

The control signal is a digital signal with:

$$V_{OH} = 10\text{ V} \quad V_{OL} = 0.1\text{ V}$$

Determine:

- The schematic of the circuit and the value of the components (including the tolerance).
- The maximum power dissipated by the transistor
- The maximum current which must be provided by the control signal at both logic output values

Exercise 7.3. MOS Switch

The output of voltage comparator which has an output voltage range of ± 9 V and maximum output current of ± 5 mA is used to drive ON and OFF a motor which has a resistance R_M of $35\ \Omega$, is powered by a $V_{DD} = 12$ V power supply and requires a minimum current I_{Mmin} of 300 mA.

1. Determine if a MOS transistor with the following parameters may be used to interface the comparator to the load.

$$i_{Cmax} = 1\text{ A} \quad P_{Dmax} = 1.3\text{ W at } T = 25^\circ\text{C} \quad \beta_n = 50\text{ mA/V}^2 \\ V_T = 4\text{ V}$$

2. If yes, then determine also the power dissipated by the transistor.

Chapter 8

Logic Circuits

Contents

8.1 Logic gates	8-2
8.1.1 Electrical parameters of logic gates	8-2
8.1.2 Logic gate compatibility	8-6
8.1.3 Noise margins	8-7
8.1.4 Fan-out	8-7
8.1.5 Logic families	8-7
8.1.6 Logic gate inputs	8-9
8.1.7 Gate output circuits	8-11
8.2 CMOS gates	8-14
8.2.1 Complementary static logic	8-14
8.2.2 nMOS-like logic	8-20
8.3 Examples	8-21
8.3.1 Fan out computation	8-21
8.3.2 LED driving	8-23
8.3.3 Wired-or	8-24
8.4 Dynamic behaviour of logic gates	8-26
8.4.1 Definitions	8-26
8.4.2 Propagation delay	8-27
8.4.3 Power analysis	8-29
8.5 Sequential circuits	8-31
8.5.1 Bistable element	8-31
8.5.2 S-R latch with NAND gates	8-31
8.5.3 SR latch with NOR gates	8-33
8.5.4 Oscillatory Metastability	8-33
8.5.5 SR Transparent Latch	8-34
8.5.6 D-latch	8-35
8.5.7 Synchronous and asynchronous circuits	8-36
8.5.8 D Flip-flop	8-37
8.5.9 Dynamic behaviour of flip flops	8-38
8.6 Exercises	8-40

THE SWITCHING TRANSISTORS are the basic components of digital circuits and first of all of logic gates. In this course we will analyze more the electrical behavior of the logic gates, both static and dynamic, than the logic functions they implement.

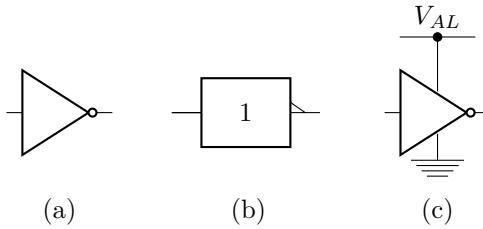


Figure 8.1: Symbols of the inverter: (a) de facto standard (b) IEEE standard (c) showing power supply interconnection.

Our analysis will be therefore more *electronic* than *logic* and in the following we will assume that the reader is familiar with the basic theorems of the boolean logic and the combinational and sequential functions. We will study elementary logic gates and in particular the inverter because results derived from the analysis of this component may be easily extended to complex gates.

8.1 Logic gates

Let us start the analysis of digital circuits from the elementary combinational gates. When considering transistors in switching mode we have already analyzed two circuits which perform the function of an inverter. In fact both the basic circuit used for the analysis of the BJT behavior and the one introduced for the MOSFET generate a low output voltage in correspondence to a high value of the input voltage and a high output voltage when the input voltage is low. Therefore the implemented logic function is the typical NOT function of the inverter.

In the following we will associate the high voltage (often referred to as H) to the logic level 1 and the low voltage level (L) to the logic 0. This association, called *positive logic* is very common but not unique. There are systems where it is more convenient to associate level H to the logic 0 and level L to the logic 1. Such association is referred to as *negative logic*.

It is always possible to represent a logic gate by the corresponding electric schematic which implements the logic function, but with many gates it would be complex to design and analyze the system. We in general prefer to use an higher level representation, abstracting from the circuit implementation.

There are several standards for drawing logic gates, among which one standardized by the IEEE which did not encounter a great success. We will use the most common, although not standardized, design style in which the inverter is represented as in figure 8.1a. Figure 8.1b shows the standardized IEEE symbol, while the symbol of figure 8.1c shows also the power supply terminals of the gate, which in most schematics of logic circuits are omitted, but remain fundamental for the device operation.

8.1.1 Electrical parameters of logic gates

The internal circuit of a logic gate (even of an inverter) is in general more complex than those analyzed in previous sections and in the following we will

study the implementation of a number of logic gates. However it is first necessary to study them as black boxes, analyzing the static input/output parameters of those circuits abstracting from their circuit implementation.

We first need a definition of input and output *high* and *low* voltages more precise than what given before. In particular we need to define the allowable voltage ranges for the input of a gate and what is expected at the output.

Let us use again the inverter as a case study and let us perform an experiment to clarify the problem. Let us power the inverter to its nominal supply voltage V_{AL} and let us put at its input a voltage generator and, as a load, another identical inverter, as shown in figure 8.2a, to emulate the normal working condition where each gate drives one or more gates of the same type. Let us vary the input voltage in the range from 0 V and V_{AL} and plot the output voltage values as a function of the input ones.

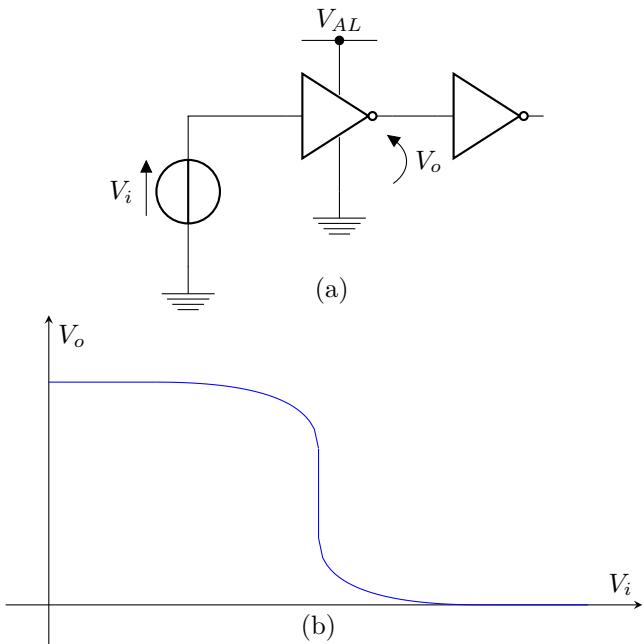


Figure 8.2: a) Inverter test circuit. b) Transcharacteristic.

We will obtain a V_o/V_i characteristic similar to that shown in figure 8.2b which corresponds to an inverter in CMOS technology and that will be analyzed later in detail. For the moment let us consider that there is a wide range of input voltages for which the output is close to the supply voltage and another range for which the output is a voltage close to 0 V. These two ranges are separated by a transition region in which the gate changes state.

If we were repeating the experiment at a different temperature or power supply or with a different gate of the same type we would obtain a similar, but not identical characteristic.

Input and output voltages

Logic circuit manufacturers, taking into account manufacturing process variability, operating temperature range, admissible range of load conditions and the electrical characteristics of gates, define the voltage ranges which are, in all conditions, interpreted as valid ‘High’ or ‘Low’ logic levels.

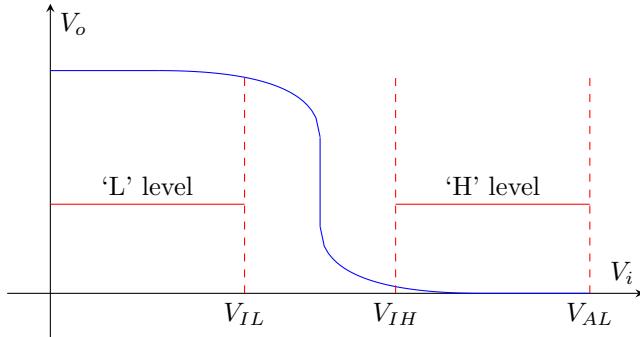


Figure 8.3: Input voltage ranges for high and low levels

Figure 8.3 shows the most significant logic levels.

- V_{IL} is the **maximum** input voltage that the manufacturer guarantees to be interpreted by the device as low (L) logic level. The range of acceptable low logic voltages is from 0 V to V_{IL} .
- V_{IH} is the **minimum** input voltage that the manufacturer guarantees to be interpreted by the device as high (H) logic level.. The range of acceptable high logic voltages is from V_{IH} to V_{AL} .

The above mentioned voltages represent the limits of two ranges. Many textbooks present formal definitions of V_{IL} and V_{IH} as points of the transfer characteristics where the derivative has unit value. These definitions allow us to understand the dependence of the values from technological parameters but do not have a practical use. In real projects we must use the parameters derived statistically by the manufacturer for a range of operating temperatures and a range of load conditions which are clearly specified on the datasheet of each device.

We have defined the input parameters and now we must characterize in the same way the output of the device and the first thing to notice is that it cannot be done without defining the load conditions. It is clear that in case of an output short circuit we may not expect an output voltage different from 0 V even if the output is supposed to be High. We must therefore associate the output voltages in the High and Low states to the currents that the device is capable to provide or absorb. As a consequence, the output voltages are defined as:

- V_{OL} : **maximum** output voltage corresponding to the LOW logic level (the minimum assumed to be 0 V) when the gate output current is lower than a maximum value referred to as I_{OL} ;
- V_{OH} : **minimum** output voltage corresponding a HIGH logic level (maximum assumed to be V_{AL}) when the output current is lower than a maximum value referred to as I_{OH} .

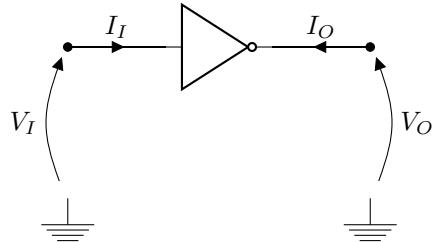


Figure 8.4: Definition of a gate electrical variables.

The output currents depend on the load attached to the gate. Manufacturers provide the output voltage values in load conditions considered to be standard for the circuits and it is not rare to have different values for different current values and this is more common for CMOS gates.

Input and output currents

In the standard case in which a gate drives other ones “compatible” with it, the load of the gate is the sum of the input currents of the driven gates. It is therefore necessary to specify for each logic gate not only the ranges of voltages admissible for the different logic levels, but also the the current absorbed by the gate in each of them.

Input and output voltages are referred to the 0 V reference and for both input and output currents the **passive sign convention** is used: therefore we consider positive a current entering the gate terminal, independently of the fact that it is an input or an output one. The definition and the conventional signs of the electrical variables are shown in figure 8.4.

The definition of the key values of input and output currents are:

- I_{IL} : the maximum current required by the input of the gate when the input voltage is at a value between 0 V and V_{IL} ;
- I_{IH} : the maximum current required by the input of the gate when the input voltage is at a value between V_{IH} and V_{AL} ;
- I_{OL} : the maximum current which may be absorbed by the gate when the output is at the low logic level and the output voltage is guaranteed to be lower than V_{OL} ;
- I_{OH} : the maximum current that the gate may provide to the load at the high logic output level, guaranteeing an output voltage higher than V_{OH} .

The sign of the above mentioned current will depend on the implementation technology. For bipolar technologies we have $I_{IL} < 0$ and $I_{IH} > 0$: for what concerns output currents, current will be sourced at the high state, that is $I_{OH} < 0$ and sunk at the low state, that is $I_{OL} > 0$.

In case of gates implemented with MOSFETs (input on the gate) the input current will be very low, when not considering protection circuits which will be analyzed later. Output currents will also be low except in the case that a gate is driving a resistive load instead of the input of another gate.

8.1.2 Logic gate compatibility

One of the reasons of defining static parameters is to be able to determine whether it is possible to connect the output of a logic gate to the input of another one and if they are compatible, that is if they understand each other correctly. As we have seen, the previously defined electrical parameters do not represent the actual voltage or current at the output/input of a gate, but only limit values and therefore we know only the possible range of the parameters in normal operating conditions. In order to determine whether two gates A and B may be connected we must determine that the logic values provided by A are correctly understood by B and that B is not a too large load for A.

In practice, for what concerns currents, it is necessary to verify that:

$$\begin{cases} |I_{IH,B}| < |I_{OH,A}| \\ |I_{IL,B}| < |I_{OL,A}| \end{cases}$$

For voltages, if both gates are powered by the same supply V_{AL} we must have:

$$\begin{cases} V_{IH,B} < V_{OH,A} \\ V_{IL,B} > V_{OL,A} \end{cases}$$

If the power supplies are different, we have to take into account the limits of the ranges: all possible values at the output of A must be comprised in the range of the admissible values for the input of B.

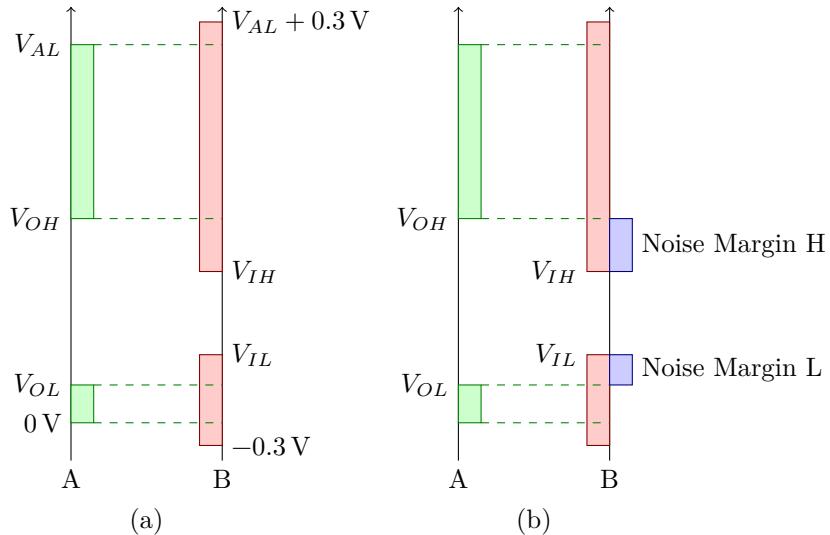


Figure 8.5: a) Pictorial representation of input and output voltage ranges of two interconnected gates; b) Noise margins.

Figure 8.5 shows pictorially the relation between the input and output voltage ranges of two interconnected compatible gates. In general a gate may accept at its input voltages slightly higher or lower than the power supply or reference ones. The connection is possible if the voltage range at the output (green) is fully comprised in the input admissible range (red).

8.1.3 Noise margins

Admissible input voltage ranges are always larger than the output ones because the voltage levels must be compatible also in the presence of electromagnetic noise on the interconnection or on the power supply lines. The differences $V_{IL} - V_{OL}$ and $V_{OH} - V_{IH}$ are referred to as *noise margins* respectively at the high and low logic states. These margins are graphically shown in figure 8.5b.

8.1.4 Fan-out

The output of a logic gate is in general connected to one or more other gates, in general built with the same technology. The *fan-out* of a gate is defined as the maximum number of inputs of other gates built with the same technology which may be connected to the output respecting all the compatibility requirements. Knowing the fanout of a gate we may immediately determine whether a given interconnection is feasible.

To compute the limit fanout when all the connected input are electrically identical and provided that the input and output voltages are compatible, it is sufficient evaluate the parameter:

$$F_O = \min \left(\left| \frac{I_{OH}}{I_{IH}} \right|; \left| \frac{I_{OL}}{I_{IL}} \right| \right)$$

F_O represents the maximum number inputs which may be connected to a given output. If the number of inputs is lower or equal to F_O then the interconnection is feasible, otherwise the schematic must be changed, possibly with the use of a buffer.

8.1.5 Logic families

Logic circuits, as we will see, are implemented using bipolar and field effect transistors cooperating in switching mode. In the design of a logic circuit it is common practice to design separately the circuit section implementing the logic function and the interface circuits to the input and output. In such a way the static input/output parameters will be independent of the logic function of the gate.

A set of logic circuits which implements different logic functions, built with the same technology and with common input-output characteristics are referred to as a *logic family*.

The first available logic families were based on bipolar transistors and were introduced since 1962 (when the first commercially available integrated circuit became available). At a later time logic families based on MOSFETs became available. Every family is characterized by an acronym and the first ones were the ECL (Emitter Coupled Logic), DTL (Diode Transistor Logic) and the RTL (Resistor Transistor Logic). RTL logic was used to implement the first CPU based on integrated circuit (used in the guide computer of the Apollo launch rocket). The most widespread logic family of the '60s and '70s has been the TTL (Transistor-Transistor Logic) family.

The LS (Low power Schottky) version of this family, introduced in 1976, has remained for several tens of years the reference standard for the static electrical characteristics of logic gates. Even now sometimes you can find a reference to

TTL-LS		$V_{AL} = (5.00 \pm 0.25) \text{ V}$		HC		$V_{AL} = 4.5 \text{ V}$	
V_{IL}	0.8 V	I_{IL}	-400.0 μA	V_{IL}	1.35 V	I_{IL}	$\pm 1 \mu\text{A}$
V_{IH}	2.0 V	I_{IH}	20.0 μA	V_{IH}	3.15 V	I_{IH}	$\pm 1 \mu\text{A}$
V_{OL}	0.5 V	I_{OL}	8.0 mA	V_{OL}	0.1 V	I_{OL}	20.0 μA
V_{OH}	2.7 V	I_{OH}	-0.4 mA	V_{OL}	0.33 V	I_{OL}	4.0 mA
				V_{OH}	4.4 V	I_{OH}	-20.0 μA
				V_{OH}	3.84 V	I_{OH}	-4.0 mA

Table 8.1: Main electrical parameters of TTL-LS and HCMOS logic families.

"equivalent TTL-LS" loads or logic levels. The typical power supply value of those bipolar logic was 5 V.

The TTL family comprised a number of integrated circuits whose acronyms were characterized by a common prefix, 74 for commercial and 54 for military ones and by 2 or 3 digit code which defined the logic functions and the circuits pinout. As an example the 7400 circuit contained 4 2-input NAND gates, while the 7404 contained 6 inverters. When the following logic families were introduced, this coding scheme was maintained implementing circuits which were functionally and topologically compatible. The acronym of the new circuits had the same 74 prefix, a few additional letters specifying the new family and the same circuit code of the corresponding TTL circuit. As an example, the 6-inverter circuit in LS technology had the acronym 74LS04 and was *pin-to-pin compatible* with the 7404.

The first logic family which was using CMOS technology was the 4000 series which was not pin-to-pin compatible with the TTL series and was characterized by an extremely low power dissipation, a large flexibility in the power supply (2 V - 18 V) and by very long switching times.

The first CMOS logic family which had dynamic performances comparable to bipolar technologies was the HC (High speed CMOS) family, introduced in 1982. The HC family maintains the functional and topological compatibility with TTL families and components have maintained the same naming scheme with the additional HC suffix (i.e. 74HC04).

During the years the amount of logic functions integrated on the same chip increased exponentially according to the well known Moore's Law and the simple logic functions of a 74 series chip were not required any more: an entire digital system is now integrated in one or a few logic circuits containing several millions of logic gates. However, often to interface complex devices it is necessary to use a few elementary gates (NOT, NAND, NOR), referred to as *glue logic*. There are still available circuits which implement these functions typically with a supply voltage far lower than the historical 5 V.

In the laboratory experiences we will use some "historical" logic gates of the LS and HC series because their parasitic parameters are easily measurable, their speed is relatively low and their physical dimensions make it easy to mount experimental circuits in the lab.

Table 8.1 shows the standard electrical parameters of LS and HC gates. In these families there are special gates such as *bus drivers* which have larger driving capabilities. In the analysis of the table note that the power supply voltage specifier for HC gates is 4.5 V and different from that of LS gates. HC gates

operate from 2 V up to 6 V and in the datasheet the parameters are specified for 2 V, 4.5 V and 6 V. Input currents of HC gates are very small and the datasheet specifies the output parameters in two conditions: when the gate is connected to other HC gates and when the gate is connected to gates which absorb a larger current. For this reason in the table there are two rows, each one for a different value of output current.

8.1.6 Logic gate inputs

Let us analyze in more detail the input characteristic of the gates. The standard input is characterized by parameters V_{IL} e V_{IH} , but for particular applications there is the possibility to have an hysteresis on the switching characteristic. In addition at the input of logic circuits there are always protection circuits independently of the function or the family technology.

Schmitt-trigger Input

When we used the operational amplifier as a comparator we have seen that the introduction on an hysteresis improves the behavior of the circuit in the presence of noise. Now the input of a logic gate is essentially a voltage comparator and the standard input has no hysteresis. If the input signal is affected by noise and has a level close to the gate switching voltage the output will vary randomly as a function of noise.

When the output of a gate is close to the input of the following one, the compatibility of logic level is guaranteed by the noise margins. However if outputs and inputs are physically far away, as in the case of boards connected by a cable, it is possible that the noise added to the digital signal is sufficiently high to induce spurious switching and block the correct operation of the system.

To solve this problem gates whose inputs have hysteresis have been introduced. They have a far better immunity to noise than standard gates but they are in general slower. Inputs with hysteresis are commonly referred to as *Schmitt-trigger* inputs.

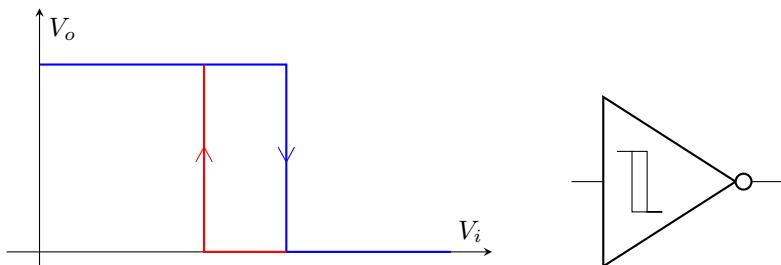


Figure 8.6: Characteristic and symbolic reference of a Schmitt-trigger inverter

Figure 8.6 shows the input-output trans-characteristic of a Schmitt-trigger inverter and the graphic symbol, used to represent it.

These devices are used not only when a large noise is added to the input signal but also when the transition of the input from one level to the other one is *very slow* and the signal remains for a long time close to the switching threshold. In this case even a little noise may generate spurious transitions in

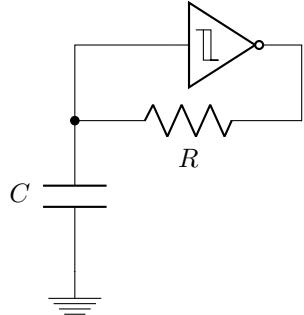


Figure 8.7: Astable multivibrator with Schmitt-trigger input.

normal gates. They are also used to create square wave generators, similar to the op-amp based astable generator analyzed in section 5.2.1. The circuit is shown in figure 8.7. It is left to the reader to compute its oscillation frequency using the gate characteristics listed in the datasheet of 74HC14 device.

Input protection

Inside all input pins of digital (and also many analog) integrated circuits there are protection circuits designed to enter into action when external events may damage permanently the device. The major dangerous event which takes place during the production and storage of the device, until its final mounting and soldering on the printed circuit board, is the creation of electrostatic charges which may accumulate on the device when it is handled and stored.

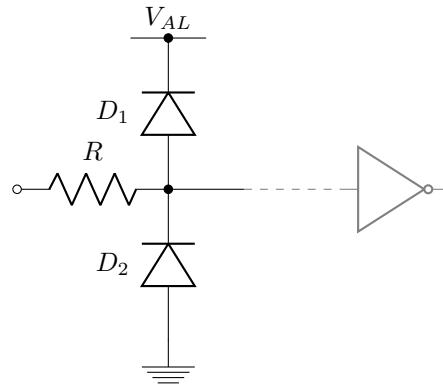


Figure 8.8: Protection circuit at the input of a digital circuit.

The input is connected to a high impedance node on which even a small amount of electrostatic charge may induce a large voltage difference. This is particularly true for CMOS components in which the input is connected to the gate of one or more MOSFETS of small size corresponding to a small capacitor. In addition, given the very limited thickness of the gate oxide, even small voltages may damage permanently the device by oxide breakdown. When the device is mounted and connected in the final circuit, the protection enters into action when the input signal goes over the limits set by the supply voltages.

The most common protection circuit is shown in figure 8.8. Resistance R has a low value in order not to limit the switching speed of the input signal. Diodes D_1 and D_2 are referred to as *clamp diodes* and in normal operation are reverse biased and do not affect (except for little leakage currents) the circuit function. When the input voltage goes either over V_{AL} or under the 0 V reference, then one of the diodes enters into conduction and limits the voltage at the input of the circuit. Before mounting of the component in the circuit, the supply pins are not connected to voltage sources and in case of accumulation of charges on the input pin, which has a small capacitance, the diodes enter into conduction and connect the pin to supply section which has an area (and capacitance) far larger, hence limiting the voltage. If the voltage further increases, then the circuit may be powered up by the parasitic charges and the active devices will enter into conduction, dissipating the charges in the circuit itself.

8.1.7 Gate output circuits

The output of a logic gate may have several different configurations, namely *totem-pole*, *tri-state* and *open-collector/open-drain*. They are characterized by different load driving capabilities and output logic levels. In fact totem-pole outputs may generate only low impedance electric levels H and L, while the other ones allow a third state, named Z, in which the output is in an high impedance state and does not define the logic state of the line to which it is connected.

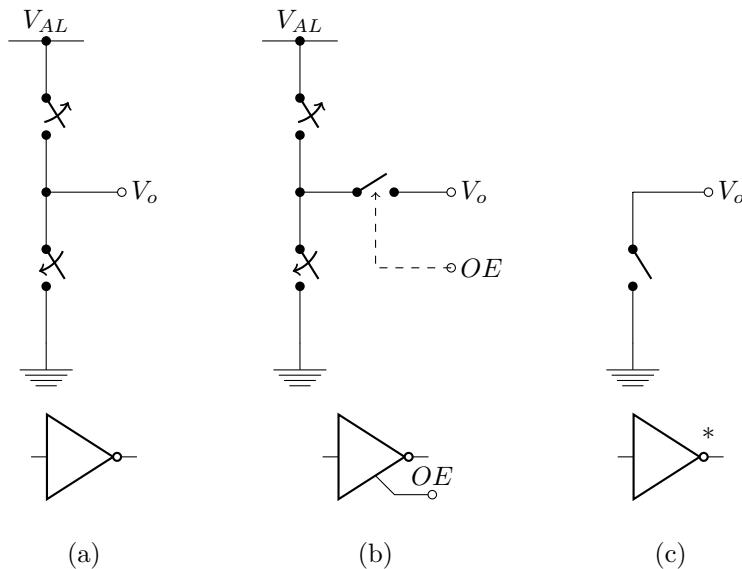


Figure 8.9: Simplified circuit models and symbols of *totem-pole* (a) *tri-state* (b) and *open-collector/open-drain* (c) outputs.

The simplified circuit model and the associated symbol for each configuration is shown in figure 8.9.

IN	OE	OUT
L	L	Z
H	L	Z
L	H	H
H	H	L

Table 8.2: Truth Table of an inverter with tri-state output.

Totem-pole Output

The typical gate output is in general totem pole and according to the gate state generates either a H or L electrical level. The simplest model of this type of output is a series of two switches, operating in counter-phase, one connected between the output and 0V and the other one between the output and V_{AL} . The name derives from the fact that with early TTL circuits the schematic of the output circuit looked somehow like a totem pole of native Americans.

Tri-State Output

When a gate must drive the line of a bus connection on which multiple “talkers” are present, it is necessary that only one among the many possible outputs connected in parallel is enabled at any time. If not, anomalous electrical condition could arise when multiple outputs would try to force *different* logic states on the line. In this case we use gates with a special output referred to with different names, tri-state, 3-state or three-state according to different manufacturers.

