Table of Contents

1.6 IDEAL TWO TERMINAL ELEMENTS 7
1.7 MODELING PHYSICAL ELEMENTS 8
1.8 SIGNAL REPRESENTATION 8
CHAPTER 2: RESISTIVE NETWORKS 8
2.3 CIRCUIT ANALYSIS: BASIC METHOD 11
CHAPTER 3: NETWORK THEOREMS 14
3.2 THE NODE VOLTAGE 14
3.3 THE NODE METHOD 15
3.4 LOOP METHOD 18
3.5 SUPERPOSITION 18
3.6 THEVENIN'S THEOREM AND NORTON'S THEOREM 19
CHAPTER 4: ANALYSIS OF NON-LINEAR CIRCUITS 22
4.1 INTRODUCTION TO NON-LINEAR ELEMENTS 22
4.2 ANALYTICAL SOLUTIONS 23
4.2 GRAPHICAL ANALYSIS 24
4.3 PIECEWISE LINEAR ANALYSIS 24
4.5 INCREMENTAL ANALYSIS 25
4.6 EXTRA PARAGRAPH: FUNNY STORY ABOUT A DIODE 29
4.7 ZENER DIODE 29
CHAPTER 5: THE DIGITAL ABSTRACTION 30
5.1 VOLTAGE LEVELS AND THE STATIC DISCIPLINE 31
5.2 BOOLEAN LOGIC 32
5.3 COMBINATIONAL GATES 32
5.4 STANDARD SUM-OF-PRODUCTS REPRESENTATION 33
5.5 SIMPLIFYING LOGIC EXPRESSIONS 33
5.6 NUMBER REPRESENTATIONS 34
CHAPTER 6: THE MOSFET SWITCH 37
6.1 THE SWITCH 37
6.2 LOGIC FUNCTIONS USING SWITCHES 37
6.3 THE MOSFET DEVICE AND ITS S MODEL 38
6.4 MOSFET SWITCH IMPLEMENTATION OF LOGIC GATES 40

6.5 STATIC ANALYSIS USING THE S MODEL 43
6.6 THE SR MODEL OF THE MOSFET 43
6.7 PHYSICAL STRUCTURE OF THE MOSFET 45
6.8 STATIC ANALYSIS USING THE SR MODEL 47
6.9 SIGNAL RESTORATION, GAIN, AND NONLINEARITY 48
6.10 POWER CONSUMPTION IN LOGIC GATES 51
CHAPTER 7: THE MOSFET AMPLIFIER 52
7.1 SIGNAL AMPLIFICATION 52
7.2 REVIEW OF DEPENDENT SOURCES 52
7.3 ACTUAL MOSFET CHARACTERISTICS 52
7.4 THE SWITCH-CURRENT SOURCE (SCS) MOSFET MODEL 54
7.5 THE MOSFET AMPLIFIER 56
7.6 LARGE SIGNAL ANALYSIS OF THE MOSFET AMPLIFIER 59
7.7 OPERATING POINT SELECTION 61
7.8 BIPOLAR JUNCTION TRANSISTOR (BJT) – From example 7.13 to … 63
CHAPTER 8: THE SMALL SIGNAL MODEL 69
8.1 OVERVIEW OF THE NONLINEAR MOSFET AMPLIFIER 69
8.2 THE SMALL SIGNAL MODEL 69
CHAPTER 9: ENERGY STORAGE ELEMENTS 80
9.1 CONSTITUTIVE LAWS 81
9.2 SERIES AND PARALLEL CONNECTIONS 86
9.3 SPECIAL EXAMPLES (Not sure is summarized everything ok) 87
9.4 SIMPLE CIRCUIT EXAMPLES 90
9.4/B ChatGPT - CONSIDERATIONS ON STEP FUNCTION u(t;T), PULSES $ddtut;T$, IMPULSES δ(t) 94
9.5 ENERGY, CHARGE, AND FLUX CONSERVATION 96
CHAPTER 10: FIRST-ORDER TRANSIENTS IN LINEAR ELECTRICAL NETWORKS 97
10.1 ANALYSIS OF RC CIRCUITS 97
10.2 ANALYSIS OF RL CIRCUITS 101
10.3 INTUITIVE ANALYSIS 103
10.4 PROPAGATION DELAY AND THE DIGITAL ABSTRACTION 105
10.5 STATE AND STATE VARIABLES 108
10.6 ADDITIONAL EXAMPLES 110
10.7 DIGITAL MEMORY 111

YOUTUBE MIT COURSE 114
LESSON 8 – 9 – AMPLIFIERS AND MOSFET 114
LESSON 9B – LARGE SIGNAL ANALYSIS 114
LESSON 10 – AMPLIFIER: SMALL SIGNAL ANALYSIS 115
LESSON 11 – SMALL SIGNAL CIRCUIT METHOD 115
LESSON 12 – CAPACITORS 116
LESSON 13 – DIGITAL CIRCUIT SPEED 117

Foundations of Analog and Digital Electronic Circuits

Electrical engineering is the purposeful use of Maxwell's Equations (or Abstractions) for electromagnetic phenomena. To facilitate our use of electromagnetic phenomena, electrical engineering creates a new abstraction layer on top of Maxwell's Equations called the lumped circuit abstraction.

1.5.1 BATTERIES

The important specifications for a battery are its nominal voltage, its total store of energy, and its internal resistance.

In this section, we will assume that the internal resistance of a battery is zero.

The second important parameter of a battery is the total amount of energy it can store, often measured in *joules*. However, if you pick up a camcorder or flashlight battery, you might notice the ratings of *ampere-hours* or *watt-hours*. Let us reconcile these ratings.

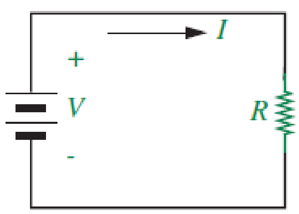


FIGURE 1.4

When a battery is connected across a resistive load in a circuit, it delivers power. The lightbulb in Figure 1.4 is an example of a resistive load. The power delivered by the battery is the product of the voltage and the current:

$$p = VI \quad (1.2)$$

Power is measured in *watts*. A battery delivers one watt of power when *V* is one volt, and *I* is one ampere. Power is the rate of delivery of energy.

Thus the amount of **energy *w* delivered by the battery is the time integral of the power**. If a constant of power *p* is delivered over an interval *T*, the energy *w* supplied is

$$w = pT \quad (1.3)$$

The battery delivers one joule of energy if it supplies one watt of power over one second. Thus, joules and watt-seconds are equivalent units.

Batteries in parallel and in series:

If one wishes to increase the current capacity of a battery without increasing the voltage at the terminals, individual cells can be connected in parallel. It is important that cells to be connected in parallel be nearly identical in voltage to prevent one cell from destroying another. If the two cells, are instead connected in series, then the current capacity will stay the same as a single cell, but the voltage will double.

Calculation examples:

Example 1:

A Lithium-Ion (Li-Ion) battery pack for a camcorder is rated as 7.2 V and 5 W-hours (this means can keep providing 5W of power over one hour). What are its equivalent ratings in mA-hours and joules? Since a joule (J) is equivalent to a W-second, 5 W-hours is the same as 5 × 3600 = 18000 J – this means the battery can provide 18000 J of energy, or 5 W-hours. Since the battery has a voltage of 7.2 V, the battery rating in ampere-hours is 5/7.2 = 0.69. Equivalently, its rating in mA-hours is 690.

Example 2:

A car battery might be rated at 12 V and 50 A-hours. This means that the battery can provide a 1-A current for 50 hours, or a 100-A current for 30 minutes. The amount of energy stored in such a battery is Energy = 12 × 50 = 600 watt-hours = 600 × 3600 = 2.16 × 106 joules.

1.5.2 LINEAR RESISTORS

Over some limited range of voltage and current, carbon, wire and polysilicon resistors obey Ohm's law:

$$v = iR$$

Resistance and Geometry:

The resistance of a piece of material depends on its geometry. As illustrated in Figure 1.14, assume the resistor has a conducting channel with cross-sectional area a, length l, and resistivity ρ.

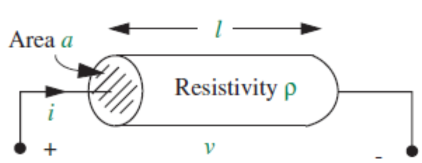


FIGURE 1.14 A cylindrical-wire shaped resistor.

$$R = \rho\frac{l}{a} \quad (1.5)$$

Equation 1.5 shows that the resistance of a piece of material is proportional to its length and inversely proportional to its cross-sectional area. Similarly, the resistance of a cuboid shaped resistor with length l, width w, and height h is given by

$$R = \rho\frac{l}{wh} \quad (1.6)$$

when the terminals are taken at the pair of surfaces with area wh.

Calculation example:

A thin poly-crystalline silicon resistor is 1 µm thick, 10 µm wide, and 100 µm long, where 1 µm is 10^{-6} m. If the resistivity of its poly-crystalline silicon ranges from 10^{-6} Ωm to 10^{2} Ωm, what is the range of its resistance? The cross-sectional area of the resistor is A = 10^{-11} m, and its length is l = 10^{-4} m. Using Equation 1.5, and the given range of resistivity, ρ, the resistance satisfies $10\ \Omega \leqslant R \leqslant 109\ \Omega$.

General resistors will be dealt with later in Chapter 9.

1.5.3 ASSOCIATED VARIABLES CONVENTION

Associated Variables Convention. **Define current to flow in at the device terminal assigned to be positive in voltage.**

Energy is pumped into an element when a positive current i is directed into the voltage terminal marked positive. Depending on the type of element, the energy is either dissipated or stored. Conversely, power is supplied by an element when a positive current i is directed out of the voltage terminal marked positive.

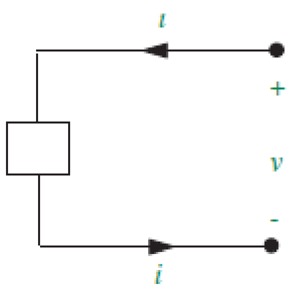


FIGURE 1.20 Definition of the terminal variables *v* and *i* for a two-terminal element under the associated variables convention.

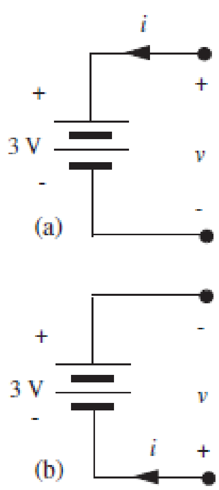


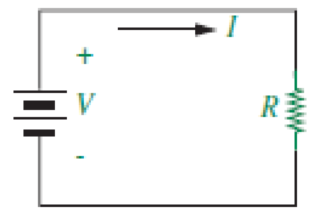
FIGURE 1.21 Terminal variable assignments for a battery.

While Figure 1.20 is quite simple, it nonetheless makes several important points. First, the two terminals of the element form a single port through which the element is addressed. Second, the current i circulates through that port. That is, the current that enters one terminal is instantaneously equal to the current that exits the other terminal. Thus, according to the lumped matter discipline, net charge cannot accumulate within the element. Third, the voltage v of the element is defined across the port. Thus, the element is assumed to respond only to the difference of the electrical potentials at its two terminals, and not to the absolute electric potential at either terminal. Fourth, the current is defined to circulate positively through the port by entering the positive voltage terminal and exiting the negative voltage terminal. Which terminal is chosen as the positive voltage terminal is arbitrary, but the relation defined between the current and voltage is not. Lastly, for brevity, the current that exits the negative voltage terminal is usually never labeled, but it is always understood to be equal to the current that enters the positive voltage terminal.

Example 1:

Figures 1.21a and b show two possible legal definitions for terminal variables for a 3 V battery. What is the value of terminal variable v in each case? For Figure 1.21a, we can see that terminal variable v = 3 V. For Figure 1.21b, however, v = −3 V. This example highlights the distinction between a terminal variable and an element property. The battery voltage of 3 V is an element property, while v is a terminal variable that we have defined. *I believe the terminal variables are variables found* *along the wires, while element variables are pertaining to a component such a battery, resistor etc.*

Example 2:



(b)

In a circuit such as that shown in figure on the right, the battery is rated at 7.2V and 10,000 J. Assume that the internal resistance of the battery is zero. Further assume that the resistance in the circuit is R = 1 kΩ. You are given that the resistor can dissipate a maximum of 0.5 W of power. (In other words, the resistor will overheat if the power dissipation is greater than 0.5 W.) Determine the current through the resistor. Further, determine whether the power dissipation in the resistor exceeds its maximum rating. The current through the resistor is given by $I = \frac{V}{R}$ and thus 7.2 mA. The power dissipation in the resistor is given by $p = I^2R$ equals to 0.052 W. Clearly, the power dissipation in the resistor is well within its capacity.

1.6 IDEAL TWO TERMINAL ELEMENTS

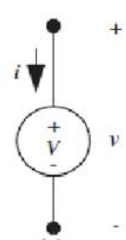
This section introduces a set of ideal two-terminal elements including voltage and current sources, and ideal wires and resistors, which form our primitives in the vocabulary of circuits. In this paragraph some basic rules are outlined, which make a simplification of reality but will be used throughout the text. Seems is a repetition of what seen already and quite intuitive concepts.

Only thing worth mentioning is the concept of conductance. It is the reciprocal of resistance, namely the conductance G having the units of Siemens (S). In this case:

$$G = \frac{1}{R} \quad (1.16)$$

1.6.2 ELEMENT LAWS

The most important characteristic of a two-terminal element is the relation between the voltage across and the current through its terminals, or v-i relationship. This relation, called the element law, represents the lumped-parameter summary of the electronic behavior of the element.



(a)

Fig. 1.29a Independent voltage source with assigned terminal variables

Element law for resistor:

V = iR

Constituent relation for the independent voltage source supplying a voltage V when its terminals defined as in fig 1.29a:

v = V

Element law for the ideal wire (or a short circuit):

$$V = 0$$

Element law for an open circuit:

$$i = 0$$

1.7 MODELING PHYSICAL ELEMENTS

-- skipped --

1.8 SIGNAL REPRESENTATION

1.8.1 ANALOG SIGNALS

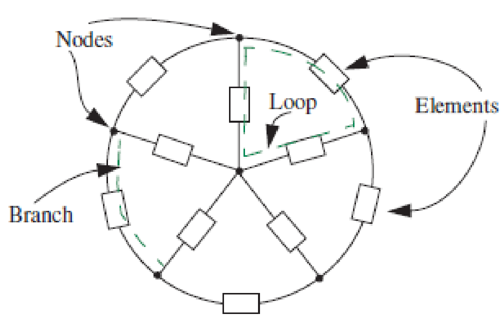
Signals in the physical world are most commonly analog, that is, spanning a continuum of values. Sound pressure is such a signal. The electromagnetic signal picked up by a mobile phone antenna is another example of an analog signal. Not surprisingly, most circuits that interact with the physical world must be able to process analog signals.

1.8.2 DIGITAL SIGNALS

Not continuum but discretization. In general, we can discretize values into any number of levels, for example, four. Thus the representation discussed thus far is a special case of value discretization called the binary representation where we discretize the voltage (or current for that matter) into two information levels: ‘‘0’’ and ‘‘1.’’

CHAPTER 2: RESISTIVE NETWORKS

Electronic access to an element is made through its terminals. An electronic circuit is constructed by connecting a collection of separate elements at their terminals, as shown in Figure 2.2. The junction points at which the terminals of two or more elements are connected are referred to as the **nodes** of a circuit. Similarly, the connections between the nodes are referred to as the edges or **branches** of a circuit. Finally, **circuit loops** are defined to be closed paths through a circuit along its branches.



Kirchhoff’s current law and Kirchhoff’s voltage law (the central topics of this chapter) describe how lumped parameter circuit elements couple at their terminals when they are assembled into a circuit.

2.2.1 KCL

The current flowing out of any node in a circuit must equal the current flowing in. KCL is applied to nodes. That is, **the algebraic sum of all branch currents flowing into any node must be zero**. Put another way, KCL states that the net current that flows into a node through some of its branches must flow out from that node through its remaining branches.

Referring to Figure 2.5, if the currents through the three branches into node a are ia, ib, and ic, then KCL states that ia + ib + ic = 0. Similarly, the currents into node b must sum to zero. Accordingly, we must have −ib − i4 = 0.

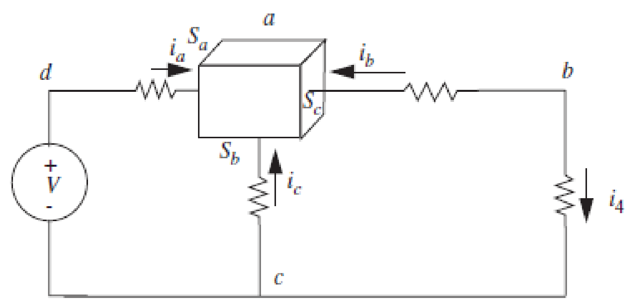


FIGURE 2.5 Currents into a node in the network.

An important simplification of KCL focuses on the two series-connected circuit elements shown in Figure 2.7. Taking KCL to state that no net current can flow into a node, the application of KCL at the node between the two elements yields

i1 − i2 = 0 ⇒ i1 = i2. (2.4)

This result is important because it shows that the branch currents passing through two series-connected elements must be the same. That is, there is nowhere for the current i1 to go as it enters the node connecting the two elements except to exit that node as i2. In fact, with multiple applications of KCL, this observation is extendible to a longer string of series-connected elements.

In figure 2.10, because of KCL, *i* must be equal to -1A.

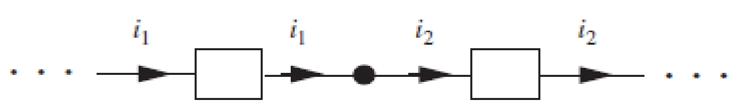


FIGURE 2.7 Two series-connected circuit elements.

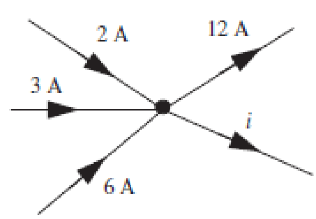


FIGURE 2.10 Five branches meeting at a node.

2.2.2 KVL

KVL is applied to circuit loops, that is, to interconnections of branches that form closed paths through a circuit. In a manner analogous to KCL, Kirchhoff's voltage law can be stated as:

The algebraic sum of the branch voltages around any closed path in a network must be zero. Alternatively, it states that the voltage between two nodes is independent of the path along which it is accumulated.

In Figure 2.14, the closed loop defined by the three circuit branches a → b, b → c, and c → a in Figure 2.14 is a closed path.

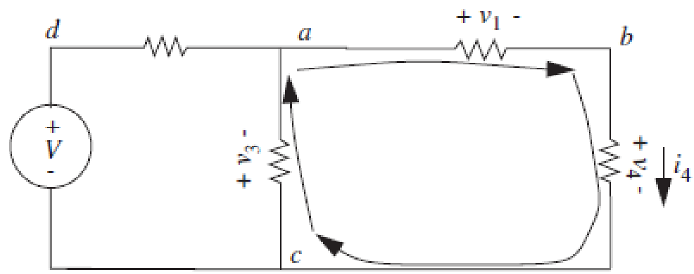


FIGURE 2.14 Voltages in a closed loop in the network.

According to KVL, the sum of the branch voltages around this loop is zero. That is,

$$vab + vbc + vca = 0 \ (v1 + v4 + v3 = 0)$$

where we have taken the positive sign for each voltage when going from the positive terminal to the negative terminal. It is important that we are consistent in how we assign polarities to voltages as we go around the loop. A helpful mnemonic for writing KVL equations is to assign the polarity to a given voltage in accordance with the first sign encountered when traversing that voltage around the loop.

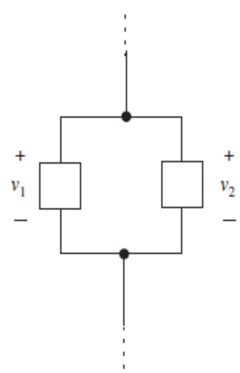


FIGURE 2.17 Two parallel-connected circuit elements.

An important simplification of KVL focuses on the two parallel-connected circuit elements shown in Figure 2.17. Starting from the upper node and applying KVL in the counterclockwise direction around the loop between the two circuit elements yields

$$v1 - v2 = 0 \Rightarrow v1 = v2. \ (2.12)$$

This result is important because it shows that the voltages across two parallel connected elements must be the same. In fact, with multiple applications of KVL, this observation is extendible to a longer string of parallel-connected elements. Such an extension would show that a common voltage appears across all parallel-connected elements in the string.

Voltage sources in series:

Two 1.5-V voltage sources are connected in series as shown in Figure 2.21. What is the voltage v at their terminals? To determine v, employ, for example, a counterclockwise application of KVL around the circuit, treating the port formed by the two terminals as an element having voltage v. In this case, 1.5 V + 1.5 V − v = 0, which has for its solution v = 3 V.

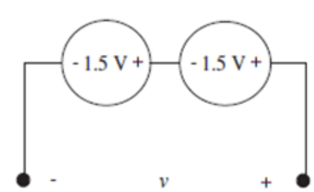


FIGURE 2.21

2.3 CIRCUIT ANALYSIS: BASIC METHOD

We are ready to introduce a systematic method for solving circuits. We saw previously that under the lumped matter discipline, Maxwell's Equations reduce to the basic element laws and the algebraic KVL and KCL. Accordingly, a systematic solution of the network involves the assembly and subsequent joint solution of two sets of equations. The first set of equations comprise the constituent relations for the individual elements in the network. The second set of equations results from the application of Kirchhoff's current and voltage laws.

The steps are the followings:

1. Define each branch current and voltage in the circuit in a consistent manner. The polarities of these definitions can be arbitrary from one branch to the next. However, for any given branch, follow the associated variables convention (see Section 1.5.3 in Chapter 1). **In other words, the branch current should be defined as positive into the positive voltage terminal of the branch**. By following the associated variables, element laws can be applied consistently, and the solutions will follow a much clearer pattern.