This type of output may be enabled or disabled according to the state of an additional logic signal called *Output Enable*, or *OE* (figure 8.9b). If *OE* is active then the gate operates as a totem-pole one, if not the gate is disabled and the output is a very large impedance, close to an open circuit. This additional logic state is referred to as Z and the gate will have 3 possible states, namely L, H and Z, and table 8.2 shows the truth table of an inverter with tri-state output. Conceptually it is as if *OE* would act on a switch in series to the output of a totem pole gate. However actual circuit implementations may be very different.

It should be noted that in many circuit implementations the *OE* signal is active at the low state (\overline{OE}), that is the gate is active when *OE* is Low and in high impedance when *OE* is High.

Open-collector (open-drain) output

A system based on tri-state gates requires a supervising unit which enables each gate at the proper time to avoid conflicts. There are however many cases when multiple gates require to set a line to a given state, independently of each other. This is the case of the interrupt line of a microprocessor, shared by many peripherals or of the control signal of a PCI bus, where in the initialization phase the number of units (boards) which are present on a bus is not known. In this case we use gates with a particular output configuration, referred to as *Open Collector*, if the technology is bipolar, or *Open Drain* if the technology is MOS based.

Open collector gates may force on the output only one logic value, typically the low level L or being in the output impedance state Z. If we compare one of them with a totem pole output, we see that there is only the switch towards the 0 V, while the connection to V_{AL} is missing.

In the schematics this gates are in general marked with a symbol (dot, asterisk, rhombus) placed close to the output pin, as shown in figure 8.9c.

The fact that the gate may impose only the low logic level or being in high impedance implies that the high logic level on the output line must be provided by an external component. This component is in most cases a simple resistor, named *pull-up resistor* connected between the output line and the power supply. When the gate is in the Z state drawing little current, the drop over the resistance is 0 and the output line will be at the supply voltage which is H for the following inputs. When the gate forces the output at L, it will absorb the current flowing in the resistance and, if it is lower than the maximum value allowed for the gate, it will guarantee a correct L level. Some open collector gates may withstand output voltages far larger than their supply voltages and may be used to interface logic gates operating at low voltages with others operating at high voltages. As an example, programmable logic controllers used in industrial environments use input operating at 12 V or 24 V are not directly compatible with standard logic gates at 5 V or 3.3 V .

A very common use of these gates is in the configuration named *wired-OR* shown in figure 8.10 where several open collector gates drive the same output line.

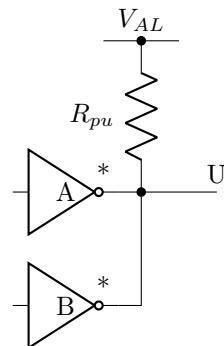


Figure 8.10: *Wired-OR connection.*

With reference to the schematic shown in the figure, if the outputs of gates A and B are both in the Z state then the electrical level of the output line will be H, due to the connection of resistor R_{pu} to V_{AL} . If at least the output of one of the gates is low, then the output will be low. In practice, the connection of two or more open collector outputs to the same output line builds a cabled logic function whose truth table is shown in table 8.3.

This configuration is referred to as *wired-OR*. Using the standard convention which associates the electrical level H to to logic 1 the function should be considered a AND. However for open collector gates which may force on the L state, this state is considered the active one historically associated to logic 1.

A	B	U
Z	Z	H
L	Z	L
Z	L	L
L	L	L

Table 8.3: Truth table of the wired-Or connection of figure 8.10.

8.2 CMOS gates

There are several methodologies for implementing logic gates in CMOS technology. Standard ones use the *Complementary static logic* implementation in which every function requires a set of pMOS transistors, referred as pull-up, which provide to the output the high logic level and a set of nMOS transistors, named pull-down, responsible for providing the low logic level. There are alternative methodologies, named *nMOS-like logic*, *dynamic logic* and *transmission gate logic*.

8.2.1 Complementary static logic

We will first analyze the simplest gate, the inverter, implemented with this methodology and we will then extend it to the realization of more complex gates.

Inverter

The design of an inverter in CMOS complementary static logic is very simple and easily understood. Let us assume the gate to have a totem pole structure and therefore to have two switches which connect the output respectively to the power supply and to 0 V as shown in figure 8.9a.

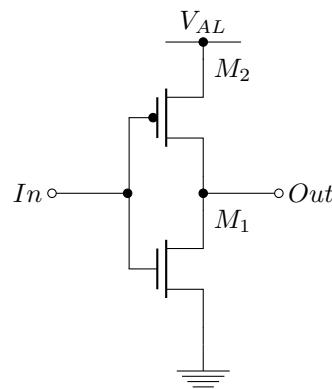


Figure 8.11: Circuit schematic of a CMOS inverter

The two switches must never be closed at the same time. The switch which connects the output to the 0 V must be active when the input is high and it is clear that a nMOS transistor with the gate connected to the input and with the

drain and source respectively connected to the output and to the 0 V implements the required function. A second switch is required to connect V_{AL} and the output when the input is low. It is easy to verify that it may be implemented by a pMOS transistor with the gate connected to the input, the source to V_{AL} and the drain to the output. The full gate is shown in figure 8.11.

Let us analyze in more detail its operation. The input is connected to transistors M_1 and M_2 and the circuit driving the gate has a purely capacitive load and must provide current only during switching transients. If all the logic system is composed of such gates then the power consumption in static conditions will be very low and current will be absorbed only at input and output state changes. However if the gates have a resistive load we will have current flowing in the MOSFETS also in static conditions.

This general statement, which will be further detailed in the chapter, is fully true for older technologies. Modern ones, which use MOS device whose size is far below the micron, require more device models different from the classical one and also the behaviour of the gate is rather different from the ideal one.

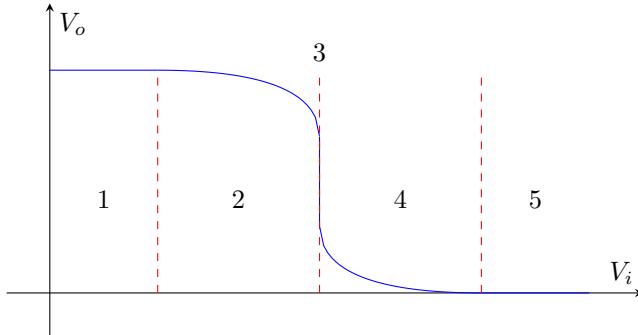


Figure 8.12: CMOS inverter transfer characteristic.

Let us now analyze the static trans-characteristic of the gate in order to study its dependencies on the devices' dimensions. At the input we have a d.c voltage V_i and lets evaluate the gate behaviour when the input has value ranging from 0 V e V_{AL} assuming that the inverter load is infinite or capacitive, as another inverter of the same type. The gate trans-characteristic is shown in figure 8.12.

There are 5 different operating zones:

1. When V_i is between 0 V and the threshold voltage V_{Tn} of M_1 , then M_1 is in cutoff ($V_{GS} < V_{Tn}$) while M_2 is in the resistive area. This because the V_{GS} of M_2 is strongly negative and the transistor is in conduction but the current in both transistors is close to 0 and the V_{DS} voltage of M_2 is very small.
2. When V_i is between V_{Tn} and a limit value referred to as V_{inv} , M_1 is in saturation because $V_{GS} - V_{Tn}$ is small while its V_{DS} is large, M_2 being in strong conduction. Increasing V_i M_2 moves from the resistive zone to the quadratic (triode) since the term $V_{DS}^2/2$ begins to be not negligible.
3. When $V_i = V_{inv}$ we are in the so called inversion zone which is very narrow and, if we use for MOSFETs a model neglecting the channel modulation factor, reduced to a single point in the input axis. In reality the width

of the zone depends on the transistor parameters but in any case a small change of the input voltage corresponds to a large change in the output voltage and in this zone the inverter behaves like an inverting amplifier with a relatively high gain. In the inversion zone both transistors M_1 and M_2 operate in saturation.

4. When V_i is between V_{inv} and $V_{AL} + V_{Tp}$ M_1 is in the linear zone and M_2 in saturation and we have a behaviour symmetric to the one of zone 2.
5. When V_i is over $V_{AL} + V_{Tp}$, then M_2 is in cutoff and M_1 in linear operation and the behaviour is symmetric to the one of zone 1.

From this analysis it is clear that the behaviour of the gate mainly depends on the positioning of the inversion zone (3 in figure 8.12). As we have seen in this zone the output rapidly switches from a high value, close to V_{AL} , to a low one, close to 0 V. Moving the position of the inversion area we also change the V_{IL} and V_{IH} values of the gate.

We have already mentioned that in zone 3 both transistors operate in saturation, that is behave as current generators. Since there are no resistive loads and the two drains are in series to each other, this condition can exist only if the two transistor currents are identical.

We may then assume that the I_D of M_1 and M_2 are equal obtaining:

$$\mu_n C_{OX} \frac{W_n}{L_n} (V_i - V_{Tn})^2 = \mu_p C_{OX} \frac{W_p}{L_p} (V_i - V_{AL} - V_{Tp})^2$$

This relation is valid for a unique value of the input voltage that we will refer to as V_{inv} or inversion voltage. Considering also the λ parameters would lead to a slightly more complex relation which would admit a solution for a narrow range of input values, in practice corresponding to the results of the simplified analysis which is precise enough for our purposes.

Note that the value of the inversion voltage depends on technological parameters such as μ_n and μ_p , V_{Tn} and V_{Tp} , and on parameters which may be set by the designer such as the transistor widths, W_n and W_p and lengths L_n and L_p , and on the value of the power supply. This means that acting on geometrical parameters it is possible to modify the inversion voltage value and therefore to design V_{inv} according to the requirements of the circuits driving the gate itself. A particularly interesting result may be obtained by setting:

$$\mu_n C_{OX} \frac{W_n}{L_n} = \mu_p C_{OX} \frac{W_p}{L_p}$$

In this case the expression simplifies to:

$$(V_{inv} - V_{Tn})^2 = (V_{inv} - V_{AL} - V_{Tp})^2$$

Taking the square root of both members, with the proper sign, we have:

$$V_{inv} - V_{Tn} = V_{AL} - V_{inv} + V_{Tp} \implies 2V_{inv} = V_{AL} + V_{Tn} + V_{Tp}$$

In general the fabrication process is set to have $V_{Tn} = -V_{Tp}$ and in this case we have:

$$V_{inv} = \frac{V_{AL}}{2}$$

This result is important because the positioning of the inversion voltage midway between 0 V and the power supply is an optimal solution when it is not necessary to interface to particular gate outputs. In fact this allows to maximize the noise margins because the high level V_{OH} is close to V_{AL} and the low one V_{OL} to 0 V. Positioning V_{inv} at midway between these two values means also to place V_{IL} and V_{IH} (which include also process tolerances) in positions symmetric to V_{inv} and as far as possible respectively from V_{OH} and V_{OL} . In addition if the voltage supply changes, then the inversion voltage and input and output logic values will automatically adapt. Please note that the condition imposed by:

$$\mu_n \frac{W_n}{L_n} = \mu_p \frac{W_p}{L_p}$$

allows additional degrees of freedom which may be exploited to design the driving capability of the gate. If it is internal to an integrated circuit and has to drive a few small gates then we will use the minimum transistor sizes allowed by the fabrication process. If, on the contrary, the gate has to drive an external large load, then the transistors will be designed to have a low output resistance and a large current capability with larger sizes, and larger semiconductor area.

NAND and NOR gates

We have seen how an inverter is implemented in complementary static logic. However to synthesize a general logic function we need to build at least a NAND or a NOR function.

In the inverter design the NMOS transistor connected between the output and the 0 V is used to provide a low level to the output and the PMOS is active when the output must be high by connecting it to the power supply. It is possible to extend this design methodology considering that for any logic function we need a NMOS transistor **network** which will be active (conducting) when the function must provide a low output and a PMOS transistor **network** which is active when the output value must be high.

Let us build such a network for a simple NOR gate with two inputs A and B. The gate output must be low if ANY of the input is high and this is simply obtained by connecting two NMOS transistors in parallel between the output and 0 V with the gates respectively connected to A and B. The output must be high if and only if BOTH inputs are low and therefore the PMOS network must create a low impedance path between the output and V_{AL} only in this condition. This may be obtained by placing two PMOS in series between the output and V_{AL} with the gates respectively connected to A and B. Figure 8.13 shows the resulting schematic.

The design of a NAND gate is similar but the roles of the NMOS and PMOS networks are reversed. In fact, being necessary to have a low output value when BOTH inputs are high, the two NMOS transistors will be in series between the output and the 0 V. The two PMOS transistors will be in parallel between the output and the power supply because a high output value is required when just one input is at the low state. The final schematic for the NAND gate is shown in figure 8.14.

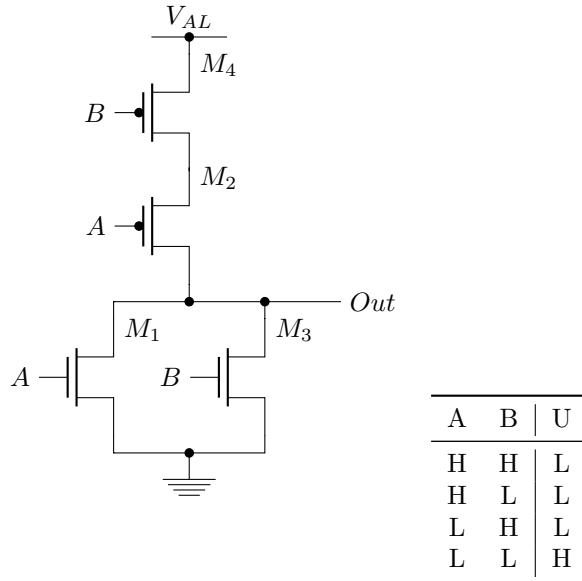


Figure 8.13: Schematic of a two input NOR gate.

The analysis of the electrical characteristics of these gates and the sizing of the transistors is beyond the purpose of this basic course and is left to following courses on Microelectronics.

The extension of the design methodology to NAND or NOR gates with more than 2 inputs is straightforward. For each additional input it is sufficient to insert a transistor in series or in parallel in the PMOS or NMOS network. For instance a 3-input NOR gate will have three NMOS in parallel and three PMOS transistors in series.

Complex logic gates

We have analyzed how to design the inverter and NAND and NOR gates. All these gates are inverting and this is a general characteristic because a non inverting gate may not be directly implemented in complementary static logic. An AND gate has to be implemented by a NAND followed by an inverter or by a NOR with two complemented inputs. This is intrinsic to the fact that NMOS transistor with a high input will conduct connecting the output to the reference voltage and therefore implementing an inverting function. The same applies, with reversed logic signal to the PMOS transistor.

There are more complex logic functions which may be synthesized directly as complementary static logic gates. Their equation must be expressible in a form referred to as *AOI*, or And-Or-Invert, that is the output must be expressed as the complement of a sequence of AND and OR functions of the non complemented inputs as:

$$U = \overline{(A + B) \cdot (C + D)}$$

NAND and NOR gates are simple examples of this category of functions.

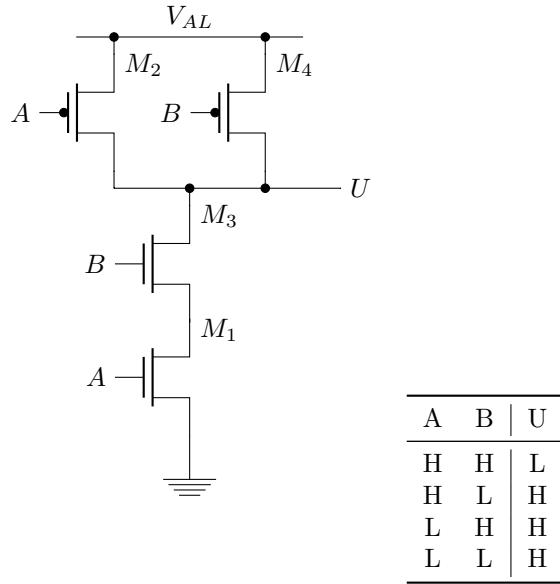


Figure 8.14: Schematic of a 2-input NAND gate.

For the synthesis of an AOI gate it is sufficient to follow a simple methodology derived from the design of NAND and NOR gates. We have to design two networks, driven by the input signals, one of NMOS transistors to generate a low output voltage and one of PMOS transistors to provide the high one by connecting the output to the supply. The two networks are operating in dual mode in the sense that for every input configuration one and one only network must be active, that is connecting the output to the corresponding voltage (0V or V_{AL}).

For the NMOS network the equation to be implemented is analyzed and for each input a transistor is added. Transistors are connected:

- in series, when a logic product (AND) is present;
- in parallel, when a logic sum (OR) is present.

One end of the designed network, referred to as *pull-down network*, must be connected to the output and one to the reference voltage 0V .

The PMOS transistor network, named *pull-up network*, is designed in a dual way: for each input we insert a transistor and each transistor is connected :

- in parallel, when a logic product (AND) is present;
- in series, when a logic sum (OR) is present.

One end of the designed network must be connected to the output and one to the supply voltage V_{AL} .

The circuit implementing the example equation will be then composed of four NMOS and four PMOS transistors, having four input signals. The two NMOS transistors with input A and B will be connected in parallel, being A and B in OR, as well as the NMOS transistors having as input C and D. The

two pairs must then be connected in series, since there is an AND function which connects the two ORs.

The PMOS network will be dual, with the PMOS transistors connected to A and B in series, as well as those connected to C and D. The two branches will be then connected in parallel. The final circuit is shown in figure 8.15.

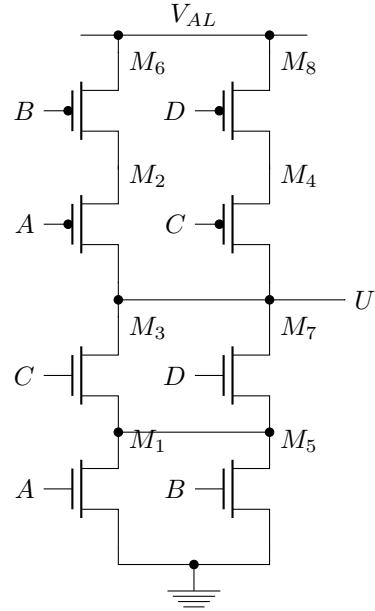


Figure 8.15: AOI implementation of $U = \overline{(A+B)} \cdot \overline{(C+D)}$.

The AOI (And Or Invert) methodology allows to implement any inverting function based on the combination of AND and OR functions. It is a powerful methodology but in general the number of inputs is never very large because the number of transistors in each networks beyond a certain limit increases the gate delay making it preferable to choose a multilevel logic solution.

8.2.2 nMOS-like logic

The main disadvantage of the complementary logic is that it requires two dual logic networks which provide, separately, the output high and low logic values. This means that it requires a silicon area double than the minimum one which would be necessary to define the logic function and that each input will have an equivalent capacitance equal to that of two transistors in parallel.

To reduce the required area we may think to implement only one of the two complementary networks which provide a logic output value and in some way guarantee the other logic value at the output when the implemented logic network is not active and conducting.

This methodology was used at the starting times of MOS technology, when it was not possible to integrate on the same chip PMOS and NMOS transistors. Typically only NMOS devices were used to implement the pull-down network connecting the output to 0 V and a pull-up resistance was added, as in open drain logic to guarantee the high logic level when the network was not active. The same

technique may be used also in CMOS circuits, where the resistance is substituted by a pMOS transistor with the gate stuck at 0 V, properly sized in order to guarantee a correct logic level both when the pull-down network is conducting and when it is not. As an example the circuit of figure 8.16 implements in nMOS-like logic the same function as the circuit of figure 8.15.

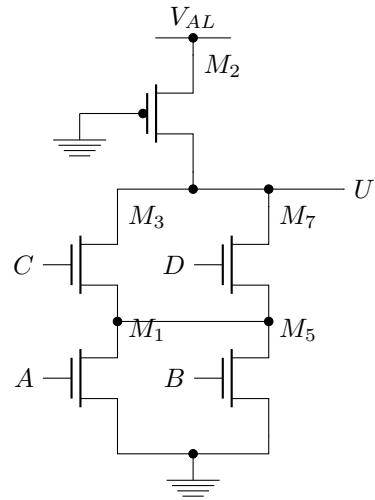


Figure 8.16: Implementation of $U = \overline{(A + B) \cdot (C + D)}$ in nMOS-like logic.

This implementation is more compact than the one in complementary logic but has an important drawback which heavily limits its use: when the output is at the logic low state currents flows from V_{AL} to 0 V through the two transistors and therefore there is **static** power dissipation that is even when the output does not change. On the contrary in static conditions no current flows in complementary logic and power dissipation is very low.

8.3 Examples

In the following paragraphs we will analyze examples of interconnection and load driving using logic gates from different technologies and with numeric value of parameters which show the typical values of relevant electric variables.

8.3.1 Fan out computation

In the schematic of figure 8.17, $R_{PD} = 2\text{ k}\Omega$ and $R_{PU} = 10\text{ k}\Omega$ represent a NON standard load connected to the output of gate 1. Power supply V_{AL} is 5 V. The output characteristics of gate 1 and the input ones of gates L are give in table a 8.4. Which is the maximum number of gates L which is possible to connect to the output of gate 1?

Solution

Let us first observe that the voltage levels of gates L are compatible with the ones of gate 1. In fact we have $V_{OL_1} < V_{IL_L}$ and $V_{OH_1} > V_{IH_L}$. Its is then

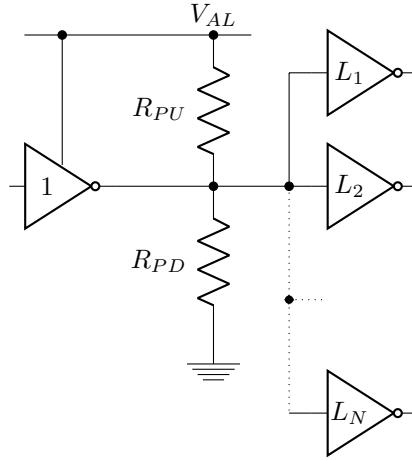


Figure 8.17: Circuit schematic of example 8.3.1.

necessary to how many gates gate 1 may drive, in addition to the non standard load, by computing the output current at both the high and to low logic level.

High output level When the output of gate 1 is at the high level at the output node we have two input currents, that is the output current of gate 1 and the current that flows in the resistance R_{PU} , and two output currents, the one that flows in R_{PD} and the sum of the input currents of the L gates. The output voltage is guaranteed to be higher than V_{OH_1} and lower than V_{AL} if we have $|I_{O_1}| < |I_{OH_1}|$.

Considering the output voltage to be equal to V_{OH_1} the current in R_{PU} can be considered as an additional contribution to the gate output current I_{O_1} while the current in R_{PD} is an additional load. I_{O_1} may be expressed as:

$$I_{O_1} = n \cdot I_{IH_L} + \frac{V_{OH_1}}{R_{PD}} - \frac{V_{AL} - V_{OH_1}}{R_{PU}} \leq |I_{OH_1}|$$

$$\Rightarrow I_O = n \cdot 20 \mu\text{A} + \frac{4.5 \text{ V}}{2 \text{ k}\Omega} - \frac{0.5 \text{ V}}{10 \text{ k}\Omega} \leq 5 \text{ mA} \Rightarrow n \leq 140$$

Low output level In this case gate 1 sinks current up to a maximum value I_{OL} equal to 5 mA. R_{PU} becomes an additional load whose current must be absorbed in addition to the one coming from the L gates input and R_{PD} helps the driving by sinking part of that current. The output voltage being V_{OL} we have:

Gate 1				Gate L			
V_{OL}	0.5 V	I_{OL}	5.0 mA	V_{IL}	1.0 V	I_{IL}	-0.4 mA
V_{OH}	4.5 V	I_{OH}	-5.0 mA	V_{IH}	3.0 V	I_{IH}	20 μA

Table 8.4: Electrical parameters of gates of type 1 and L .

$$I_O = n \cdot |I_{IL}| + \frac{V_{AL} - V_O}{R_{PU}} - \frac{V_O}{R_{PD}} \leq I_{OL}$$

$$I_O = n \cdot 0.4 \text{ mA} + \frac{4.5 \text{ V}}{10 \text{ k}\Omega} - \frac{0.5 \text{ V}}{2 \text{ k}\Omega} \leq 5 \text{ mA} \implies n \leq 12$$

Therefore the fanout is the lower of the two evaluated limits and the maximum number of L gates which may be connected is 12.

8.3.2 LED driving

We want to drive a LED diode with a TTL-LS inverter assuming the following specifications:

- Supply voltage: $V_{AL} = 5 \text{ V}$.
- LED parameters: $V_{ON} = 1.7 \text{ V}$; current required to obtain a bright diode: $I_{ON} = 5 \text{ mA}$.
- Gate output parameters: $V_{OH} = 2.7 \text{ V}$; $V_{OL} = 0.5 \text{ V}$; $I_{OH} = -400 \mu\text{A}$; $I_{OL} = 8 \text{ mA}$.

Solution

LEDs are built with semiconductors different from silicon (GaAs, GaP, GaAsP, SiC, GaInP, ...). The voltage current characteristic of a LED is similar to a silicon diode except for the fact that the threshold voltage is larger and depends on the substrate material and doping which determine the colour of the emitted light. The current necessary to obtain a given brightness depends also on the diode type. The given parameters refer to a typical red LED used to show the state of digital signals.

It is not possible to directly connect the LED directly between the output of a logic circuit and 0 V or V_{AL} because it would not be possible to control the current through it and it is necessary to insert in series a resistance with a proper value. Two possible circuits, with the LED connected either to 0 V or V_{AL} are shown in figure 8.18.

If the driving capability of the gate is symmetric (that is if $|I_{OH}| \simeq |I_{OL}|$), then both circuits may be used, provided that the current required by the LED is lower than the maximum output gate current. If not, then we must determine which circuit may be used and we have to analyze both of them with the output both at the high and the low state.

1. With the first topology (fig. 8.18.a) there are no problems at the low level: $V_{OL} = 0.5 \text{ V}$ and the LED is in this case off. However at the high logic value the maximum current that the gate may provide is $I_{OH} = -0.4 \text{ mA}$ and is not enough to turn on the LED.
2. Using the second solution (fig. 8.18.b) at the low level current may flow in the LED because it is absorbed by the gate which has a maximum voltage V_{OL} . The maximum value of the output current is compatible with the current I_{ON} which must flow in the LED. The voltage across the LED is its threshold voltage V_{ON} and the drop across the resistance will be

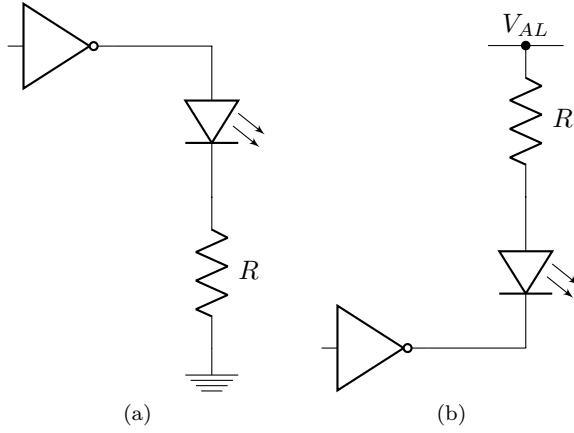


Figure 8.18: Two possible connections of the LED diode.

$V_{AL} - V_{ON} - V_{OL}$. It is therefore sufficient to size the resistance in such a way that with that voltage the current will be equal to I_{ON} that is:

$$R = \frac{V_{AL} - V_{ON} - V_{OL}}{I_{ON}} = 560 \Omega$$

What happens at the high level? At a first glance it would appear as if current would flow in the LED since $V_{OH} = 2.7\text{ V}$ and the drop across the diode and the resistance is larger than the threshold V_{ON} . In reality this does not happen because the gate may only source and not sink current and the output voltage will rise to a value which will not allow the flow of current in R .

The asymmetric driving capability of TTL gate, analyzed in this exercise, forced the designer to drive loads at the low output voltage level. Current technologies, which have symmetric driving capabilities at the high and low output states allow to drive loads referred both to 0 V or to V_{AL} .