2. Assemble the element laws for the elements. These element laws will specify either the branch current or branch voltage in the case of an independent source or specify the relation between the branch current and voltage in the case of a resistor. Examples of these element laws were presented in Section 1.6.

3. Apply Kirchhoff's current and voltage laws as discussed in Section 2.2.

4. Jointly solve the equations assembled in Steps 2 and 3 for the branch variables defined in Step 1.

2.3.1 SINGLE-RESISTOR CIRCUITS

Some simple examples in the book.

In general, a circuit having B branches will have 2B unknown branch variables: B branch currents and B branch voltages. To find these variables, 2B independent equations are required, B of which will come from element laws, and B of which will come from the application of KVL and KCL. Moreover, if the circuit has N nodes, then N−1 equations will come from the application of KCL and B−N+1 equations will come from the application of KVL.

The physical results of the analysis of the circuit in Figure 2.25, and of any other circuit for that matter, cannot depend on the polarities of the definitions of the branch variables.

2.3.3 ENERGY CONSERVATION

Once the branch variables of a circuit have been determined, it is possible to examine the flow of energy through the circuit. Among other things, such an examination should show that energy is conserved in the circuit.

We will use two energy-based approaches in this book:

1. One energy approach equates the energy supplied by a set of elements in a circuit to the energy absorbed by the remaining set of elements in a circuit. Usually, this method involves equating the power generated by the devices in a circuit to the power dissipated in the circuit. Some examples are shown in the book: basically we take two elements and equate the power generated and dissipated by them, which must be equal, and thus find the missing variable.
2. Another energy approach equates the total amount of energy in a system at two different points in time (assuming that there are no dissipative elements in the circuit). Illustrated in chapter 9 section 9.5.

2.3.4 VOLTAGE AND CURRENT DIVIDERS

VOLTAGE DIVIDERS

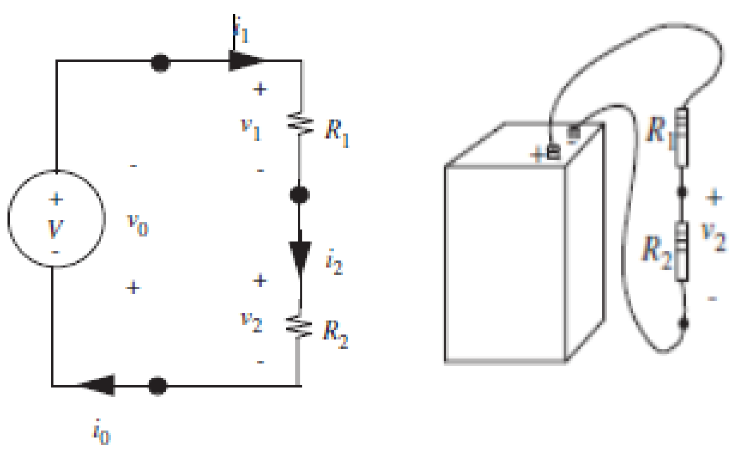


FIGURE 2.31 Assignment of branch variables to the voltage divider.

A voltage divider is an isolated loop that contains two or more resistors and a voltage source in series as in picture 2.31. If we solve the circuit in figure, we found that:

$$i1 = i2 = i0 = \frac{1}{R1 + R2} V \ (2.45)$$

$$v1 = \frac{R1}{R1 + R2} V \ (2.47)$$

$$v2 = \frac{R2}{R1 + R2} V \ (2.48)$$

Notice that v2 is some fraction of the voltage V, as desired.

Having determined its branch variables, we can now examine the flow of energy through the two-resistor voltage divider. See the book for this, basically we prove that energy is conserved as power generated by the voltage source is dissipated in the two resistors.

RESISTORS IN SERIES

The analysis of the voltage divider shows that two resistors in series act as a single resistor having a resistance RS equal to the sum of the two individual resistances R1 and R2. In other words, series resistances add.

RS = R1 + R2

Notice that, from equation 2.45, the more RS increases, the lower the current flowing through the circuit.

CURRENT DIVIDERS

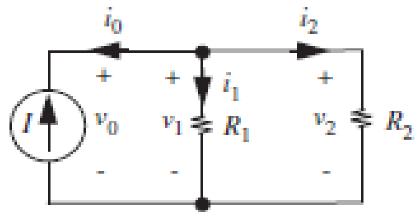


FIGURE 2.37 A current divider with two resistors.

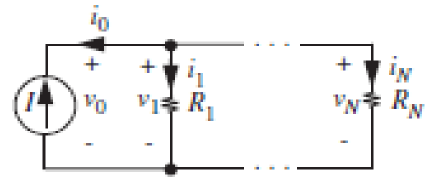


FIGURE 2.38 A current divider with N resistors.

A current divider is a circuit with two nodes joining two or more parallel resistors and a current source. Two current dividers are shown in Figures 2.37 and 2.38, the first with two resistors and the second with N resistors. Solving the circuit we obtain:

$$v2 = v0 = v1 = \frac{R1R2}{R1 + R2} I \ (2.85)$$

$$i1 = \frac{R2}{R1 + R2} I \ (2.83)$$

$$i2 = \frac{R1}{R1 + R2} I \ (2.84)$$

This means that the current is divided between the two branches according to the conductance of the branches, and less current will flow into the branch with more conductance (less resistance) and the opposite in the other branch.

Having determined its branch variables, we can now examine the flow of energy. See the book for this, basically we prove that energy is conserved as power generated by the current source is dissipated in the two resistors.

RESISTORS IN PARALLEL

The equivalent resistance of two resistors in parallel is given by:

$$Rp = \frac{R1R2}{R1 + R2} \quad (2.96)$$

This follows from what we have seen above that individual conductance for two resistors in parallel add together. Which means that another way to write equation 2.96 is:

$$\frac{1}{Rp} = \frac{1}{R1} + \frac{1}{R2}$$

The short notation for two resistors in parallel is R1||R2 = Rp

This result can be generalized for N results in parallel, the total resistance will be:

$$\frac{1}{Rp} = \frac{1}{R1} + \ \dots \ \frac{1}{R(N)}$$

This result in term of resistances is:

$$Rp = \ \frac{1}{\frac{1}{R1} + \frac{1}{R2} + \cdots + \frac{1}{RN}}$$

And where N resistors have the same value then:

$$Rp = \frac{R}{N}$$

2.3.5 A MORE COMPLEX CIRCUIT

Check page 85 in the book.

CHAPTER 3: NETWORK THEOREMS

We will develop several network theorems which will greatly simplify circuit analysis. The first one of these techniques is the node method. The node method works with a set of variables called the node voltages.

3.2 THE NODE VOLTAGE

Earlier we worked with <u>branch voltage, which is the potential difference across the element in a branch.</u> **A node voltage instead is the potential difference between the given node and some other node that has been chosen as a reference node.** The reference node is called the ground. The potential at this node is defined to be zero V. Ground is arbitrary, but it is usually picked as the node with maximum number of circuit elements connected to it.

If we know the node voltages, we can find the branch voltages V by difference according to KVL, and assuming we know that the elements are resistors with resistance R, then finding the current I can be easily done using the formula I = V/R.

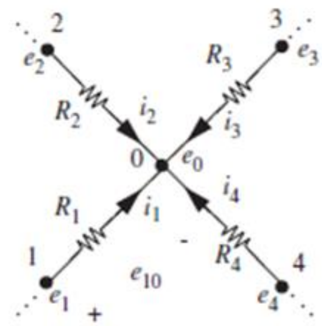


FIGURE 3.7

It follows that we can also write KCL for each of the nodes in a network in terms of node voltages and the element parameters.

In the case of fig. 3.7, we have:

i1 = (e1-e0)/R1

i2 = (e2-e0)/R2

i3 = (e3-e0)/R3

i4 = (e4-e0)/R4

and thus, KCL for Node 0 in terms of node voltages and element values is:

$$(e1-e0)/R1 + (e2-e0)/R2 + (e3-e0)/R3 + (e4-e0)/R4 = 0$$

In summary, a voltage is always defined as the potential difference between a pair of points: the two branch terminals for a branch voltage, and two nodes for a node voltage. What explained above already hints at how the node method works, which we see more in details in the next paragraph.

3.3 THE NODE METHOD

The steps of the node method can be written as:

1. Select a reference node (ground) from which all other voltages will be measured. Define its potential to be 0 V.
2. Label the potentials of the remaining nodes with respect to the ground node. Any node connected to the ground node through either an independent or a dependent voltage source should be labeled with the voltage of that source. The voltages of the remaining nodes are the primary unknowns and should be labeled accordingly. Since there are generally far fewer nodes than branches in a circuit, there will be far fewer primary unknowns to determine in a node analysis.
3. Write KCL for each of the nodes that has an unknown node voltage (in other words, the ground node and nodes with voltage sources connected to ground are excluded), using KVL and element laws to obtain the currents directly in terms of the node voltage differences and element parameters. Thus, one equation is written for each unknown node voltage.
4. Solve the equations resulting from Step 3 for the unknown node voltages.
5. Back-solve for the branch voltages and currents. More specifically, use node voltages and KVL to determine branch voltages as desired. Then, use the branch voltages, the element laws, and KCL to determine the branch currents, again as desired.

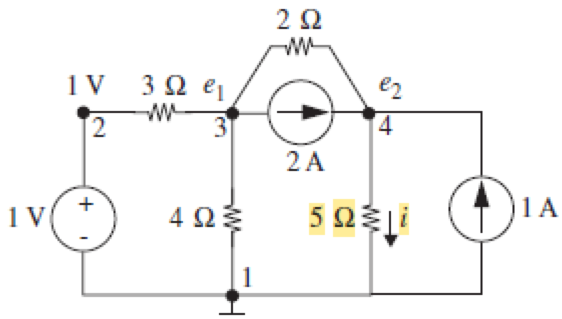


FIGURE 3.14 Determining the unknown current i.

Example 3.5 is shown here. Find the current i through the 5-Ω resistor. Following the procedure above, we define Node 1 as our ground node. Per step 2, we label node 2 as 1V, node 3 as e1, node 4 as e2. Per step 3, we write KCL equations for node 3 and 4.

$$\frac{e1-1}{3}+\frac{e1}{4}+\frac{e1-e2}{2}+2=0$$

$$-2+\frac{e2-e1}{2}+\frac{e2}{5}-1=0$$

Step 4, we solve the equations and find the values.

Final note. The method is inherently simple; in case of difficulties grasping the concepts making some exercises helps a lot.

3.3.2 FLOATING INDEPENDENT VOLTAGE SOURCES

A floating independent voltage source is a voltage source that has neither terminal connected to ground, neither directly nor through one or more other independent voltage sources. See example Figure 3.20.

Node analysis as described above does not work for circuits that contain floating independent voltage sources. The reason is that the element law for an independent voltage source does not relate its branch current to its branch voltage. Therefore, it is not possible to complete Step 3 of node analysis. In this case, it is necessary to modify the node analysis slightly (essentially is the same method).

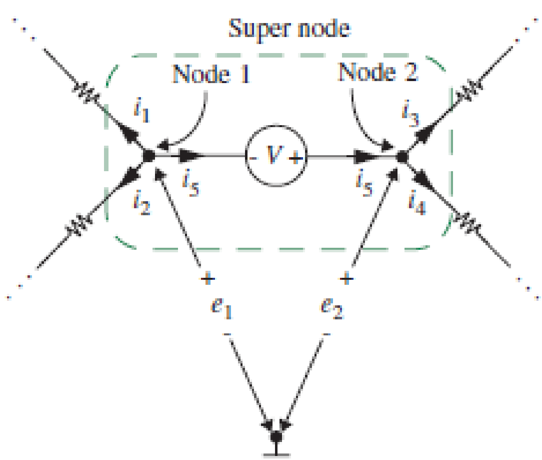


FIGURE 3.20 A floating independent voltage source and its treatment as a super node.

To apply node analysis to a circuit containing a floating voltage source we must realize that <u>the node voltages at the terminals of the source are directly related by the element law for that source</u>. For example, the application of KVL to the circuit in Figure 3.20 shows that

e2 = V + e1 (3.32)

Because of this, the number of unknown node voltages in the circuit can be immediately reduced by one since e1 and e2 can be determined directly from each other using Equation 3.32.

Consequently, the number of independent statements of KCL needed to determine the unknown node voltages can similarly be reduced by one. Thus, Nodes 1 and 2 in Figure 3.20 must together contribute one statement of KCL to the first part of Step 3 of the node analysis (namely, writing KCL for each of the nodes that has an unknown node voltage). Further, this single statement of KCL should not involve i5 since i5 cannot be determined from the element law of the voltage source in the second part of Step 3 (namely, using KVL and element laws to obtain the currents directly in terms of the node voltage differences and element parameters). To derive the desired statement of KCL for Nodes 1 and 2, we draw a surface around both nodes, enclosing what is referred to as a super node in the process. Then, we write KCL for the super node. In the case of Figure 3.20, KCL applied to the super node yields equation 3.33 for the first part of Step 3:

i1 + i2 + i3 + i4 = 0 (3.33)

Following this, in the second part of Step 3, the currents are eliminated by substituting node voltages and element parameters in their place (as usual in node method). In our example, i1 and i2 are determined using e1 and the parameters of the elements through which i1 and i2 flow. Similarly, i3 and i4 are determined using e1 + V and the parameters of the elements through which i3 and i4 flow, with e1 serving as the one unknown node voltage. Alternatively, i1 and i2 can be determined using e2−V, and i3 and i4 can be determined using e2, with e2 serving as the one unknown node voltage. Finally, it should be recognized that a floating string of independent voltage sources is handled in the same manner as a floating isolated independent voltage source.

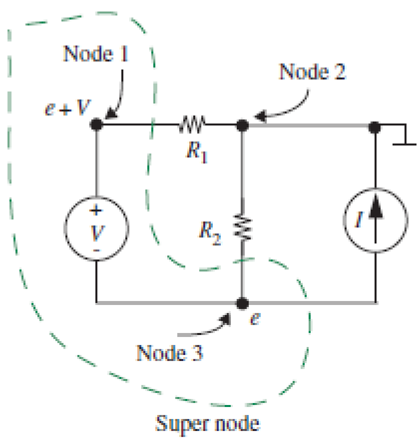


FIGURE 3.21 A circuit with a floating independent voltage source.

In summary, step 1 and step 2 are the same as for the node method we saw previously, step 3 instead is different because we have one less unknown variable to find. This is because the two nodes in the super node are related and have only one unknown variable.

An example in figure 3.21 illustrates this. By setting ground and labelling the nodes we have completed step 1 and step 2. Next, we perform step 3 for the super node. This yields:

$$\frac{e+V}{R1} + \frac{e}{R2} + I = 0 \quad (3.36)$$

Following step 4 we find the solution to the equation and solve the problem. As can be seen from this example, essentially, we are applying the node method, but node 1 and node 3 only have one unknown variable instead of two. So, if we have a independent voltage source the process gets simplified.

3.3.3 DEPENDENT SOURCES AND THE NODE METHOD

A dependent source will also complicate the node analysis previously described when its element law does not easily relate its branch current to its branch voltage. In this case, it will again not be possible to complete Step 3, and so it is again necessary to modify the node analysis slightly.

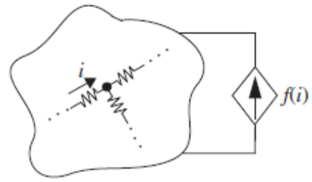


FIGURE 3.23 A circuit containing a dependent current source.

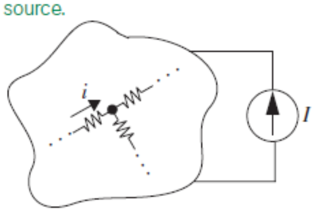


FIGURE 3.24 Replacing the dependent current source with an independent current source with an assumed current *I*.

Since there are four types of dependent sources, and the branch currents and voltages that control them can appear through or across many different types of elements, we present here a general, despite not the most efficient method that treats all cases of dependent sources and illustrate how this method can be made more efficient in a few illustrative cases (see textbook but it is really common sense).

The general method:

Figure 3.23 is the circuit at support of the explanation. Our method of applying node analysis to a circuit containing dependent sources begins by assuming that we know the value of each dependent source. This assumption allows us to treat each dependent source as an independent source and carry out a node analysis of the circuit as described in the previous subsections.

Suppose that the expression for i is some function of the assumed current I and is of the form:

$$i = g(I) \quad (3.40)$$

We substitute this expression for the branch variable into the element law for the dependent current source as

$$I = f(i) = f(g(I)) \quad (3.41)$$

and solve for I. The solution for I will not contain the variable i. Finally, we back-substitute the actual values of the dependent sources in other words, the solution for I into the original node analysis, thereby completing the analysis in total.

3.4 LOOP METHOD

(Blank paragraph in the book)

3.5 SUPERPOSITION

The superposition theorem states that in a <u>linear network</u> with a number of independent sources, the response can be found by summing the responses to each independent source acting alone, with all other independent sources set to zero. These individual responses can be found very readily by forming subcircuits in which all independent sources except one are set to zero.

In simpler terms:

Problem: Circuits with several independent sources can be complex to analyze.

Solution: Superposition method.

How: Analysis is easier by analyzing one independent source at a time.

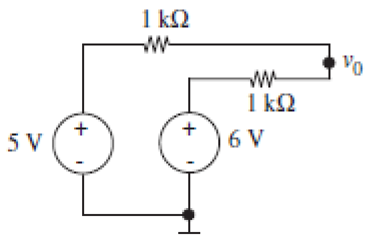


FIGURE 3.36 Circuit for performing superposition analysis.

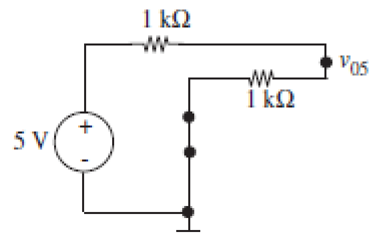


FIGURE 3.37 Circuit with 5-V source acting alone.

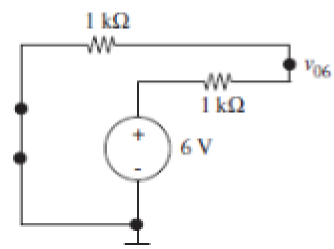


FIGURE 3.38 Circuit with 6-V source acting alone.

The superposition method for linear networks can be stated as follows:

1. For each independent source, form a subcircuit with all other independent sources set to zero. Setting a voltage source to zero implies replacing the voltage source with a short circuit, and setting a current source to zero implies replacing the current source with an open circuit.

2. From each subcircuit corresponding to a given independent source, find the response to that independent source acting alone. This step results in a set of individual responses.

3. Obtain the total response by summing together each of the individual responses.

Example:

Show that the node voltage v0 in the circuit shown in Figure 3.36 is the average of the two input voltages using the method of superposition. By the method of superposition, the voltage v0 can be determined by summing the responses of each of the sources acting alone. We will first obtain v05, the response of the 5-V source acting alone. The subcircuit corresponding to the 5-V source acting alone is shown in Figure 3.37. By the voltage divider action, we can write v05 = 5/2 V. Next, we obtain v06, the response of the 6-V source acting alone. By the voltage divider action, we can write v06 = 6/2 V. We now sum the two partial responses to obtain v0 = 6/2 V + 5/2 V = 5.5 V.

3.5.1 SUPERPOSITION RULES FOR DEPENDENT SOURCES

What do we do about dependent sources? A practical way is to leave all the dependent sources in the circuit. The network can then be solved for one independent source at a time by setting all other independent sources to zero and summing the individual responses.

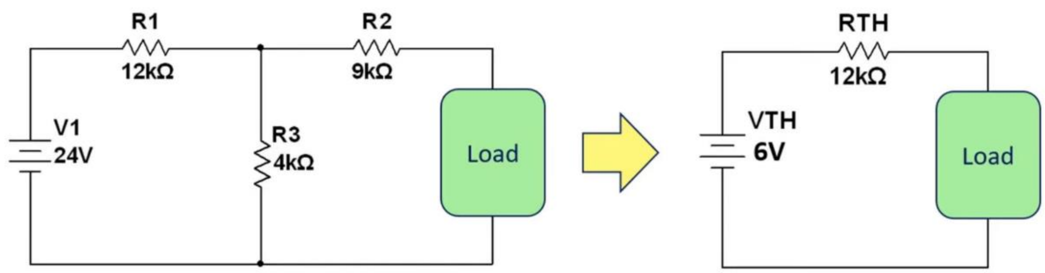
3.6 THEVENIN'S THEOREM AND NORTON'S THEOREM

A simple extension of the concept of superposition yields two additional network theorems of great power, which allow us to suppress a lot of detail in circuit analysis and focus attention only on that part of a network we are really interested in.

The part of the °network we are interested in° is any one pair of terminals we are interested in. We need to find the relationship between v and i at the terminals.

We will show that no matter how complex a circuit is; any collection of voltage sources, current sources, and resistors can be represented at any one pair of terminals by one voltage source and one resistor (Thevenin), or by one current source and one resistor (Norton).

To explain why these two theorems of circuit analysis are important and the reason we use them let's consider the figure below. On the left there is a circuit where there is connected a load RL. We can simplify the circuit, in this case thanks to Thevenin Theorem, to the circuit on the right, which contains only one voltage source and one resistor.



3.6.1 THE THÉVENIN EQUIVALENT NETWORK

Following Figure 3.55 below, in Figure 3.55b, we have chosen to apply a test current source to the terminals. To calculate the response vt by superposition, first set all the internal independent sources to zero, as in Figure 3.55c, and calculate the voltage va. Then set itest to zero, as in Figure 3.55d, and calculate vb. The desired value of vt is the sum va +vb. From Figure 3.55c, va = itestRt.

Rt is the net resistance measured between the two terminals when all internal independent sources are set to zero. Resistance Rt is called the Thévenin Equivalent Resistance. From Figure 3.55d, vb is obviously just the voltage appearing at the terminals of the original network when no current is flowing; we call this the open-circuit voltage. That is, vb = voc. Now by superposition, vt = va + vb = voc + itestRt.

This simple relation between voltage and current at a pair of terminals applies regardless of the complexity of the network, provided only that the network is linear.