8.3.3 Wired-or

We want to size the pull-up resistance R_{PU} of the circuit shown in schematic 8.19. The symbol * on the logic gate specifies that their output is of the *open collector* type. The electrical characteristic of the gates are:

- LS05 open collector gates:

$$V_{OH} \leq 5.5\text{ V}; I_{OH} = 100 \mu\text{A}; V_{OL} = 0.5\text{ V}; I_{OL} = 8 \text{ mA};$$

- HC CMOS gates:

$$V_{IH} = 3.15\text{ V}; I_{IH} = 1 \mu\text{A}; V_{IL} = 1.35\text{ V}; I_{IL} = -1 \mu\text{A};$$

- Standard TTL LS gates:

$$V_{IH} = 2\text{ V}; I_{IH} = 20 \mu\text{A}; V_{IL} = 0.8\text{ V}; I_{IL} = -400 \mu\text{A};$$

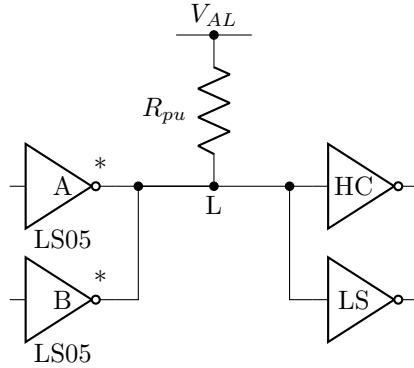


Figure 8.19: Schematic of example 8.3.3.

Solution

The output is at the high logic level when both A and B have their output in the high impedance state and is at the low logic level in all other cases. To correctly size R_{PU} we analyze the behaviour of the circuit separately for the high and low states of the output.

High output level Looking at the datasheet of LS05 gates we may notice that a V_{OH} and a I_{OH} parameter are defined and since at the high state the gate is ideally an open circuit at high impedance state it is necessary to understand the meaning of these parameters. V_{OH} is defined as the maximum voltage which may be applied to the gate output which in the high state, that is the maximum voltage to which the pull-up resistance may be connected. There are open collector gates for which this parameter has a far higher value than the gate power supply. I_{OH} represents the maximum leakage current that the gate sinks when the output is in the high impedance high state and this current is the dominant component in R_{PU} when the output is high.

In fact the current that flows in the pullup resistance in the high state is::

$$I_{PU} = 2I_{OH} + I_{IH,HC} + I_{IH,LS} = 0.221 \text{ mA}$$

And the output voltage will be:

$$V_L = V_{AL} - V_{PU} = V_{AL} - R_{PU}I_{PU}$$

Voltage V_L must be higher or equal to the maximum high input logic voltage of HC and LS gates, that is:

$$V_L \geq \max(V_{IH,HC} = 3.15 \text{ V}, V_{IH,LS} = 2 \text{ V})$$

$$R_{PU} \leq \frac{V_{AL} - V_{IH,HC}}{I_{PU}} = \frac{5 \text{ V} - 3.15 \text{ V}}{221 \mu\text{A}} \simeq 8.2 \text{ k}\Omega$$

It is generally preferable to have a minimum high output voltage slightly larger than 3.15 V in order to guarantee some noise margin. Assuming a value of 3.7 V then the maximum value of the pullup resistance becomes 5.6 kΩ.

Low output level The output the system may be at the low level for three different sets of values of the outputs of namely: $A = L, B = L$; $A = Z, B = L$; $A = L, B = Z$. In the first set both gates are active and may sink current, while in the other sets only one gate is active and has to sink the current from the pull-up resistance and the inputs of the drive gates. These are the worst case conditions for which we will size the circuit and we will not consider the I_{OH} of the not active gate, first because it is a parasitic component and secondly because it would increase the sinking capability of active gate and not considering it leads to a worst case . The equation at node L is then:

$$I_O = \frac{V_{AL} - V_O}{R_{PU}} + I_{IL,HC} + I_{IL,LS} \leq I_{OL}$$

The maximum voltage V_O at the output is V_{OL} , and therefore:

$$\frac{5\text{ V} - 0.5\text{ V}}{R_{PU}} + 0.4\text{ mA} \leq 8\text{ mA}$$

$$R_{PU} \geq 592\Omega$$

The minimum value of R_{PU} in the E12 series is 680Ω .

Final considerations The analysis of the circuit has defined a allowable range of R_{PU} values of approximately a decade. The optimal design choice depend on a number of additional parameters, among which the most important are:

- **Power Dissipation.** A low value resistance corresponds to a higher current in the pull-up resistance when the output is at the low state and this increase the power dissipation of the circuit. (At the high level the current depends lightly in the resistance value).
- **Switching Speed.** The High to Low transition of the output is due to the activation of one of the gates and the current available to discharge output parasitic capacitances does not depend on the resistance value. However in the Low to High transition the capacitances are charged through R_{PU} and a low resistance decreases the RC time constant of the circuit leading to faster High to low transitions.

According to the system requirements and to the characteristics of the logic signals, that is of the *environment* in which the circuit is inserted, the optimal pullup resistance value is chosen.

8.4 Dynamic behaviour of logic gates

8.4.1 Definitions

Until now we have analyzed the static behaviour of logic gates and now let us analyze their behaviour when switching from one logic level to the other one.

Logic gates, being real circuits cannot react instantaneously to a change of the input signals and the change in the output cannot take place in zero time. It is important to define clearly how to measured the delay of signals and the

switching times. The time required for a logic signal to switch from the high to the low level is called *fall time* and is defined as the time required to the signal to go from the 90% of the logic swing to its 10% as shown in figure 8.20a. In a similar way is defined as *rise time* the time required to switch from the low to the high level.

The time required to propagate the signal inside the logic circuit to output is called *propagation delay* and is measured from the instant in which the input signal crosses the midpoint of the input swing and the time at which the output crosses the same value as shown in figure 8.20b. The propagation delay may be different for a high to low output transition (t_{pdHL}) and low to high one (t_{pdLH}).

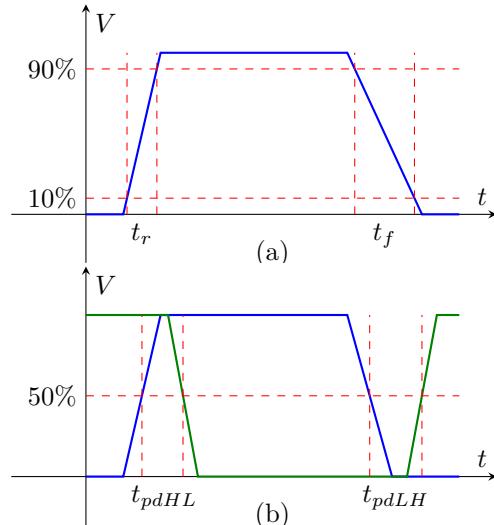


Figure 8.20: a) Rise and fall times; b) inverter propagation delays

8.4.2 Propagation delay

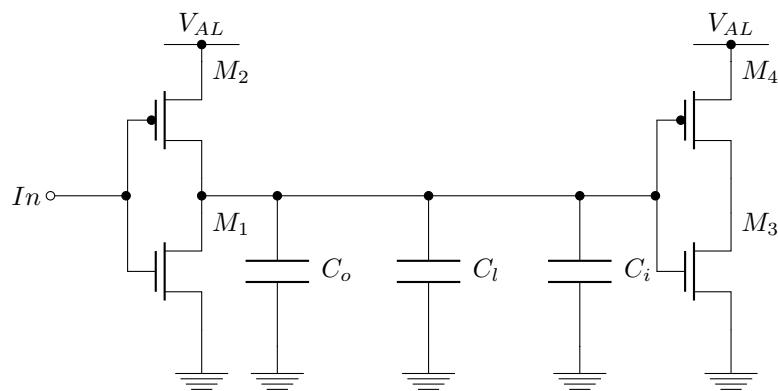


Figure 8.21: Parasitic capacitances in the connection between logic gates

To characterize the propagation delay in a digital circuit we may use a very

simple schematic, representing an inverter whose output is connected to another device of the same type. The circuit shown in figure 8.21 represents two cascaded inverters.

In the figure capacitor C_o represents the parasitic capacitance associated to the output of the first inverter while capacitor C_i represents the input capacitance of the second inverter. The additional capacitor C_l represents the parasitic capacitance of the interconnection line. In our analysis we will assume that the propagation delay associated to the length of the interconnection line is negligible with respect to the switching speed of the gates. Otherwise the analysis would be more complex and considering a full transmission line.

The second inverter switches when its input voltage crosses the logic threshold voltage (typically at half of the power supply). Let us suppose initially the output of the first inverter to be low. MOS transistor M_1 will conduct in the resistive region and the output voltage will be close to 0 V.

Let us now change the input voltage of the first inverter to logic one in order to switch off M_1 and set M_2 into conduction. An initial time t_0 is required before M_1 is fully switched off and M_2 is switched on.

When M_2 is in conduction it varies the output voltage by charging all the capacitances. Assuming that the voltage transition at the input of the first inverter is fast, then M_2 will operate in the saturation region, at constant current I_{Dp} for most of the output transition. To compute the propagation delay we have to compute the time required to cross the logic threshold V_{inv} of the second inverter. Assuming I_{Dp} being constant we have:

$$V_{inv} = \frac{I_{Dp}}{C_o + C_l + C_i} t_1$$

and t_1 will be given by:

$$t_1 = \frac{V_{inv}}{I_{Dp}} (C_o + C_l + C_i)$$

The propagation delay will be the sum of the initial time t_0 and time t_1 required to reach voltage V_{inv} , that is:

$$t_{pd} = t_0 + \frac{V_{inv}}{I_{Dp}} (C_o + C_l + C_i)$$

According to the model there is a linear relationship between the propagation delay and the *total capacitance* driven by the gate. In a real circuits often the connections are not point-to-point and a gate may drive several other ones. Then the *total* delay depends linearly on the number of driven gates, each one adding a parasitic capacitance to the output.

The manufacturer provides on the datasheet the values of the rise time, fall time and the propagation delay considering a standard capacitive load, typically of 50 pF. Given the linear dependency between capacitive load and delays an expression of this type is also provided:

$$t_{pLH} = t_{pLH}|_{50\text{ pF}} + k \frac{\text{ps}}{\text{pF}}$$

The relation is composed of two terms:

1. The propagation delay on a standard 50 pF load considered as an input of one driven gate plus the interconnection. The input capacitance of the input is specified in the circuit datasheet and is typically of the order of 15 pF. The remaining 35 pF are assigned to the interconnection capacitance approximating the parasitic capacitance of a printed circuit board line between two components not very far away.
2. For every additional pF we will have an additional delay of k ps. The common practice is to count the driven gates, after the first one, and to add their input capacitances to compute the final delay. If the interconnection is longer than usual (but not long enough to be considered an interconnection line) an additional capacitance term for it may be added.

The same analysis may be carried on for a high to low output transition and the NMOS transistor current. In this case the t_{pHL} parameter is evaluated.

It is therefore clear that the speed of a logic circuit depends on the fanout of each gate. If we analyze again CMOS technology, using as an example HC technology whose static parameters are given in table 8.1, we may conclude that from the static point of view each gate may drive thousands of other ones. However if for the same logic family we compute the propagation delay as a function of the driven gates, then the dynamic performances degrade rapidly with the load and the maximum number of driven gates is limited to 10 or 20 when considering applications with average operating frequency. Therefore the maximum fanout is determined more by the dynamic characteristics than by the static ones.

8.4.3 Power analysis

We have seen that CMOS gates, when connected to other gates of the same type, do not dissipate any power when the output level is constant, apart from a very little component due to a slight *quiescent current*. We have also seen that the switching between logic levels requires the charge/discharge of the parasitic capacitances associate to the output node and these operations require the sink of a current from the power supply and therefore a certain power consumption.

In order to evaluate the power which must be provided by the power supply and correctly size it, it is necessary to compute the energy that is dissipated in a transition. Referring again to figure 8.21 the parasitic capacitance of the output is $C = C_o + C_l + C_i$ and when the output is switched to the logic 1 is charged through transistor M_1 to voltage V_{AL} storing a charge $Q = C \cdot V_{AL}$. If switching takes place in time t_r the average current and the average power provided by the power supply during the switching will be respectively:

$$I_{ave} = \frac{C \cdot V_{AL}}{t_r}$$

$$P_{ave} = I_{ave} \cdot V_{AL} = \frac{C \cdot V_{AL}^2}{t_r}$$

The total energy absorbed from the power supply will be :

$$\mathcal{E} = P_{ave} \cdot t_r = \frac{C \cdot V_{AL}^2}{t_r} \cdot t_r = CV_{AL}^2$$

Since the energy stored in the output capacitor is :

$$\mathcal{E}_c = \frac{1}{2}CV_{AL}^2$$

then an equivalent amount of energy will be dissipated in transistor M_1 .

When the output switches back to the low level, then the capacitor discharges, dissipating its energy in transistor M_2 and the total energy dissipated in a full charge/discharge cycle will be $\mathcal{E}_T = CV_{AL}^2$.

Suppose that the circuit is synchronous and the input is a clock at frequency f_{CLK} . Each inverter or gate connected to this signal will go through a charge/discharge cycle at that frequency and will dissipate a power:

$$P = \frac{\mathcal{E}_T}{T} = \mathcal{E}_T \cdot f = f_{CLK} \cdot CV_{AL}^2$$

Let's now consider all the gates of the circuit and define as C the average capacitance associated to a gate input. If n is the average number of gates whose output switches at each clock cycle we may evaluate the power dissipated by the circuit as:

$$P_d = n \cdot f_{CLK} \cdot C \cdot V_{AL}^2$$

This relation is very important because it clearly shows the parameters on which to act to control the power consumption of a logic circuit. Let us consider separately each term:

- **Supply voltage:** the power consumption depends on the **square** of this parameter and it is therefore clear that it is important to reduce its value to the minimum which guarantees the logic gate operation. In fact the internal circuits of complex integrated circuits such as processors and memories operate at low voltages, typically of the order of 1 V.
- Equivalent **capacitance** of the gates. This parameter basically depends on the area of the devices and therefore decreases by reducing the physical size of the circuit.
- **Number** of gates which switch at every clock cycle. This parameter depends on the complexity of the circuit but it is possible to reduce the average number of switching gates by appropriate architectural choices. As an example in the design of a microprocessor proper choices in the coding of instructions may lead to strong reductions in the power consumption of the circuit.
- **Clock frequency.** Clearly the performances of a microprocessor depend on the clock frequency and this parameter is in many cases under designer or user control. In portable devices frequency is lowered when the device is in standby to increase battery life, while in high performance systems it is increased when high computational power is required.

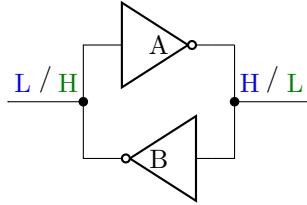


Figure 8.22: Inverter base bistable circuit.

8.5 Sequential circuits

The logic circuits analyzed until now are referred to as *combinational* because the outputs at any given time depend only on the values of the inputs (apart from propagation delays). If on the other hand the *state* of the circuit depends not only on the current value of the inputs but also on their history, that is on the temporal sequence of the inputs then the circuit is referred to as *sequential*. In other words a circuit is sequential if it contains memory elements which determine, together with current inputs, the current and future state of the system and as a consequence the current and future value of the outputs.

8.5.1 Bistable element

The circuit shown in figure 8.22 is the simplest sequential one which may be implemented with logic gates. Both left and right nodes are connected to an output and there is no real input, but this is the base of a static memory cell and in the following we will see how to set and change the internal value.

For the moment let us assume that the left node is at the low level. Inverter A output will be at the high level and is connected to the input of inverter B. The output of this latter will be at the low level maintaining the original value of the left node. Hence the logic value is maintained through the feedback loop of inverters A and B and the same happens when assuming a high initial logic value. The feedback loop remembers the initial value and maintains it indefinitely. This circuit is called a **bistable** and is the basis of static memory circuits.

8.5.2 S-R latch with NAND gates

The simplest sequential circuit with inputs is the *Set-Reset* or *S-R latch*, shown in figure fig. 8.23b which is both used directly in some application and as the base of the more complex flip-flop which is the fundamental component of synchronous sequential circuits. Conceptually the latch is just an evolution of the previously analyzed bistable.

In fact a NAND gate with an input at H level is equivalent to an inverter and therefore the circuit of figure 8.23a is equivalent to a bistable. If we allow to inputs to go to the low state we obtain the SR latch where the internal state may be set from the inputs. The same result may be obtained using NOR gates with inputs kept at the L level.

Since the outputs of the sequential circuits depend also on internal state, that is the temporal sequence of the inputs, we do not define truth tables, as

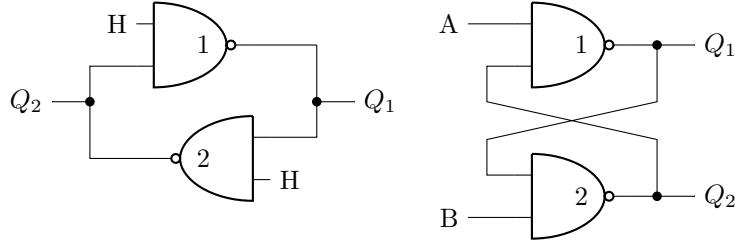


Figure 8.23: From the bistable to the latch. The first circuit is the bistable, the second the *latch*

A	B	Q	\bar{Q}
L	L	H	H
L	H	H	L
H	L	L	H
H	H	Q_{-1}	\bar{Q}_{-1}

Table 8.5: Excitation table of the NAND latch. The first configuration is the forbidden one and the last is the memory one.

for combinational functions, but *excitation tables* in which the internal state is considered an input. For the latch the state is associated to the current value of the output and the future value of the output depends also on the current one. The latch is an *asynchronous* circuit because any change of the inputs at any time may cause a change of the circuits state. We will see that it is simple to design *synchronous* circuits where inputs are sampled (and become active) only at given times (typically at the active transition of a clock signal).

In normal operation the two digital outputs Q_1 are Q_2 complementary, are identified as Q e \bar{Q} and the NAND latch operates as a bistable, that is a single bit memory.

If one of the two inputs, for instance A , is brought to the L state and B remains at the H state, then logic gate 1 sets its output Q to the high state H and gate 2 inverts it setting \bar{Q} to the L state. The same behaviour, with inverted roles, takes place exchanging A and B .

Since NAND gates have a H output when an input is L then it is clear that the complementarity of the outputs Q and \bar{Q} is violated when both A and B are at the L state. In the NAND latch the configuration with both inputs at the low state is referred to as *forbidden* not only because the outputs are not complementary but also for other reasons which will be more clear in the following.

SET and RESET inputs Table 8.5 shows the excitation table for the latch. The Q output is high when input A is low and is low when input B is low.

For this reason we may give a logic meaning to each input according to the effect on output Q : input A is the input *inverted SET* (\bar{S}) and input B is the *inverted RESET* (\bar{R}) input because they are active at the low state. The schematic and the latch symbol are shown in figure 8.24.

Labeling latch input and outputs is made according to the following con-

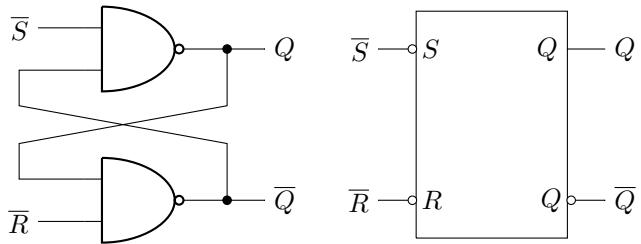


Figure 8.24: Circuit schematic and symbol of the *Set-Reset latch* with NAND gates

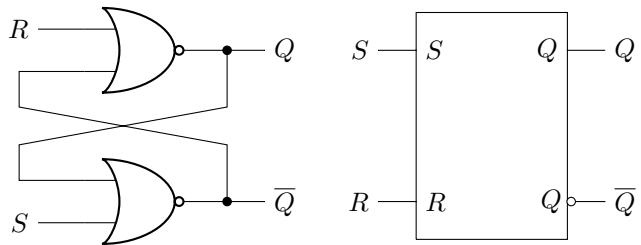


Figure 8.25: Circuit schematic and symbol of the *lSet-Reset latch* with NOR gates.

vention, used in throughout the notes: when a signal name is inside a symbol, the name is not inverted, even if the signal is active at the low state. Active low signal are indicated by a small circle at the connection between the signal line and the body of the symbol. If the signal name is on the connection wire and the signal is active low then the name has an horizontal bar over it. Other conventions exist such as preceding active low signals with a / or the letter *N* (/SET, NSET).

8.5.3 SR latch with NOR gates

It is possible to implement an SR latch using NOR gates instead of NAND ones, as shown in figure 8.25 because a NOR gate with one input at the L state is equivalent to the inverter. In this case the memory input configuration will be the (L, L) and the forbidden one is (H, H) . The upper input signal is active high and *Resets* the output Q while the lower one set the output to H and has therefore implements the *Set* function.

8.5.4 Oscillatory Metastability

The main problem of the latch arises when there is a transition of the inputs directly from the forbidden to the memory configuration.

Let us analyze the behaviour of a NAND based latch, assuming that the propagation delays of the two gates are equal. Initially (inputs (L, L)) the outputs are both high and then the inputs become both high at the same time, after a propagation delay, both outputs switch at the low state. When these values arrive at the inputs of the latch they force the outputs to the high state after another propagation delay and the system is back in the starting state.

At this point the cycle repeats itself indefinitely and the gates oscillate between the states (H, H) and (L, L) as shown in figure 8.26.

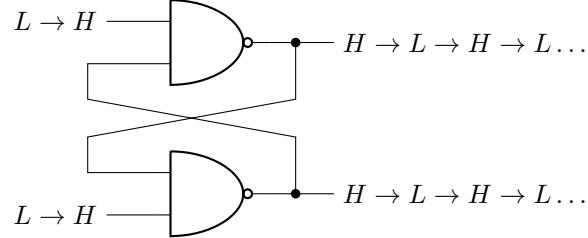


Figure 8.26: Oscillatory metastability in a NAND gates based SR latch.

The latch becomes a square wave generator with a period equal to the sum of the two propagation delays of the NAND gates because there are two switching per cycle. This phenomenon is named *oscillatory metastability* because it is not permanent, as an stable system, but, due to minimal differences in the gates propagation delays, after some time ends in one of the two stable states (H, L) or (L, H) . However the duration of the transient is not known a priori and may be described only deterministically. To avoid this phenomenon it is necessary to avoid the prohibited configuration of the inputs.

8.5.5 SR Transparent Latch

It is possible to modify the SR latch by adding a CLK synchronization control signal which makes the latch insensitive to variations of the inputs in correspondence to one of the two levels of the control signal. If the latch is NAND gates based we may add two other NAND gates at the inputs of the latch as shown in figure 8.27. In the same way for NOR based latches two NOR gates are added to the inputs. The circuit is referred to as *SR Transparent Latch*.

When CLK is in the H state the NAND gates at the input forward the complemented value of the A and B inputs of the latch at the right. The latch outputs will assume the state corresponding to the excitation table 8.5 .

When CLK is in the L state, the outputs of the NAND gates on the left will be high and the latch will be operating in memory mode. The excitation table of the complete circuit is given in table 8.6.

The oscillatory metastability problem has not yet been solved because if the

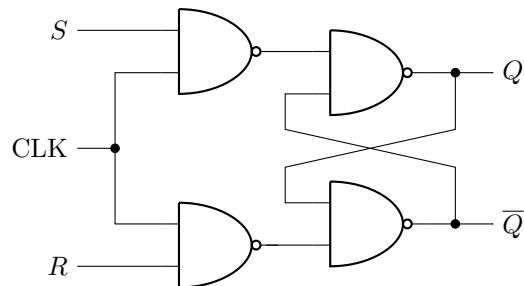


Figure 8.27: Schematic of the SR transparent latch.

CLK	S	R	Q	\bar{Q}
L	X	X	Q_{-1}	\bar{Q}_{-1}
H	L	L	H	H
H	L	H	L	H
H	H	L	H	L
H	H	H	Q_{-1}	\bar{Q}_{-1}

Table 8.6: Excitation table of the transparent SR latch with NAND gates.

S and R input signals are both at the H level when CLK is also H, then the inputs of the internal latch are both low in the forbidden configuration. If in this condition CLK switches to the L state then the internal latch go in the memory state and oscillatory metastability will occur.

8.5.6 D-latch

To use the latch as a memory cell without timing problems we are required to avoid the forbidden configuration. This may be achieved by connecting the inputs of the transparent latch to an input signal and to its complement (through an inverter) as shown in figure 8.28.

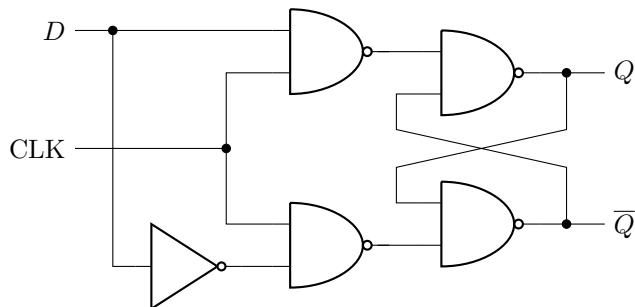


Figure 8.28: Circuit schematic of the *D-Latch*.

The single input D will set the state of the latch when CLK is at the high level and the latch will be in the memory state when CLK is low.

In summary the *D-type transparent latch*, or shortly the *D-latch* reads data D when the CLK is High and transfers it to the output, keeping it in memory when CLK is Low. When CLK is High any change in D will propagate to the output, delayed by the total delay of the internal gates and this is the *transparent* mode of operation. The device performs a sampling of D which lasts the whole positive half period of CLK as shown in figure 8.29. Please note that metastability is still possible: if the falling edge of CLK is close to the $L \rightarrow H$ transition of D , then the upper gate will see a H input and this value will be still present on the lower one for all the propagation delay of the gates. If, before this time has elapsed, the CLK signal is Low at both gates input, then on the inputs of the internal SR-latch there is exactly the same sequence that induces oscillatory metastability. From this analysis it is clear that D may not change within a given time interval around the $H \rightarrow L$ transition of CLK and this will be better defined when analyzing the flip-flop.

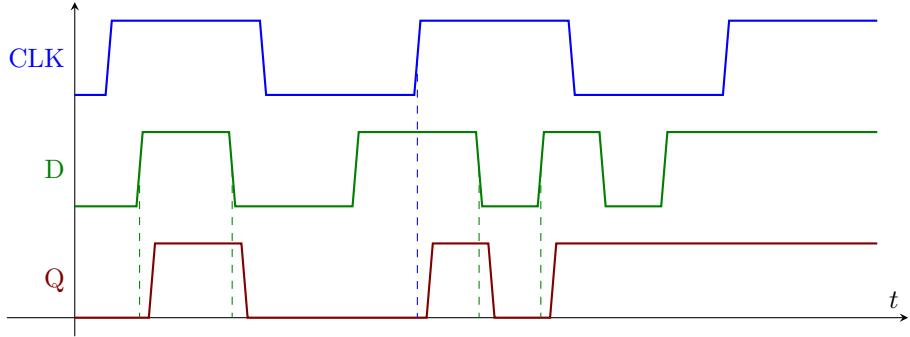


Figure 8.29: Example of input and output signas of a D-latch

The circuit symbol of a D-latch is shown in figure 8.30. The rectangle in the body of the symbol in correspondence to the CLK signal means that this signal is active at the high level. In some cases the control signal name is referred to as *EN* (Enable) or *LD* (load), reserving the CLK name for flip flops. It is possible to implement D-Latches with the control signal active at the low level and in this case the symbol of the circuit will show a small circle in correspondence to the CLK signal.

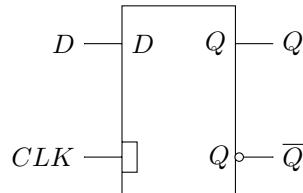


Figure 8.30: Simbolo circuitale del D-latch.

8.5.7 Synchronous and asynchronous circuits

The SR latch is an *asynchronous* circuit: in fact it is a sequential circuit since it has memory, but an input change at any time may induce a change of the output. The positive side of it is that the SR latch may react immediately to a stimulus, but it makes harder the work of the designer when the circuit which drives the inputs is a complex combinational logic. When studying the synthesis methodologies of combinational logic functions we may verify that, unless designing the circuit with significant redundancies, spurious output spikes or *hazards* may be generated before the output settles to the final state. These spikes at the inputs of a SR latch may generate an unwanted change of state. Although it is possible to build complex machines in asynchronous logic, designers prefer to use a common synchronization signal, commonly named *clock*, to sample the inputs of the sequential components only at given times, when hazards or other transients are finished. The transparent latch is an hybrid system since it has a clock signal but the sampling time is long and during it spurious hazards at the input are passed to the SR latch which is its core.