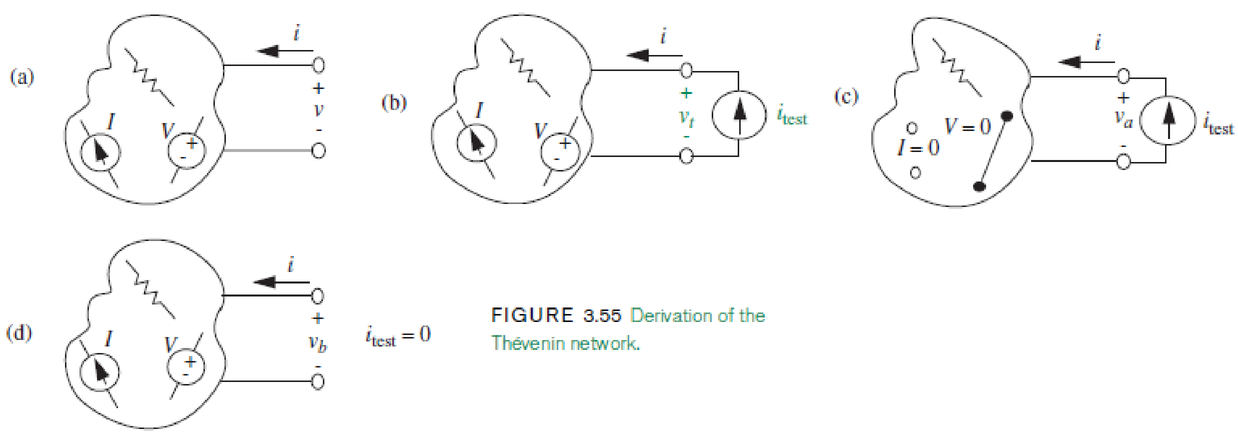


FIGURE 3.55 Derivation of the Thévenin network.

A Method for Determining the Thévenin Equivalent Circuit.

The Thévenin equivalent circuit for any linear network at a given pair of terminals consists of a voltage source vTH in series with a resistor RTH. The voltage vTH and resistance RTH can be obtained as follows:

1. vTH can be found by calculating or measuring the open-circuit voltage at the designated terminal pair on the original network.
2. RTH can be found by calculating or measuring the resistance of the open-circuit network seen from the designated terminal pair with all independent sources internal to the network set to zero. That is, with independent voltage sources replaced with short circuits, and independent current sources replaced with open circuits. (Dependent sources must be left intact, however.)

3.6.2 THE NORTON EQUIVALENT NETWORK

Recall that our goal is to find the v-i relation for the network in Figure 3.75a so that we can replace the network with a simple equivalent circuit that yields the same v-i relation as the original network. To find the v-i relationship, this time we apply a test voltage v_test to the circuit, as in Figure 3.75b, and find the resultant current i_t. Using superposition, the two subcircuits needed to find it are shown in Figure 3.75c and 3.75d. In 3.75c, v_test is set to zero and we measure i_a (-i_sc). In 3.75d, all independent sources are set to zero and we measure ib (v_test/R_t). Then, i_t = ia + ib.

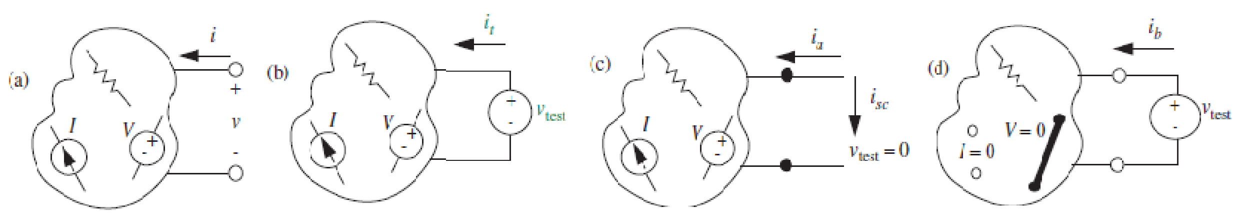


FIGURE 3.75 Derivation of Norton network

To complete the derivation, we find by superposition it = ia + ib = −isc + vtest/Rt (3.121)

As in the Thévenin derivation, this equation can be interpreted in terms of a circuit. It states that the terminal current is the sum of two components: a current source isc and a resistor current vtest/Rt. Hence the Norton equivalent network, Figure 3.76, has a current source in parallel with a resistor.

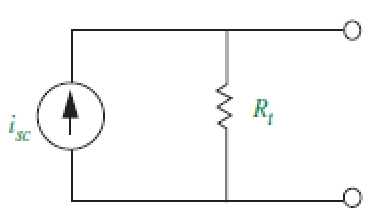


FIGURE 3.76 The Norton equivalent network.

Examination of either the two equations, Equations 3.121 and 3.115, or the two figures, Figures 3.56 and 3.76 show that there is a simple relation between voc and isc. Working from the figures, we can calculate the open-circuit voltage of each circuit to find voc = iscRt. (3.122) Thus it is a simple matter to change from one of these equivalent networks to the other.

In summary, the Norton method allows us to abstract the behavior of a linear network at a given pair of terminals as a current source in parallel with a resistor. The current source in parallel with the resistor is called the Norton equivalent circuit of the network.

A Method for Determining the Norton Equivalent Circuit. The Norton equivalent circuit for any linear network at a given pair of terminals consists of a current source iN in parallel with a resistor RN. The current iN and resistance RN can be obtained as follows:

1. iN can be found by applying a short at the designated terminal pair on the original network and calculating or measuring the current through the short circuit.
2. RN can be found in the same manner as RTH, that is, by calculating or measuring the resistance of the open-circuit network seen from the designated terminal pair with all independent sources internal to the network set to zero; that is, with voltage sources replaced with short circuits, and current sources replaced with open circuits.

CHAPTER 4: ANALYSIS OF NON-LINEAR CIRCUITS

4.1 INTRODUCTION TO NON-LINEAR ELEMENTS

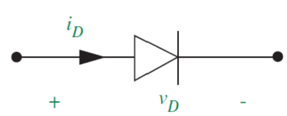


FIGURE 4.1 The symbol for a diode.

The first of the elements we see is the diode. In Fig. 4.1 the symbol for a diode. The diode is a two-terminal, nonlinear resistor whose current is exponentially related to the voltage across its terminals. An analytical expression for the nonlinear relation between the voltage vD and the current iD for the diode is the following:

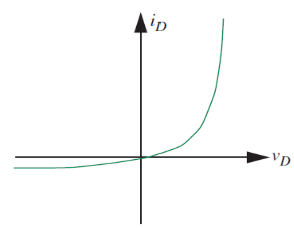


FIGURE 4.2 v–i characteristics of a silicon diode.

$$iD = Is\left(e^{\frac{vD}{VTH}} - 1\right) \ (4.1)$$

For silicon diodes the constant Is is typically 10^{-12} A and the constant VTH is typically 0.025 V. This function is plotted in Figure 4.2.

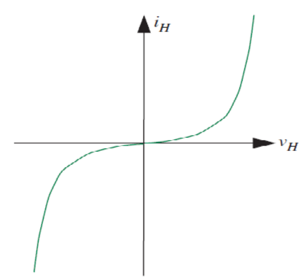


FIGURE 4.3 Another nonlinear v–i characteristics.

A second analytical expression for the relationship between voltage vH and current iH for another hypothetical nonlinear device is shown in Equation 4.2. In the equation, IK is a constant. The relationship is plotted in Figure 4.3.

$$iH = IKvH^3 \ (4.2)$$

The v-i relationship for yet another two-terminal nonlinear device is shown in equations below and figure 4.4.

$$iDS = \frac{K(vDS - VT)^2}{2}\ for\ vDS \geq VT\ (4.3)$$

$$iDS = 0\ for\ vDS < VT$$

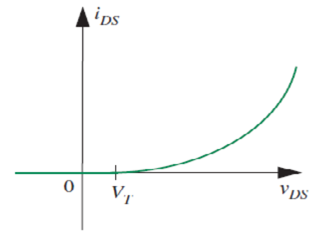


FIGURE 4.4 The v–i characteristics for a square law device.

A remark to make is that since the v-i relationship is not linear, it means the resistance of the diode also cannot be found using ohm's law. But, if I have the values at a point for V and I across a diode, then I can find the resistance of the diode in that point as R = V/I. This is called static resistance.

Given the above analytical expressions for the characteristic of a nonlinear device, how can we calculate the voltages and currents in a simple circuit? We will now discuss four methods for solving such nonlinear circuits.

4.2 ANALYTICAL SOLUTIONS

We first try to solve the simple nonlinear resistor circuit in Figure 4.9 by analytical methods. Assume that the hypothetical nonlinear resistor in the figure is characterized by the following v-i relationship:

$$iDS = KvD^2$$

$$iDS = 0\ for\ vD \leq 0$$

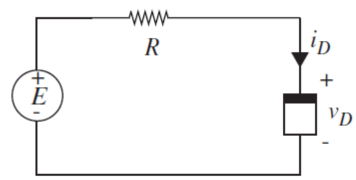


FIGURE 4.9 A simple circuit with a nonlinear resistor.

About linearity:

This circuit is amenable to a straightforward application of the node method. Recall that the node method and its foundational Kirchhoff's voltage and current laws are derived from Maxwell's Equations with no assumptions about linearity. Note, however, that the superposition method, the Thévenin and Norton methods do require a linearity assumption.

Method 1 – Direct use of node method:

Steps to solve the circuit:

1. Decide the ground node.

2. Write KCL for the node that has an unknown node voltage. As prescribed by the node method, we will use KVL and the device relation to obtain the currents directly in terms of the node voltage differences and element parameters. For the node with voltage vD:

$$\frac{vD - E}{R} + iD = 0$$

Then we substitute iD with the element law of the nonlinear resistor KvD^2. Solving the equation we can solve the circuit.

Method 2 – Thevenin/Norton plus node method:

A second example is when we have multiple linear devices and one nonlinear resistor. We can isolate the nonlinear resistor and solve first the linear part of the circuit applying the Thevenin method. Once we have the Thevenin equivalent circuit we are back to method 1: one voltage source, one linear resistor and one nonlinear resistor.

4.2 GRAPHICAL ANALYSIS

Many nonlinear circuits cannot be solved analytically. Graphical analysis can give us some help for some circuits, always given the <u>constraint that there is only one nonlinear element in the circuit</u>. This is because no matter how complicated the circuit, we can always reduce it, thanks to Thevenin and Norton, to the circuit we are about to see in fig. 4.16.

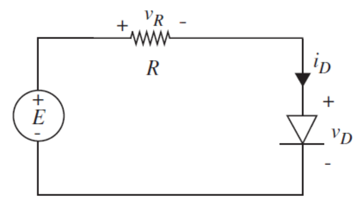


FIGURE 4.16 A simple non-linear

The circuit in figure 4.16 has a diode; we will try to solve it. By applying the analytical method just studied we arrive at the following two expressions that are not solvable without a computer.

$$\frac{vD - E}{R} + iD = 0 \quad \rightarrow \quad iD = -\frac{vD - E}{R}$$

$$iD = Is(e^{\frac{vD}{VTH}} - 1)$$

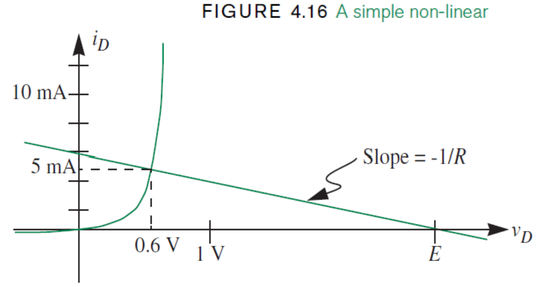


FIGURE 4.20 Graphical solution for diode circuit.

Using graphical solution though we can solve it, discounting some accuracy but getting much more insight. For concreteness E = 3V and R = 500Ω.

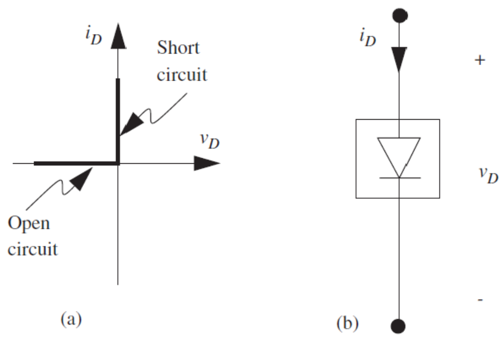
Plotting the equations above we find the point of intersection as in fig. 4.20.

It is easy to see from the construction that if E were made three times as large, the voltage across the diode would increase by only a small amount, perhaps to about 0.65 V. This illustrates the kind of insight available from graphical analysis.

4.3 PIECEWISE LINEAR ANALYSIS

With this method, we represent the nonlinear v-i characteristics of each nonlinear element by a succession of straight-line segments, then make calculations within each straight-line segment using the linear analysis tools already developed.

Simple example – the ideal diode model.



As we have seen the diode lets current pass exponentially as the input voltage increases, with a turning point usually at 0.7 volts. An approximation that preserves this dichotomy is the characteristic shown in Figure a: two linear segments intersecting at the origin, one of zero slope, indicating the behavior of an open circuit, the other infinite, indicating a short circuit. The abstraction is of sufficient use that we give it a special symbol, as shown in Figure b. This is yet another primitive in our vocabulary, called an ideal diode.

Considering the problem in figure 4.16 (4.24 here), and assuming case 1 is E=-5V and case 2 E=+3V, with R=500Ω in both cases, we can solve the two scenarios with the rules applied to linear circuits considering diode ON (short circuit) for case 2, and diode OFF (open circuit) for case 1. The result is a good approximation, and this is because the diode has a huge spike around 0.7 volts, resembling an open circuit.

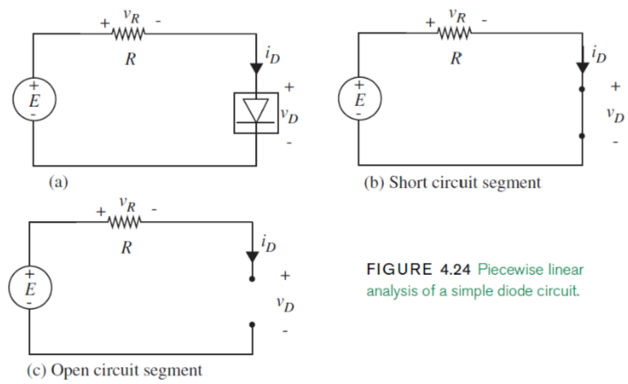


FIGURE 4.24 Piecewise linear analysis of a simple diode circuit.

A critical point of the method is to figure out the segment of operation (ON or OFF) associated with each of the nonlinear devices. This was not too hard with a single nonlinear device such as an ideal diode, but can be challenging when there are a number of nonlinear devices. It turns out that the approach that we discussed in this example generalizes to the method of assumed states, which will be discussed in more detail in Chapter 16.

Example 4.11, page 209 for a full numerical example.

Example 4.12, page 211 example with more complicated circuits – several independent sources and linear resistors, but only one nonlinear resistor as before. Method in a nutshell → Simplify the circuit to a linear circuit by using the piecewise method, and then solve the linear circuit with the methods studied.

4.5 INCREMENTAL ANALYSIS

There are many applications in electronic circuits where nonlinear devices are operated only over a very restricted range of voltage or current, as in many sensor applications. In such cases, it makes sense to find a piecewise linear device model in a way that ensures maximum accuracy of fit over that narrow operating range. This process of linearizing device models over a very narrow operating range is called incremental analysis or small-signal analysis (mainly applied to MOSFET). The benefit of incremental analysis is that the

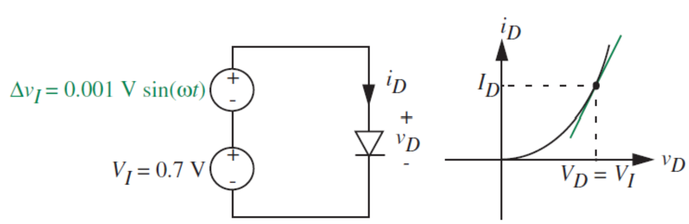


FIGURE 4.37 Incremental analysis.

incremental variables (ΔvD, ΔiD in our future examples) satisfy KVL and KCL, as well as linear v-i relations over the narrow operating range.

Using an example, let's say we need to find the diode current iD in figure 4.37 above. The diode is subject to a constant DC current plus a small AC current, so the diode operates under a narrow voltage range.

ANALYTICAL APPROACH: BAD IDEA

Using the analytical approach**,** we would have a complicated equation like this below, and thus we will abandon this approach.

$$iD = Is\left(e^{\frac{0.7V+0.001V\sin(wt)}{VTH}} - 1\right) \quad (4.53)$$

INCREMENTAL ANALYSIS – SMALL SIGNAL ANALYSIS

Using incremental analysis to find iD, it means that we want to approximate iD in the small operating range, which is in the vicinity of ID. This is done with the line in the graph in fig. 4.37. To do this, and find the approximation function, the first step is to approximate iD function using the Taylor Series expansion formula. Pay attention that the function we want to approximate with the Taylor Series expansion is not the original diode v-i equation, but an ID+ΔiD, which is:

$$iD = ID + \Delta\mathrm{iD} = \mathrm{Is}(\mathrm{e}^{\frac{\mathrm{VD}+\Delta\mathrm{vD}}{\mathrm{VTH}}} - 1)$$

Pages 215-218 explain the mathematical process behind the Taylor Series expansion, which is a way to express a complicated function as a sum of simpler terms, specifically polynomials. In other words, it helps create a "polynomial version" of a function that is easier to work with. In our example we will not consider second and higher order terms, and so we will end up with this equation:

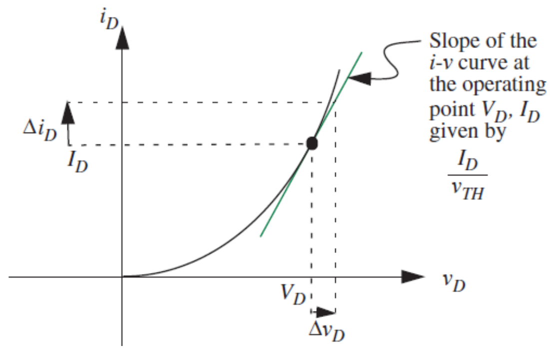
$$iD = ID + \Delta iD = IS\left(e^{\frac{VD}{VTH}} - 1\right) + \left(ISe^{\frac{VD}{VTH}}\right)\left(\frac{1}{VTH}\Delta vD\right) \; (4.61)$$

Equating corresponding DC terms and incremental terms:

$$ID = IS\left(e^{\frac{VD}{VTH}} - 1\right) \; (\mathbf{4.62})$$

$$\Delta iD = \left(ISe^{\frac{VD}{VTH}}\right)\left(\frac{1}{VTH}\Delta vD\right) \; (\mathbf{4.63})$$

Equation 4.62 is simply the diode equation and the point VD,ID of reference. Equation 4.63 is the straight line passing through the VD, ID point and tangent to the curve, as shown in fig. below. Equation 4.63 is linear with respect to ΔvD, which means I can make the small signal change, and I will move along that straight line. If the small signal is not far from VD, ID the approximation should be good enough.



We can also approximate equation 4.63 further and obtain:

$$\Delta iD = ID \frac{1}{VTH} \Delta vD \ (4.65)$$

For the future, skipping the mathematical work, we can find the relationship between the incremental parameters ΔiD and ΔvD directly using equation 4.71 below (always derived from the Taylor Series above):

$$\boxed{\Delta i_D = \left.\frac{df}{dv_D}\right|_{V_D} \Delta v_D. \quad (4.71)}$$

In words, the incremental change in the current is equal to df/dvD evaluated at vD = VD, multiplied by the incremental change in the voltage. In even simpler words, I take the original v-i equation for the nonlinear device and derive it with respect to vD at point VD and multiply the result by the incremental change in the voltage.

The same result can be developed graphically from fig 4.38 above (the figure just above). The incremental current ΔiD is simply the product of ΔvD and the slope of the iD versus vD curve at the point ID, VD. the slope of the iD versus vD curve at the point ID, VD is:

$$\left.\frac{df}{dv_D}\right|_{V_D}$$

SOME FUTHER SIMPLIFICATION:

Although this process yielded quickly the form of iD, a bit of insight will simplify the process even further by enabling the use of linear circuit techniques to solve the problem as promised in the introduction of this section. We proceed by drawing attention to Equation 4.61. Equation 4.61 is certainly nonlinear. But an important interpretation central to all incremental arguments allows us to solve the problem by linear circuit methods. Note from Equation 4.62 that the first term in Equation 4.61, the DC current ID, is independent of ΔvD. It depends only on the circuit parameters and the DC voltage VD which is the same as the DC source voltage VI. Thus, ID can be found with $\Delta vD = 0$. On this basis, the second term in Equation 4.61 is linear in ΔvD, because we have shown that there is no hidden ΔvD dependence in ID. Hence the second term, the change in the current i (ΔiD) is linearly proportional to the change in v (ΔvD), can be found from a linear circuit. In general, the incremental behavior of a nonlinear device is that of a linear resistor, whose value rd is given by:

$$r_d = \frac{1}{\left.\frac{df}{dv_D}\right|_{v_D=V_D}}. \quad (4.75)$$

Or for a diode, rd = VTH/ID.

SUMMARY

We began our analysis with the goal of determining the current (iD) through the diode when an input voltage in the form of a DC value (VD) plus a small time-varying component (ΔvD) is applied across it. Equation 4.61 shows that the resulting diode current is made up of two terms, a DC term, ID, which depends only on the DC voltage applied VD, and a small-signal or incremental term ΔiD, which depends on the small-signal voltage and also on the DC voltage VD. But for fixed VD, the incremental current ΔiD is linearly related to ΔvD. The constant of proportionality is a conductance "gd" given by Equation 4.73 (the slope of the iD vD curve in VD,ID). The reverse of this conductance is the resistance rd specified in equation 4.75 above. Because the incremental circuit model of Figure 4.39 correctly represents the relationship between ΔiD and ΔvD, this linear circuit can be used to solve for ΔiD. In many situations, only the incremental change in the output is of interest, but in case we can also sum the two components to get the total.