8.5.8 D Flip-flop

The transparent latch is the basic building block of the most common synchronous circuit, that is the D-type flip-flop. It differs from the D latch by the fact that the input signal is sampled only in correspondence to the transition of the clock signal from one level signal to another (active edge) which may be $L \rightarrow H$ or $H \rightarrow L$ according to the circuit implementation.

The DFF (acronym of D Flip-Flop) is built by cascading two D-latches as shown in figure 8.31. The first one, named *master*, is connected to the input and its output is the input of the second one, named *slave*, whose output is the output of the circuit. The control symbols of the two latches are complemented.

The leftmost latch is in transparent mode when CLK is at the high level. In this phase signal D propagates to the output of the latch (output Q_1) and is stopped at the input of the second latch which is in memory mode because its CLK input is low. The leftmost D-latch switches to the memory mode when CLK becomes low thus memorizing the input value present at the falling edge of CLK. When CLK is low the rightmost D-latch is in transparent mode propagating to the output the value sampled by the leftmost one and keeping the output constant until the following falling edge of CLK. An example temporal sequence is shown in figure 8.32.

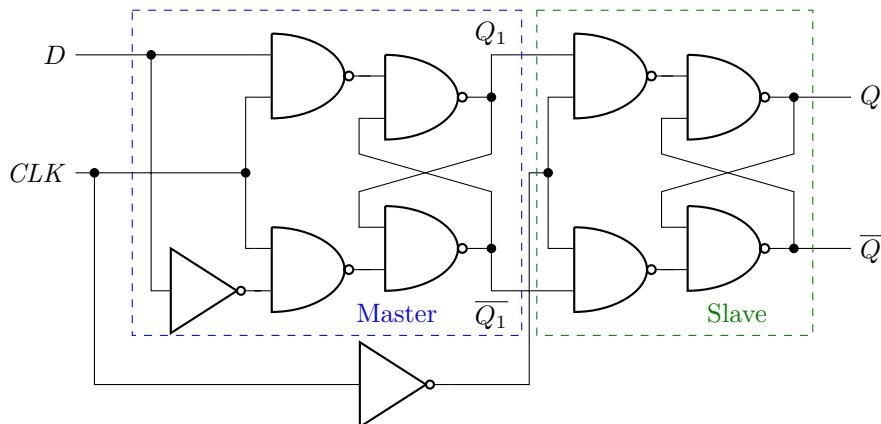


Figure 8.31: Circuit schematic of the D-type Flip Flop

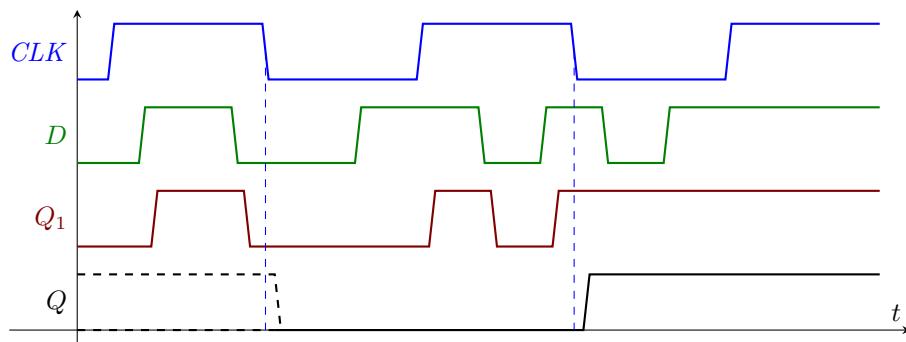


Figure 8.32: Example temporal sequence for the DFF of figure 8.31.

The DFF samples the input signal at the falling edges of the clock because the first latch is transparent with the CLK high and the second one with the CLK low. Hence inverting the clock signals of the two latches, that is changing the position of the inverter, it is possible to obtain the sampling on the rising edge of the clock.

This schematic is the most common implementation of a DFF but many other ones include circuitry handling two additional signals named respectively *PRESET* and *CLEAR* which act asynchronously with respect to CLK and allow to force the output either to the high or low states independently of the active edge of CLK or of the state of the input. They must be used in particular cases, typically at power-up or at reset, and to them apply the same considerations already made for asynchronous latches. The circuit symbol of a complete DFF with Preset and Clear is shown in figure 8.33. The triangular symbol in the body of the DFF in correspondence to the CLK signal indicates an activation on the rising edge. The addition of a small dot to the triangle indicates activation on the falling edge. Asynchronous control signal are typically placed on the top or bottom of the symbol.

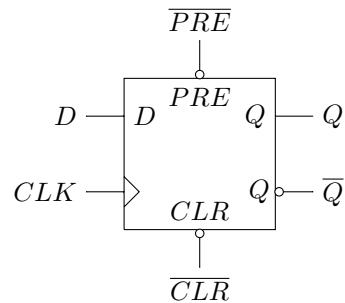


Figure 8.33: Circuit symbol of a DFF with asynchronous Preset and Clear.

8.5.9 Dynamic behaviour of flip flops

When studying latches we have analyzed their metastability problems. The D flip-flop is affected by similar problems but only in a limited time interval around the active edge of the clock. If we consider the left most D latch in transparent mode the internal nodes, which should have always complementary states, every time input D changes for a short interval may be both at the same logic level or in some intermediate state due to inverters' delay. This is not a problem if the remains transparent until the transient ends, however if at the same time the CLK has an active transition to the memory state this may induce either a metastable state in the master or, depending on the implementation, oscillatory metastability. Since the slave is in the transparent mode, this behaviour transfers to the output.

To avoid any metastability problem input signal D must remain stable in a time interval around the active edge of the clock. The manufacturers provide two parameters which indicate the time intervals before and after the clock edge during which a transition of D is not allowed as shown in figure 8.34.

- the **Setup time** t_{su} defines for how long the input must be stable before

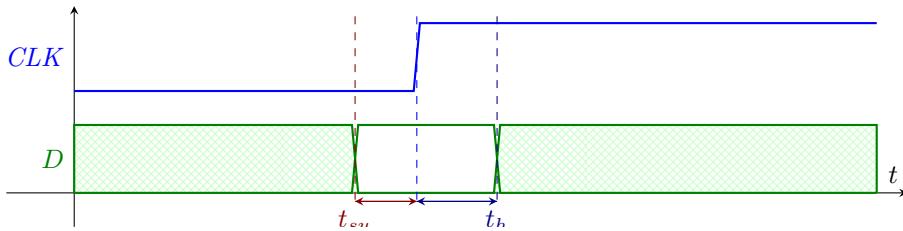


Figure 8.34: Setup and Hold times

the active clock edge.

- the **Hold time** t_h defines for how long the input must continue to be stable after the active clock edge.

Therefore a change in the input in violation of the setup and hold times may induce metastability in the flip flop. The probability of the transition into a metastable state depends on the architecture of the device, the process parameters, the dimensional parameters of the transistors from which the gate is built, the operating temperature, but may never be considered negligible.

Setup and Hold times are defined and must be respected also for D latches, with respect to the transition of the clock which brings the latch from the transparent to the memory mode.

Inside a synchronous system with a single clock signal it is not difficult that all flip flop input signals respect setup and hold signals. However if the system is connected to a non synchronized input it is not possible to avoid metastable behaviours.

As all physical systems also the flip-flop is affected by propagation delays. Time t_{pdCKQ} is defined as the time interval between the active clock transition and the corresponding change of the output.

8.6 Exercises

Exercise 8.1. CMOS inverter

A CMOS inverter has the following parameters:

$\mu_n C_{ox} = 1 \text{ mA V}^{-2}$	$\mu_p C_{ox} = 0.3 \text{ mA V}^{-2}$
$V_{Tn} = 0.5 \text{ V}$	$V_{Tp} = -0.5 \text{ V}$
$V_{AL} = 2.5 \text{ V}$	$L_n = L_p = 100 \text{ nm}$
$W_n = 100 \text{ nm}$	$W_p = 150 \text{ nm}$

1. compute the inversion voltage
2. compute a new width of the pMOS device to obtain $V_{inv} = 1.25 \text{ V}$

Exercise 8.2. CMOS Logic

Design in CMOS logic the following logic function:

$$U = \overline{(A + B) \cdot C}$$

1. Draw the schematic of the circuit

Exercise 8.3. Open drain circuit

Design a circuit that connects the outputs of three inverters type 74HC05 to the input of a single HC gate. The supply voltage should be $V_{AL} = 5 \text{ V}$. The circuit should work in the entire commercial temperature range (0°C to 85°C).

1. Draw the schematic of the circuit.
2. Find the value of the pull-up resistor to obtain the fastest possible low to high transition.
3. Find the value of the pull-up resistor to obtain the lowest possible power consumption of the circuit.

Exercise 8.4. CMOS delay

The standard propagation delay of a logic gate is $t_{pLH_{std}} = 8 \text{ ns}$. The incremental delay is $\Delta t_{pLH} = 20 \text{ ps pF}^{-1}$. The average input capacitance of a logic gate based on the same technology is $C_{in} = 10 \text{ pF}$.

1. What is the propagation delay of the low to high transition at the output of the gate if it is connected to 20 inputs?

Exercise 8.5. CMOS power

A synchronous CMOS logic circuit is clocked at $f_{CLK} = 10 \text{ MHz}$. The average capacitance of a gate of the circuit is $C = 1 \text{ pF}$. The power supply voltage is $V_{DD} = 2.5 \text{ V}$.

1. Compute the average power dissipated by the circuit if the average number of gates switching in a clock period is $n = 10^5$.
2. Compute the dissipated power if the power supply is reduced to $V_{DD} = 1.8 \text{ V}$.
3. Compute the dissipated power if, additionally, the clock frequency is reduced to $f_{CLK} = 1 \text{ MHz}$

Chapter 9

Passive devices

Contents

9.1 Resistors	9-1
9.1.1 Nominal value and tolerance	9-2
9.1.2 Power dissipation and packaging	9-3
9.1.3 Maximum voltage and temperature coefficient	9-4
9.1.4 Material and parasitic parameters	9-4
9.1.5 Noise	9-5
9.2 Capacitors	9-5
9.2.1 Parasitic parameters	9-6
9.2.2 Ceramic capacitors	9-6
9.2.3 Plastic film capacitors	9-8
9.2.4 Electrolytic capacitors	9-9
9.3 Inductors	9-11
9.3.1 Parasitic parameters	12

PASSIVE devices include resistors, capacitors, inductors and more complex modules like transformers, crystals and ceramic filters. We only focus on the first three items. It is important to understand the parameters of passive devices, specially the parasitic parameters, to be able to choose the right component for a certain application.

9.1 Resistors

Resistors are the most widely used passive components and the easiest to build. When we need to select a resistor for our circuit, we have to take into account several different parameters. The most important are:

- Nominal resistance value
- Tolerance
- Maximum power dissipation
- Physical dimensions and package type
- Maximum voltage

- Temperature coefficient
- Parasitic capacitance
- Series inductance
- Noise

Let us analyze in details these elements.

9.1.1 Nominal value and tolerance

When we need a resistor, obviously the first parameter we take into account is the nominal value of the resistance. In most applications, we compute the value we need and then we have to choose among a set of standard values the one which is closest to the one we computed. Depending on the application, we can tolerate to use a value which is not exactly the one we need but is not so different. If instead we need a very precise value, we can normally obtain it but at the expenses of a much higher cost with respect to a low precision device.

The least expensive resistors nowadays have tolerances in the range of 5% to 10%. The most common set of standard values is the series so called E12. This series of standard values was introduced for resistors with 10% tolerance. E12 means that in the range of one decade (that means from 1 to 9.999) there are 12 different possible values of resistance. Starting from 1, if the tolerance is 10%, it makes no sense in offering other values in the range 1 to 1.1, because a resistor with nominal value for example $1.05\text{ k}\Omega$, with 10% tolerance, can have an effective value lower than that of a resistor with nominal value 1. The first reasonable value higher than 1 is 1.2, so that the range of possible real values of a resistor with value 1.2 is touching but not too much overlapped with that of a resistor with nominal value 1, see figure 9.1.

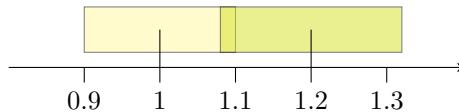


Figure 9.1: Range of real values versus nominal for components with 10% tolerance

The E12 of standard values, according to international standard IEC60063, is listed in table 9.1, together with series E24 for components with 5% of tolerance. Other series are defined for more precise components: E48 for 2% tolerance

E12	1	1.2	1.5	1.8	2.2	2.7
	3.3	3.9	4.7	5.6	6.8	8.2
E24	1	1.1	1.2	1.3	1.5	1.6
	1.8	2.0	2.2	2.4	2.7	3.0
	3.3	3.6	3.9	4.3	4.7	5.1
	5.6	6.2	6.8	7.5	8.2	9.1

Table 9.1: Standard series of values E12 and E24.

devices, E96 for 1% and E192 for 0.5%.

The E12 series is the most used, even now that the most inexpensive resistors have 5% tolerance. Laboratory stock devices are normally all resistors from $10\ \Omega$ to $10\ M\Omega$ in the E12 series.

9.1.2 Power dissipation and packaging

Another very important parameter for resistors is the maximum power dissipation. The ability to dissipate power is directly linked to package type and dimensions. Low power devices can be very small while high power ones are necessarily large and in some cases need convection cooling, that is a fan creating an air flow on the resistor to ease power dissipation and cooling.

In the chapter on printed circuit boards we will cover more in details the packaging of electronic devices. For now, figure 9.2 shows examples of resistors with different packaging.

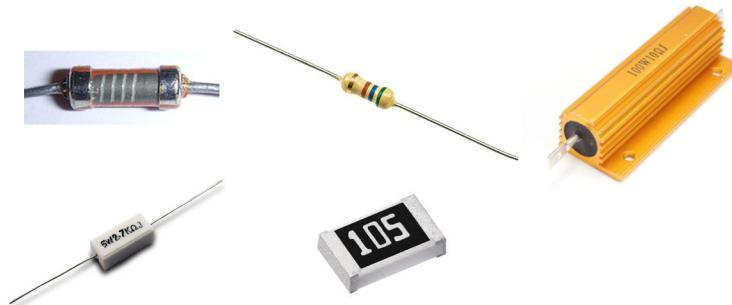


Figure 9.2: Different resistors

In the figure we can note resistors with colored bands and leads, two power resistors, one with leads, the other with wire solder pins and one surface mount leadless chip resistor. High power resistors are identified by a text string including resistor value, tolerance, maximum power dissipation and other information. Leaded low power resistors, with power range between $0.125\ W$ and $1\ W$ are normally identified by a set of four or five colored bands, defining value and tolerance, while surface mount devices only bear a printed number identifying resistor value. For devices with colored band, the information can be decoded using table 9.2.

The first two bands indicate the first two digits of the value in ohm. The third band is the number of zeros following the two digits and the fourth band is the tolerance. For surface mount resistors the code is numeric but the meaning is similar: the first two digits are the first two digits of the value in ohm, the third digit is the number of zeros following the first digits. No indication of tolerance is given in this case. For example, the code 105 printed on the SMD resistor of figure 9.2 means $1\ M\Omega$: 10 followed by 5 zeros, in ohm.

Color	Band 1	Band 2	Band 3	Band 4
Black	0	0	10^0	—
Brown	1	1	10^1	—
Red	2	2	10^2	—
Orange	3	3	10^3	—
Yellow	4	4	10^4	—
Green	5	5	10^5	—
Blue	6	6	10^6	—
Violet	7	7	10^7	—
Grey	8	8	10^8	—
White	9	9	10^9	—
Silver	—	—	—	10%
Gold	—	—	—	5%

Table 9.2: Color codes of resistors.

9.1.3 Maximum voltage and temperature coefficient

For large valued resistors, maximum power dissipation is not the only limit in using the device, but also maximum voltage drop becomes significant. Every type of package and every resistive material can sustain a maximum level of electric field, beyond which the material breaks and high current can start flowing.

Another very important parameter is temperature coefficient. Most resistors are designed to have minimum variation of resistance with temperature, but there are some resistors which are designed to have a very well defined variation with temperature. They are known as thermistors. We call PCT resistors those who increase their value with temperature. These are used either as temperature sensors, like the well known PT100 standard ($100\ \Omega$ at room temperature), or as power limiting devices: an increase of current causes an increase of power dissipation, which in turn increases resistor temperature and resistance value, which increases power dissipation but also normally reduces voltage drop in the load and restores normal operating conditions.

There are also NTC resistors whose value decreases with temperature. These are normally used as temperature sensors.

9.1.4 Material and parasitic parameters

The resistors are manufactured using different kinds of resistive material. Most low power devices are manufactured starting from a deposition of a film of resisting material over an insulating substrate. The film is then etched in spiral, as visible in the upper-left of figure 9.2, or in serpentine shape, as in most SMD devices.

The material used for leaded cylindrical resistors like the ones we have in the lab is carbon or nickel chromium film, while SMD devices often use different types of ceramic and/or metallic films (cermet) like tantalum nitrate or ruthenium oxide, lead oxide, bismuth ruthenate or others.

High power devices are normally based on a conductive wire, for example nickel chromium, wound on a ceramic substrate, see figure 9.3¹.

¹By Zureks - Own work, Public Domain, <https://commons.wikimedia.org/w/index.php?curid=8952040>

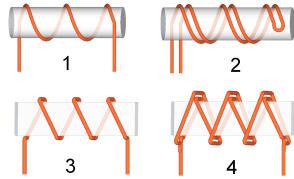


Figure 9.3: Wirewound resistors

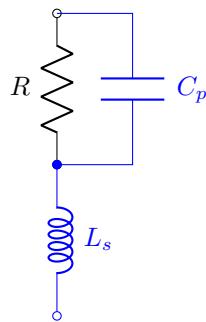


Figure 9.4: Resistor and parasitic devices

The kind of conductive material, the package and the manufacturing process determine the amount and the value of parasitic series inductance and parallel capacitance, see figure 9.4. To understand why we have parasitic devices, we can focus on a wirewound resistor. In this case it is evident that the wire which is wound in spiral around an insulating core, is also an inductor. All other types of resistors show a certain level of parasitic inductance and capacitance. These are normally completely negligible at low frequency but are very important at radio frequency. SMD devices show the smallest values of parasitic devices.

9.1.5 Noise

Every resistor, even an ideal resistor, introduces thermal noise in the circuit. This is called Johnson-Nyquist noise and is unavoidable. Circuits very sensitive to noise have to be operated at very low temperature. Noise is an advanced topic and we will not cover it in this course.

9.2 Capacitors

Capacitors are more difficult to build than resistors and can be rather big, specially for high capacitance and high voltage devices. Manufacturing techniques highly impact on device performance. We will give a brief introduction on capacitor types, but, before that, we will consider the main parameters that a designer has to take into account when dealing with capacitors.

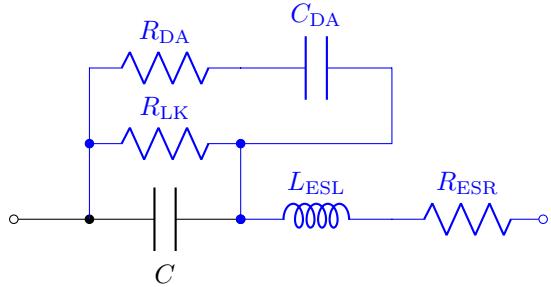


Figure 9.5: Capacitor and parasitic devices

9.2.1 Parasitic parameters

Figure 9.5 represent a capacitor together with its parasitic devices. The term L_{ESL} , equivalent series inductance, is higher on leaded devices and can normally be neglected for SMD devices. An ideal capacitor should be completely reactive, while a real device will dissipate some power due to resistance of the package and losses in dielectric. This is represented by the equivalent series resistance (ESR). When a capacitor is charged, it should retain the charge for an infinite amount of time, while a real device will discharge more or less slowly due to dielectric losses. We take this into account by inserting R_{LK} , leakage resistance. If we short circuit a charged capacitor for a short time and then we leave it unconnected, the voltage across the capacitor should remain zero. This is not always true, certain high voltage capacitor can exhibit an even dangerous level of voltage across the terminals after some time. This is due to a complex mechanism called dielectric absorption and is represented by R_{DA} and C_{DA} in our model.

Not all these parameters are important for all capacitors in all applications. For example, C and L_{ESL} form a resonant circuit. For some capacitors the self resonating frequency (SRF) can be in the GHz range, so it is perfectly negligible for audio applications, but not for radio frequency ones.

Other very important parameters of a capacitor are the rated voltage and maximum RMS current. If a voltage above the rated is applied to the capacitor, the electric field in the device can rise above the maximum allowed, causing a breakdown in the dielectric and the destruction of the device. The ESR of the capacitor is responsible for power dissipation of the device. Currents above the maximum RMS current cause an increase of temperature of the capacitor that can lead to device failure.

9.2.2 Ceramic capacitors

Nowadays ceramic represents the most diffused dielectric in capacitors in the range from picofarad to a few microfarad. This is due to the properties of having very small losses, high dielectric constant, being solid and having the possibility of creating layers with a minimum thickness of $0.5 \mu\text{m}$.

The dielectric is composed of finely ground powder of paraelectric or ferroelectric materials, mixed with others to achieve the target characteristics. The powders are sintered at high temperatures and sandwiched between two metal layers to obtain a capacitor.



Figure 9.6: Ceramic capacitors

To obtain high capacitance values, several layers of dielectric and metal electrodes are stacked to create a so called MLCC (Multi Layer Ceramic Capacitor). This allows to have low voltage capacitors with values in excess of $200 \mu\text{F}$, a value that up to few years ago was obtained only by electrolytic capacitors.

Ceramic capacitors exist in very different packages, both leaded and SMD. A small sample is depicted in figure 9.6. Markings on the package vary widely with package type. Small SMD devices do not bear any marking, unlike resistors. Axial devices normally have markings. For small devices, the marking is a sequence of three numbers that can be read as in the case of resistors using as unit the picofarad. The disc capacitor in the upper left corner of the picture is a 100nF device. The numbers can be followed by other letter-number combinations, indicating temperature dependence, maximum voltage, type of dielectric or similar, but there is no well defined standard.

There are many different ceramic dielectrics, each with different properties. The composition of the dielectric influences the amount of losses, the equivalent series resistance, the self resonating frequency and the temperature dependency. This last parameter is very important because it is difficult to obtain a capacitor very stable with temperature.

Two different industry standards classify the dielectric of a ceramic capacitor: IEC/EN 60384-1/8/9/21/22 and EIA RS-198. Both standards define three classes, of which only the first two are of practical interest. Class 1 devices are the best for applications where low losses and high stability are mandatory, while class 2 devices are for general use, less demanding applications and offer a higher density (higher capacitance vs volume) with respect to class 1 devices.

The best behavior with temperature is obtained by class 1 capacitors named NP0 in the IEC standard and C0G in the EIA classification. They are the ones to be used when very high stability is required.

EIA classification is the most used in class 2 devices and consists in a three characters letter-number-letter combination. The first letter defines the lower temperature limit and can be $X = -55^\circ\text{C}$, $Y = -30^\circ\text{C}$ or $Z = 10^\circ\text{C}$. The second number is the upper temperature limit. The most used numbers are $5 = 85^\circ\text{C}$, $7 = 125^\circ\text{C}$ and $8 = 150^\circ\text{C}$. The third letter is the maximum capacitance change over the temperature range defined by the first two characters. The most common are $P = \pm 10\%$, $R = \pm 15\%$, $U = +22\% / -56\%$ and $V = +22\% / -82\%$. A good MLCC has dielectric code X7R or, better but much more expensive, X8P. Very high value capacitors can only be built at Z5U or Y5V specifications, which are very poor. Figure 9.7² shows typical capacitance variation with temperature for the above mentioned dielectrics compared to NP0.

²Credits: Wikipedia https://en.wikipedia.org/wiki/Ceramic_capacitor

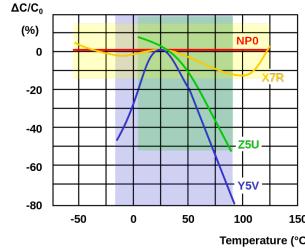


Figure 9.7: Typical capacitance variation with temperature for MLCCs

In MLCCs, the parasitic parameters R_{LK} , R_{DA} and C_{DA} are normally very good (can be neglected for most applications). R_{ESR} is very low and is generally lower in capacitors based on better dielectrics for the same nominal capacitance (i.e. X7R has lower ESR than Y5V). Inductance is lower in smaller devices. As an example, three capacitors with the same nominal value of capacitance but with different SMD cases are compared in figure 9.8.

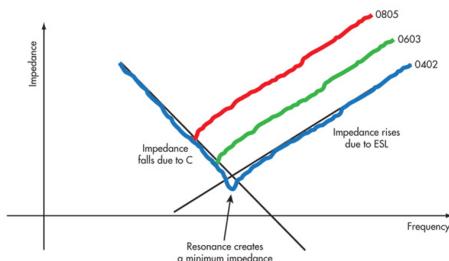


Figure 9.8: Self resonating frequency versus case

Another parasitic parameter of MLCCs is the dependence of the capacitance from the applied DC voltage. One of the most common applications for ceramic capacitors is the bypass of power supply lines to avoid propagation of high frequency noise generated by pulsed current absorption of digital devices during transitions. In this case a high DC voltage is constantly applied to the capacitor. Under these conditions the effective capacitance can be reduced by 50% even for X7R devices.

9.2.3 Plastic film capacitors

Film capacitors are manufactured using plastic film as dielectric between two metallic electrodes. Different kind of plastic material can be used, the most common being polypropylene (PP) and polyester (PET), accounting together for 90% of total film capacitors production.

Figure 9.9 shows some examples of film capacitors. They find applications specially in high voltage and high power applications, where a typical use is as spike suppressors or snubbers. The electrical permittivity of the dielectric is lower, compared with ceramic capacitors, so they are usually bigger, but they exhibit better stability with temperature and precision.

The capacitance range is from 100 pF to 10 µF. Equivalent series resistance and equivalent series inductance are typically low. They can be used up to the

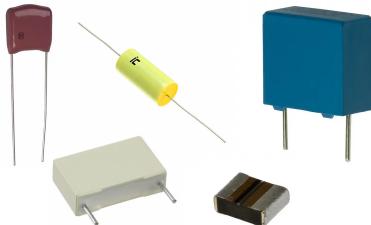


Figure 9.9: Plastic film capacitors

MHz frequency range and do not suffer from reduction of capacitance when used with high constant DC voltage polarization.

9.2.4 Electrolytic capacitors

Electrolytic capacitors are manufactured using a totally different approach with respect to the previously analyzed types. They are polarized. This means that one plate of the capacitor, the anode, has to be at a higher potential than the other plate, called cathode. Reversing the polarity of the applied voltage can result in the destruction of the device. This is potentially hazardous, with risk of explosion or fire.

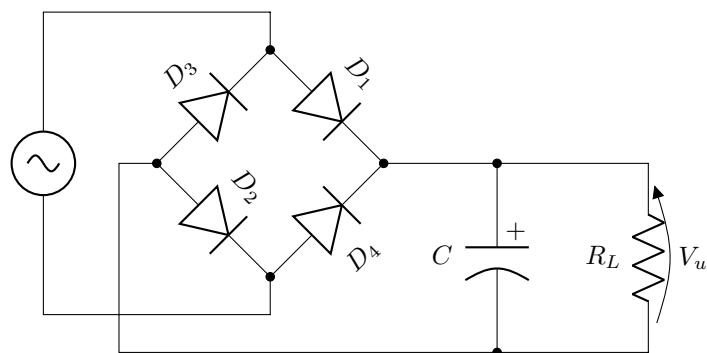


Figure 9.10: Example of electrolytic capacitor as filter in power supply circuit

The positive plate of the capacitor is a metal foil that is processed through an electrolytic process, called anodization, that grows a thin oxide layer on the surface. This oxide layer is the dielectric of the capacitor. Being very thin, it is possible to obtain much larger capacitance values per volume with respect to ceramic or film capacitors. The oxide layer is covered by a solid or liquid electrolyte constituting the cathode of the capacitor.

The two main families of electrolytic capacitors are aluminum and tantalum capacitors. The former use aluminum as anode, aluminum oxide as dielectric and, in most cases, liquid electrolyte as cathode, the latter use tantalum as anode, tantalum pentoxide as dielectric and, normally, solid electrolyte as cathode.

Aluminum capacitors are the most inexpensive and have capacitance range from the microfarad range to more than one farad. The rated voltage ranges from few volts to several hundreds volts. Figure 9.11 shows some examples of this



Figure 9.11: Aluminum electrolytic capacitors

type of capacitors. Liquid cathode devices have self healing properties, because the liquid can restore oxygen to altered dielectric, but suffer from evaporation so that the capacity and other properties will vary with time and the device will fail after a limited amount of service hours. Broken electrolytic capacitors is the main cause of failures in power supply units.



Figure 9.12: Tantalum electrolytic capacitors

Tantalum capacitors are more compact than aluminum ones and are normally SMD devices, except for high voltage applications. They are more reliable than aluminum capacitors and show better characteristics, but are more expensive. Figure 9.12 shows some examples of this type of capacitors.

Parasitic parameters

Parasitic parameters of electrolytic capacitors have high impact on possible fields of application. In practice, this type of capacitors is working as such only up to a few kHz. Above this frequency, the ESR is dominant and from the megahertz range the device becomes an inductor.

To show this, we can try to find the best 1000 μF , 25 V rated, aluminum electrolytic capacitor from a online reseller (DigiKey) in 2020. The capacitor with lowest ESR has $R_{\text{ESR}} = 41 \text{ m}\Omega$. The same capacitor has $L_{\text{ESL}} = 20 \text{ nH}$. If we plot on log-log axes the magnitude of the impedance of this device in the frequency range 1 Hz - 10 MHz, we obtain the graph of figure 9.13. If we think that most capacitors of the same value have ESR up to 0.5Ω , we can see that we cannot neglect parasitic parameters at frequencies above the kilohertz limit. Tantalum capacitors exhibit similar curves but are generally smaller than corresponding aluminum ones, and one order of magnitude more expensive.

Electrolytic capacitors are used in many low frequency applications ranging from decoupling DC polarization in audio amplifiers to filters in power supply circuits, like in figure 9.10. For high power applications, great care must be taken in considering the maximum allowable RMS current, sometimes called

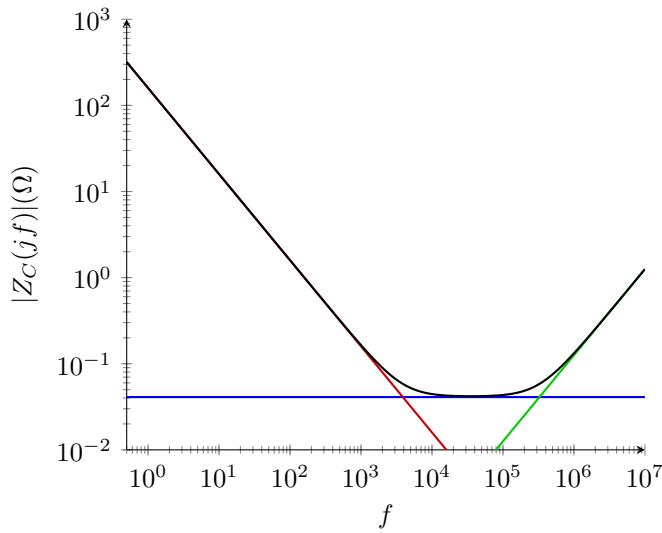


Figure 9.13: Magnitude of capacitor impedance: ESR (blue), ideal capacitor (red), ESL (green), total impedance (black)

ripple current, in the capacitor. Exceeding current limits can lead to dangerous failure of the capacitor.

9.3 Inductors

Inductors, also called coils or chokes, are not normally used in low power and low frequency circuits, mainly because useful values of inductance result in huge and heavy devices, compared to capacitors and resistors.

Sometimes it is more convenient to simulate the impedance of an inductor by using two operational amplifiers and other five passive components, resistors and capacitors, than inserting a physical inductor.

In low frequency circuits, inductors are indeed used in high power applications, such as power supply circuits or filters on the output of audio amplifiers as for example crossover filters for speakers.

High frequency circuits on the other hand heavily use inductors, because small value devices can be of size comparable to resistors or capacitors. Also, parasitic parameters of small value inductors are not as important as those of high value one. An inductor is normally composed of an isolated wire wound as a



Figure 9.14: Different inductor types: air, laminated iron, ferrite core.

coil around a core. There are two categories of cores: high magnetic permeability material or air. Figure 9.14 shows some examples of inductors.

The high magnetic permeability cores are built using ferromagnetic materials. At low frequencies laminated iron is used, while at higher frequencies the preferred materials are different kinds of ferrites.

Air cores can be made of plastics, ceramics or other non-magnetic materials. They can even be pure air, if the wire used for the coil has sufficient mechanical strength.

The main difference between ferromagnetic cores and air is in the value of inductance obtainable, much higher in the former case, and the fact that the magnetic field can be concentrated in the core, thus reducing emitted noise. On the other hand, ferromagnetic materials suffer from saturation effects, that is if the magnetic field becomes higher than a certain limit, the core will not accept more energy and the differential value of inductance will drop to that of the air equivalent.

9.3.1 Parasitic parameters

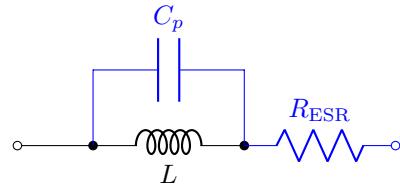


Figure 9.15: Inductor parameters

A real inductor, like any physical device, is affected by a number of parasitic parameters. A simplified schematic is depicted in figure 9.15.

One of the most important is the equivalent series resistance, ESR. In an inductor the ESR is not constant but generally increases with frequency, because of the skin effect and of the core frequency response. Iron cores in particular suffer from the Eddy currents (or Foucault currents) which limit drastically their usage at high frequency.

The adjacent turns of the coil create a distributed capacitor, which is represented (in lumped model) with C_p in figure 9.15. For this reason, every inductor resonates at a certain Self Resonating Frequency (SRF). For high frequency applications, the inductor's datasheet lists the quality factor of the device, which is defined as $Q = \omega L / \text{ESR}$ at a given frequency. The higher the Q the better the inductor for applications at that frequency.

Other parameters are the maximum allowed peak and RMS current. The former is related to core saturation phenomena, the latter to the maximum power dissipation due to ESR. Some inductors are also characterized by their behavior in saturation zone.

Appendix A

Solution to exercises

A.1 Solution to chapter 1 exercises

A.1.1 Solution to exercise 1.1

1. The voltage gain of the amplifier is

$$G_V = \frac{V_o}{V_g} = \frac{R_i}{R_i + R_g} \frac{R_L}{R_L + R_o} A_V = \frac{10 \text{ k}\Omega}{11.2 \text{ k}\Omega} \frac{1 \text{ k}\Omega}{1.1 \text{ k}\Omega} \cdot 100 = 81.2 \Rightarrow 38.2 \text{ dB}$$

2. The power drawn from the input generator is

$$P_g = \frac{V_g^2}{R_i + R_g}$$

The power transferred to the load is:

$$P_L = \frac{V_o^2}{R_L} = \frac{(Vg \cdot G_V)^2}{R_L}$$

The power gain is then

$$G_P = \frac{P_L}{P_g} = G_V^2 \cdot \frac{(R_i + R_g)}{R_L} = 6593 \Rightarrow 48.7 \text{ dB}$$

A.1.2 Solution to exercise 1.2

1. The current gain of the amplifier is

$$\frac{I_L}{I_g} = A_i \frac{R_g}{R_i + R_g} \frac{R_o}{R_L + R_o} = \frac{50 \text{ k}\Omega}{51 \text{ k}\Omega} \frac{100 \text{ k}\Omega}{101 \text{ k}\Omega} \cdot 20 = 19.4 \Rightarrow 25.8 \text{ dB}$$

A.1.3 Solution to exercise 1.3

1. The transconductance of the amplifier is

$$\frac{I_L}{V_g} = G_m \frac{R_i}{R_i + R_g} \frac{R_o}{R_L + R_o} = \frac{10 \text{ k}\Omega}{15 \text{ k}\Omega} \frac{100 \text{ k}\Omega}{102 \text{ k}\Omega} \cdot 50 \text{ mA V}^{-1} = 32.7 \text{ mA V}^{-1}$$

A.1.4 Solution to exercise 1.4

- To find the input resistance we connect a test ideal voltage generator V_x at the input port and we measure the corresponding input current I_x , as in figure A.1

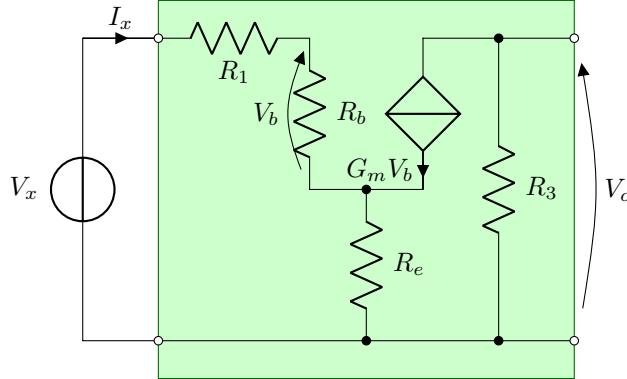


Figure A.1: Input resistance.

We can write KVL at the input loop:

$$V_x = I_x(R_1 + R_b) + (I_x + G_m V_b)R_e$$

but $V_b = I_x R_b$, so we can collect I_x :

$$V_x = I_x[R_1 + R_b + R_e(1 + G_m R_b)]$$

$$R_i = \frac{V_x}{I_x} = R_1 + R_b + R_e(1 + G_m R_b)$$

$$R_i = 11 \text{ k}\Omega + 1 \text{ k}\Omega \cdot (1 + 10 \text{ mA V}^{-1} \cdot 10 \text{ V mA}^{-1}) = 112 \text{ k}\Omega$$

- To find the gain, we apply an ideal voltage source to the input port and compute the resulting voltage at the output port. We can refer to the same schematic of figure A.1. The output voltage is the voltage drop on R_3 . The current in R_3 is $-G_m V_b$, but $V_b = I_x R_b$ and the relationship between I_x and V_x is the one we found in the previous point, so:

$$V_o = -G_m V_b R_3 = -G_m I_x R_b R_3 = -\frac{G_m R_b R_3 V_x}{R_1 + R_b + R_e(1 + G_m R_b)}$$

We can therefore compute the gain:

$$A_V = \frac{V_o}{V_x} = -\frac{100 \cdot 10 \text{ k}\Omega}{112 \text{ k}\Omega} = 8.92 \implies 19 \text{ dB}$$

- to find R_o we must short the input port, place an ideal voltage generator at the output port and measure the current. If we short the input port V_b will be null and the current generator will be an open circuit. Therefore the output port only includes R_3 . Finally:

$$R_o = R_3$$

A.1.5 Solution to exercise 1.5

- To find the asymptotic behavior in DC we have to substitute C_1 with an open circuit. The gain is in this case:

$$G_V = \frac{V_o}{V_g} = \frac{R_i}{R_i + R_g} \frac{R_L}{R_L + R_o} A_V = \frac{10 \text{ k}\Omega}{15 \text{ k}\Omega} \frac{2 \text{ k}\Omega}{2.1 \text{ k}\Omega} \cdot 10 = 6.35 \implies 16 \text{ dB}$$

We can find the high frequency gain by substituting C_1 with a short circuit. In this case we have $V_i = 0 \text{ V}$ and the output is null, so the gain is zero.

- To compute $V_o(s)/V_g(s)$ we first introduce

$$Z_1 = R_i \parallel C_1 = \frac{R_i}{sC_1 R_i + 1}$$

We can now find the value of V_i :

$$V_i = V_g \cdot \frac{Z_1}{Z_1 + R_g} = \frac{R_i}{R_i + R_g} \cdot \frac{1}{sC_1(R_i \parallel R_g) + 1}$$

Thus we obtain:

$$\frac{V_o}{V_g} = \frac{R_i}{R_i + R_g} \cdot \frac{R_L}{R_L + R_o} \cdot A_V \cdot \frac{1}{sC_1(R_i \parallel R_g) + 1} = \frac{6.35}{sC_1(R_i \parallel R_g) + 1}$$

- By inspecting the above equation, we have only one pole at frequency:

$$f_{p1} = \frac{1}{2\pi C_1(R_i \parallel R_g)} = 477 \text{ Hz}$$

At lower frequency, the diagram is flat and the gain is the one we found in the first point, 16 dB. The Bode diagram is in figure A.2.

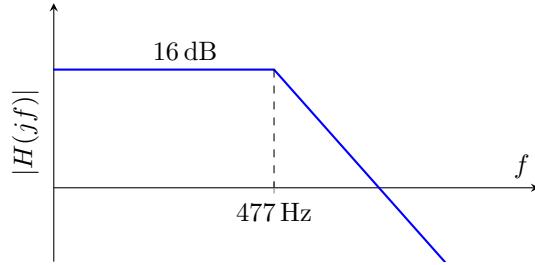


Figure A.2: Transfer function Bode diagram

A.1.6 Solution to exercise 1.6

- To find the asymptotic behavior in DC we have to substitute C_1 with an open circuit.

In this case R_L is disconnected from the rest of the circuit and the gain is zero.

We can find the high frequency gain by substituting C_1 with a short circuit.

The gain in this case is:

$$G_V = \frac{V_o}{V_g} = \frac{R_i}{R_i + R_g} \frac{R_L}{R_L + R_o} A_V = \frac{100 \text{ k}\Omega}{110 \text{ k}\Omega} \frac{5 \text{ k}\Omega}{5.1 \text{ k}\Omega} \cdot 50 = 44.6 \Rightarrow 33 \text{ dB}$$

- To compute $V_o(s)/V_g(s)$ we first introduce

$$Z_o = \frac{1}{sC_1} + R_o = \frac{sC_1 R_o + 1}{sC_1}$$

We can now find the value of V_o :

$$V_o = A_V V_i \cdot \frac{R_L}{R_L + Z_o} = A_V V_i \cdot \frac{sC_1 R_L}{sC_1 (R_L + R_o) + 1}$$

Thus we obtain:

$$\frac{V_o}{V_g} = \frac{R_i}{R_i + R_g} \cdot A_V \cdot \frac{sC_1 R_L}{sC_1 (R_L + R_o) + 1}$$

- By inspecting the above equation, we have one zero in DC and one pole at frequency:

$$f_{p1} = \frac{1}{2\pi C_1 (R_L + R_o)} = 312 \text{ Hz}$$

At lower frequency, the diagram is flat and the gain is the one we found in the first point, 16 dB. The Bode diagram is in figure A.3.

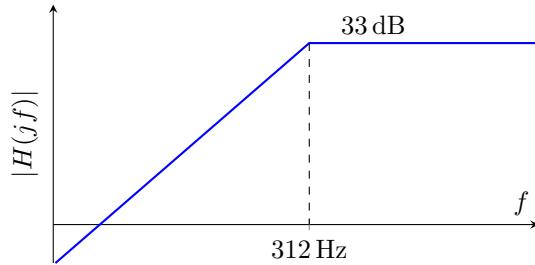


Figure A.3: Transfer function Bode diagram

A.2 Solution to chapter 2 exercises

A.2.1 Solution to exercise 2.1

Resistors R_1 and R_3 are in parallel and may be considered as a single resistor $R_{eq} = R_1 \parallel R_3 = 1.35 \text{ k}\Omega$

1. The gain assuming an ideal operational amplifier is given by:

$$\frac{v_u}{v_s} = \frac{R_2 + R_{eq}}{R_{eq}} = \frac{1}{\beta} = 21$$

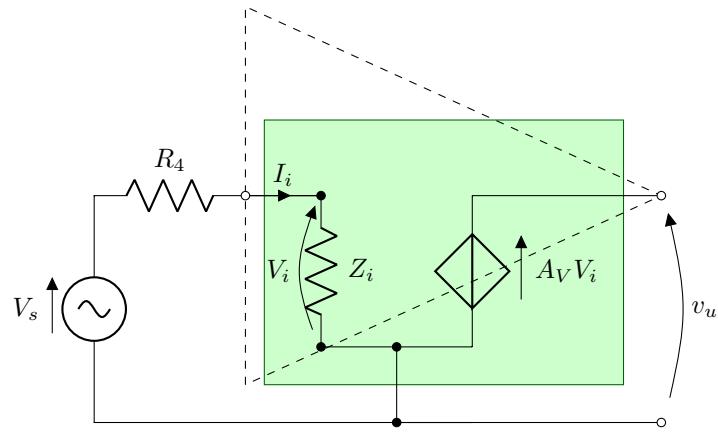


Figure A.4: Equivalent voltage amplifier

2. If the operational amplifier gain is $A_d = 1000$, the loop gain is given by::

$$T = A_d \beta = 1000 \cdot \frac{1}{21} = 47.62$$

and the input impedance at the non inverting pin will be:

$$Z_i = R_{id} \cdot (1 + A_d \beta) = 2.431 \text{ M}\Omega$$

We can then compute the gain A_V of the equivalent voltage amplifier (see figure A.4):

$$\frac{v_u}{V_i} = \frac{1}{\beta} \cdot \frac{1}{1 + \frac{1}{T}} = 21 \cdot 0.979 = 20.57$$

The voltage gain, considering the OA non ideality, is therefore given by:

$$\frac{v_u}{v_s} = \frac{Z_i}{Z_i + R_4} \cdot A_V = 20.57 \cdot 0.996 = 20.49$$

A.2.2 Solution to exercise 2.2

With an ideal operational amplifier the gain of the circuit is:

$$A_{\infty} = -\frac{R_2}{R_1} = -\frac{220 \text{ k}\Omega}{68 \text{ k}\Omega} = -3.235$$

Due to the finite gain of the operational amplifier the loop gain $A_d\beta$ is also finite and equal to:

$$A_d\beta = 1000 \cdot \frac{68 \text{ k}\Omega}{68 \text{ k}\Omega + 220 \text{ k}\Omega} = 236.1$$

Therefore the real gain A of the circuit will be:

$$A = A_{\infty} \frac{A_d\beta}{1 + A_d\beta} = -3.235 \cdot \frac{236.1}{237.1} = -3.22$$

A.2.3 Solution to exercise 2.3

1. By applying the superposition principle we have:

$$v_u = 0.4 \text{ V} \left(1 + \frac{47 \text{ k}\Omega}{1 \text{ k}\Omega} \right) - 0.2 \text{ V} \frac{47 \text{ k}\Omega}{1 \text{ k}\Omega} = 19.2 \text{ V} - 9.4 \text{ V} = 9.8 \text{ V}$$

2. Considering a finite gain and differential resistance we have

$$A\beta = 1000 \cdot \frac{1}{48} = 20.83$$

and the input impedance at the non inverting pin will be

$$Z_i = R_{id} (1 + A\beta) = 50 \text{ k}\Omega \cdot 21.83 = 1.092 \text{ M}\Omega$$

therefore the output voltage will be given by:

$$\begin{aligned} v_u &= 0.4 \text{ V} \frac{1.09 \text{ M}\Omega}{1.09 \text{ M}\Omega + 1 \text{ M}\Omega} \cdot 48 \cdot \frac{20.83}{1 + 20.83} - 0.2 \text{ V} \cdot 47 \cdot \frac{20.83}{1 + 20.83} = \\ &= 0.4 \text{ V} \cdot 0.52 \cdot 48 \cdot 0.954 - 0.2 \text{ V} \cdot 47 \cdot 0.954 = 9.52 \text{ V} - 8.97 \text{ V} = 0.55 \text{ V} \end{aligned}$$

3. Let us assume I_B positive entering the input pins and v_{off} positive from the non inverting to the inverting pin.

$$\begin{aligned} v_u &= 0.4 \text{ V} \cdot 48 - 0.2 \text{ V} \cdot 47 - (\pm 3 \text{ mV}) \cdot 48 - 100 \text{ nA} \cdot 1 \text{ M}\Omega \cdot 48 + 100 \text{ nA} \cdot 47 \text{ k}\Omega = \\ &= 0.4 \text{ V} \cdot 48 - 0.2 \text{ V} \cdot 47 \mp 0.14 \text{ V} - 4.8 \text{ V} + 4.7 \text{ mV} = 5.005 \text{ V} \mp 0.14 \text{ V} \end{aligned}$$

A.2.4 Solution to exercise 2.4

- If the operational amplifier is ideal, the gain for signal V_S does not depend on resistor R_4 because the inverting input pin is at virtual ground. Therefore the gain is given by:

$$\frac{v_u}{V_S} = -\frac{R_2}{R_1} = -\frac{47 \text{ k}\Omega}{6.8 \text{ k}\Omega} = -6.91$$

- The β factor is given by:

$$\beta = \frac{R_1||R_4}{R_1||R_4 + R_2} = \frac{3.4 \text{ k}\Omega}{3.4 \text{ k}\Omega + 47 \text{ k}\Omega} = 0.0675$$

and we have

$$A\beta = 2000 \cdot 0.0675 = 135$$

The amplifier gain, considering a non infinite gain operational amplifier, is given by:

$$\frac{v_u}{V_S} = -\frac{R_2}{R_1} \cdot \frac{A\beta}{1 + A\beta} = -6.91 \cdot \frac{135}{136} = -6.86$$

- For the offset voltage the gain is:

$$A_{off} = 1 + \frac{R_2}{R_1||R_4} = 14.82$$

and therefore the output voltage for $V_S = 0$ is

$$v_u = \pm 10 \text{ mV} \cdot 14.82 = \pm 148.2 \text{ mV}$$

In reality since we do know neither the *maximum* value of the offset voltage nor its sign an equally correct expression would be:

$$v_u = \mp 148.2 \text{ mV}$$

A.2.5 Solution to exercise 2.5

As usual, we can apply the superposition effect and consider separately the contribution of the three effects.

1. Contribution of v_{off}

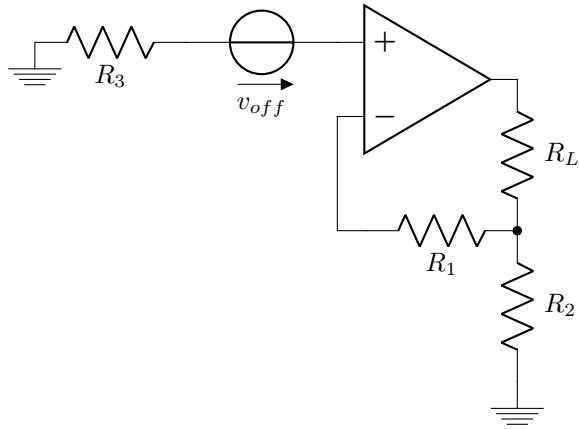


Figure A.5: Circuit for offset voltage calculation

Referring to figure A.5, we can see that the offset voltage can be considered an independent voltage source applied in series to R_3 . Being the operational amplifier ideal, no current will flow into both inputs of the amplifier, so no voltage drop will develop across R_3 and R_1 . Therefore the voltage drop on R_2 will be V_{off} . The current in R_L will also flow in R_2 .

$$I'_L = \frac{\pm v_{off}}{R_2} = \frac{\pm 6 \text{ mV}}{100 \Omega} = \pm 60 \mu\text{A}$$

2. Contribution of I_B (assumed positive entering the input pins) Referring to figure A.6, we can see that the bias current can be considered as two independent current sources in parallel to the two inputs and with the same direction. We will compute separately the two contributions.

- At the non inverting pin, the current will flow in R_3 , so that the voltage at the pin will be $-R_3 \cdot I_B$. The same voltage applies to the inverting pin. No current will flow in R_1 and the current in the load will also flow in R_2 .

$$I''_L^+ = -I_B \frac{R_3}{R_2} = -200 \text{ nA} \frac{1.2 \text{ k}\Omega}{100 \Omega} = -2.4 \mu\text{A}$$

- At the inverting pin, the current will flow in R_1 . No current will flow in R_3 , so the inverting pin is at 0 V. The voltage drop across R_2 is the same, so the current in R_L is the sum of the current in R_1 and the current in R_2 . The current in R_1 is I_B and the current in R_2 is $I_B \cdot R_1/R_2$. Therefore:

$$I''_L^- = I_B \left(1 + \frac{R_1}{R_2}\right) = 200 \text{ nA} \left(1 + \frac{1 \text{ k}\Omega}{100 \Omega}\right) = 200 \text{ nA} \cdot 11 = 2.2 \mu\text{A}$$

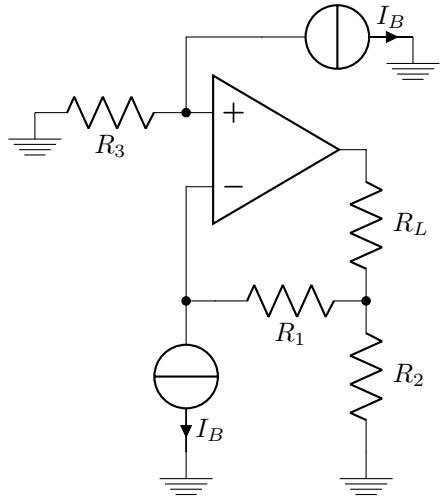


Figure A.6: Circuit for bias current calculation

$$I''_L = I''_L^+ + I''_L^- = -2.4 \mu\text{A} + 2.2 \mu\text{A} = -0.2 \mu\text{A}$$

3. Contribution of I_{off}

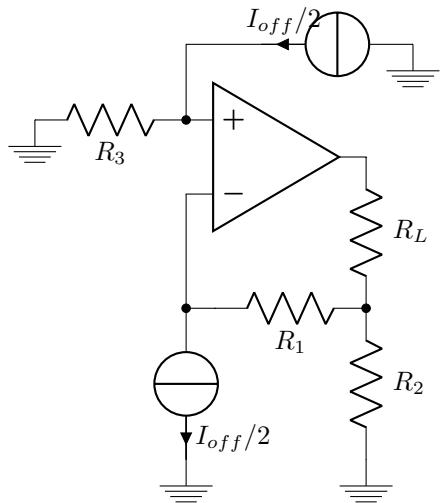


Figure A.7: Circuit for offset current calculation

Referring to figure A.7, we can see that the offset current behaves like the bias current, but the direction of one of the two generators is reversed. Therefore:

- At the non inverting pin

$$I''_L^+ = \frac{I_{off}}{2} \frac{R_3}{R_2} = \pm 50 \text{ nA} \frac{1.2 \text{ k}\Omega}{100 \Omega} = \pm 1.1 \mu\text{A}$$

- At the inverting pin

$$I_L'''- = \frac{I_{off}}{2} \left(1 + \frac{R_1}{R_2} \right) = \pm 50 \text{ nA} \left(1 + \frac{1 \text{ k}\Omega}{100 \Omega} \right) = \pm 50 \text{ nA} \cdot 11 = \pm 0.55 \mu\text{A}$$

$$I_L''' = I_L'''+ + I_L'''- = \pm 1.1 \mu\text{A} \pm 0.55 \mu\text{A} = \pm 1.65 \mu\text{A}$$

In summary

$$I_L = I_L' + I_L'' + I_L''' = \pm 60 \mu\text{A} - 0.2 \mu\text{A} \pm 1.65 \mu\text{A} = -0.2 \mu\text{A} \pm 61.65 \mu\text{A}$$

Note that we do not know the sign of offset voltage and offset current and they are two independent sources, so we have to sum the two contributions disregarding the sign.