CONCLUSION AND 2 STEPS METHOD:

Thus, based on the preceding discussion, a systematic procedure for finding incremental voltages and currents for a circuit with a nonlinear device characterized by the v−i relation iD=f(vD) is as follows:

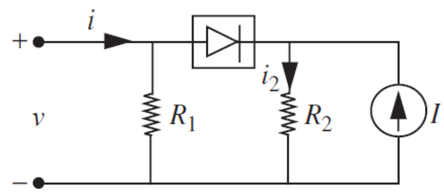
1. Find the DC operating variables, ID and VD, using the subcircuit derived from the original circuit by setting all small-signal sources to zero. Any of the methods of analyzing nonlinear circuits discussed in the preceding sections—analytical, graphical, or piecewise linear—is appropriate.
2. Find the incremental output voltage and incremental nonlinear device current (the change away from the DC variables calculated in Step 1) by forming an incremental subcircuit in which the nonlinear device is replaced by a resistor of value rd (computed as shown in Equation 4.75), and all DC sources are set to zero. (That is, voltage sources are replaced by short circuits, and current sources by open circuits.) The incremental subcircuit is linear, so incremental voltages and currents can be calculated by any of the linear analysis techniques developed in Chapter 3, including superposition, Thévenin, etc.

Example 4.19 applies the method we just laid out here.

4.6 EXTRA PARAGRAPH: FUNNY STORY ABOUT A DIODE

What we learned so far it is that a diode is forward biased, meaning that it lets current pass only in one direction. All good except that we stumbled on exercise 4.6 here below.

EXERCISE 4.6 For the circuit in Figure 4.52, find the input characteristic, i versus v, and the transfer characteristic i_2 versus v. I is fixed and positive. Express your results in graphs, labeling all slopes, intercepts, and coordinates of any break points.



When the diode is ON, the current is divided between the two resistors as it is normal for a current divider circuit. But the current that gets divided is the sum of the currents coming from I and i. In the case of i2 we have:

$$i2 = (I + i)\frac{R1}{R1 + R2}$$

4.7 ZENER DIODE

Forward Bias: Like a regular diode conducts current in the forward direction when the voltage across it exceeds a certain threshold (0.7V for a silicon diode).

Reverse Bias: In reverse bias (when the voltage is applied in the opposite direction), a regular diode blocks current flow, acting like an open circuit. A Zener diode is instead designed to conduct in the reverse direction when the reverse voltage reaches a specific value, known as the Zener breakdown voltage (or Zener voltage). Once the Zener breakdown voltage is reached, even if the voltage source voltage increases dramatically, the voltage across the Zener diode will not change. This characteristic makes it useful for voltage regulation and protection in circuits.

The Voltage Regulation Mechanism:

A Zener diode is placed in reverse bias across the load you want to regulate. A resistor R is placed in series with the Zener diode to limit the current flowing through the diode. This resistor is crucial because it helps drop the excess voltage when the input voltage exceeds the Zener voltage.

When the input voltage Vin is below the Zener breakdown voltage Vz, the Zener diode does not conduct, and the output voltage Vout is just Vin minus the small drop across the resistor. When Vin exceeds Vz, the Zener diode begins to conduct. The Zener diode clamps the voltage across itself (and therefore across the load) to the Zener voltage Vz.

Example:

Zener Voltage (Vz): Let's say the Zener diode is rated for 5V.

Input Voltage (Vin): Suppose your input voltage can vary from 6V to 12V.

Current-Limiting Resistor (R): A resistor is used to limit the current to safe levels.

Case 1: Vin=6V (Just Above Zener Voltage)

The Zener diode will start conducting when Vin reaches slightly above 5V.

The Zener diode clamps Vout at 5V.

The current through the resistor R will be $I = \frac{Vin - Vz}{R}$, in our case $I = \frac{6-5}{R} = \frac{1}{R}$.

Case 2: Vin=12V (Much Higher than Zener Voltage)

The Zener diode still clamps the output voltage Vout at 5V.

The excess voltage (12V - 5V = 7V) is dropped across the resistor R.

The current through the resistor increases to $I = \frac{12-5}{R} = \frac{7}{R}$.

Key Points:

Voltage Clamping: No matter how much the input voltage increases beyond the Zener voltage, the Zener diode keeps the output voltage at the Zener voltage level by allowing more current to flow through the resistor.

Power Dissipation: The resistor R dissipates the extra power as heat. The Zener diode's job is to maintain a stable voltage across the load, and it does this by increasing the current through R to accommodate the higher input voltage.

CHAPTER 5: THE DIGITAL ABSTRACTION

Value discretization forms the basis of the digital abstraction. The idea is to lump signal values that fall within some interval into a single value.

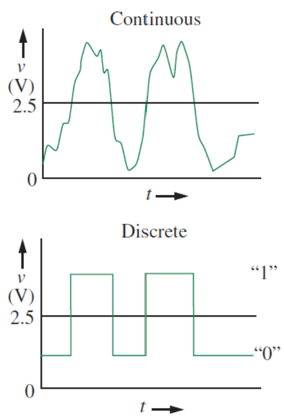


FIGURE 5.1 Value discretization into two levels.

An example of value discretization is in Figure 5.1, where a voltage signal is discretized into two levels. An observed voltage value between 0 volts and 2.5 volts is treated as a ''0,'' and a value between 2.5 volts and 5 volts as a ''1''. The discrete signal shown in Figure 5.1 comprises the sequence of values 0-1-0-1-0.

Although the digital approach seems wasteful of signal dynamic range, it has a significant advantage over analog transmission in the presence of noise. Clearly, unless the error (noise) causes the signal to cross the 2.5 volts boundary, the value will be transmitted correctly.

Most of this chapter deals exclusively with signals that can take on one of two values. Each of these two-level signals is communicated over a single wire. When each digit takes on one of two values, the digit is called a **binary digit, or bit**. Much as the familiar decimal system uses multiple digits to represent numbers other than 0 through 9, the binary system uses multiple bits to represent numbers other than 0 or 1. Multi-bit signals are commonly transmitted by allocating multiple wires, one for each bit. The two levels in the binary representation are variously called (a) TRUE or FALSE, (b) ON or OFF, (c) 1 or 0, (d) HIGH or LOW. Digital signals are commonly implemented using voltage levels, for example, 0 V to represent FALSE, and 5 V to represent TRUE.

The two-level representation is commonly known as the binary representation. Virtually all digital circuits use the binary representation because two-level circuits are much easier to build than multilevel circuits.

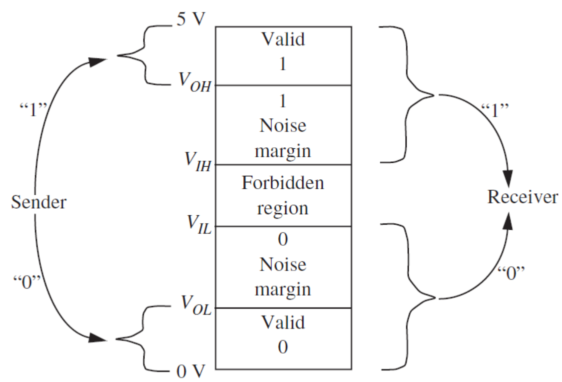
5.1 VOLTAGE LEVELS AND THE STATIC DISCIPLINE

The static discipline is a specification (set of rules) for digital devices. The static discipline requires devices to adhere to a common representation, and to guarantee that they interpret correctly inputs that are valid logical signals according to the common representation, and to produce outputs that are valid logical signals, provided they receive valid logical inputs. By adhering to a common representation, digital devices based on different technologies or built by different manufacturers can communicate with each other.

Our representation in figure 5.8 below divides a voltage range into different intervals. The goal is to associate a logic value, 0 or 1, to a voltage level sent by a sender and received by a receiver, considering the possibility of noise. The following considerations are important and describe figure 5.8:

- Signals in the forbidden regions are invalid. Forbidden region allows for a buffer zone.
- Signals > 5V and < 0V are correctly interpreted so there are no higher or lower boundaries.
- Setting different limits to the values that senders can send and receivers can receive is a way to limit the possibility of error due to noise. VOL (VOH) are the max (min) values a sender can send, and VIL (VOH) are the max (min) values a receiver can interpret. Noise margin for a logical 0 is given by NM(0) = VIL – VOL, and the noise margin for a logical 1 is given by NM(1) = VOH − VIH. The region between VIL and VIH is the forbidden region. If NM(1) = NM(0) the noise margins are symmetric.

FIGURE 5.8 A mapping between voltage levels and logical signals that provides noise margins. For a logical high, senders must output values in the V_{OH} to 5-V range. For a logical low, senders must output values in the 0 V to V_{OL} range. Receivers must correspondingly interpret values greater than V_{IH} as a logical high, and output values lower than V_{IL} as a logical low.



5.2 BOOLEAN LOGIC

The binary representation has a natural correspondence to logic, and therefore digital circuits are commonly used to implement logic procedures.

OPERATOR	SYMBOL
AND	·
OR	+
NOT	~

TABLE 5.2 Some logic operations and their symbols.

X	Y	Z
0	0	0
0	1	0
1	0	0
1	1	1

TABLE 5.3 Truth table for $Z = X \cdot Y$.

As an example, let's consider the statement "If X is TRUE AND Y is TRUE then Z is TRUE else Z is FALSE", which can be summarized as:

$$Z = X \text{ AND } Y \quad / \quad Z = X \cdot Y \quad / \quad Z = XY$$

We often find it convenient to use a truth table representation of Boolean functions. A truth table enumerates all possible input value combinations and the corresponding output values. In table 5.3 the truth table for the statement just presented.

5.3 COMBINATIONAL GATES

Another possible representation of Boolean functions makes use of the combinational gate abstraction. The output of combinational gates:

- is purely a function of their inputs.
- Therefore, combinational functions can always be enumerated using truth tables.
- Combinational gates follow the set of rules that form the static discipline.
- Provided they are given inputs that fall within valid input levels, they will produce outputs that satisfy valid output thresholds.

Figure 5.14 shows several useful gate symbols and their truth tables in table 5.8.

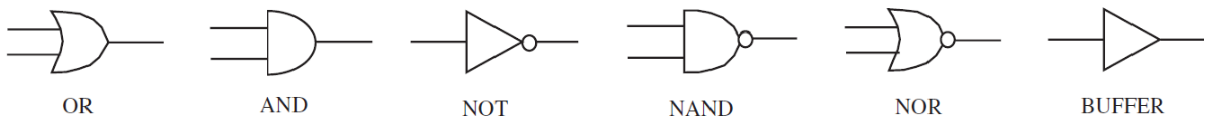


FIGURE 5.14 Gate symbols.

INPUTS		AND	OR	NAND	NOR
B	C	$B \cdot C$	$B + C$	$\overline{B \cdot C}$	$\overline{B + C}$
0	0	0	0	1	1
0	1	0	1	1	0
1	0	0	1	1	0
1	1	1	1	0	0

TABLE 5.8 Truth table for several two-input functions.

As we might expect, gates are implemented using lumped circuit elements, such as resistors and current sources. In other words, a gate representing a function F(x) is simply an abstraction for a circuit that performs the function F(x).

5.4 STANDARD SUM-OF-PRODUCTS REPRESENTATION

Here we introduce a standard or canonic form of writing logic expressions called the sum-of-products form. As the name implies, logic expressions in the sum-of-products form are **represented using two levels of operations as a set of product terms (AND)**, each comprising one or more variables in their true forms (for example, A) or complement forms (e.g., A), **combined using the OR function**.

<u>How to write a sum-of-products expression from a truth table:</u>

We can write a sum-of-products expression from a truth table representation by first writing a product term for each row in the truth table with a 1 in its output column, and then summing these product terms. Each product term comprises an AND function of all the input variables. A variable will appear in its true or complement form in a product term corresponding to a given row in the truth table depending on whether it appears as a 1 or a 0 in that row.

X	Y	Z
0	0	0
0	1	1
1	0	1
1	1	1

TABLE 5.4 Truth table for $Z = X + Y$.

Thus, for example, a logic expression for the truth table in Figure 5.4 is Z = $\bar{X}$Y + X$\bar{Y}$ + XY. By construction, this expression is in a sum-of-products form.

5.5 SIMPLIFYING LOGIC EXPRESSIONS

Logic expressions can often be simplified to reduce their complexity. Some primitive rules (5.9 to 5.22) and the Morgan's Laws (5.23, 5.24) to help in the simplification are below:

$A \cdot \overline{A} = 0$ (5.9)	$A + \overline{A} = 1$ (5.13)
$A \cdot A = A$ (5.10)	$A + A = A$ (5.14)
$A \cdot 0 = 0$ (5.11)	$A + 0 = A$ (5.15)
$A \cdot 1 = A$ (5.12)	$A + 1 = 1$ (5.16)
$A + \overline{A}B = A + B$ (5.17)	$A + B = B + A$ (5.20)
$A(B + C) = AB + AC$ (5.18)	$(AB)C = A(BC)$ (5.21)
$AB = BA$ (5.19)	$(A + B) + C = A + (B + C)$ (5.22)
$\overline{A \cdot B} = \overline{A} + \overline{B}$ (5.23)	$\overline{A + B} = \overline{A} \cdot \overline{B}$ (5.24)

All these rules can be verified with truth tables.

5.6 NUMBER REPRESENTATIONS

The bit can represent two numbers, 0 and 1. How do we represent other numbers? In general, the value of the binary number A(n)A(n−1) . . . A(2)A(1)A(0) is given by

$$\sum_{i=0}^{i=n} Ai2^{i}$$

Thus, the binary number 10 corresponds to the decimal number 2, the binary number 11 corresponds to the decimal number 3, and the binary number 101 corresponds to the decimal number 5.

How do we represent negative numbers? One simple alternative is to interpret the leading bit as a sign bit: a 0 denotes a positive number and a 1 denotes a negative number. Therefore, the number 110 represents −2, and the number 010 represents 2.

Operations can be performed on binary numbers following the same logic as decimal numbers. For example, to add the binary number 10 and 11, we first add 0 and 1, which is 1, then we add 1 and 1 which is 0; but there is a rest of 1 that goes into the third place, so the result is 101. Translating these numbers into decimal numbers we have 2+3=5. 5 in binary is indeed 101.

Example 5.19. Suppose we want to add a pair of two-bit positive numbers A: A1A0 and B: B1B0. We will implement a two-bit adder using two techniques.

The first method starts by writing the truth table below from which we obtain the sum-of-products in figure. Figure 5.28 displays a gate-level implementation. I guess we have S2,S1,S0 because the sum of A1A0+B1B0 is a binary number of three digits; we indeed can try to sum A+B by row and we obtain S.

A_1	A_0	B_1	B_0	S_2	S_1	S_0
0	0	0	0	0	0	0
0	0	0	1	0	0	1
0	0	1	0	0	1	0
0	0	1	1	0	1	1
0	1	0	0	0	0	1
0	1	0	1	0	1	0
0	1	1	0	0	1	1
0	1	1	1	1	0	0
1	0	0	0	0	1	0
1	0	0	1	0	1	1
1	0	1	0	1	0	0
1	0	1	1	1	0	1
1	1	0	0	0	1	1
1	1	0	1	1	0	0
1	1	1	0	1	0	1
1	1	1	1	1	1	0

TABLE 5.12
Truth table for the two-bit adder.

each of S_0, S_1, and S_2 as follows:

$$\begin{aligned}
S_0 = {} & \bar{A}_1\bar{A}_0\bar{B}_1B_0 + \bar{A}_1\bar{A}_0B_1B_0 \\
& + \bar{A}_1A_0\bar{B}_1\bar{B}_0 + \bar{A}_1A_0B_1\bar{B}_0 \\
& + A_1\bar{A}_0\bar{B}_1B_0 + A_1\bar{A}_0B_1B_0 \\
& + A_1A_0\bar{B}_1\bar{B}_0 + A_1A_0B_1\bar{B}_0 \qquad (5.39) \\
= {} & \bar{A}_0B_0 + A_0\bar{B}_0 \qquad (5.40)
\end{aligned}$$

$$\begin{aligned}
S_1 = {} & \bar{A}_1\bar{A}_0B_1\bar{B}_0 + \bar{A}_1\bar{A}_0B_1B_0 \\
& + \bar{A}_1A_0\bar{B}_1B_0 + \bar{A}_1A_0B_1\bar{B}_0 \\
& + A_1\bar{A}_0\bar{B}_1\bar{B}_0 + A_1\bar{A}_0\bar{B}_1B_0 \\
& + A_1A_0\bar{B}_1\bar{B}_0 + A_1A_0B_1B_0 \qquad (5.41) \\
= {} & A_1A_0B_1B_0 + A_1\bar{B}_1\bar{B}_0 \\
& + A_1\bar{A}_0\bar{B}_1 + \bar{A}_1B_1\bar{B}_0 \\
& + \bar{A}_1A_0\bar{B}_1B_0 + \bar{A}_1\bar{A}_0B_1 \qquad (5.42)
\end{aligned}$$

$$\begin{aligned}
S_2 = {} & \bar{A}_1A_0B_1B_0 + A_1\bar{A}_0B_1\bar{B}_0 \\
& + A_1\bar{A}_0B_1B_0 + A_1A_0\bar{B}_1B_0 \\
& + A_1A_0B_1\bar{B}_0 + A_1A_0B_1B_0 \qquad (5.43) \\
= {} & A_1B_1 + A_1A_0B_0 + A_0B_1B_0 \qquad (5.44)
\end{aligned}$$

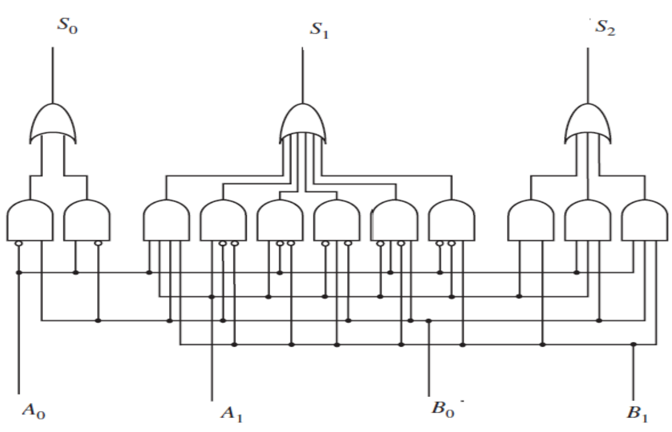


FIGURE 5.28 Direct implementation of two-bit adder.

The second method develops a two-bit adder circuit by composing two one-bit adders, and illustrates an important engineering technique called Divide-And-Conquer. The one-bit adders are called full adders. A full adder takes three inputs: two one-bit numbers to be added (Ai and Bi), and one carry bit Ci from a lower digit. The full adder produces two outputs: a sum bit Si and a carry bit Ci+1 to a higher digit. The truth table for the one-bit full adder is depicted in Table 5.13. From the table, we derive the logic expression for the sum bit Si and the carry bit Ci+1 in the figure below.

We can create a two-bit adder out of a pair of one-bit full adders by feeding the carry-out bit (Ci+1) of one adder to the carry-in bit (Ci) of another adder. The carry-in bit of the low-digit adder is set to 0. The two bit-adder is shown in Figure 5.31. Because the carry bit ripples through the adders, this type of adder circuit is called a ripple-carry adder. Similarly, we can use n one-bit full adders to construct an n-bit adder.

A_i	B_i	C_i	C_{i+1}	S_i
0	0	0	0	0
0	0	1	0	1
0	1	0	0	1
0	1	1	1	0
1	0	0	0	1
1	0	1	1	0
1	1	0	1	0
1	1	1	1	1

TABLE 5.13 Truth table for the one-bit full adder.

$$S_i = \bar{A}_i\bar{B}_iC_i + \bar{A}_iB_i\bar{C}_i + A_i\bar{B}_i\bar{C}_i + A_iB_iC_i \tag{5.45}$$

$$C_{i+1} = \bar{A}_iB_iC_i + A_i\bar{B}_iC_i + A_iB_i\bar{C}_i + A_iB_iC_i. \tag{5.46}$$

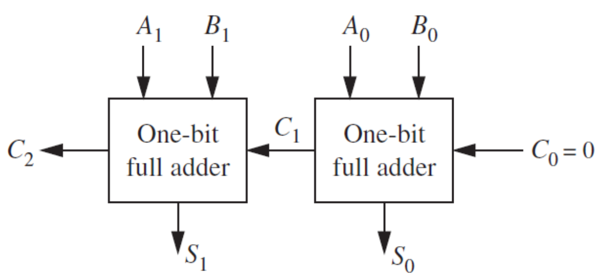


FIGURE 5.31 A two-bit ripple-carry adder using two one-bit full adders.

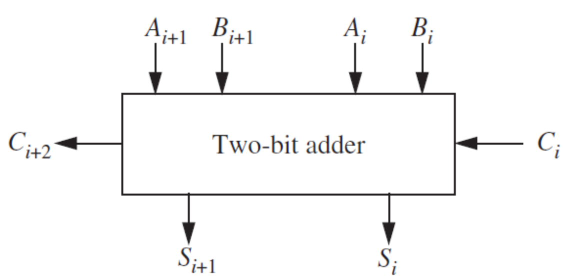


FIGURE 5.32 A two-bit adder block.

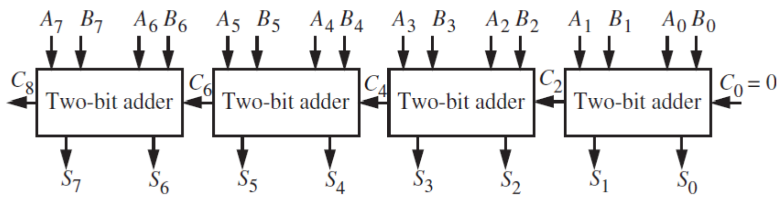


FIGURE 5.33 An eight-bit adder circuit.