A.2.6 Solution to exercise 2.6

The schematic of the circuit is given in figure A.8

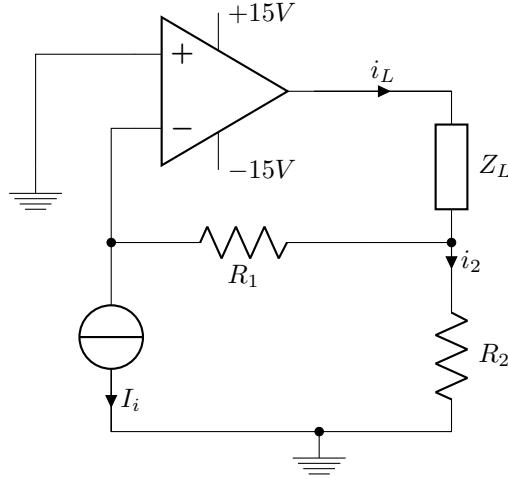


Figure A.8: Current Amplifier

Since the operational amplifier is ideal there is no current in the input pins and no voltage drop between them. Therefore:

$$i^+ = i^- = 0$$

$$V^- = 0 \text{ V}$$

We will then have:

$$V_{R1} = R_1 I_i = V_{R2} = R_2 i_2$$

$$i_2 = \frac{R_1}{R_2} I_i$$

$$i_L = i_2 + I_i = \left(1 + \frac{R_1}{R_2}\right) I_i$$

$$1 + \frac{R_1}{R_2} = 10 \rightarrow \frac{R_1}{R_2} = 9$$

$$R_1 = 9R_2$$

Let us now consider the effects of V_{off} and currents I_b and I_{off} .

Offset voltage V_{off}

Referring to figure A.9 since there is no current in R_1 and $V^+ = V^-$ we have:

$$V_{R2} = V_{off}$$

and the value of the output current i_L due to V_{off} will be given by :

$$i_L|_{V_{off}} = I_2 = \frac{V_{off}}{R_2}$$

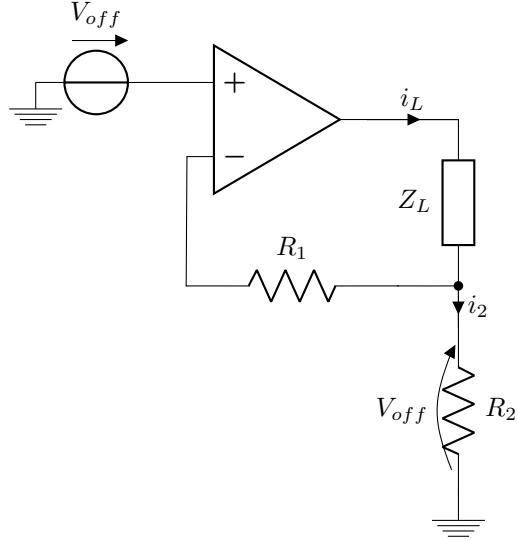


Figure A.9: Current amplifier with V_{off}

Offset and bias currents

Referring to figure A.10 we must compute the value of $i_L \Big|_{I_b - \frac{I_{off}}{2}}$

$$V^+ = V^- = 0$$

Since R_d is a large resistance and $v_d \approx 0$ there will be no current in R_1 and we have:

$$i_L \Big|_{I_b - \frac{I_{off}}{2}} = 0$$

Referring to figure A.11 we may compute the contribution due to $i_L \Big|_{I_b + \frac{I_{off}}{2}}$:

$$V_{R1} = \left(I_b + \frac{I_{off}}{2} \right) R_1 = V_{R2}$$

$$V_{R2} = R_2 i_2$$

$$i_2 = \left(I_b + \frac{I_{off}}{2} \right) \frac{R_1}{R_2}$$

$$i_L \Big|_{I_b + \frac{I_{off}}{2}} = \left(I_b + \frac{I_{off}}{2} \right) \left(1 + \frac{R_1}{R_2} \right)$$

Superposition of effects

We may now add the three components in order to determine the total output current I_L due to the input currents and to offset and bias components:

$$I_L = \left(1 + \frac{R_1}{R_2} \right) I_i + \frac{V_{off}}{R_2} + \left(I_b + \frac{I_{off}}{2} \right) \left(1 + \frac{R_1}{R_2} \right)$$

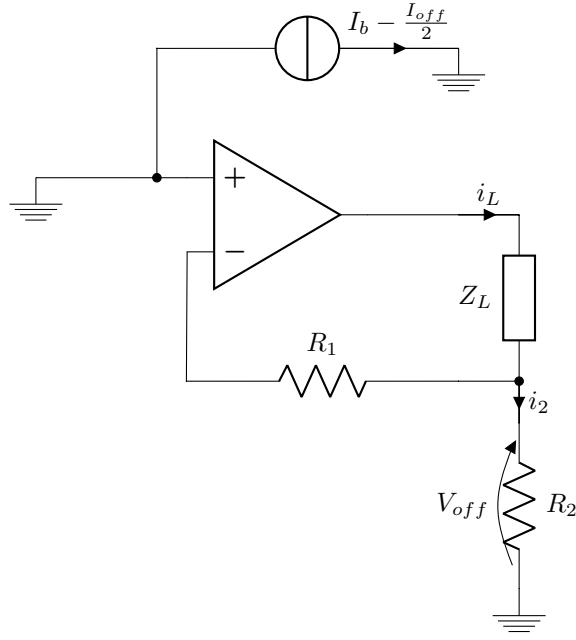


Figure A.10: Analysis of current ($I_b - \frac{I_{off}}{2}$)

Obviously it is not possible to modify the gain $\left(1 + \frac{R_1}{R_2}\right)$ therefore to minimize the offset contribution we must choose a value of R_2 large enough to make the contribution $\frac{V_{off}}{R_2}$ negligible.

There are two possible optimizations we may make.

The first one does not consider the range of the input current and minimizes the contribution of the offset voltage with respect to the offset and bias currents. In this case we have:

$$\begin{aligned} \frac{V_{off}}{R_2} &<< \left(1 + \frac{R_1}{R_2}\right) \left(I_b + \frac{I_{off}}{2}\right) \\ \left|\frac{V_{off}}{R_2}\right| &< \left|\frac{1}{10} \left(1 + \frac{R_1}{R_2}\right) \left(I_b + \frac{I_{off}}{2}\right)\right| \\ \frac{V_{off}}{R_2} &< 225 \text{ nA} \Rightarrow \frac{3 \text{ mV}}{R_2} < 125 \text{ nA} \Rightarrow R_2 > \frac{3 \text{ mV}}{125 \text{ nA}} = 24 \text{ k}\Omega \end{aligned}$$

This value is clearly incompatible with the output voltage range of the amplifier because with input current of 1 mA the output voltage should be equal to $10 \cdot 1 \text{ mA} \cdot 24 \text{ k}\Omega = 240 \text{ V}$! **Therefore we cannot use this solution.**

The second optimization considers that the contribution of the bias and offset current is already far lower than the range of the input current and then the contribution of the input offset voltage should be made small with respect to that of the minimum input current which is 0.1 mA. We have in this case:

$$\frac{V_{off}}{R_2} \ll \left(1 + \frac{R_1}{R_2}\right) I_{i_{\text{MIN}}}$$

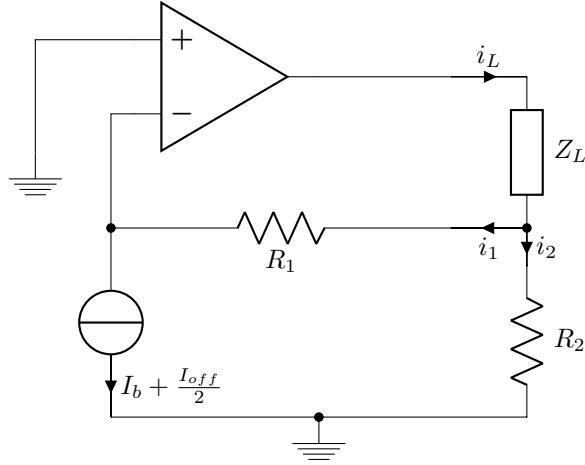


Figure A.11: Analysis of currents ($I_b + \frac{I_{off}}{2}$)

$$\left| \frac{V_{off}}{R_2} \right| < \left| \frac{1}{10} \left(1 + \frac{R_1}{R_2} \right) I_{i_{\text{MIN}}} \right|$$

$$\frac{V_{off}}{R_2} < 0.1 \text{ mA} \Rightarrow \frac{3 \text{ mV}}{R_2} < 0.1 \text{ mA} \Rightarrow R_2 > \frac{3 \text{ mV}}{0.1 \text{ mA}} = 30 \Omega$$

In order to remain in linearity the output voltage must remain within the output voltage range of the amplifier, that is we must have $V_{out,max} = 13V$ when the amplifier provides the maximum current of 10 mA which gives:

$$R_2 \cdot 9 \text{ mA} + Z_{L,max} \cdot 10 \text{ mA} < 13V$$

If we want to maximize the range of Z_L that can be used by the amplifier, we can say for example

$$R_2 \cdot 9 \text{ mA} < 3 \text{ V} \Rightarrow R_2 < 333 \Omega$$

Choosing $R_2 = 330 \Omega$ we obtain:

$$Z_{L,max} \cdot 10 \text{ mA} < 10 \text{ V}$$

$$Z_{L,max} < \frac{10V}{10 \text{ mA}} = 1 \text{ k}\Omega$$

We will then have:

$$R_2 = 330 \Omega$$

$$R_1 = 3.0 \text{ k}\Omega$$

from the E24 series.

A.2.7 Solution to exercise 2.7

The circuit shown in figure A.12 is one of the class which is usually referred to as Howland current amplifiers. If we assume the operational amplifier to be

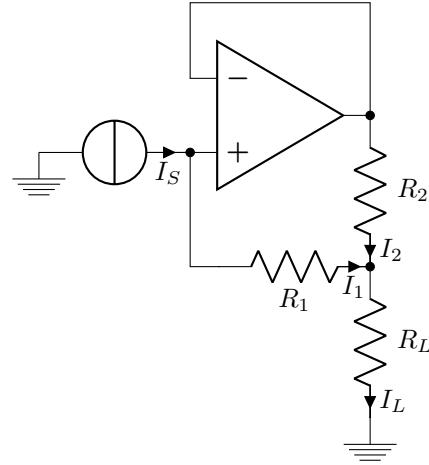


Figure A.12: Howland current Amplifier

ideal the voltage drop across the input pins will be 0 and the output voltage will be the same as the one on the non inverting pin. Assuming for currents I_1 and I_2 in resistors R_1 and R_2 the signs shown in the figure we may write:

$$I_1 R_1 = I_S R_1 = I_2 R_2$$

from which we have

$$I_2 = I_S \cdot \frac{R_1}{R_2}$$

and

$$I_L = I_S + I_2 = I_S \left(1 + \frac{R_1}{R_2} \right)$$

A.2.8 Solution to exercise 2.8

Ideally the gain of the voltage follower would be 1 if the operational amplifier were ideal. We may consider the real value of its parameters as *small* disturbances from its ideal behavior and apply the superposition principle considering each one of them separately and then summing their effects.

Limited Gain

The effect of the limited gain may be computed by the relation

$$v_u = \frac{A_d}{1 + A_d\beta} v_s = \frac{A_d}{1 + A_d} v_s = \frac{1}{1 + \frac{1}{A_d}} v_s \approx (1 - \frac{1}{A_d}) v_s$$

Substituting the numerical values we have:

$$v_u \approx (1 - \frac{1}{80000}) v_s = (1 - 1.25 \cdot 10^{-5}) v_s$$

Corresponding to a gain error $\epsilon_1 = -1.25 \cdot 10^{-5}$.

Common mode gain

The effect of the common mode gain may be evaluated by considering the $A_c v_c$ which it generates at the output of the amplifier as an *additive disturbance* (such as noise or distortion) which is reduced by the negative feedback by a factor $(1 + A_d\beta)$ (in our case $(1 + A_d)$).

Considering the differential signal v_d very small, the value of the common mode signal is $v_c \approx v_s$ and since A_d is very large the total effect on the amplifier output is

$$v_u = \frac{A_c v_s}{(1 + A_d)} \approx \frac{1}{CMRR} v_s$$

Since 100dB correspond to a ratio of 10^5 the error introduced by the common mode gain is:

$$\epsilon_2 = \pm 1 \cdot 10^{-5}$$

because we do not know the sign of the common mode gain.

Input Impedance

The voltage partition between the internal resistance of the generator and the input impedance of the amplifier introduces a further error. The nominal value of the input impedance Z_{in} is $Z_{in} = R_{id}(1 + A_d\beta) = R_{id}(1 + A_d) = 80 \text{ k}\Omega \cdot 80000 = 6.4 \text{ G}\Omega$. The effect of the input partition is given by:

$$v_u = v_s \frac{Z_{in}}{R_s + Z_{in}} = v_s \frac{1}{1 + \frac{R_s}{Z_{in}}} \approx v_s \left(1 - \frac{R_s}{Z_{in}} \right)$$

The last approximation holds if $R_s \ll Z_{in}$ and, substituting the numerical value, the error introduced by R_{id} is given by:

$$\epsilon_3 = -\frac{R_s}{Z_{in}} = -\frac{470 \text{ k}\Omega}{6.4 \text{ G}\Omega} = -7.34 \cdot 10^{-5}$$

Output Impedance

The output impedance of the amplifier is:

$$Z_o = \frac{R_o}{1 + A_d \beta} = \frac{R_o}{1 + A_d} = \frac{400 \Omega}{80000} = 5 \text{ m}\Omega$$

The effect of the output partition is given by:

$$v_u = v_s \frac{R_L}{R_L + Z_o} = v_s \frac{1}{1 + \frac{Z_o}{R_L}} \approx v_s \left(1 - \frac{Z_o}{R_L}\right)$$

The last approximation holds if $R_L \gg Z_o$ and substituting the numerical values, the error introduced by non zero R_o is given by:

$$\epsilon_4 = -\frac{Z_o}{R_L} = -\frac{5 \text{ m}\Omega}{1 \text{ k}\Omega} = -0.5 \cdot 10^{-5}$$

Total Error

The total error is the worst case sum of all the four error , that is:

$$\epsilon_{TOT} = \epsilon_1 + \epsilon_2 + \epsilon_3 + \epsilon_4 = -1.25 \cdot 10^{-5} - 1 \cdot 10^{-5} - 7.34 \cdot 10^{-5} - 0.5 \cdot 10^{-5} = -1.09 \cdot 10^{-4}$$

and the real nominal gain is 0.999899.

A.3 Solution to chapter 3 exercises

A.3.1 Solution to exercise 3.1

- Let us first observe that resistance R_3 has no influence on the circuit, being in parallel with a voltage generator. The transfer function of the circuit is:

$$\frac{V_u}{V_S} = \left(1 + \frac{Z_2}{Z_1} \right)$$

where $Z_2 = R_2 \parallel [1/(sC_2)]$ and $Z_1 = R_1 + 1/(sC_1)$. Substituting these values we obtain an expression which may be easily be put in the form:

$$\frac{V_u}{V_S} = \frac{\left(1 + \frac{s}{\omega_{z1}} \right) \cdot \left(1 + \frac{s}{\omega_{z2}} \right)}{\left(1 + \frac{s}{\omega_{p1}} \right) \cdot \left(1 + \frac{s}{\omega_{p2}} \right)}$$

where ω_{p1} and ω_{p2} are given respectively by:

$$\begin{aligned}\omega_{p1} &= \frac{1}{R_1 C_1} \\ \omega_{p2} &= \frac{1}{R_2 C_2}\end{aligned}$$

and ω_{z1} and ω_{z2} may be evaluated either by solving the equation at the numerator of the expression of the transfer function or by considering that for $\omega \rightarrow 0$ and for $\omega \rightarrow \infty$ the absolute value of the transfer function is 1 (or 0 dB in logarithmic scale). You can prove this assertion by substituting the capacitors with a short circuit for $\omega \rightarrow \infty$: $Z_2 \rightarrow 0$ and the gain is 1. If you substitute the capacitors with an open circuit for $\omega \rightarrow 0$, you obtain $Z_2 \rightarrow \infty$ and the gain is 1 again. If $\omega_{p1} \ll \omega_{p2}$ the Bode diagram will be the one shown in figure A.13 :

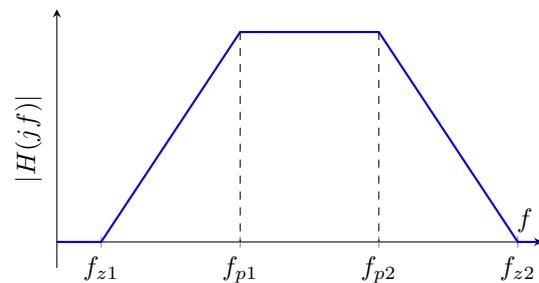


Figure A.13: Bode diagram of the amplifier transfer function

substituting the numerical values we obtain

$$\begin{aligned}f_{p2} &= \frac{1}{2\pi \cdot 100 \text{ k}\Omega \cdot 180 \text{ pF}} = \frac{1}{2\pi \cdot 100 \cdot 10^3 \cdot 180 \cdot 10^{-12}} = 8.8 \text{ kHz} \\ f_{p1} &= \frac{1}{2\pi \cdot 1 \text{ k}\Omega \cdot 1.8 \text{ }\mu\text{F}} = \frac{1}{2\pi \cdot 10^3 \cdot 1.8 \cdot 10^{-6}} = 88 \text{ Hz}\end{aligned}$$

The in band gain A between f_{p1} and f_{p2} is given by

$$A = 1 + \frac{R_2}{R_1} = 101 \rightarrow 40.6dB$$

You can find this result by observing that, starting one decade above f_{p1} , C_1 behaves as a short circuit while, until one decade below f_{p2} , C_2 behaves as an open circuit.

The frequencies of the two zeroes may be easily obtained as:

$$\begin{aligned} f_{z1} &= \frac{f_{p1}}{A} = \frac{88\text{ Hz}}{101} = 0.87\text{ Hz} \\ f_{z2} &= f_{p2} \cdot A = 8.8\text{ kHz} \cdot 101 = 889\text{ kHz} \end{aligned}$$

2. If the input signal $V_S = 0$ then at the output there is the contribution of the offset signal. For the offset voltage the circuit operates as a voltage follower, while the bias currents of the two inputs flow respectively in R_1 and in $(R_3||0) = 0$ because the input voltage generator is shorted.

We then have:

$$V_u = v_{off} \cdot 1 + I_b \cdot R_2 - I_b \cdot 0 = \pm 3\text{ mV} + 50\text{ nA} \cdot 100\text{ k}\Omega = (\pm 3 + 5)\text{ mV} = 2\text{ mV} \div 8\text{ mV}$$

A.3.2 Solution to exercise 3.2

We can compute the transfer function in two steps:

$$\frac{V_u}{V_S} = \frac{V^+}{V_S} \cdot \frac{V_u}{V^+}$$

We can do this because no current flows in the non inverting input of the operational amplifier (except for the bias current, which can be neglected for transfer function calculations).

With easy calculations, we find that:

$$\begin{aligned}\frac{V^+}{V_S} &= \frac{sC_3R_3}{sC_3R_3 + 1} \\ \frac{V_u}{V^+} &= \frac{R_1 + R_2}{R_1} \frac{sC_2R_1 \parallel R_2 + 1}{sC_2R_2 + 1}\end{aligned}$$

and finally:

$$\frac{V_u}{V_S} = \frac{R_1 + R_2}{R_1} \frac{(sC_2R_1 \parallel R_2 + 1) \cdot sC_3R_3}{(sC_2R_2 + 1)(sC_3R_3 + 1)}$$

1. The transfer function is of the type shown in figure A.14: one zero in DC, two poles and a second zero at high frequency. The high frequency gain is 0 dB, and that can be seen either by computing the limit for $s \rightarrow \infty$ of the transfer function or, more evidently, by substituting two short circuits to the capacitors.

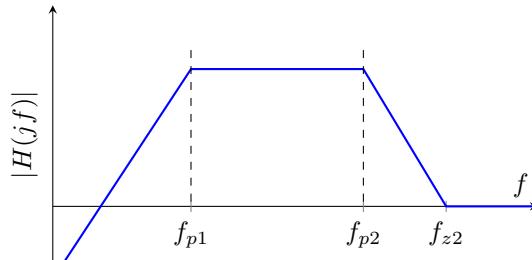


Figure A.14: Transfer function Bode diagram

and we have:

$$\begin{aligned}f_{p2} &= \frac{1}{2\pi \cdot R_2 C_2} = \frac{1}{2\pi \cdot 120 \text{ k}\Omega \cdot 220 \text{ pF}} = 6.03 \text{ kHz} \\ f_{p1} &= \frac{1}{2\pi \cdot R_3 C_3} = \frac{1}{2\pi \cdot 330 \text{ k}\Omega \cdot 27 \text{ nF}} = 17.9 \text{ Hz}\end{aligned}$$

The in band gain A is given by:

$$A = \left(1 + \frac{R_2}{R_1}\right) = 1 + \frac{120 \text{ k}\Omega}{8.2 \text{ k}\Omega} = 15.63 \rightarrow 23.9 \text{ dB}$$

Considering that for $f \rightarrow \infty$ the gain is 1, the frequency of the zero may be easily computed as:

$$f_{z1} = f_{p1} \cdot A = 6.03 \text{ kHz} \cdot 15.63 = 94.3 \text{ kHz}$$

2. The gain for the offset voltage generator v_{off} is also A and its contribution to the output voltage is:

$$v'_u = -v_{off} \cdot A = \pm 3 \text{ mV} \cdot 15.63 = \pm 46.9 \text{ mV}$$

The contribution of the bias current I_B is:

$$v''_u = -I_B \cdot R_3 \cdot A + I_B \cdot R_2 = -50 \text{ nA} \cdot 330 \text{ k}\Omega \cdot 15.63 + 50 \text{ nA} \cdot 120 \text{ k}\Omega = -252 \text{ mV}$$

The total is:

$$v_u = \pm 46.9 \text{ mV} - 252 \text{ mV} = -205 \text{ mV} \div -299 \text{ mV}$$

A.3.3 Solution to exercise 3.3

1. In our case we see that $R_1 = R_3$ and $R_2 = R_4$, so the amplifier is a pure differential amplifier, so that:

$$A_d = \frac{R_2}{R_1} = \frac{330 \text{ k}\Omega}{12 \text{ k}\Omega} = 27.5$$

$$A_C = 0$$

2. If the resistors have 1% tolerance, we can compute the gains $A_1 = V_u/V_1$ and $A_2 = V_u/V_2$:

$$A_1 = \frac{R_4}{R_3 + R_4} \cdot \frac{R_1 + R_2}{R_1}$$

$$A_2 = -\frac{R_2}{R_1}$$

Let us start considering maximum and minimum value for A_2 . We have

$$A_{2\text{MAX}} = -\frac{333.3 \text{ k}\Omega}{11.88 \text{ k}\Omega} = -28.06$$

$$A_{2\text{MIN}} = -\frac{326.7 \text{ k}\Omega}{12.12 \text{ k}\Omega} = -26.955$$

The case of A_1 is more complex because we can play with the combination of four elements. By trying all the combinations we can find that the minimum value is achieved when $R_2 = 326.7 \text{ k}\Omega$, $R_1 = 12.12 \text{ k}\Omega$, $R_3 = 11.88 \text{ k}\Omega$, $R_4 = 333.3 \text{ k}\Omega$ and the computed value is $A_1 = 26.993$.

The maximum value is found when all the resistors are at the opposite end of their tolerance range. In that case $A_1 = 28.016$.

Obviously R_2 and R_1 contribute to both values, so we have to pair the values of the two gains to the same values of the resistors.

Once we have the two gains, we can convert to differential and common mode gain:

$$A_{d\text{MAX}} = \frac{A_{1\text{MAX}} - A_{2\text{MAX}}}{2} = 28.036$$

$$A_C = A_{1\text{MAX}} + A_{2\text{MAX}} = -0.039$$

On the other end of the tolerance range we have:

$$A_{d\text{MIN}} = \frac{A_{1\text{MIN}} - A_{2\text{MIN}}}{2} = 26.974$$

$$A_C = A_{1\text{MIN}} + A_{2\text{MIN}} = 0.038$$

To compute CMRR, we note that the ratio A_d/A_C is the same in both cases. The minimum CMRR is then

$$CMRR = 20 \log_{10} \left(\frac{A_d}{A_C} \right) = 20 \log_{10}(710) = 57 \text{ dB}$$

A.3.4 Solution to exercise 3.4

1. We start computing the voltages at the inverting and non inverting inputs of the differential amplifier when we apply the stretching to the strain gauge. The resulting value for R_B is:

$$R_B = R_{B_{\text{NOM}}} + \Delta R_B = 1005 \Omega$$

We use this value and obtain:

$$\begin{aligned} V^- &= \frac{R_A}{R_A + R_A} \cdot 5 \text{ V} = 2.5 \text{ V} \\ V^+ &= \frac{R_B}{R_B + R_A} \cdot 5 \text{ V} = 2.50623 \text{ V} \end{aligned}$$

From those values we can then compute differential and common mode voltages at the input of the differential amplifier:

$$\begin{aligned} V_C &= \frac{V^+ + V^-}{2} = 2.5031 \text{ V} \\ V_d &= V^+ - V^- = 6.23 \text{ mV} \end{aligned}$$

To obtain $V_u = 5 \text{ V}$ we need a differential gain:

$$A_d = \frac{5 \text{ V}}{6.23 \text{ mV}} = 802$$

The common mode gain should be very low not to influence the output value, given that the input common mode voltage is very high compared to the differential voltage.

A.3.5 Solution to exercise 3.5

1. If $R_{S1} = R_{S2} = 0$ the function implemented by the circuit is:

$$V_u = V_{S1} \left(1 + \frac{R_2}{R_1} \right) - V_{S2} \frac{R_2}{R_1} = V_{S1} \left(1 + \frac{120 \text{ k}\Omega}{6.8 \text{ k}\Omega} \right) - V_{S2} \frac{120 \text{ k}\Omega}{6.8 \text{ k}\Omega} = \\ = 18.65V_{S1} - 17.65V_{S2}$$

2. If $R_{S1} \neq 0$ and $R_{S2} \neq 0$ the function is given by:

$$V_u = V_{S1} \left(1 + \frac{R_2}{R_1 + R_{S2}} \right) - V_{S2} \frac{R_2}{R_1 + R_{S2}} = V_{S1} \left(1 + \frac{120 \text{ k}\Omega}{8.8 \text{ k}\Omega} \right) - V_{S2} \frac{120 \text{ k}\Omega}{8.8 \text{ k}\Omega} = \\ = 14.64V_{S1} - 13.64V_{S2}$$

A.3.6 Solution to exercise 3.6

Recalling the theory, in our case we have $A_N = 2 + 1 = 3$ and $A_P = 3$. Since we have $A_P \neq A_N + 1$ we must add a resistor on the non-inverting side of the adder, so that we balance the gains:

$$V_u = 3V_1 - 2V_2 - V_3 + 1 \cdot 0 \text{ V}$$

The adder circuit which implements the desired function is shown in figure A.15 .

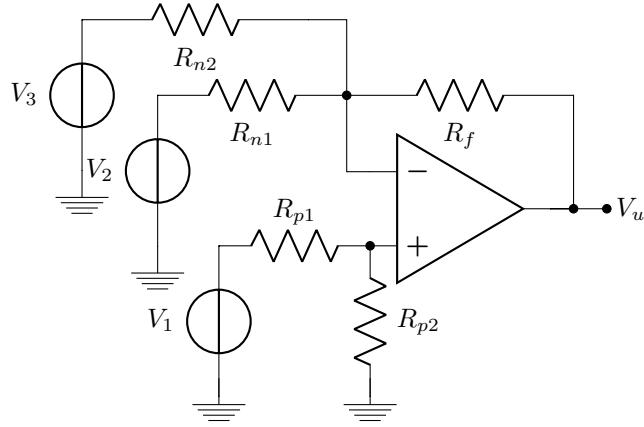


Figure A.15: Adder circuit

On the inverting side we have: $R_{n1} = R_f/2$ and $R_{n2} = R_f$. On the non-inverting side: $R_{p1} = R_f/3$ and $R_{p2} = R_f$.

The gain for generator v_{off} is 4. Therefore for $v_1 = v_2 = v_3 = 0$ the output voltage wil be :

$$V_u = \pm 4 \text{ mV} \cdot 4 = \pm 16 \text{ mV}$$

A.3.7 Solution to exercise 3.7

From the theory we know that the differential gain of the instrumentation amplifier is:

$$A_d = \left(1 + \frac{R_a + R_b}{R}\right) \cdot \frac{R_2}{R_1}$$

We also recall that the conditions to make the third amplifier differential are $R_4 = R_2$ and $R_3 = R_1$.