CHAPTER 6: THE MOSFET SWITCH

6.1 THE SWITCH

In Figure 6.1 a simple circuit with a switch. The switch is normally off and behaves like an open circuit. When pressure is applied to the switch, it closes and behaves like a wire and conducts current. Accordingly, the switch can be modeled as a three-terminals device which includes a control terminal, an input terminal, and an output terminal. When the control terminal has a TRUE or a logical 1 signal on it, the input is connected to the output through a short circuit, and the switch is said to be in its ON state. Otherwise, there is an open circuit between the input and the output, and the switch is said to be in its OFF state.

The V/I characteristics of a switch can be expressed in algebraic form as:

- if Control = 0 → I = 0
- if Control = 1 → V = 0

Although the switch is a nonlinear device, circuits containing a switch and other linear devices can be analyzed by considering two linear subcircuits; one when the switch is in its ON state and one for the switch in its OFF state. Thus, standard linear techniques can be applied to each subcircuit.

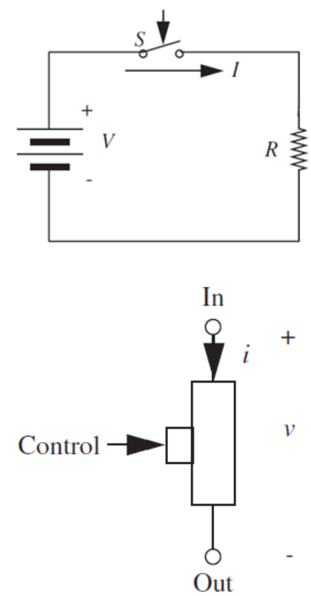
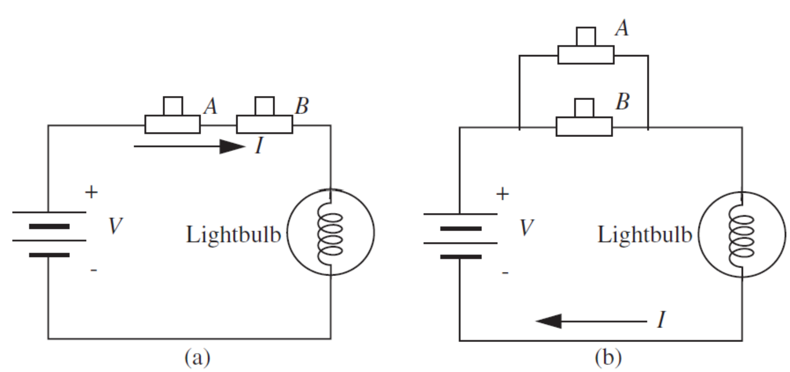


FIGURE 6.1 Circuit with a switch & three-terminal switch representation

6.2 LOGIC FUNCTIONS USING SWITCHES

The picture below is self-explanatory. Switches can be combined in AND-OR configurations to implement more complicated functions.

FIGURE 6.6 (a) The lightbulb circuit with switch in an *AND* configuration; (b) the lightbulb circuit with switches in an *OR* configuration.



The switches we looked at so far require mechanical pressure to turn on and off. Preferably, a three-terminal switch device that responded to voltages would enable the construction of switching circuits using voltages alone. The MOSFET is one such device that can be implemented cheaply in VLSI technology.

> *"VLSI stands for Very Large-Scale Integration. It refers to the process of creating integrated circuits (ICs) by combining thousands to millions of transistors onto a single semiconductor chip. This technology is fundamental to modern electronics, enabling the creation of complex electronic systems within compact devices."* CHATGPT

6.3 THE MOSFET DEVICE AND ITS S MODEL

6.3.1 THE MOSFET DEVICE

The MOSFET belongs to a class of devices called transistors. The MOSFET is a three-terminal device with:

- a control terminal, called its gate G.
- an input terminal, called its drain, D.
- an output terminal, called its source, S.

Current flows from the drain (terminal with the higher voltage) to the source.

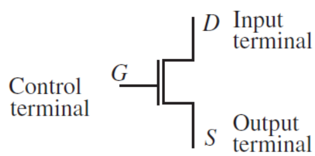


FIGURE 6.8 The MOSFET circuit symbol.

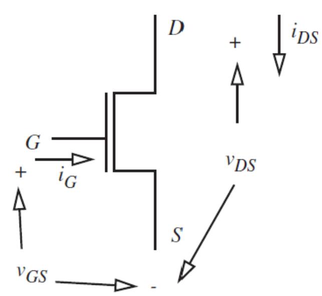


FIGURE 6.9 Definitions of v_{GS}, v_{DS}, and i_{DS}.

As depicted in Figure 6.9:

- voltage across the gate and source of the MOSFET is vGS.
- voltage across the drain and the source is vDS.
- current through G terminal is iG, and through the D terminal is iDS.

6.3.2 THE MOSFET'S SWITCH MODEL (S MODEL)

A simple circuit model for a specific type of MOSFET device called the n-channel MOSFET is depicted in Figure 6.10. This model based on the simple switch is called the MOSFET's Switch Model, or S Model.

- When vGS > vT, then the device is ON. Usually, VT is 0.7 volts.

- In the ON state, the S model approximates the connection between the drain and the source as a short circuit. In practice, there is some nonzero resistance between the drain and the source, but we ignore it now. In the OFF state, an open circuit exists between the drain and the source.
- An open circuit exists between the gate and the source, and between the gate and the drain at all times. Thus, iG = 0 always. *

We can summarize the S model for the MOSFET in algebraic form by stating its v-i characteristics as follows:

- If vGS < vT, iDS = 0
- if vGS ⩾ vT, vDS = 0

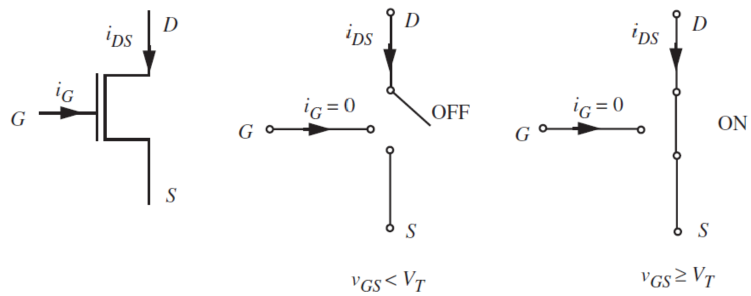


FIGURE 6.10 The S model of the MOSFET.

* **Why iG=0 always?** *"The gate and source terminals are separated by a thin layer of insulating material (typically silicon dioxide, SiO_2). Because of this insulation:*

- *No current flows directly between the gate and the source under normal operating conditions.*
- *Instead, the gate voltage (vGS) controls the formation of a conductive channel between the source and drain.*

In a MOSFET, the gate-source combination behaves like a capacitor. When you apply a voltage between the gate and source (vGS > vT), you induce an electric field across the insulating layer, but no actual current flows between the gate and source because they are insulated from each other.

This is why iG = 0 at all times—there is no conductive path for current through the gate. The MOSFET turns "on" when vGS ⩾ vT because the electric field attracts charge carriers (electrons or holes) to form a channel between the drain and source, allowing current (ID) to flow between them." CHATGPT

6.4 MOSFET SWITCH IMPLEMENTATION OF LOGIC GATES

Let us now build logic gates using MOSFETs. Consider the circuit shown in Figure 6.14.

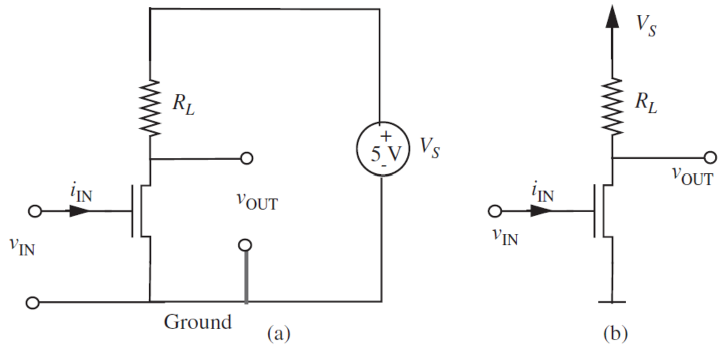


FIGURE 6.14 (a) The MOSFET inverter; (b) the same inverter circuit using the shorthand notation for power and ground.

Let us analyze the behavior of the circuit by replacing the MOSFET with its equivalent S model. Figure 6.17 displays the equivalent model for the circuit shown in Figure 6.14. Assumption is that a logical high is represented using 5 V and a logical low using 0 V.

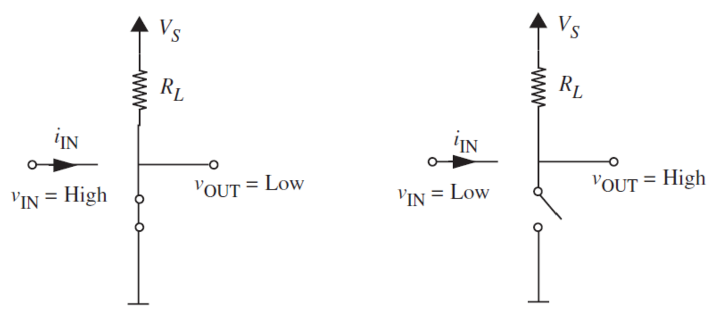


FIGURE 6.17 The S circuit model of the n-channel MOSFET inverter.

When the input vIN is high, the MOSFET is in the ON state (if vIN > VT), thereby pulling the output voltage to a low value. When vIN is low, the MOSFET is off, and the output is raised to a high value by RL.

The purpose of the load resistor RL is to provide a logical 1 output when the MOSFET is off. When the MOSFET is on, RL limits the flow of current; and this is why RL is usually chosen to be large.

Because the resistance between the gate-to-source and the drain-to-source ports of the MOSFET is infinite in the S model, the current iIN is 0.

> ***Why does it work like this?*** *I explain through a circuit I built, according to fig. 6.14, to which I linked a LED between vOUT and GND. If the switch is off, the LED lights up, as it is the only path for current to flow to GND. If the switch is on, the current has two paths that can choose to GND, one path with the LED, which has a small resistance, and one path through the switch, with no resistance at all. As we know, current flows through the path of less resistance and thus through the switch, effectively shutting down the LED. From my point of view RL is needed because if the switch is on, without RL I would have a short circuit. This is the most important point. So if switch is on, the LED is off, if switch is off the LED is on: behavior of an inverter.*

The logical values IN and OUT represented by vIN and vOUT exhibit the behavior of an inverter. Figure 6.18 shows a sample input waveform and the corresponding output waveform for our circuit and the truth table of the circuit.

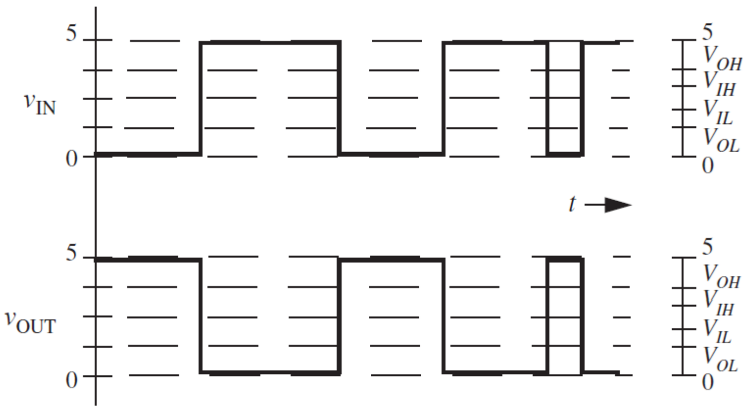


FIGURE 6.18 Sample input-output waveforms for the inverter.

IN	OUT
0	1
1	0

TABLE 6.1 Truth table for the MOSFET circuit.

We can also construct other gates in a like manner. Figure 6.20 shows a NAND gate circuit and Figure 6.21 shows its equivalent S circuit model. Using intuition from the two-input NAND circuit, we can build multiple input NAND and NOR circuits. Figure 6.22a shows an n-input NOR gate and 6.22b shows an n-input NAND gate. In the multiple input NOR gate, the output is pulled to ground when any of the inputs is high. Correspondingly, in the NAND gate, the output remains high if even one input is low.

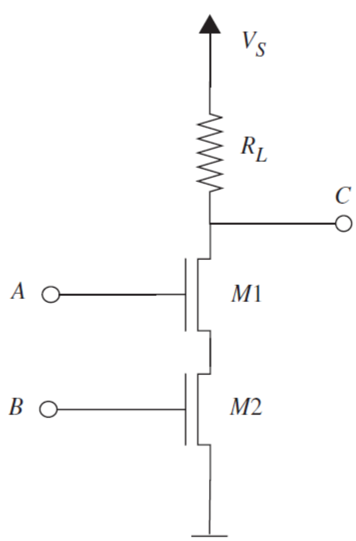


FIGURE 6.20 The circuit for a NAND gate.

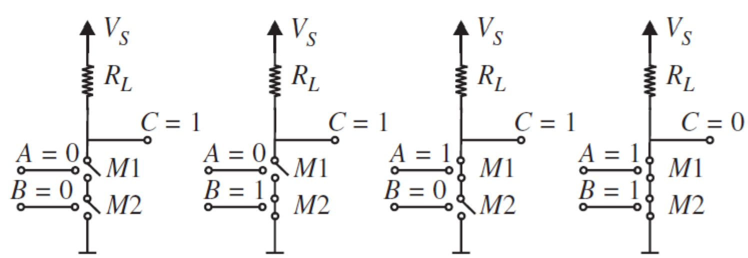


FIGURE 6.21 The S circuit model for a NAND gate.

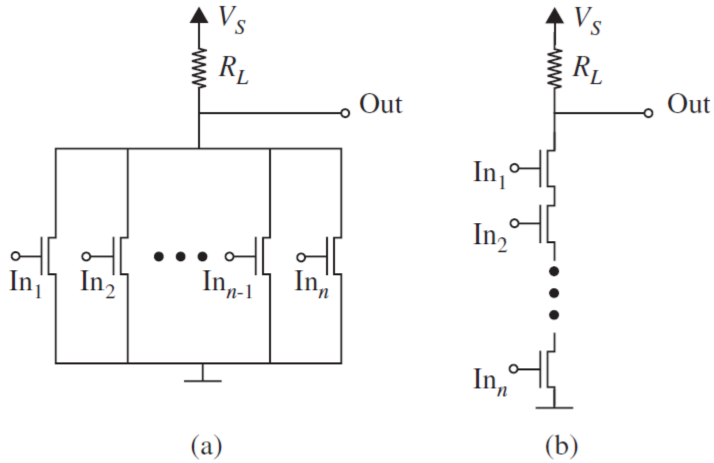


FIGURE 6.22 Multiple-input NOR and NAND gates.

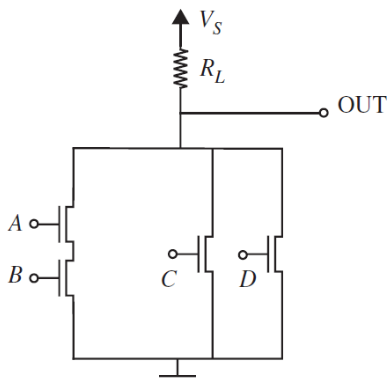


FIGURE 6.23 Transistor-level implementation of $\overline{AB + C + D}$.

Practical example: Recall the combinational logic expression whose gate-level implementation we had seen earlier:

$$\overline{AB + C + D}$$

Previously, we had implemented the expression using several abstract logic gate elements. Now that we understand how gates are constructed using MOSFETs and resistors, we can construct a single compound combinational logic gate using MOSFETs and resistors that implement each of these functions. Using the intuition that switches in series implement the AND property and switches in parallel implement the OR property, we can implement the expression as in Figure 6.23. By checking the circuit against its truth table, we see that the circuit does work as desired.

It is important to point out two key properties of the MOSFET that make it an ideal component for building gates:

1. First, notice that we could compose multiple gate components into more complicated circuits without worrying about the internal circuit of the gates. The reason we can do so is that the output of the MOSFET has no effect on its inputs. In other words, although the input voltage at G impacts the behavior of the MOSFET at its D and S terminals, the voltages or currents at its D and S terminals have no impact on G.

2. Second, the infinite resistance seen at the gate (G terminal) of a MOSFET makes it have no effect on the output of another gate driving its input. This feature of the MOSFET allows us to build systems containing many gates without worrying about how each gate affects the logical properties of other gates to which it is connected. This property of a gate is called composability. Imagine if the MOSFET input had zero resistance. In that case, we would not be able to connect the output of one inverter to the input of another and expect the first inverter to satisfy the static discipline.

6.5 STATIC ANALYSIS USING THE S MODEL

To respect a static discipline, outputs must meet the output constraints specified, provided they receive inputs that meet the input constraints.

Let's determine whether the inverter studied in the previous paragraph satisfies a given static discipline. Based on the inverter characteristics here below, we can determine whether the static discipline is satisfied, for example for these voltage thresholds: V(OH)=4.5V, V(OL)=0.5V, V(IH)=4V, V(IL)=0.9V.

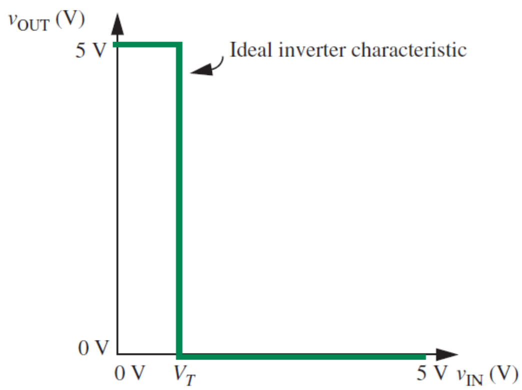


FIGURE 6.27 The transfer characteristic of the inverter.

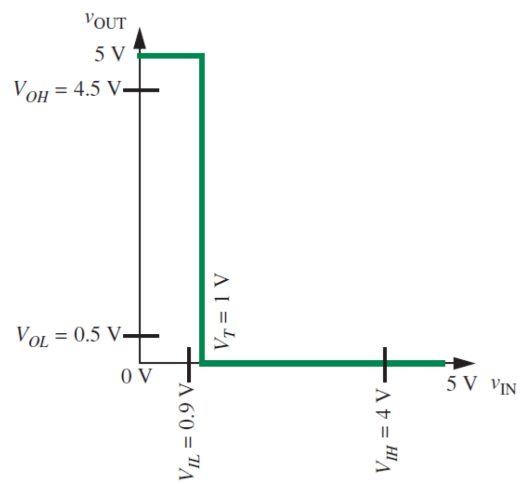


FIGURE 6.28 A mapping between logic values and voltage levels corresponding to a static discipline appropriate for the inverter.

V(OH): Since inverter produces output high (5V) > 4.5V, static discipline is respected.

V(OL): Since inverter produces an output low (0V) < 0.5V, static discipline is respected.

V(IH): Since V(IH) = 4V, our inverter must interpret any V>4 as logical 1. Since the inverter turns on when V > VT(1V) and pulls the output to low voltage, the condition is met.

V(IL): Inverter must interpret any V < 0.9V as a logical 0. Our inverter is off when input V < 1V, so condition is met.

6.6 THE SR MODEL OF THE MOSFET

The S model for the MOSFET is a great simplification of its actual properties. In particular, a practical MOSFET, when it is in its on state, displays a non-zero resistance between its D and S terminals.

A more accurate model for the MOSFET uses a resistance R(ON) in place of the short between D and S when the MOSFET is on. As before there is an open circuit between the gate and source terminals and the gate and drain terminals so iG=0. See fig. 3.999.

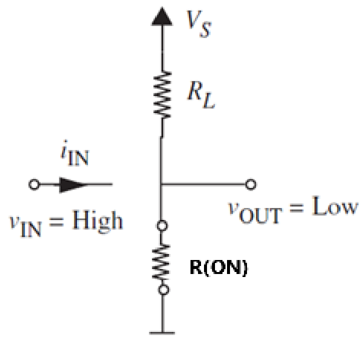


Fig. 3.999

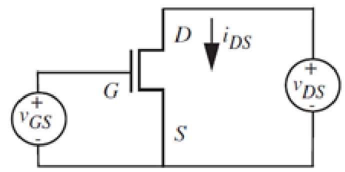


FIGURE 6.30 Setup for observing MOSFET characteristics.

There are still several big simplifications in the SR models though (see fig. 6.30 for reference):

- Variable resistance. Even if SR model displays resistive behavior when vDS ≤ vGS – vT, the resistance R(ON) is not fixed but should be a function of vGS.
- Resistor becomes a current source. When vDS ≥ vGS – VT, the drain-to-source behavior is not resistive at all, rather it is that of a current source.

The SR model does not solve the two simplifications just outlined, but some constraints are put on the model to live with them:

1. The SR model is valid only when vDS ≤ vGS – vT. By doing so we do not have to worry about the resistor R(OL) becoming a current source.
2. There is only one value for the gate voltage when the input is high (ie vGS=VS). By doing so we can have a single fixed value for R(OL). This allows the RS model to be useful to analyze some aspects of digital circuits because the gate can only be high or low – which translate to resistor R(OL) or open circuit.

The characteristics of the MOSFET according to the SR model are graphically displayed in Figure 6.31.

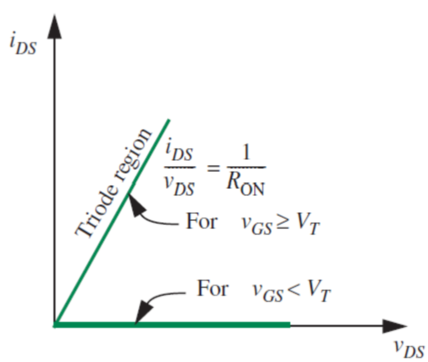


FIGURE 6.31 Characteristics of the MOSFET according to the SR model. The top-left, bottom-left, and the bottom-right quadrants are not shown for the reasons given in Section 6.3. As discussed in more detail in Chapter 7, the region in which the MOSFET displays resistive behavior is within the triode region of MOSFET operation.

The SR model for the MOSFET can also be expressed as:

iDS = vDS/R(ON) for vGS ≥ VT;

iDS = 0 for vGS < VT.