1. There are infinite solutions to obtain $A_d = 5$. Probably the simplest one is to set $R_a = R_b = R$. In this case the contribution of the first two amplifiers is 3 and the expression of A_d becomes:

$$A_d = 3 \cdot \frac{R_2}{R_1} = 5 \implies \frac{R_2}{R_1} = \frac{5}{3}$$

Therefore we have: $R_1 = R_3 = 30 \text{ k}\Omega$; all other resistors are $50 \text{ k}\Omega$.

2. If we add a resistor R_p in parallel to R , the expression of the gain changes to:

$$A_d = \left(\frac{R_a + R_b}{R} + 1\right) \cdot \frac{R_2}{R_1} + \frac{R_a + R_b}{R_p} \cdot \frac{R_2}{R_1}$$

given the previous choices we made, we can rework the expression to obtain:

$$R_p = \frac{10}{3} \frac{R}{A_d - 5}$$

To obtain $A_d = 10$, $R_p = 33.3 \text{ k}\Omega$.

3. We use the same formula as above. To obtain $A_d = 50$, we need $R_p = 3.7 \text{ k}\Omega$.

A.3.8 Solution to exercise 3.8

1. To compute DC asymptotic value of the gain we substitute an open circuit to the capacitor C_1 . Resistors R_2 and R_3 are in series, so the DC gain of the circuit is:

$$A_{V0} = -\frac{R_2 + R_3}{R_1} = -\frac{39 \text{ k}\Omega + 120 \text{ k}\Omega}{12 \text{ k}\Omega} = -13.25 \Rightarrow 22.4 \text{ dB}$$

In high frequency, capacitor C_1 is a short circuit and R_3 is shorted out of the circuit, so we have

$$A_{V\infty} = -\frac{R_2}{R_1} = -\frac{39 \text{ k}\Omega}{12 \text{ k}\Omega} = -3.25 \Rightarrow 10.2 \text{ dB}$$

2. We easily find the transfer function considering $Z_3 = R_3 \parallel C_1$ in our equation:

$$\frac{V_u(s)}{V_i(s)} = -\frac{1}{R_1} \left(R_2 + \frac{R_3}{sC_1R_3 + 1} \right) = -\frac{R_2 + R_3}{R_1} \cdot \frac{s(C_1R_3 \parallel R_2) + 1}{sC_1R_3 + 1}$$

3. We need to find the frequencies of the pole and of the zero in the transfer function:

$$f_{p1} = \frac{1}{2\pi R_3 C_1} = 603 \text{ Hz}$$

$$f_{z1} = \frac{1}{2\pi C_1 R_3 \parallel R_2} = 2458 \text{ Hz}$$

The Bode plot is available in figure A.16.

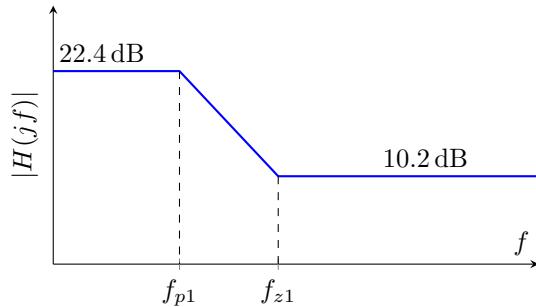


Figure A.16: Transfer function Bode diagram

A.3.9 Solution to exercise 3.9

1. The figure A.17 shows the simplified circuit. In details:

- Resistors R_8 and R_9 are loads for the two operational amplifiers but do not influence the transfer function so they can be removed (open circuit).
- Resistor R_1 is connected between two ideal voltage generators, so it draws current from the generators but it does not change the transfer function.
- resistor R_{10} is in parallel to an ideal voltage source. It draws current from it but does not modify the transfer function.
- resistors R_2 and R_3 are in series with the non inverting input of an operational amplifiers, so, given that no currents flows through them, they can be substituted by short circuits when computing the transfer functions. They have a role when considering offset currents, but they can be neglected otherwise.

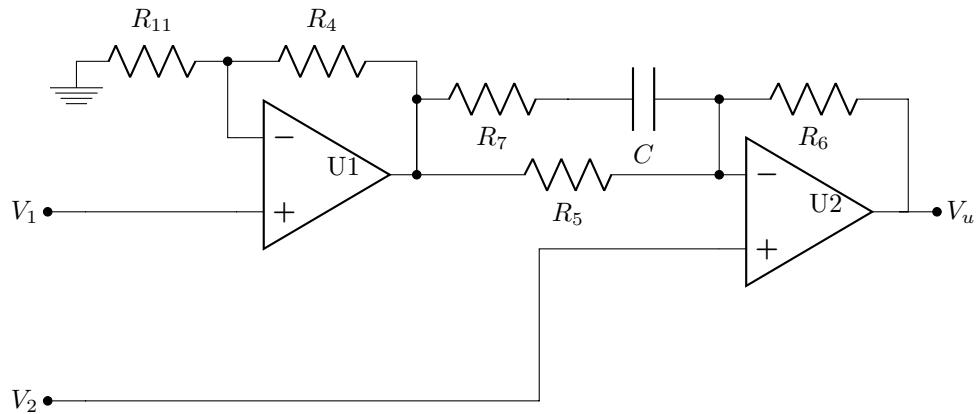


Figure A.17: Circuit schematic diagram

2. To compute DC asymptotic value of the gain we substitute an open circuit to the capacitor C . Resistor R_7 is removed, so we have:

$$\frac{V_u}{V_1} = \left(1 + \frac{R_4}{R_{11}}\right) \cdot \left(-\frac{R_6}{R_5}\right) = -\left(1 + \frac{12 \text{ k}\Omega}{12 \text{ k}\Omega}\right) \cdot \frac{120 \text{ k}\Omega}{120 \text{ k}\Omega} = -2 \Rightarrow 6 \text{ dB}$$

$$\frac{V_u}{V_2} = 1 + \frac{R_6}{R_5} = 1 + \frac{120 \text{ k}\Omega}{120 \text{ k}\Omega} = 2 \Rightarrow 6 \text{ dB}$$

3. In high frequency, capacitor C is a short circuit and R_7 is in parallel to R_5 , so we have

$$\frac{V_u}{V_1} = \left(1 + \frac{R_4}{R_{11}}\right) \cdot \left(-\frac{R_6}{R_5 \parallel R_7}\right) = -\left(1 + \frac{12 \text{ k}\Omega}{12 \text{ k}\Omega}\right) \cdot \frac{120 \text{ k}\Omega}{9.23 \text{ k}\Omega} = -26 \Rightarrow 28.3 \text{ dB}$$

$$\frac{V_u}{V_2} = 1 + \frac{R_6}{R_5 \parallel R_7} = 1 + \frac{120 \text{ k}\Omega}{9.23 \text{ k}\Omega} = 14 \Rightarrow 22.9 \text{ dB}$$

4. To ease the computation of transfer functions we introduce:

$$Z_5 = R_5 \parallel \left(\frac{1}{sC} + R_7 \right) = \frac{R_5(sCR_7 + 1)}{sC(R_5 + R_7) + 1}$$

$$K = \left(1 + \frac{R_4}{R_{11}} \right) = 2$$

With these definitions, we can write:

$$\frac{V_u}{V_1} = K \cdot \left(-\frac{R_6}{Z_5} \right) = -K \cdot \frac{R_6}{R_5} \cdot \frac{sC(R_5 + R_7) + 1}{sCR_7 + 1}$$

$$\frac{V_u}{V_2} = 1 + \frac{R_6}{Z_5} = 1 + \frac{R_6}{R_5} \cdot \frac{sC(R_5 + R_7) + 1}{sCR_7 + 1} = \left(1 + \frac{R_6}{R_5} \right) \cdot \frac{sC \frac{R_5R_6 + R_5R_7 + R_6R_7}{R_5 + R_6} + 1}{sCR_7 + 1}$$

5. We need to find the frequencies of the pole and of the zeros in the transfer functions:

$$f_{p1} = \frac{1}{2\pi R_7 C} = 3386 \text{ Hz}$$

$$f_{z1} = \frac{1}{2\pi C(R_5 + R_7)} = 260.5 \text{ Hz}$$

$$f_{z2} = \frac{1}{2\pi C \frac{R_5R_6 + R_5R_7 + R_6R_7}{R_5 + R_6}} = 483.8 \text{ Hz}$$

We have an alternative way of finding the zeros, by using the ratio of the high and low frequency gains for the two transfer functions:

$$f_{z1} = \frac{f_{p1}}{13} = 260.5 \text{ Hz}$$

$$f_{z2} = \frac{f_{p1}}{7} = 483.8 \text{ Hz}$$

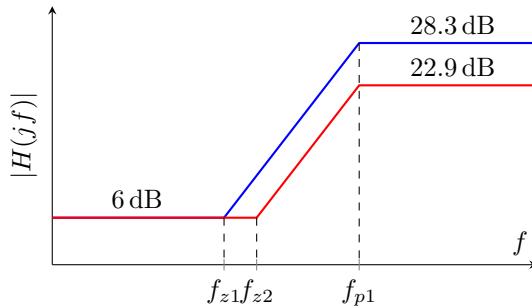


Figure A.18: Bode diagram of V_u/V_1 (blue) and V_u/V_2 (red)

The Bode plots are available in figure A.18.

A.4 Solution to chapter 4 exercises

A.4.1 Solution to exercise 4.1

The schematic of the circuit is the one depicted in figure 4.4. We recall that

$$V_{PA} = -V_{REF} \frac{R_1}{R_3}$$

Given that $V_{PA} < 0$ we use the positive power supply, $V_R = 12\text{ V}$, so that:

$$-3\text{ V} = -12\text{ V} \frac{R_1}{R_3} \implies \frac{R_1}{R_3} = \frac{3}{12} \implies R_3 = 4R_1$$

The gain of the circuit when the rectifier is active is

$$A_V = -\frac{R_2}{R_1} = -4 \implies R_2 = 4R_1$$

If we set $R_2 = R_3 = 27\text{ k}\Omega$ we compute $R_1 = 6.75\text{ k}\Omega$ which is very close to the E12 standard resistor $6.8\text{ k}\Omega$.

A.4.2 Solution to exercise 4.2

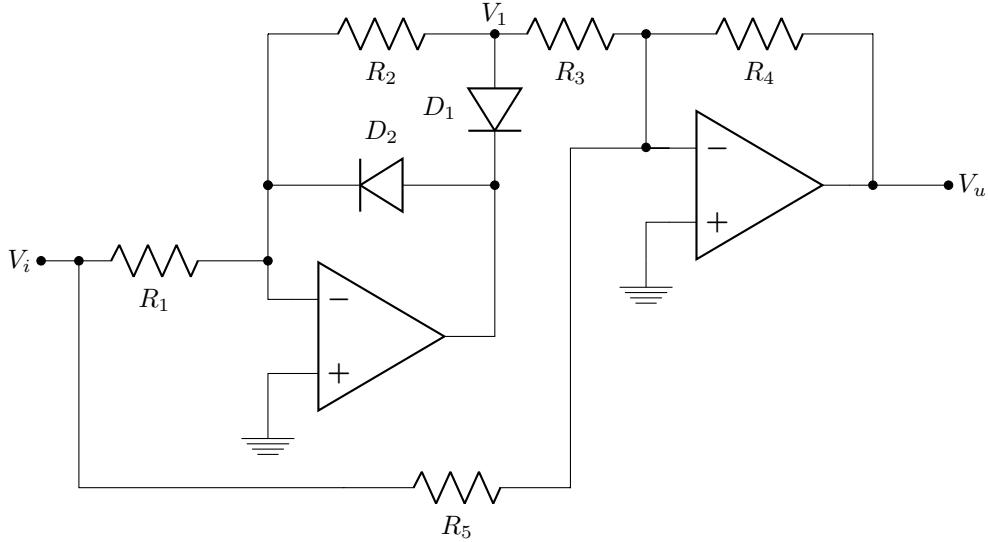


Figure A.19: Schematic of the full wave rectifier

$$V_u = -V_1 \frac{R_4}{R_3} - V_i \frac{R_4}{R_5}$$

$$V_1 = -V_i \frac{R_2}{R_1}$$

$$V_u = V_i \frac{R_2}{R_1} \frac{R_4}{R_3} - V_i \frac{R_4}{R_5} = V_i \left(\frac{R_2}{R_1} \frac{R_4}{R_3} - \frac{R_4}{R_5} \right)$$

When:

$$V_i < 0 \Rightarrow V_1 = 0$$

We have:

$$V_u = -V_i \frac{R_4}{R_5}$$

$$\frac{R_4}{R_5} = 2 \implies R_4 = 2R_5$$

When

$$V_i > 0 \Rightarrow V_1 \neq 0$$

We have:

$$V_u = V_i \left(\frac{R_2}{R_1} \frac{R_4}{R_3} - \frac{R_4}{R_5} \right)$$

$$\frac{R_2}{R_1} \frac{R_4}{R_3} - \frac{R_4}{R_5} = 2$$

$$\frac{R_2}{R_1} \frac{R_4}{R_3} = 4 \Rightarrow \begin{cases} \frac{R_2}{R_1} = 4; \frac{R_4}{R_3} = 1 \\ \frac{R_2}{R_1} = 1; \frac{R_4}{R_3} = 4 \end{cases}$$

Choice number 1 : $\frac{R_2}{R_1} = 4$; $\frac{R_4}{R_3} = 1$

This choice optimizes the contribution of the offset voltages, reducing the effect of the components of the second amplifier. However it reduces the input voltage range of the system and is used only when high precision is required.

Choice number 2 : $\frac{R_2}{R_1} = 4$; $\frac{R_4}{R_3} = 1$

This choice is the most common one and maximizes the input voltage range of the circuit.

We choose then:

$$R_2 = R_1 = 10 \text{ k}\Omega$$

$$R_3 = 11 \text{ k}\Omega$$

$$R_4 = 4R_3 = 44 \text{ k}\Omega$$

$$R_5 = 22 \text{ k}\Omega$$

A.4.3 Solution to exercise 4.3

The first step is to compute the slopes of the two lines of the characteristic:

$$A_V^- = \frac{3V - 1V}{-3V + 2V} = -2$$

$$A_V^+ = \frac{1V - 0V}{-2V - 4V} = -\frac{1}{6}$$

The schematic of the circuit is the one drawn in figure 4.9.

From the computed slopes we can find the following relations:

$$A_V^- = -\frac{R_4}{R_5} = -2 \implies R_5 = \frac{1}{2}R_4$$

$$A_V^+ = \frac{R_4R_2}{R_3R_1} - \frac{R_4}{R_5} = -\frac{1}{6}$$

If we choose $R_2 = R_1$ to have a unity gain precision rectifier we have:

$$\frac{R_4}{R_3} = -\frac{1}{6} + 2 = \frac{11}{6} \implies R_3 = \frac{6}{11}R_4$$

Now we need to fix reference voltages and associated resistors. The corner input voltage is -2 V. This implies

$$V_{PA} = -V_{R1} \frac{R_1}{R_6} = -2V \implies V_{R1} = +V_{AL} = 5V$$

$$\frac{R_1}{R_6} = \frac{2}{5} \implies R_6 = \frac{5}{2}R_1$$

Next we consider the characteristic at the left of the corner point:

$$V_u = -V_i \frac{R_4}{R_5} - V_{R2} \frac{R_4}{R_7} = -2V_i - V_{R2} \frac{R_4}{R_7}$$

By entering the coordinates of the corner point $(-2,1)$ we obtain:

$$1V = 4V - V_{R2} \frac{R_4}{R_7} \implies V_{R2} \frac{R_4}{R_7} = 3V$$

Again using the supply voltages as reference voltages we get $V_{R2} = +V_{AL} = 5V$ so that

$$\frac{R_4}{R_7} = \frac{3V}{5V} \implies R_7 = \frac{5}{3}R_4$$

If we fix reasonable values for R_4 and R_2 we can set the values of all resistors. For example:

$$\begin{aligned} R_2 &= 22\text{k}\Omega & R_1 &= 22\text{k}\Omega & R_6 &= 55\text{k}\Omega \\ R_4 &= 33\text{k}\Omega & R_5 &= 16.5\text{k}\Omega & R_3 &= 18\text{k}\Omega \\ R_7 &= 55\text{k}\Omega \end{aligned}$$

If we need to use E12 resistors then we have to approximate the values of $R_6 = R_7 = 56\text{k}\Omega$ and $R_5 = 18\text{k}\Omega$, all other resistors being already in the series.

A.5 Solution to chapter 5 exercises

A.5.1 Solution to exercise 5.1

The two thresholds of the comparator are respectively given by:

$$V_{s2} = -V_{OH} \frac{R_1}{R_2} = -10 \text{ V} \frac{18 \text{ k}\Omega}{47 \text{ k}\Omega} = -3.82 \text{ V}$$

and by:

$$V_{s1} = -V_{OL} \frac{R_1}{R_2} = 8 \text{ V} \frac{18 \text{ k}\Omega}{47 \text{ k}\Omega} = 3.06 \text{ V}$$

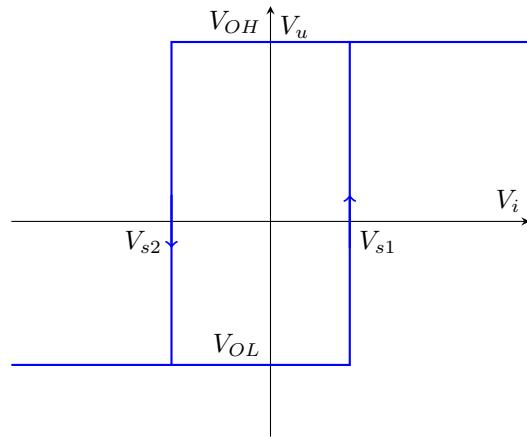


Figure A.20: Non inverting comparator

A.5.2 Solution to exercise 5.2

We may substitute the voltage divider at the non inverting pin with its Thevenin equivalent obtaining the circuit of figure A.21. The circuit has now the canonical

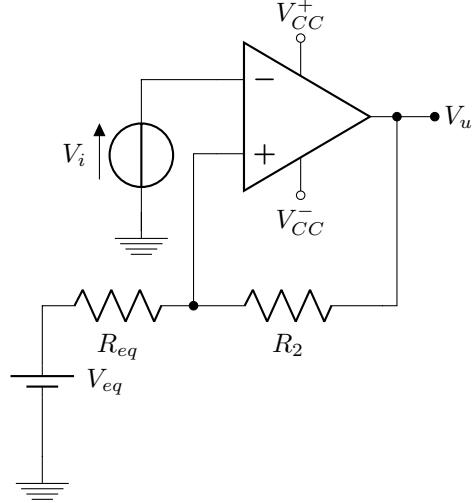


Figure A.21: Simplified circuit

form of an inverting voltage comparator with hysteresis where $v_{umax} = -12\text{ V}$ and $v_{umin} = 0\text{ V}$. We have:

$$V_{eq} = V_{CC}^+ \frac{R_3}{R_1 + R_3} = 1.09\text{ V}$$

$$R_{eq} = R_1 \parallel R_3 = 16.36\text{ k}\Omega$$

The threshold voltages are given respectively by:

$$\begin{aligned} v_{s1} &= V_{OH} \frac{R_{eq}}{R_{eq} + R_2} + V_{eq} \frac{R_2}{R_{eq} + R_2} = \\ &= 10\text{ V} \frac{16.36\text{ k}\Omega}{68\text{ k}\Omega + 16.36\text{ k}\Omega} + 1.09\text{ V} \frac{68\text{ k}\Omega}{68\text{ k}\Omega + 16.36\text{ k}\Omega} = 2.81\text{ V} \\ v_{s2} &= V_{OL} \frac{R_{eq}}{R_{eq} + R_2} + V_{eq} \frac{R_2}{R_{eq} + R_2} = \\ &= 0\text{ V} \frac{16.36\text{ k}\Omega}{68\text{ k}\Omega + 16.36\text{ k}\Omega} + 1.09\text{ V} \frac{68\text{ k}\Omega}{68\text{ k}\Omega + 16.36\text{ k}\Omega} = 0.88\text{ V} \end{aligned}$$

The complete transfer characteristic is given in figure A.22

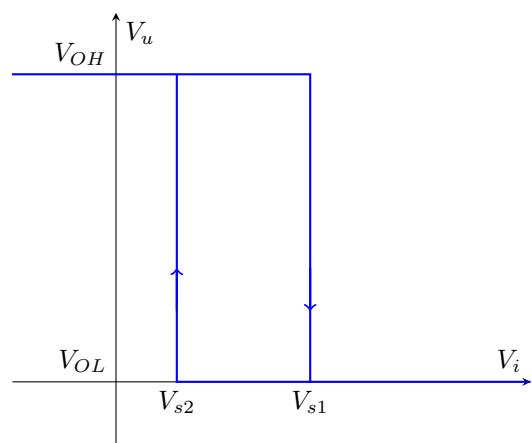


Figure A.22: Input/Output transfer characteristic

A.5.3 Solution to exercise 5.3

Since $V_{umax} = -V_{umin}$ the frequency of the waveform is given by the relation:

$$f = \frac{1}{4RC \frac{R_1}{R_2}} = \frac{1}{4 \cdot 10 \text{ k}\Omega \cdot 33 \text{ nF} \cdot \frac{18 \text{ k}\Omega}{82 \text{ k}\Omega}} = 3.541 \text{ kHz}$$

A.6 Solution to chapter 6 exercises

A.6.1 Solution to exercise 6.1

- In the case of ideal diode, given that its cathode is connected to ground via R_2 , the diode is a short circuit when anode voltage is positive. So when the input voltage is V_H , the output voltage is:

$$V_{oH} = V_H \cdot \frac{R_2}{R_2 + R_1} = 3 \text{ V} \cdot \frac{2 \text{ k}\Omega}{3 \text{ k}\Omega} = 2 \text{ V}$$

When the input voltage is negative, the diode does not conduct and $V_{oL} = 0 \text{ V}$.

- in case of a real diode with $V_\gamma = 0.7 \text{ V}$, the diode will conduct only if the input voltage is $V_i > V_\gamma$. There will be a voltage drop across the diode $v_d = V_\gamma$.

When the input voltage is V_H the diode will conduct and the output voltage will be:

$$V_{oH} = (V_H - V_\gamma) \cdot \frac{R_2}{R_2 + R_1} = 2.3 \text{ V} \cdot \frac{2 \text{ k}\Omega}{3 \text{ k}\Omega} = 1.53 \text{ V}$$

When the input voltage is negative, the diode does not conduct and $V_{oL} = 0 \text{ V}$.

A.6.2 Solution to exercise 6.2

- In the case of ideal diode, given that its anode is connected to ground via R_2 , the diode is a short circuit when cathode voltage is negative. So when the input voltage is V_L , the output voltage is:

$$V_{oL} = V_L \cdot \frac{R_2}{R_2 + R_1} = -6 \text{ V} \cdot \frac{2 \text{ k}\Omega}{3 \text{ k}\Omega} = -4 \text{ V}$$

When the input voltage is positive, the diode does not conduct and $V_{oL} = 0 \text{ V}$.

- in case of a real diode with $V_\gamma = 0.7 \text{ V}$, the diode will conduct only if the input voltage is $V_i < -V_\gamma$. There will be a voltage drop across the diode $v_d = V_\gamma$.

When the input voltage is V_L the diode will conduct and the output voltage will be:

$$V_{oL} = (V_L + V_\gamma) \cdot \frac{R_2}{R_2 + R_1} = -5.3 \text{ V} \cdot \frac{2 \text{ k}\Omega}{3 \text{ k}\Omega} = -3.53 \text{ V}$$

When the input voltage is positive, the diode does not conduct and $V_{oH} = 0 \text{ V}$.

A.6.3 Solution to exercise 6.3

- When V_Z is positive, the zener diode behaves like an open circuit if $V_i < V_{Z0}$, so $V_Z = V_i$ in that region.

When $V_i > V_{Z0}$, the equivalent circuit changes to the one of figure A.23

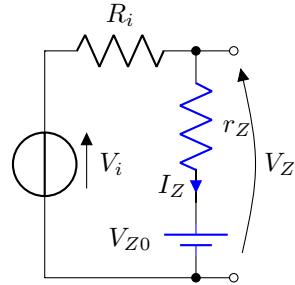


Figure A.23: Equivalent circuit in breakdown region

In this region, we can compute V_Z with the following equation:

$$V_Z = (V_i - V_{Z0}) \cdot \frac{r_Z}{R_i + r_Z} + V_{Z0} = \frac{1}{11} \cdot V_i + 4.55 \text{ V}$$

When $V_Z < 0 \text{ V}$, The diode behaves like a normal silicon diode, so it will be an open circuit if $V_Z > -V_\gamma$, a short circuit otherwise.

The complete transfer function is in figure A.24

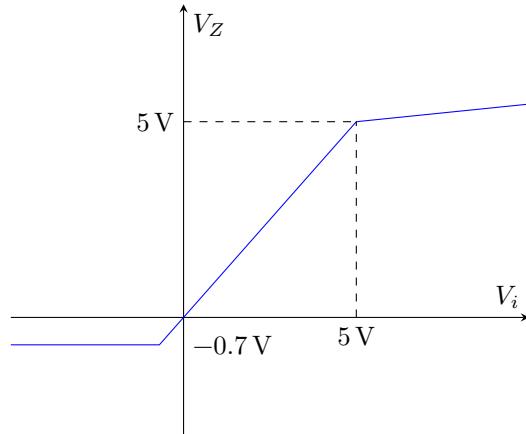


Figure A.24: DC transfer function

- The maximum value of V_Z derives from the maximum value of V_i

$$V_{Z_{\text{MAX}}} = 5 \text{ V} \cdot \frac{1}{11} + 5 \text{ V} = 5.45 \text{ V}$$

Input and output voltages are coincident in the region from $V_i = -0.7 \text{ V}$ to $V_i = 5 \text{ V}$. The time diagram is in figure A.27.

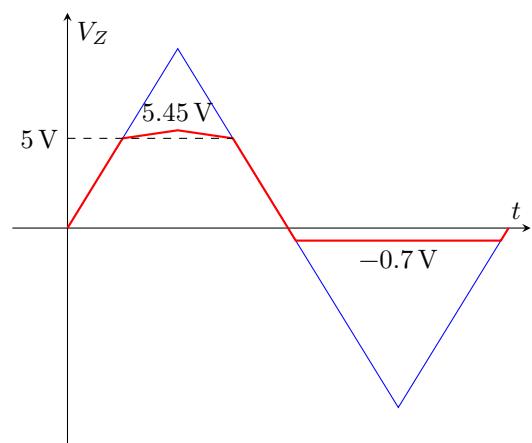


Figure A.25: Input waveform (blue), V_z (red)

A.6.4 Solution to exercise 6.4

- The two zener diodes are in anti-series configuration: when $V_O > 0 \text{ V}$, D_{Z1} is forward biased and D_{Z2} is reverse biased, the other way around when $V_O < 0 \text{ V}$. Given that the diodes are identical, they will only conduct if $|V_O| > V_{Z0} + V_\gamma$. In that case they will block any further increase of the magnitude of V_O . In the text of the exercise we specified $r_Z = 0 \Omega$, so the curve will be flat. In case we have to take into account r_Z , we have to use a more complex equivalent circuit for the diodes and take into account a resulting voltage divider. If $|V_O| < V_{Z0} + V_\gamma$ the diodes will be open circuit and will not influence the value of V_O .

V_O depends on V_i :

$$V_O = \frac{R_L}{R_L + R} V_i = \frac{9}{10} V_i$$

The output voltage will be limited to $|V_O| \leq \pm(V_{Z0} + V_\gamma) = \pm 4.7 \text{ V}$. This corresponds to

$$V_i = \pm 4.7 \text{ V} \cdot \frac{10}{9} = \pm 5.22 \text{ V}$$

The resulting DC transfer function is available in figure A.26.

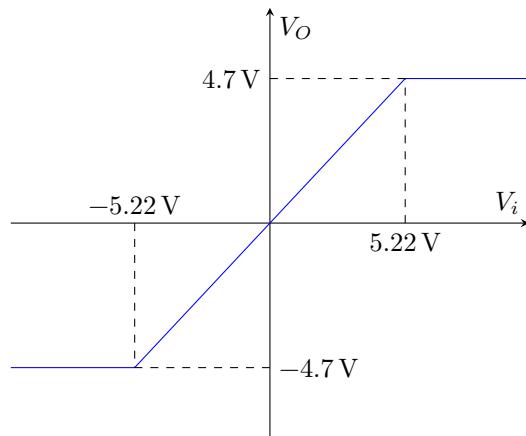


Figure A.26: DC transfer function

- we can draw the time diagram according to the just found DC transfer function. The resulting waveforms are in figure ??

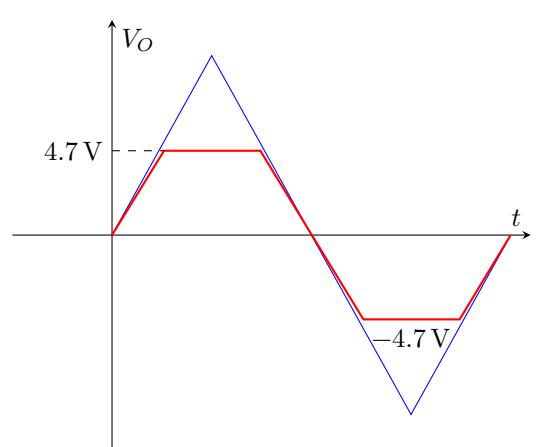


Figure A.27: Input waveform (blue), output waveform (red)

A.6.5 Solution to exercise 6.5

- To find the quiescent operating point we can use the corresponding model for the BJT. The model is drawn in blue in figure A.28.

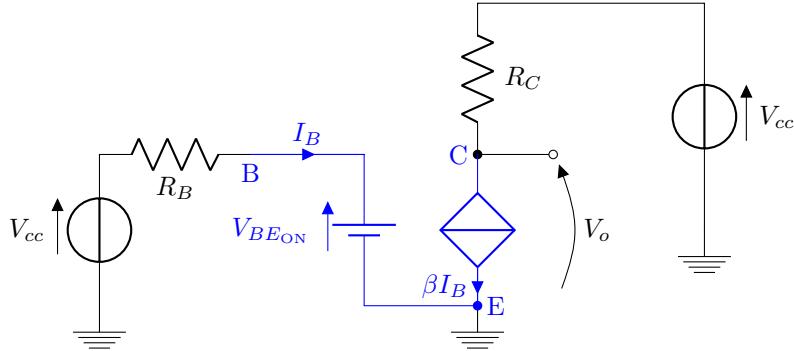


Figure A.28: Equivalent circuit for q.o.p.

We can easily find all parameters:

$$I_B = \frac{V_{cc} - V_{BE_{ON}}}{R_B} = \frac{14.3 \text{ V}}{470 \text{ k}\Omega} = 30.4 \mu\text{A}$$

$$I_C = \beta I_B = 100 \cdot 30.4 \mu\text{A} = 3.04 \text{ mA}$$

$$V_{CE} = V_o = V_{cc} - I_C R_C = 15 \text{ V} - 3.04 \text{ mA} \cdot 2.2 \text{ k}\Omega = 8.31 \text{ V}$$

- To compute the small signal gain of the circuit we need to use the small signal model of the BJT. It is the circuit in blue in figure A.29.

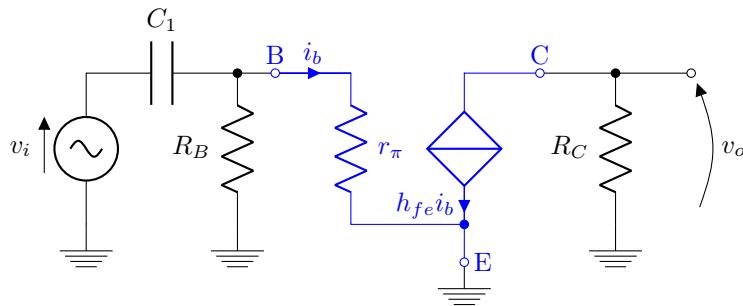


Figure A.29: Equivalent circuit for small signal analysis.

We know that:

$$r_\pi = \frac{\beta V_T}{I_{C0}} = \frac{100 \cdot 26 \text{ mV}}{3.04 \text{ mA}} = 855 \Omega$$

We have:

$$i_b = \frac{v_i}{r_\pi}$$

$$v_o = -h_{fe} i_b R_C$$

$$\frac{v_o}{v_i} = -\frac{h_{fe} R_C}{r_\pi} = -\frac{100 \cdot 2.2 \text{ k}\Omega}{855 \Omega} = -257 \implies 48.2 \text{ dB}$$

3. To compute the frequency response of the circuit, we take into account the capacitor:

$$v_b = v_i \cdot \frac{R_B \parallel r_\pi}{1/sC_1 + R_B \parallel r_\pi} = v_i \cdot \frac{sC_1(R_B \parallel r_\pi)}{1 + sC_1(R_B \parallel r_\pi)}$$

$$i_b = \frac{v_b}{r_\pi}$$

$$v_o = -h_{fe} i_b R_C$$

$$\frac{v_o}{v_i} = -\frac{h_{fe} R_C}{r_\pi} \cdot \frac{sC_1(R_B \parallel r_\pi)}{1 + sC_1(R_B \parallel r_\pi)}$$

$$f_p = \frac{1}{2\pi C_1(R_B \parallel r_\pi)} = 18.7 \text{ Hz}$$

The Bode plot of the transfer function is in figure A.30.

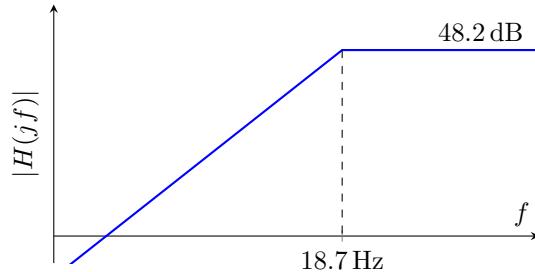


Figure A.30: Transfer function Bode diagram

A.6.6 Solution to exercise 6.6

- To find the quiescent operating point we can use the corresponding model for the BJT. The model is drawn in blue in figure A.31.

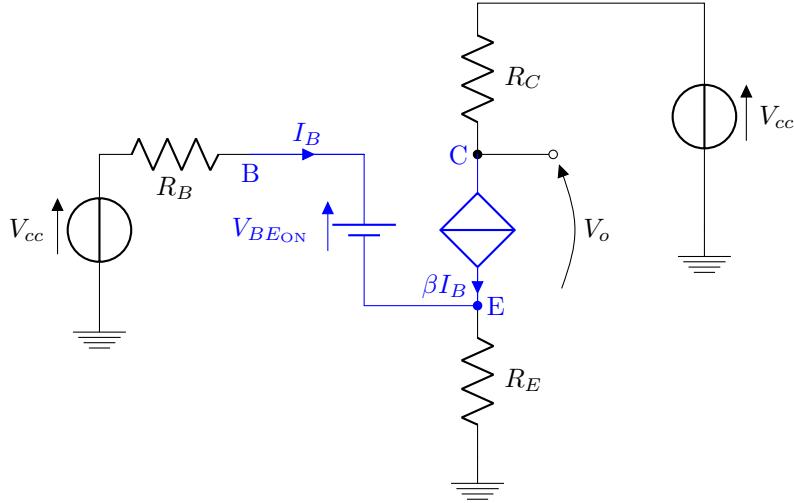


Figure A.31: Equivalent circuit for q.o.p.

To find I_B we use KVL:

$$V_{cc} - I_B R_B - V_{BE_{ON}} - I_B(1 + \beta)R_E = 0$$

$$I_B = \frac{V_{cc} - V_{BE_{ON}}}{R_B + (1 + \beta)R_E} = \frac{14.3 \text{ V}}{492.2 \text{ k}\Omega} = 29.1 \mu\text{A}$$

We can then easily find all other parameters:

$$I_C = \beta I_B = 100 \cdot 29.1 \mu\text{A} = 2.91 \text{ mA}$$

$$V_{CE} = V_{cc} - I_C R_C - I_C \cdot \frac{\beta + 1}{\beta} R_E = 15 \text{ V} - 2.91 \text{ mA} \cdot (2.2 \text{ k}\Omega + 222.2 \Omega) = 7.96 \text{ V}$$

- To compute the small signal gain of the circuit we need to use the small signal model of the BJT. It is the circuit in blue in figure A.32.

We know that:

$$r_\pi = \frac{\beta V_T}{I_{C0}} = \frac{100 \cdot 26 \text{ mV}}{2.91 \text{ mA}} = 895 \Omega$$

We have:

$$i_b = \frac{v_i}{r_\pi + (1 + h_{fe})R_E}$$

$$v_o = -h_{fe} i_b R_C$$

$$\frac{v_o}{v_i} = -\frac{h_{fe} R_C}{r_\pi + (1 + h_{fe})R_E} = -\frac{100 \cdot 2.2 \text{ k}\Omega}{895 \Omega + 22.2 \text{ k}\Omega} = -9.52 \implies 19.6 \text{ dB}$$

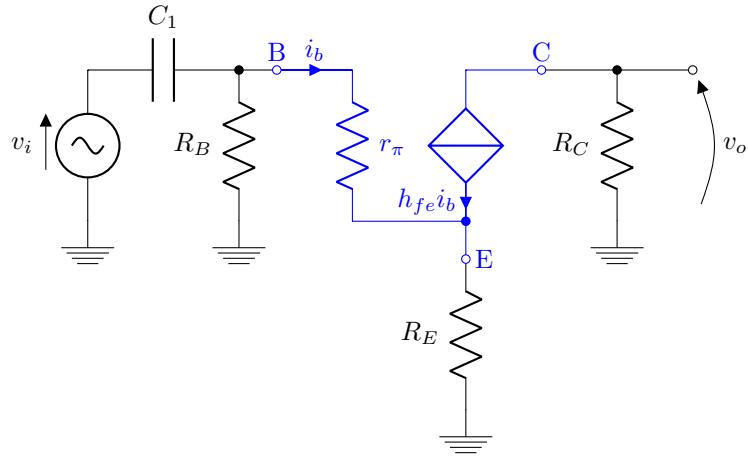


Figure A.32: Equivalent circuit for small signal analysis.

3. To compute the frequency response of the circuit, we take into account the capacitor:

$$\begin{aligned}
 v_b &= v_i \cdot \frac{R_B \parallel [r_\pi + (1 + h_{fe})R_E]}{1/sC_1 + R_B \parallel [r_\pi + (1 + h_{fe})R_E]} \\
 v_b &= v_i \cdot \frac{sC_1\{R_B \parallel [r_\pi + (1 + h_{fe})R_E]\}}{1 + sC_1\{R_B \parallel [r_\pi + (1 + h_{fe})R_E]\}} \\
 i_b &= \frac{v_b}{r_\pi + (1 + h_{fe})R_E} \\
 v_o &= -h_{fe}i_b R_C \\
 \frac{v_o}{v_i} &= -\frac{h_{fe}R_C}{r_\pi + (1 + h_{fe})R_E} \cdot \frac{sC_1\{R_B \parallel [r_\pi + (1 + h_{fe})R_E]\}}{1 + sC_1\{R_B \parallel [r_\pi + (1 + h_{fe})R_E]\}} \\
 f_p &= \frac{1}{2\pi C_1 R_B \parallel [r_\pi + (1 + h_{fe})R_E]} = 0.69 \text{ Hz}
 \end{aligned}$$

The Bode plot of the transfer function is in figure A.33.

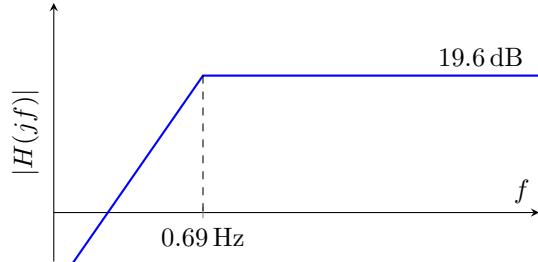


Figure A.33: Transfer function Bode diagram

A.6.7 Solution to exercise 6.7

- To find the quiescent operating point we can use the corresponding model for the MOSFET. The model is drawn in blue in figure A.34.

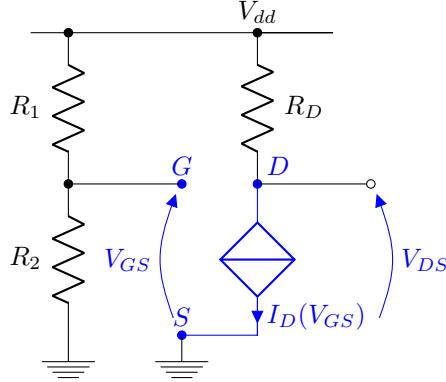


Figure A.34: Equivalent circuit for q.o.p.

To find V_{GS} we calculate the voltage divider:

$$V_{GS} = V_{dd} \frac{R_2}{R_1 + R_2} = 4 \text{ V}$$

We can then easily find all other parameters:

$$I_D = \frac{1}{2} k_n (V_{GS} - V_{th})^2 = 1 \text{ mA V}^{-2} \cdot (1.5 \text{ V})^2 = 2.25 \text{ mA}$$

$$V_{DS} = V_{dd} - I_D R_D = 12 \text{ V} - 2.25 \text{ mA} \cdot 2 \text{ k}\Omega = 7.5 \text{ V}$$

We obtain $V_{DS} > V_{GS} - V_{th}$: that confirms that we used the correct equation and the MOSFET is in channel saturation region.

- To compute the small signal gain of the circuit we need to use the small signal model of the MOSFET. It is the circuit in blue in figure A.35, where $R_{eq} = R_1 \parallel R_2$.

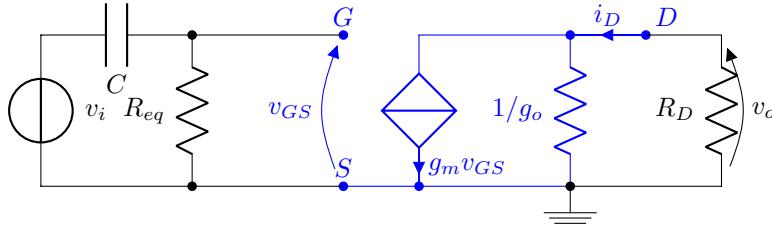


Figure A.35: Equivalent circuit for small signal analysis.

When we consider the capacitor a short circuit, $v_{GS} = v_i$.

We know that

$$g_m = \sqrt{2k_n I_{D0}} = \sqrt{4 \text{ mA V}^{-2} \cdot 2.25 \text{ mA}} = 3 \text{ mA V}^{-1}$$

where I_{D0} is the value of I_D computed in the quiescent operating point.
The expression of g_o is:

$$g_o = \lambda I_{D0} = 22.5 \mu\text{A V}^{-1}$$

We have:

$$v_o = -g_m v_{GS} \frac{R_D}{1 + R_D g_o} = -g_m v_i \frac{R_D}{1 + R_D g_o}$$

$$A_V = \frac{v_o}{v_i} = -g_m \frac{R_D}{1 + R_D g_o} = -3 \text{ mA V}^{-1} \cdot \frac{2 \text{ k}\Omega}{1 + 0.045} = -5.74 \Rightarrow 15.2 \text{ dB}$$

3. To compute the frequency response of the circuit, we take into account the capacitor:

$$v_{GS} = v_i \cdot \frac{R_{eq}}{1/sC + R_{eq}} = v_i \cdot \frac{sCR_{eq}}{1 + sCR_{eq}}$$

$$v_o = -g_m v_{GS} \frac{R_D}{1 + R_D g_o} = -g_m \frac{R_D}{1 + R_D g_o} \cdot v_i \cdot \frac{sCR_{eq}}{1 + sCR_{eq}}$$

$$\frac{v_o}{v_i} = -g_m \frac{R_D}{1 + R_D g_o} \cdot \frac{sCR_{eq}}{1 + sCR_{eq}}$$

$$f_p = \frac{1}{2\pi C R_{eq}} = 47.7 \text{ Hz}$$

The Bode plot of the transfer function is in figure A.36.

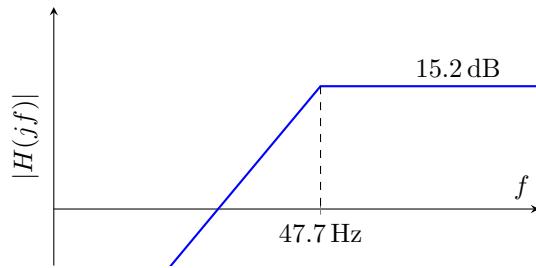


Figure A.36: Transfer function Bode diagram

A.7 Solution to chapter 7 exercises

A.7.1 Solution to exercise 7.1

The schematic of the switching circuit is shown in figure A.37

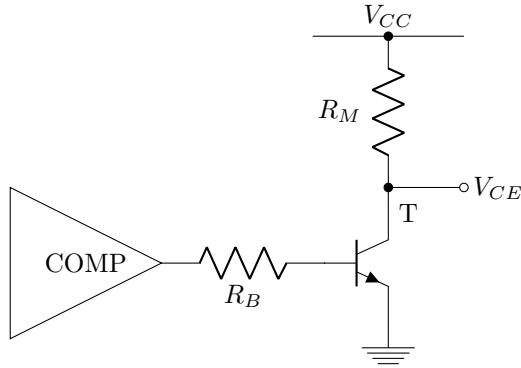


Figure A.37: Basic schematic of the switching circuit

The maximum voltage drop across the transistor acting as switch is determined by the minimum current which must flow in the load:

$$\frac{V_{CC} - V_{CEmax}}{R_M} \geq I_{Mmin}$$

from which we obtain:

$$V_{CEmax} = V_{CC} - R_M \cdot I_{Mmin} = 12 \text{ V} - 35 \Omega \cdot 320 \text{ mA} = 0.8 \text{ V}$$

It is therefore necessary for the transistor to be in saturation.

To guarantee saturation we must have

$$\beta_F = \frac{I_{Csat}}{I_{Bsat}} \ll \beta_{min}$$

Assuming $\beta_F = \beta_{min}/2 = 35$ the base current is

$$I_{Bsat} = \frac{320 \text{ mA}}{35} = 9.14 \text{ mA}$$

The maximum value of R_B is given by:

$$R_B \leq \frac{V_{OH} - V_{BE}}{I_{Bsat}} = \frac{9 \text{ V} - 0.7 \text{ V}}{9.14 \text{ mA}} = 908 \Omega$$

The closest normalized value is 820Ω and the current I_B will be greater than the previously computed value and in any case **larger** than the maximum value which may be provided by the comparator. Therefore the transistor specified **may not** be used to interface the motor load.

A.7.2 Solution to exercise 7.2

The schematic of the switching circuit is shown in figure A.38.

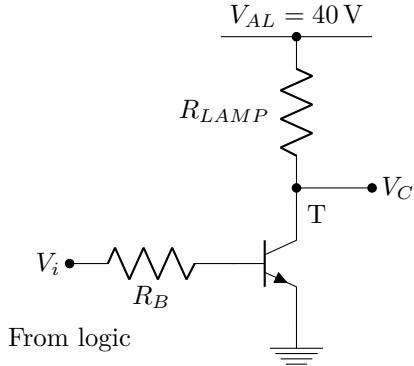


Figure A.38: Basic schematic of the switching circuit

A power of 10 W with a power supply of 40 V corresponds to a load resistance of

$$R_{LAMP} = \frac{V_{AL}^2}{P_D} = \frac{1600}{10} = 1600 \Omega$$

The maximum current in the load and in the transistor is

$$I_{max} = \frac{V_{AL}}{R_{LAMP}} = 250 \text{ mA}$$

The maximum power dissipated by the transistor in saturation is

$$P_{Dmax} = I_{max} \cdot V_{CEsat} = 250 \text{ mA} \cdot 0.3 \text{ V} = 75 \text{ mW}$$

In order to guarantee that the transistor operates in saturation we must have:

$$\beta_F = \frac{I_{max}}{I_{Bsat}} \ll \beta_{min} = 60$$

If we assume $\beta_F = 25$ we then have:

$$I_{Bsat} = \frac{I_{max}}{\beta_F} = 10 \text{ mA}$$

When $v_i = V_{OH}$ we have :

$$\frac{10 \text{ V} - 0.7 \text{ V}}{R_b} \geq I_{Bsat}$$

from which we obtain:

$$\frac{10 \text{ V} - 0.7 \text{ V}}{I_{Bsat}} = 930 \Omega \geq R_b$$

The closest (lower) normalized value is 820Ω and with this value the current which must be provided by the control logic at the high state is

$$\frac{10 \text{ V} - 0.7 \text{ V}}{820 \Omega} = 11.3 \text{ mA}$$

At the low state the transistor is OFF and the current is 0.

A.7.3 Solution to exercise 7.3

A MOS transistor operating in the linear region may be approximated by a resistance R_D . The maximum value of the MOS resistance is determined by the minimum current which must flow in the load:

$$\frac{V_{DD}}{R_M + R_D} \geq I_{Mmin}$$

Therefore we have:

$$R_D \leq \frac{V_{DD}}{I_{Mmin}} - R_M = \frac{12 \text{ V}}{300 \text{ mA}} - 35 \Omega = 40 \Omega - 35 \Omega = 5 \Omega$$

In the linear region the resistance of the transistor is:

$$R_D = \frac{1}{\beta_n (V_{GS} - V_T)} = \frac{1}{50 \text{ mA/V}^2 \cdot 5 \text{ V}} = 4 \Omega$$

Since R_D is lower than the limit set by the load current, then the transistor may be used to interface the load. The maximum current in the transistor will be:

$$I_{Dmax} = \frac{V_{DD}}{R_M + R_D} = \frac{12 \text{ V}}{35 \Omega + 4 \Omega} = 308 \text{ mA}$$

and the maximum power dissipated is:

$$P_D = I_{Dmax}^2 \cdot R_D = 0.308^2 \cdot 4 = 0.38 \text{ W}$$

Since both values (at $25^\circ C$) are lower than the maximum ones for the device, then it is compatible with the application.

A.8 Solution to chapter 8 exercises

A.8.1 Solution to exercise 8.1

1. To find V_{inv} we need to solve the quadratic equation:

$$\mu_n C_{OX} \frac{W_n}{L_n} (V_{inv} - V_{Tn})^2 = \mu_p C_{OX} \frac{W_p}{L_p} (V_{inv} - V_{AL} - V_{Tp})^2$$

By entering the values of the parameters we obtain:

$$0.55V_{inv}^2 + 0.8V_{inv} - 1.55 = 0$$

Only the positive solution is valid: $V_{inv} = 1.1$ V

2. Given that $V_{Tp} = -V_{Tn}$ we can obtain $V_{inv} = V_{AL}/2 = 1.25$ V if:

$$\mu_n C_{OX} \frac{W_n}{L_n} = \mu_p C_{OX} \frac{W_p}{L_p}$$

Given that $L_p = L_n$, solving for W_p we obtain:

$$W_p = \frac{\mu_n C_{OX} W_n}{\mu_p C_{OX}} = \frac{1}{0.3} \cdot 0.5 \text{ } \mu\text{m} = 1.67 \text{ } \mu\text{m}$$

A.8.2 Solution to exercise 8.2

The schematic of the CMOS gate is shown in figure A.39

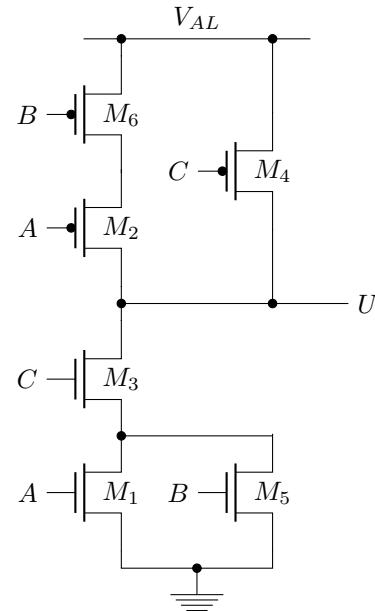


Figure A.39: Schematic of the gate

A.8.3 Solution to exercise 8.3

1. The circuit able to connect together three open drain outputs performs the wired-or function as explained in the theory. Open drain output need a pull-up resistor to be able to output a logic high level. The circuit is in figure A.40.

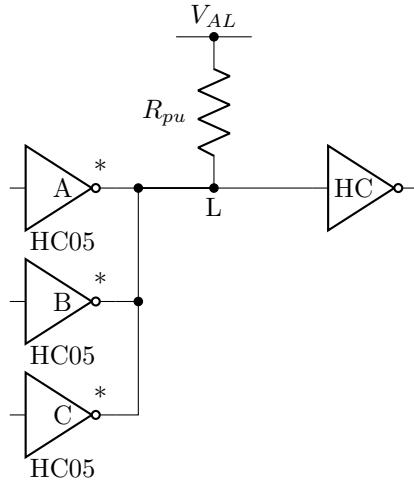


Figure A.40: Schematic of wired-or circuit.

2. We refer to the parameters that we can find in the datasheet of 74HC05. We need the output characteristics for the three outputs. There is no specification at $V_{AL} = 5$ V, so we consider $V_{AL} = 4.5$ V which is very close. When the device is active, we read $V_{OL} = 0.33$ V at a rated $I_{OL} = 4$ mA. When inactive, the maximum leakage current flowing in the output is $I_{OH} = \pm 5$ μ A and the voltage should be limited at $V_{AL} + 0.5$ V. We also have to read the input parameters, because the single gate connected to the line has to receive correct voltage values: we have $V_{IL} = 1.35$ V, $V_{IH} = 3.15$ V. The maximum static leakage current in the input is $I_I = \pm 1$ μ A.

To find the valid range of values for the pull-up resistor we can estimate the behavior of the circuit when the line is at low level and at high level.

There are different combinations of output logic levels that produce a low level on the line: one, two or three outputs can be low and (if any) the rest at high impedance. If more than one output is low, then the current in the pull-up resistor will be shared by the active gates, so the worst case is when only one gate is active. The maximum current in an output is $I_{OL} = 4$ mA. It is possible to neglect the contribution of the input leakage current, being very small. We can simply write:

$$\frac{V_{AL} - V_{OL}}{R_{pu}} < I_{OL} \implies R_{pu} > 1.17 \text{ k}\Omega$$

When the line is at high level, all the output are in high impedance state. Only the leakage currents flow in the outputs and in the input. The voltage

drop on R_{pu} should be such that the voltage on the line is $V_H > V_{IH}$:

$$V_{AL} - R_{PU}(3 \cdot I_{OH} + I_I) > V_{IH} \implies R_{PU} < \frac{V_{AL} - V_{IH}}{3 \cdot I_{OH} + I_I} = 116 \text{ k}\Omega$$

By choosing the lowest standard value in the range we obtain the minimum rise time of the system: $R_{pu} = 1.2 \text{ k}\Omega$

3. By choosing the maximum value in the range we obtain the lowest power consumption: $R_{pu} = 100 \text{ k}\Omega$.

A.8.4 Solution to exercise 8.4

1. The standard delay is the typical delay we measure when only one input is connected to the output of the logic gate. We have 19 more inputs. Each input will increase the delay by a factor:

$$\Delta t_1 = \Delta t_{pLH} \cdot C_{in} = 0.2 \text{ ns}$$

We have to multiply this value by the number of excess inputs (19) and add it to the standard value:

$$t_{pLH} = t_{pLH_{std}} + 19 \cdot \Delta t_1 = 11.8 \text{ ns}$$

A.8.5 Solution to exercise 8.5

1. The formula we found for power dissipation is:

$$P_D = n \cdot f_{clk} \cdot C \cdot V_{AL}^2 = 10^5 \cdot 10 \text{ MHz} \cdot 1 \text{ pF} \cdot (2.5 \text{ V})^2 = 6.25 \text{ W}$$

2. The reduction in power supply voltage brings a reduction in power:

$$P_D = 10^5 \cdot 10 \text{ MHz} \cdot 1 \text{ pF} \cdot (1.8 \text{ V})^2 = 3.24 \text{ W}$$

3. Given that the dissipated power is proportional to the clock frequency, the resulting power is:

$$P_D = 10^5 \cdot 1 \text{ MHz} \cdot 1 \text{ pF} \cdot (1.8 \text{ V})^2 = 0.324 \text{ W}$$