6.7 PHYSICAL STRUCTURE OF THE MOSFET

Physical structure:

Here is a simple breakdown of how MOSFETs are physically constructed.

- Starting Material: The MOSFET is built on a single-crystal silicon wafer, which serves as a base. This wafer is flat and can be quite large, though the individual MOSFETs on it are extremely small, often less than a square micrometer. One micrometer (µm), or micron micron (µ) is 10^{-6} meters.
- Layering: MOSFET construction involves creating multiple layers on the silicon wafer:
    - Insulating Layers: Made from silicon dioxide (SiO_2), created by oxidizing parts of the silicon surface.
    - Conducting Layers: Often metals like aluminum or copper, or sometimes polycrystalline silicon, which is a specific form of silicon.
    - Semiconducting Layers: Silicon that has been doped with impurities to add free electrons or holes, making it conductive.
- Doping: This is the process of adding impurities to silicon, either by diffusion or ion implantation, to create two types:
    - n-type: Silicon doped to be rich in electrons (negative charge carriers).
    - p-type: Silicon doped to be rich in holes (positive charge carriers).
    - These doped regions are also called diffusion regions.
- Layer Connections: The conducting layers are separated by insulating layers. To connect these layers, tiny holes are etched into the insulation, and metal is “poured” through these holes, establishing electrical connections.

Figures 6.33 shows two views of the physical structure of an n-channel MOSFET:

- The substrate: It is the main body, and it is made of p-type silicon. This is the single-crystal silicon wafer we mentioned above.
- Source & drain. Two n+ doped regions (n+ means it is heavily doped) separated by a small distance are the source and the drain.
- Channel region: It is the region separating the source and the drain.
- Gate Oxide: A thin insulating layer of silicon dioxide (SiO_2), known as gate oxide, separates the gate from the channel region.
- Gate (Polysilicon): On top of the gate oxide is a conductive layer of polysilicon, labeled G.

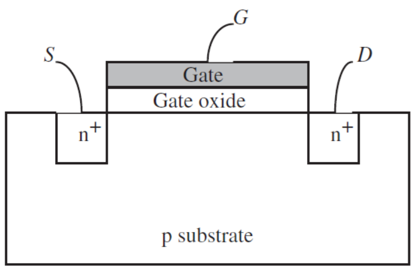
How the MOSFET works:

In simple terms, when voltage is applied to the gate what happens? Current cannot flow through the gate oxide, because of its insulating characteristics but an electric field is formed. This electric field controls the flow of electrons through the channel, and thus current flows.

In more specific terms, let’s connect the gate and the source of the MOSFET to ground, as illustrated in Figure 6.36.

- If no voltage is applied to the gate, vGS = 0. Because the n+ doped source and drain are separated by a p-type layer, they will not conduct any current when a voltage is applied across them (vDS > 0).

- When a positive voltage is applied at the gate of the device (vGS > 0), negative charges are attracted to the surface from the nearby negative charge-rich source region (as shown in Figure



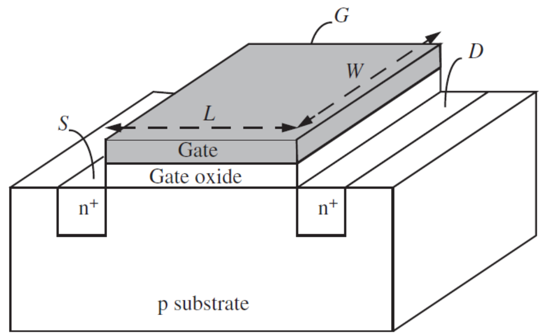


FIGURE 6.33

6.36) and positive charges are repelled from the surface. Of course, no current flows between the gate and the substrate because of the insulating gate oxide layer. As the gate voltage increases, more negative charges are attracted to the surface until they form an n-type conducting channel that connects the source and the drain. The conducting channel forms when the gate voltage crosses a threshold voltage VT (in other words, vGS > VT). A current begins to flow between the drain and the source when a positive voltage is applied across the drain and the source (vDS > 0). The MOSFET in our example is called an n-channel device because of the n-type channel that is formed.

Relation between on-resistance and geometry of the MOSFET:

The conducting n-channel that is formed in the MOSFET discussed here is not an ideal conductor and has some resistance RON. Also notice that the resistance of the gate is related to the geometry of the channel. Let the channel length be L and the channel width be W. Then, the resistance is proportional to L/W. If Rn is resistance per square of the n-channel MOSFET in its on state, then the resistance of the channel is given by:

$$R(ON) = Rn\frac{L}{W} \; (6.4)$$

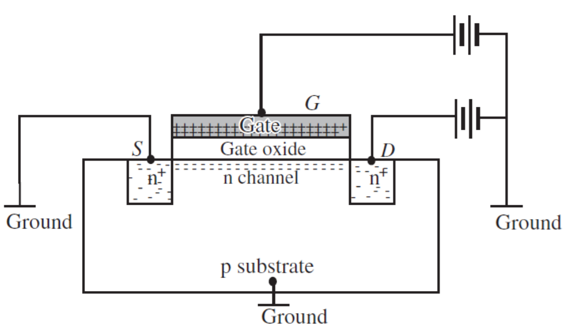


FIGURE 6.36 MOSFET operation when a positive gate voltage is applied.

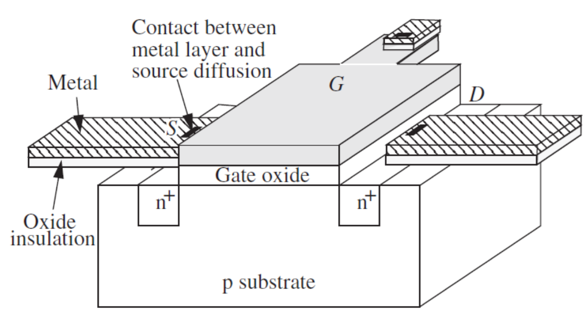


FIGURE 6.37 Connecting to a MOSFET.

6.8 STATIC ANALYSIS USING THE SR MODEL

We analyze our inverter circuit using the SR model of the MOSFET (Figure 6.39). The input-output transfer characteristics:

- When input is low, the MOSFET is off, and the output is raised to a high value.
- When input is high and > vT, the MOSFET is on and there is a resistance R(ON) between D and S, thereby pulling the output voltage lower. The value of the output voltage is given by the voltage-divider relationship:

$$vOUT = VS\frac{R(ON)}{R(ON) + RL} \quad (6.5)$$

I <u>personally</u> argue though, that if I connect a Load between the output and GND terminals, this Load has a non-zero resistance R(Load), and thus the vOUT is:

$$vOUT = VS\frac{R(ON)|R(Load)}{(R(ON)|R(Load)) + RL} \quad (6.5.Matteo)$$

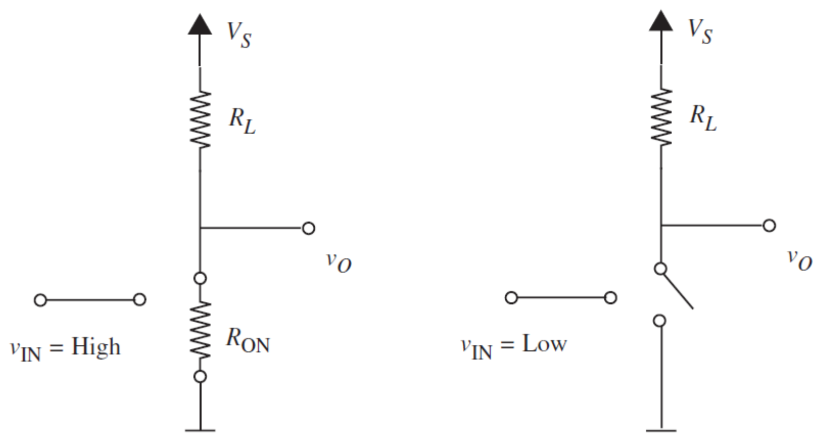


FIGURE 6.39 Circuit model of the n-channel MOSFET inverter using the SR MOSFET model.

Simple electrical switching analysis for intuition:

When a sending inverter drives a receiving inverter, the sender must be able to switch the MOSFET in the receiving inverter in its ON state when the sender produces a high voltage. Similarly, the sender must be able to switch the MOSFET in the receiving inverter into its OFF state when the sender produces a low voltage. Since our inverter produces a high output of VS, as long as VS > VT, no problem in achieving the first; but when it comes to the second it's tricky because the (sender) inverter produces a non-zero output. In simple terms, even if the sender inverter sends a low output, so the receiving inverter will be in its off state, care needs to be taken because the low output of the sender inverter is close to 0V, but not 0V. So, it is important that:

$$VS\frac{R(ON)}{R(ON) + RL} < VT \quad (6.6)$$

Of course, the left side of the equation refers to the sender inverter, since that is in its ON state, and the vT refers to the receiver inverter.

Making things a little more complicated:

Commonly it is not enough to meet the just mentioned criteria, because it must also provide for adequate noise margins by satisfying a static discipline. Here below the example 6.6, found at page 309 of the book.

Suppose our static discipline is:

- V(OH) = 4.5V
- V(IH) = 4V
- V(IL) = 0.9V
- V(OL) = 0.2V

Our inverter rules are as follows (based on some parameters of the resistors and voltage source):

- vT = 1V, so voltages above 1V are interpreted as HIGH and below 1V as LOW
- Output High = 5V
- Output Low = 0.33V

When outputting a logical 1, the voltage produced must be greater than 4.5V. Since 5 > 4.5, the condition is satisfied. When outputting a logical 0, the voltage produced must be less than 0.2V. Since 0.33 > 0.2, this condition is not met. At the inputs, voltages greater than 4V must be recognized as logical 1. Since our inverter recognizes voltages > 1 as logical 1, condition is met. At the inputs, voltages less than 0.9V must be recognized as logical 0. Since our inverter recognizes voltages < 1 as logical 0, this condition is met.

How to redesign our inverter to satisfy the given static discipline? Because the output voltage of the inverter for a high input is given by equation 6.5, we have three methods to reduce the output voltage:

1. Reduce VS. Not good, as will also reduce the output high voltage.
2. Reduce R(ON) by increasing the W/L ratio of the MOSFET. We find R(ON) < 0.58 kilo-ohms, so W/L > 8.62.
3. Increase RL. We obtain RL > 24 kilo-ohms

We can also analyze other gates in the like manner; see the book for more examples.

6.9 SIGNAL RESTORATION, GAIN, AND NONLINEARITY

We saw that the provision of noise margins enables error free communication in the presence of noise. Here we want to demonstrate that logic devices must incorporate both gain and nonlinearity to provide nonzero noise margins.

6.9.1 SIGNAL RESTORATION AND GAIN

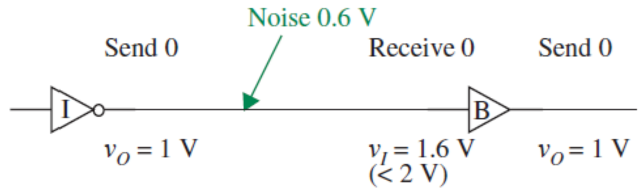


FIGURE 6.44 Noise margins and signal transmission.

In Figure 6.44 we have the inverter I and the buffer B. Like the inverter, the buffer has a single input and a single output. It performs the identity function. Assume that both our logic gates adhere to a static discipline with the following voltage levels:

VIL = 2 V VIH = 3 V VOL = 1 V VOH = 4 V

The inverter sends a 0 by placing vOUT = 1 V (VOL) on the wire. Figure 6.44 shows 0.6 V of noise being added to the signal by the transmission channel. However, the buffer is able to correctly interpret the received value as a 0 because the received value of 1.6 V is within the low input voltage threshold of VIL = 2 V. The buffer, in turn, performs the identity logical operation on the signal and produces a logical 0 at its output. According to the static discipline, the voltage level at the buffer's output is 1 V corresponding to V(OL).

In Figure 6.44, notice that to obey the static discipline, the buffer must convert the 1.6V signal at its input to a 1V value at its output. In fact, the buffer must restore any voltage up to 2 V at its input to voltage of 1 V or lower at its output. This restoration property is key to our being able to compose multiple logic devices together.

A static discipline that provides for nonzero noise margins requires logic devices that provide a minimum gain. This means that nonzero noise margins require that V(IL) > V(OL) and V(OH) > V(IH). The magnitude of the change in the voltage for an input transition from V(IL) to V(IH) is given by ΔVI = V(IH)-V(IL); for the output it is ΔVO=V(OH)-V(OL). Therefore, the gain of a device that can convert a V(IL) to V(IH) transition at its input to a V(OL) to V(OH) transition to its output is given by

$$Gain = \frac{\Delta VO}{\Delta VI} = \frac{V(OH) - V(OL)}{V(IH) - V(IL)}$$

And since VOH – VOL > VIH – VIL then:

$$Gain = \frac{V(OH) - V(OL)}{V(IH) - V(IL)} > 1$$

6.9.2 SIGNAL RESTORATION AND NONLINEARITY

You might have realized that although logic devices must demonstrate a gain greater than unity when they transition from VIL to VIH, they must also attenuate the signal at other times. For example, Figure 6.48 shows the signal from Figure 6.47 with some noise superimposed on it. It should be clear from Figure 6.48 that to obey the static discipline the buffer has reduced the 0-V to 2-V noise excursions at the input to 0-V to 1-V noise excursions at its output.

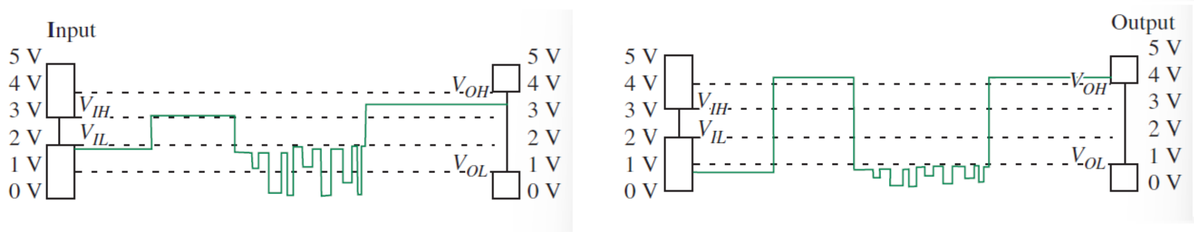
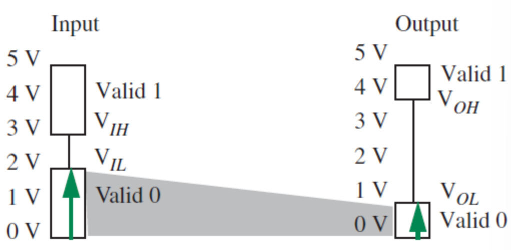


FIGURE 6.48 Input waveform and restored output waveform in the presence of noise.

We can also verify this fact using the basic noise-margin inequalities in Equations 6.8 and 6.9. Equations 6.8 implies that any voltage between 0 and VIL at the input must be attenuated to a voltage between 0 and VOL at the output (see Figure 6.49). Since VIL > VOL according to Equation 6.8, it follows that voltage transfer ratio must be less than unity. In other words,

$$\frac{V(OL) - 0}{V(IL) - 0} = \frac{V(OL)}{V(IL)} < 1$$

FIGURE 6.49 Signal restoration and attenuation.



The same reasoning applies to valid high voltages. Because VIH < VOH,

$$\frac{5 - V(OH)}{5 - V(IH)} < 1$$

6.9.3 BUFFER TRANSFER CHARACTERISTICS AND THE STATIC DISCIPLINE

The presence of gain and nonlinearity in the buffer becomes abundantly clear if we look at its transfer characteristic. Figure 6.50 graphically plots the transfer characteristic of a logic device that can serve as a valid buffer. The shaded region depicts the valid region for the buffer transfer curve. The x-axis shows input voltages and the y-axis output voltages. We can make several interesting observations from this graph:

1. Amplification occurs in the forbidden region; slope of the curve in that section is greater than 1.
2. Input voltages result in valid output voltages.
3. The curve for input between 0 and V(IL) or higher than V(IH) must have an overall gain < 1.

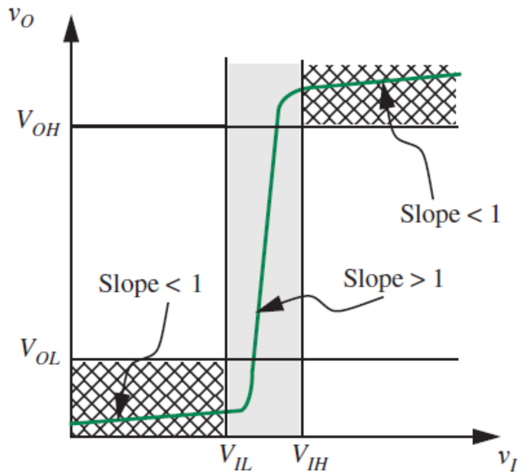


FIGURE 6.50 The buffer characteristic.

6.9.4 INVERTER TRANSFER CHARACTERISTICS AND THE STATIC DISCIPLINE

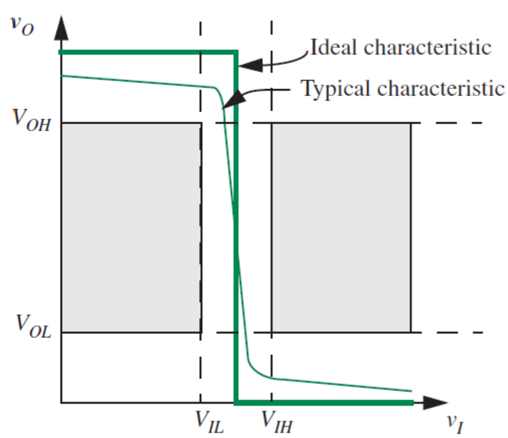


FIGURE 6.51 The inverter characteristic.

6.10 POWER CONSUMPTION IN LOGIC GATES

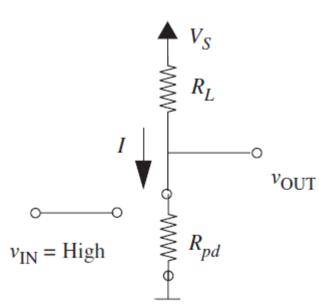


FIGURE 6.52

We can use the SR model to calculate the maximum power consumed by logic gates. Referring to Figure 6.52, the power consumed by a logic gate is given by:

$$Power = VsI = \frac{Vs^2}{RL + Rpd}$$

CHAPTER 7: THE MOSFET AMPLIFIER

7.1 SIGNAL AMPLIFICATION

Amplification, or gain, is key to both analog and digital processing of signals. Section 6.9.2 discussed how gain is employed in digital systems to achieve immunity to noise. This chapter will focus on the analog domain.

Amplifiers can be represented as in figure 7.1. Three-ported devices with a control input port, an output port and a power port. Each port comprises two terminals. An amplifier can amplify the input current, voltage, or both. When the V × I product of the output exceeds that of the input, a power gain results. The power supply provides the necessary power for the resulting power amplification. A device must provide power gain to be called an amplifier. In practical amplifier designs, the input and the output signals commonly share a reference ground connection. Furthermore, the power port is commonly not shown explicitly. See figure 7.2.

Amplifiers are useful for signal transmission in the presence of noise; by amplifying the signal they make it less subject to noise. They are also useful for buffering.

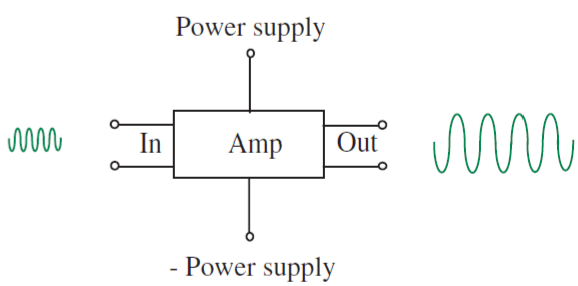


FIGURE 7.1 Signal amplification.

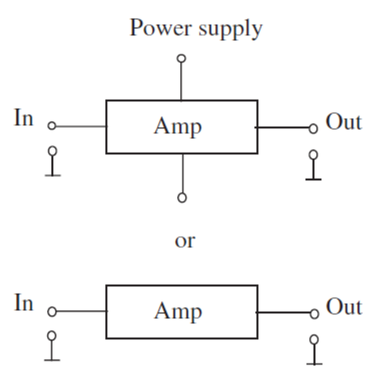


FIGURE 7.2 Reference ground and implicit power connections.

7.2 REVIEW OF DEPENDENT SOURCES

Amplifiers are naturally modeled using dependent sources.

7.3 ACTUAL MOSFET CHARACTERISTICS

The SR model is a reasonable representation of MOSFET behavior only when $v_{DS} < v_{GS}-v_T$. Accordingly, the SR model is useful to design and analyze digital circuit gates because a common mode of operation for the MOSFET within digital gates is one in which the gate voltage is high, and the drain voltage is relatively low. For example, we might have $V_{OH} = 4V$ applied as a logical-high input to the gate of a MOSFET in an inverter (assume $V_T = 1V$), which might produce as the output a corresponding logical-low drain voltage $V_{OL} = 1V$. With these values, $v_{DS} = 1$ V, $v_{GS} = 4$ V. Since $V_T = 1$ V, the constraint $v_{DS} < v_{GS}-v_T$ is satisfied.

However, there are other situations demanding higher drain voltages in which we wish to use the MOSFET in an ON state, and for which the SR model of the MOSFET is inappropriate.

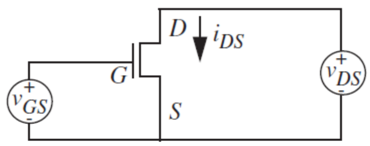


FIGURE 7.7 Setup for observing MOSFET characteristics.

For our explanation, we will use the setup shown in Figure 7.7. We apply a fixed, high gate-to-source voltage such that vGS ≥ vT, and observe that iDS increases more or less linearly as vDS increased from 0V, accordingly, as SR model prescribes: vDS/iDS = R(ON). The region where vDS < vGS-vT is called the triode (Figure 7.10).

Now, keeping vGS at the same value, we increase vDS further. As we can see in Figure 7.10, vDS approaches the value of vGS – vT and the curve flattens out; in other words iDS saturates and vDS begins exceeding vGS – vT.

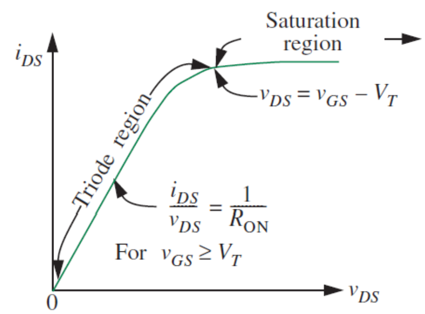


FIGURE 7.10 The saturation region of MOSFET operation.

The region where vDS ≥ vGS-vT is called the saturation region of MOSFET operation. In the saturation region, because iDS does not change as vDS increases, the MOSFET behaves like a current source.

So far, we kept vGS constant at some value > VT. It turns out that the iDS curve saturates at a different value for different values of vGS. Thus, as illustrated in Figure 7.11, we get a different iDS versus vDS curve for each setting of vGS. This family of curves represents the actual MOSFET characteristics. Notice that the slope of each of the curves in the triode region also varies somewhat with vGS. The actual MOSFET characteristics with the triode, saturation, and cutoff regions marked are shown in Figure 7.12. The dashed line represents the locus of the points for which vDS = vGS – vT.

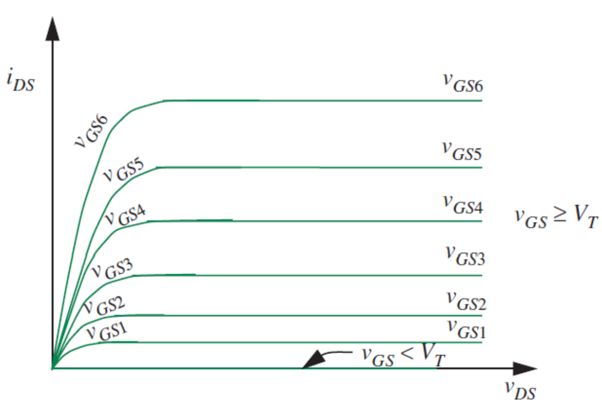


FIGURE 7.11 Actual characteristics of the MOSFET. Each setting of v_{GS} results in a separate i_{DS} versus v_{DS} curve.

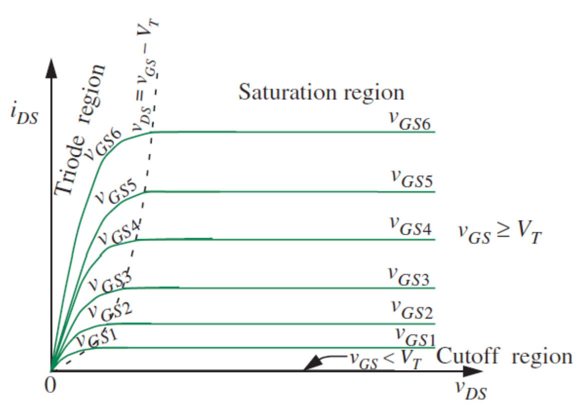


FIGURE 7.12 Actual characteristics of the MOSFET showing the triode, saturation, and cutoff regions.

In summary, there are three regions of operation for the MOSFET:

1) Cutoff. When vGS<vT.
2) Triode. When vDS < vGS – vT and vGS ≥ vT.
3) Saturation. When vDS ≥ vGS – vT and vGS ≥ vT.

Applying Piecewise Linear Analysis:

Given the roughly straight-line behavior of the iDS versus vDS curves within the triode region and the saturation region, it is natural to seek a piecewise-linear model for the MOSFET. Figure 7.13 shows our choice of straight-line segments that model the actual MOSFET characteristics.

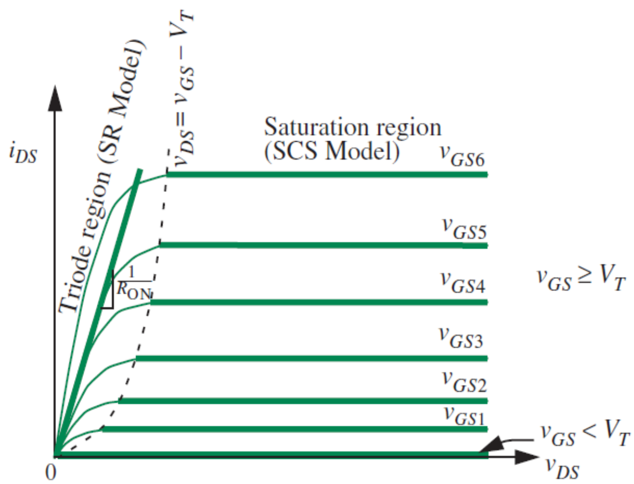


FIGURE 7.13 SR and SCS models.

To the right of the vDS = vGS − VT boundary (saturation region), we use a set of horizontal straight-line segments (one for each value of vGS) to represent the actual MOSFET characteristics. The circuit interpretation of each of the horizontal straight-line segments is a current source. Furthermore, because the value of the current depends on the value of vGS, the behavior is that of a voltage-controlled current source. This behavior, captured by the switch current source (SCS) model of the MOSFET.

To the left of the vDS = vGS − VT boundary (triode region), one possible modeling choice uses a single straight-line segment to approximate the iDS versus vDS curve for a given value of vGS. This is our familiar SR model of the MOSFET.

Application of the two models:

As we said, SR model is appropriate to use in digital circuits. Conversely, in the design of amplifiers, we will establish the saturation discipline, which will constrain amplifier designs to operate MOSFETs exclusively in their saturation region, thereby allowing the use of the SCS model.

7.4 THE SWITCH-CURRENT SOURCE (SCS) MOSFET MODEL

The SCS model is depicted in Figure 7.15 below:

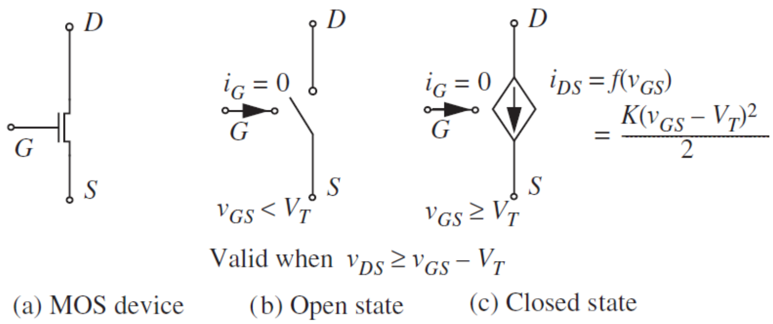


FIGURE 7.15 The switch-current source model of the MOSFET.

iG into the gate terminal is always zero, reflecting an open circuit both between the gate and the source, and the gate and the drain. When vGS ⩾ VT, and vDS ⩾ vGS − VT, the amount of current provided by the source is given by

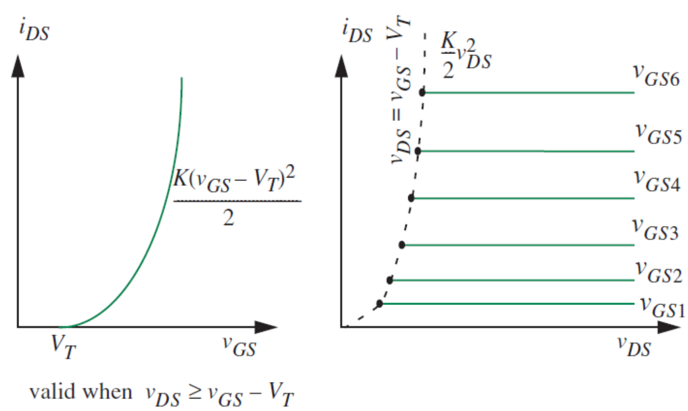
$$\boldsymbol{iDS = \frac{K(vGS - vT)^2}{2} \quad (7.8)}$$

where K is a constant having units of A/V^2 . The value of K is related to the physical properties of the MOSFET. The parameter K is related to the physical structure of the MOSFET as follows: $K = Kn\frac{W}{L}$, where W is the MOSFET gate width and L is the gate length. Kn is a constant related to other MOSFET properties such as the thickness of its gate oxide.

The characteristics of the MOSFET in the saturation region according to the SCS model are summarized graphically in Figure 7.16. The constraint curve separating the triode and saturation regions given by vDS = vGS − VT can also be rewritten in terms of iDS and vDS by substituting vDS = (vGS − VT) as follows (essentially same as Eq. 7.8):

$$iDS = \frac{K}{2}v^2DS \quad (7.11)$$

FIGURE 7.16 Characteristics of the MOS device in the saturation region.



What does it mean? It means that just by changing vGS, I can control the amount of current flowing through. And this is very useful, for example, to control a motor.

7.5 THE MOSFET AMPLIFIER

A MOSFET amplifier circuit is shown in Figure 7.19. Notice this circuit is identical to the inverter circuit we saw earlier! Unlike the inverter circuit, however, the input and output voltages of the MOSFET amplifier must be carefully chosen so that the MOSFET operates in its saturation region.

In the saturation region of operation, the SCS model can be used to analyze the MOSFET amplifier. Let's then replace the MOSFET in Figure 7.19 with its SCS circuit model as illustrated in Figure 7.20.

From previous paragraphs, we know the conditions on the circuit for the MOSFET to be in saturation mode. We also remember equation 7.8.

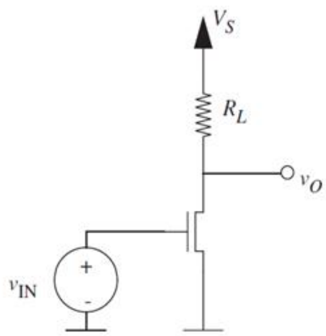


FIGURE 7.19 The MOSFET amplifier.

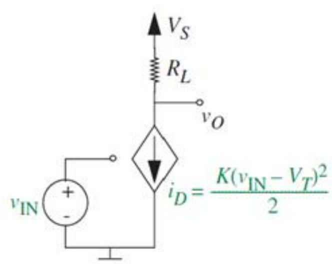


FIGURE 7.20 SCS circuit model of the MOSFET amplifier.

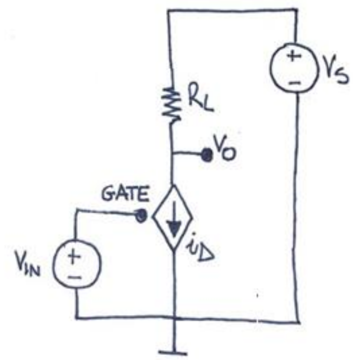


FIGURE 7.21

What is the relationship between the amplifier output vO and its input vIN?

This relationship will describe the gain of the amplifier. We will begin by formulating the output voltage vO as a function of the input voltage vIN. Any of the methods described in Chapters 2 and 3 can be used to analyze this circuit. We will use the node method here. Since the current into the MOSFET gate is zero, the node with voltage vO is the only interesting node in the circuit. Figure 7.21, which essentially is figure 7.20, perhaps makes it easier to understand how to apply the node method. Writing the node equation, we get:

$$iD = \frac{vS - vO}{RL}$$

And rearranging the terms we get:

$$vO = vS - iD * RL$$

When vIN ⩾ VT and vO ⩾ vIN − VT, we know that the MOSFET is in saturation and the SCS model for the MOSFET applies. Substituting for iD from Equation 7.8, we get the transfer function of the amplifier given by

$$\boldsymbol{vO = vS - K\frac{(vIN - vT)^2}{2}RL} \quad \mathbf{(7.14)}$$

The transfer function relates the value of the output voltage to that of the input voltage. Accordingly, the gain of the amplifier is given by

$$\frac{vO}{vIN} = \frac{vS - K\frac{(vIN - vT)^2}{2}RL}{vIN} \quad (7.15)$$

Figure 7.21 plots vO versus vIN for the MOSFET amplifier. This decidedly nonlinear relationship is called the **transfer function of the amplifier**. When vIN < VT, the MOSFET is off, and the output voltage is VS. In other words, iD = 0 when vIN < VT. As vIN increases beyond the threshold voltage VT, so does the current sustained by the MOSFET. Therefore, vO rapidly decreases as vIN increases. The MOSFET operates in the saturation region until the output voltage vO falls one threshold below the gate voltage, at which point the MOSFET enters the triode region (shown as a dashed line in Figure 7.21), and the saturation model and Equation 7.14 are no longer valid. The magnitude of the slope of certain regions of the curve within the saturation region is greater than one, thereby amplifying input signals that fall within this region.

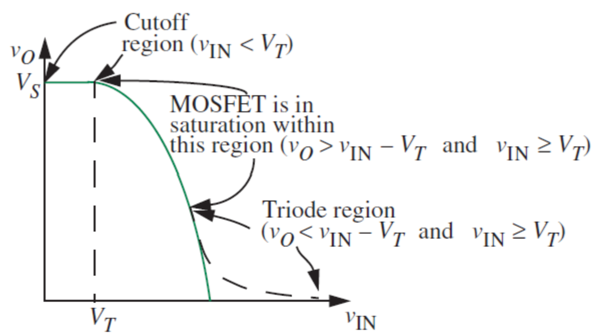


FIGURE 7.21 v_O versus v_{IN} curve for the amplifier.

7.5.1 BIASING THE MOSFET AMPLIFIER

As said, the MOSFET is in saturation only within a certain region of the amplifier transfer curve. The MOSFET circuit works as a reasonable amplifier only within this region. To ensure that the amplifier

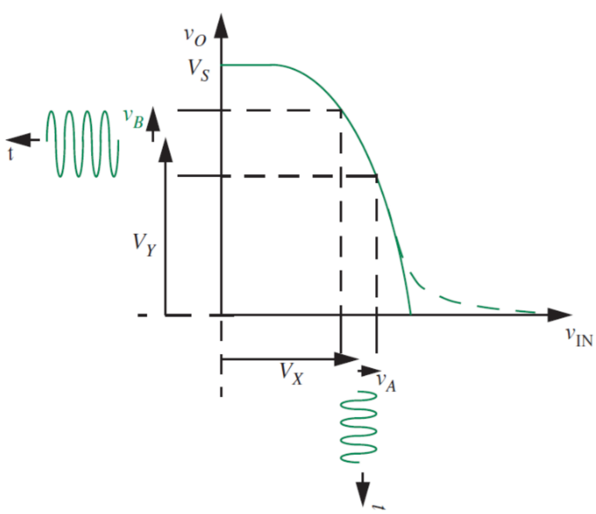


FIGURE 7.25 Boosting the input signal of interest with a suitable DC offset so that the MOSFET operates in its saturation region for the entire range of input signal excursions.

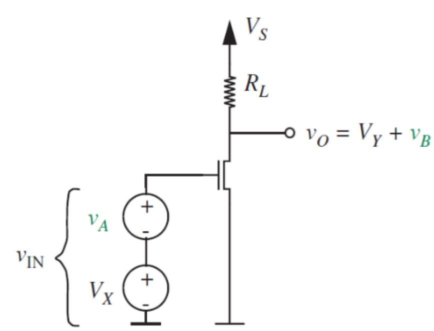


FIGURE 7.26 Circuit for boosting the input signal of interest (v_A) with a suitable DC offset (V_X) so that the MOSFET operates in its saturation region for the entire range of input signal excursions.

operates within this region of the curve, we must transform the input voltage appropriately. As illustrated in Figure 7.25, one way of doing so is to boost the signal that we want to amplify (for example, vA) with a DC offset (say, VX) so that the amplifier operates in its saturation region even for negative excursions of the input signal. Figure 7.26 shows the corresponding circuit that adds an offset to the input signal by connecting a DC voltage source (VX) in series with the input signal source (vA). In other words, we have vIN = VX + vA where vA is the desired input signal. Notice in Figure 7.25 that the corresponding output voltage vO also contains a DC offset VY added to the time varying output signal vB. vB is an amplified version of the input signal vA.

The DC offset is also called a DC bias. The use of the DC offset voltage at the input establishes an operating point for the amplifier. The operating point is also referred to as the bias point (VX and VY, in our example).

We make one final observation about our amplifier. Although vB is an amplified version of the input signal vA when the input signal is boosted with a DC offset, vB is not linearly related to vA. Notice from Equation 7.14 that our amplifier is nonlinear even when the MOSFET operates in the saturation region. Fortunately, the MOSFET amplifier behaves as an approximately linear amplifier for small signals; in other words, when the desired input signal vA is very small (Chapter 8 topic). In the rest of this chapter though, we will not assume that the input is a small signal. Rather, we will assume that the input vIN that is fed into the amplifier comprises both the signal component of interest to this user (which may be a large valued signal), and a DC offset (or DC bias). For simplicity, all calculations will be performed on this boosted signal.

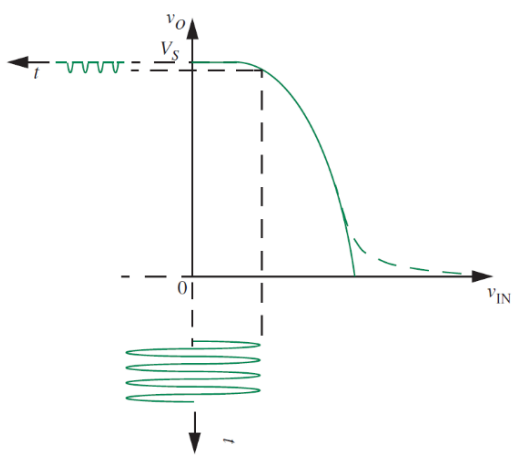


FIGURE 7.27

<u>What happens if we do not provide a DC offset?</u>

In Figure 7.27 the input signal source is applied directly without offsets. The MOSFET operates in its cutoff region for most of the input signal, and the output is highly distorted, bearing little resemblance to the input. The form of distortion suffered by the signal in the example in Figure 7.27 is called clipping.

7.5.2 THE AMPLIFIER ABSTRACTION AND THE SATURATION DISCIPLINE

We would like the user of a MOSFET amplifier to be able to treat it as the abstract entity depicted in Figure 7.28, ignoring the internal details of the circuit. This abstract amplifier has vIN and iIN at its input port and vO and iO at its output port, and provides power gain. Details such as the power supply and the like are hidden from the user. The amplifier shown in Figure 7.28 uses ground as an implicit second terminal for both the input port and the output port. This form of amplifier is also called the single-ended amplifier.

Much like the gate abstraction went hand in hand with the static discipline which dictated the valid range for applied inputs and expected outputs the amplifier abstraction is associated with the saturation discipline, which prescribes constraints on the valid set of applied input signals and expected output signals. The saturation discipline simply says that the amplifier be operated in the saturation region of the MOSFET. As we shall see shortly, we choose this definition of the saturation discipline, because the amplifier provides a good amount of power gain in the saturation region, thereby operating well as an amplifier.

Specification of the saturation discipline serves two purposes: First, it prescribes constraints on how the device can be used; and second, it establishes a set of design criteria for the device. The amplifier abstraction and its associated usage discipline can be likened to procedural abstractions in software systems. Software procedures are an abstraction for the internal function they implement. Procedures are also associated with a usage discipline often articulated as comments at the head of the procedure. Section 7.6 will be concerned with identifying valid usage ranges under the saturation discipline.

7.6 LARGE SIGNAL ANALYSIS OF THE MOSFET AMPLIFIER

Large signal analysis deals with how the amplifier behaves for large changes in the input voltage. Large signal analysis also determines the range of inputs for which the amplifier operates under the saturation discipline (DC bias discussed above). In other words, large signal analysis answers the following questions related to the design of the amplifier:

1. What is the relationship between the amplifier output vO and its input Vin in the saturation region? Equation 7.14, developed using the analytical method, summarized the answer to this question. Here we use the graphical method to determine the same relationship.

2. What is the range of valid input values for the amplifier under the saturation discipline? What is the corresponding range of output values?

In figures 7.29 and 7.30 our setup of reference for this paragraph.

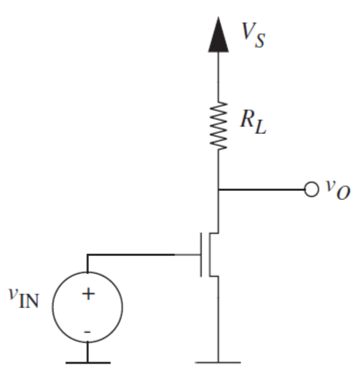


FIGURE 7.29

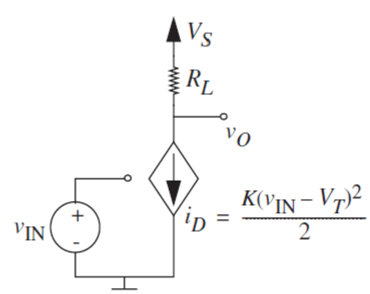


FIGURE 7.30

<u>Answer to question 1. Vin versus Vout in the saturation region.</u>

Two constraints to respect; the first one is the node method to find vO. As before, applying the node method we obtain $vDS = vS - iDS * RL$ which can be rewritten as

$$iDS = \frac{vS}{RL} - \frac{vDS}{RL} \quad (7.21)$$

Second constraint is equation 7.8 (rewritten here), since we operate in saturation region of the MOSFET:

$$iDS = \frac{K(vGS - vT)^2}{2}.$$

Combining the two above, we can graphically solve for the behavior of the output voltage by overlaying the two graphs in Figure 7.32. The iDS versus vDS relationship from equation 7.21 is plotted as a black straight line, called the "load line". The slope of the line, -1/RL, is inversely proportional to the load resistance. Equation 7.8 is plotted as a dotted black line.

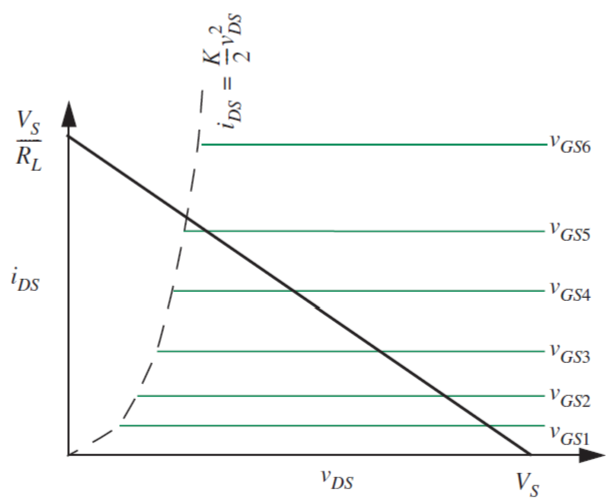


FIGURE 7.32 Load line super-imposed on the characteristic curves of the MOSFET.

<u>Answer to question 2. Valid input and output voltage ranges.</u>

A valid voltage range is the range of input voltages (and the resulting range of output voltages) for which the MOSFET operates in the saturation region.

- Lowest valid input voltage. From Figure 7.36, we can see that vIN must be greater than vT for the MOSFET to exit its cutoff region. When vIN is exactly equal to vT, since the conditions to operate in saturation are vGS ≥ vT and vDS ≥ vGS – vT, any positive vDS allows the MOSFET to be in saturation region. So, the lower limit on input voltages is VT, at (x) point in the graph below. The

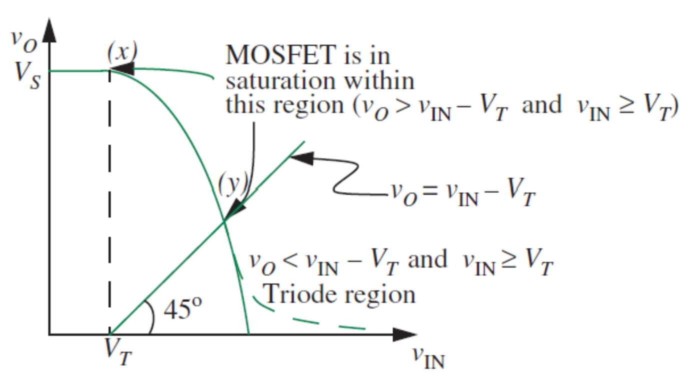


FIGURE 7.36 The lowest valid input voltage under the saturation discipline is marked by the point (*x*), and the highest valid input voltage under the saturation discipline is marked by the point (*y*).

corresponding value of the output voltage is VS, because the MOSFET only “starts” to be on and is not truly conductive of current yet.

- Highest valid input voltage. The higher limit is the point just before vO ≥ vIN – VT starts to be true. That is the edge of the triode region. To build intuition, we first determine graphically the input voltage for which the output crosses into the triode region as follows. Referring to Figure 7.36, the straight line drawn at 45◦ to the vIN axis and intersecting it at VT reflects the set of points in the vIN versus vO plane for which

  vO = vIN – VT (7.27)

  assuming, of course, that vIN and vO use the same scale. Thus, the point (y) at which this 45◦ line intersects the vIN versus vO transfer curve marks the upper limit of the valid input range. *It is easy to see this, since at those points vO is always equal to vIN – vT, you can plot it in a graph to see.*
  We can also determine analytically the value of this upper limit by solving for the intersection of the straight line in Figure 7.36 represented by Equation 7.27, and the transfer curve determined by Equation 7.14. Again, through some algebraic passages (see book page 358), we find point (y) in Figure 7.36, and we obtain the following equation:

$$vIN = \frac{-1 + \sqrt{1 + 2 * vS * RL * K}}{RL * K} + vT \quad (7.32)$$

Taking the min and max input voltage values, we can find the corresponding min and max output values:

- Output of min input: VS (as we said above).
- Output of max input: $\frac{-1+\sqrt{1+2*vS*RL*K}}{RL*K}$

The corresponding drain current range for the valid input voltage range is:

- Current of min input: 0
- Current of max input: $\frac{K}{2}(vIN - vT)^2$

<u>Alternative method for valid input and output voltage ranges</u>

This paragraph finds the same results as the one just above, starting from the load line and the MOSFET device characteristics. See the book for more details.

7.7 OPERATING POINT SELECTION

We are often interested in amplifying time-varying signals. Because the amplifier turns off for input voltages less than VT, it is important to add an appropriate DC offset voltage to the time-varying input signal so that the amplifier remains in the saturation region for the entire range of input voltage variation. This input DC offset voltage defines the operating point of the amplifier.

The DC offset must be chosen carefully, for if it is too large, the amplifier will be pushed into the triode region, and if it is too low, the amplifier will slide into the cutoff region. How do we choose this operating point?

Since the MOS amplifier is nonlinear, we define the output offset as the value of vO when the DC input offset voltage is the only signal applied at the input.

The input offset voltage is also called the input bias voltage or the input operating voltage. The corresponding output voltage and the output current define the output operating point of the amplifier. Together, the input bias voltage, and the corresponding output voltage and the output current, define the operating point of the amplifier. We denote the operating point values of vIN, vO, and iD as VIN,VO, and ID, respectively.

There are several factors that can govern our choice of the operating point. For example,

1) The operating point dictates the maximum dynamic range of the input signal for both positive and negative excursions for which the MOSFET operates in saturation.
2) As can be seen from Equation 7.15, the operating point value of the input voltage also governs the signal gain of the amplifier.

This section will focus on selecting an operating point based on maximizing the useful input signal range. We will have more to say about the relationship between the gain of the amplifier and its operating point in Chapter 8.

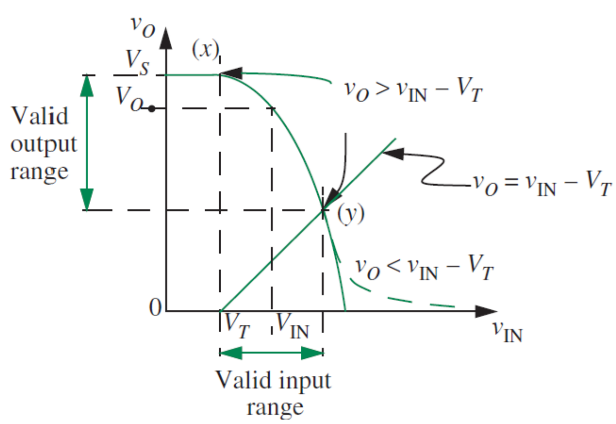


FIGURE 7.42 Selection of the input operating point.

We will assume an equal magnitude for both the positive and negative excursions of the time-varying signal from the DC offset. To obtain maximum useful input signal range, we might choose the input bias voltage VIN to be at the center of the valid range of input voltages for the amplifier, as illustrated in Figure 7.42.

Using the amplifier parameters that we have been using thus far (RL = 10 kΩ | K = 1 mA/V^2 | VS = 5V | VT = 1V), because our amplifier operates under the saturation discipline for input voltages in the range 1 V → 1.9 V (calculations in example in book), we might choose an input operating point voltage at the center of this range, namely VIN = 1.45 V. This choice is illustrated in Figure 7.43, and in Figure 7.44. As we expect, the output will vary between 0.9 V and 5 V as the input varies between 1 V and 1.9 V.

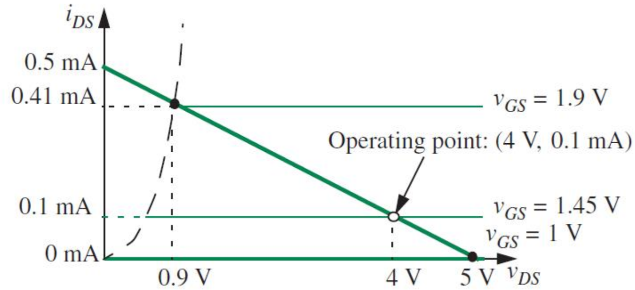


FIGURE 7.43 Operating point and valid input and output voltage ranges.

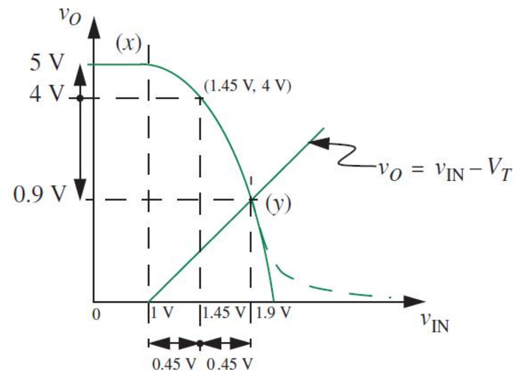


FIGURE 7.44 Input operating point.

Let us take a closer look at the behavior of the amplifier for the given input bias voltage by determining the corresponding output operating point. For a given VIN, we can determine VO from Equation 7.14, and ID from the MOSFET SCS model given in Equation 7.8. Substituting for the circuit parameters in Equation 7.14 we get the value for our operating point, being VIN = 1.45V | VO = 4V | ID = 0.1mA.

This operating point maximizes the peak-to-peak input voltage swing for which the amplifier operates under the saturation discipline. The operating point for our amplifier, along with the valid input and output voltage ranges, is shown in Figure 7.43. For this choice of the operating point, the maximum input voltage swing for positive excursions is 1.45 V → 1.9 V, and the maximum input voltage swing for negative excursions is 1.45 V → 1 V. The corresponding output voltage swings are 4 V → 0.9 V and 4 V → 5 V. Although we chose the input operating point to be at the center of the valid input range, notice the asymmetry of the output voltage range about the output operating voltage. The asymmetry arises from the nonlinearity of the gain of the MOSFET amplifier.

7.8 BIPOLAR JUNCTION TRANSISTOR (BJT) – From example 7.13 to …

Figure 7.47 depicts another three-terminal device, called the bipolar junction transistor (BJT), that is commonly used in VLSI circuits. A BJT has three terminals called base (B), collector (C), and emitter (E).

The actual characteristics of a BJT:

The actual characteristics of a BJT (iC versus vCE for various values of iB) are shown in Figure 7.48. The horizontal nature of the iC versus vCE curves indicates that the device operates like a dependent current source when:

- the base current iB > 0
- vCE > 0.2V

The current supplied by the current source is typically about 100 times the base current.

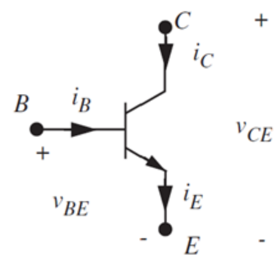


FIGURE 7.47 A bipolar junction transistor.

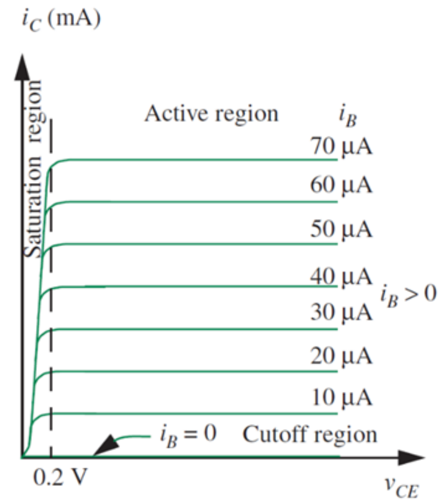


FIGURE 7.48 Actual characteristic curves for a bipolar junction transistor.

Although these curves are qualitatively similar to those of a MOSFET, there are also some differences.

1. Notice that we have chosen the BJT's base current iB as our control parameter (the control parameter was the gate-to-source voltage for the MOSFET, and the gate current was zero).
2. The collector current is linearly related to the base current (when the MOSFET operated as a current source, its drain current was quadratically related to the gate-to-source voltage).

The BJT characteristics show three regions of operation:

1. When iB > 0 and vCE > 0.2 V, the BJT is in the **active region of operation**. Here the horizontal collector current curves behave like a current source.
2. When iB = 0, the BJT is in the **cutoff region**.
3. Finally, when iB > 0 and vCE ⩽ 0.2 V, the collector current drops sharply, and the BJT is in the **saturation region**.

Figure 7.49b shows a model for the BJT containing a current-controlled current source and a pair of diodes. The current supplied by the dependent source is β times iB’. The constant β typically has a value of around 100.

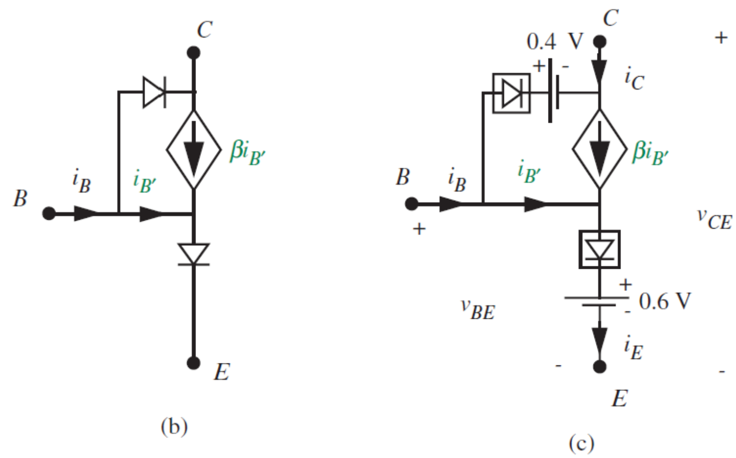


FIGURE 7.49 (b) a model for the BJT; (c) a piecewise-linear model for the BJT.

<u>Piecewise linear model for BJT:</u>

Although we can analyze circuits directly with the model in Figure 7.49b, our analysis can significantly be simplified by using simple piecewise-linear models for the diodes. Figure 7.49c depicts such a piecewise-linear model for the BJT, in which we have replaced the diodes with simple piecewise-linear diode models comprising an ideal diode in series with a voltage source. In the model in Figure 7.49c, the dependent current source models the horizontal active region curves of the BJT.

The states of the two diodes (both ON, both OFF, and one OFF and one ON) result in distinct piecewise linear regions of BJT operation.

1. Both diodes in their OFF state | Cut-off region:

If iB = 0, then both diodes are OFF, and so is the current source. Since the base-to-collector diode is off, iB=iB’.

2. Emitter diode ON, and collector diode OFF | Active region:

Required conditions for active region (emitter diode ON, collector diode OFF):

- iB > 0
- vCE > vBE − 0.4 V (7.48)

Here the ideal diode between B and E turns ON and appears as a short circuit. The 0.6V source models the corresponding 0.6V diode drop. Observe further that iB = iB’ in the active region because the base-to-collector diode is off. In the active region, BJTs display a more-or-less constant voltage drop of about 0.6V between their base and emitter terminals when the base current iB > 0 (a fact not evident from the characteristic curves in Figure 7.48).

The condition $v_{CE} > v_{BE} - 0.4$ V ensures that the base-collector diode stays OFF. The condition states that the collector voltage must not fall below the base voltage by more than 0.4 V, because if it did, the base-collector diode would turn ON. In the active region, the dependent current source amplifies the current supplied by the base by a factor β, so that the collector current is:

$$i_C = \beta i_B$$

(recall, $i_B = i_B'$ in the active region), and the emitter current is

$$i_E = i_B(\beta + 1)$$

If the functioning of the active region is unclear, here is an explanation from ChatGPT of how it works and specifically how is it possible that the BC diode is OFF, and the BE turns ON.

The key lies in the voltages applied across the diodes and their orientation within the transistor. In the piecewise linear model of a Bipolar Junction Transistor (BJT), the transistor is represented by two diodes:

- *Base-Emitter (BE) Diode: Forward-biased with a threshold voltage of 0.6 V.*
- *Base-Collector (BC) Diode: Forward-biased with a threshold voltage of 0.4 V.*

In the active region of the BJT operation:

- *The BE junction is forward-biased, meaning the base is at a higher potential than the emitter by at least the diode's forward voltage (0.6 V).*
- *The BC junction is reverse-biased or at least not forward-biased enough to conduct. This means the collector is at a higher potential than the base, resulting in a negative voltage across the BC diode when measured from base to collector.*

Here's why the 0.6 V BE diode conducts while the 0.4 V BC diode does not:

1. Voltage Across Diodes:

- BE Diode: The voltage VBE is greater than 0.6 V, which forward-biases the diode, allowing current to flow from the base to the emitter.

- BC Diode: The voltage VBC is negative in the active region because the collector voltage VC is higher than the base voltage VB. Therefore, $V_{BC} = V_B - V_C$ is negative, reverse-biasing the BC diode.

2. Diode Orientation:

-The BE diode is oriented so that when the base is more positive than the emitter by at least 0.6 V, it conducts.

- The BC diode would require the base to be more positive than the collector by at least 0.4 V to conduct. In the active region, this condition is not met because $V_C > V_B$.

Answering Your Question:

The BC diode remains off because it is reverse-biased due to the voltage conditions in the active region. The lower threshold voltage of the BC diode does not cause it to conduct because the base is not at a higher potential than the collector by 0.4 V. Therefore, the BE diode being on does not inherently cause the BC diode to turn on; each diode conducts based on its own voltage conditions and orientation within the circuit.

Both diodes ON | Saturation region:

The base-to-collector diode also turns ON when:

- iB > 0
- vCE = vBE − 0.4 V or equivalently, if vBC = 0.4V or vCE = 0.2V (condition of Equation 7.48 is violated)

In the BJT's saturation region, the BJT model stops looking like a current source, and instead displays a pair of very low resistance paths from the base into the collector and emitter (due to the pair of forward-biased diodes). Because of their low resistance, the path currents are determined by external circuit constraints. By summing voltages along the path E, B, C, we see that the collector-to-emitter voltage is pinned at 0.2 V, irrespective of the current iC.

Emitter diode is OFF, and the collector diode is ON:

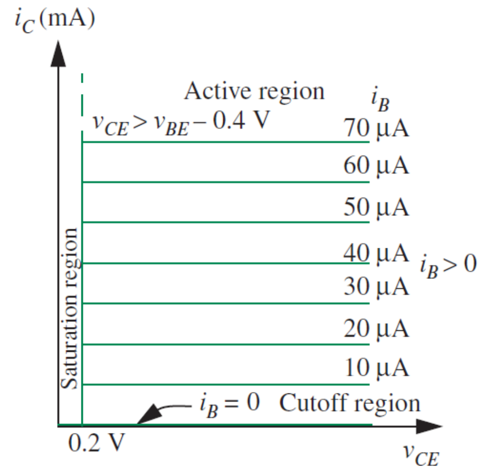
When the BC voltage is 0.4 V, and the BE voltage is less than 0.6 V. This region of operation is called the reverse injection region. In this region, the behavior of the BJT is that of a forward biased diode between the base and the collector, and an open circuit at the emitter. We do not study this state!!!

<u>Comparison of piecewise linear model for BJT with measured characteristics</u>

The results of the piecewise linear analysis are plotted in Figure 7.51. By comparing this graph with the one in Figure 7.49, it shows the accuracy of the piecewise linear analysis.

As a final thought, although our piecewise linear model for the BJT seems a bit complicated at first glance, analog circuits are commonly designed such that BJT always operates in its active region, and the base-to-collector diode is always OFF. We can achieve the desired effect by ensuring that the base-to-collector

FIGURE 7.51 Characteristic curves for a bipolar junction transistor as predicted by the piecewise linear model.



voltage never exceeds 0.4V during normal operation (that is, vBC < 0.4 V, or equivalently, vCE > vBE−0.4V). This assumption will be made in all the BJT circuits in this book, so the collector diode can be safely ignored. The resulting, simplified BJT model is depicted in Figure 7.52.

The simplified model, our model of reference

As a final thought, although our piecewise linear model for the BJT seems a bit complicated at first glance, analog circuits are commonly designed such that BJT always operates in its active region, and the base-to-collector diode is always OFF. We can achieve the desired effect by ensuring that the base-to-collector voltage never exceeds 0.4V during normal operation (that is, vBC < 0.4 V, or equivalently, vCE > vBE−0.4V).

This assumption will be made in all the BJT circuits in this book, so the collector diode can be safely ignored. The resulting, simplified BJT model is depicted in Figure 7.52.

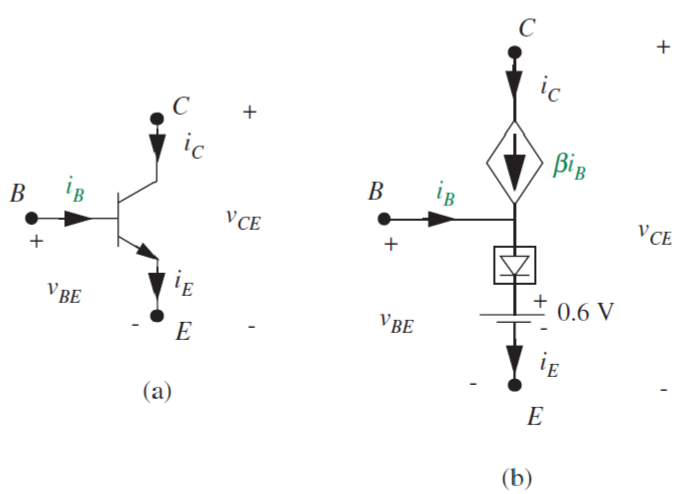


FIGURE 7.52 A simpler BJT model suitable for the cutoff and active regions.

Large signal analysis of the BJT Amplifier (example 7.16 in the book)

CHAPTER 8: THE SMALL SIGNAL MODEL

8.1 OVERVIEW OF THE NONLINEAR MOSFET AMPLIFIER

The nonlinear input-output relationship of the MOSFET makes it difficult for us to analyze and to build circuits using the amplifier.

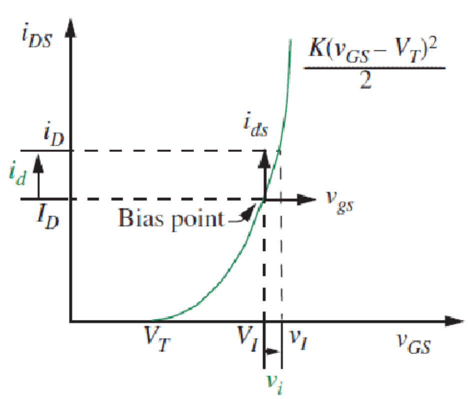
8.2 THE SMALL SIGNAL MODEL

If we want a linear amplifier, with a constant input/output gain, we can use small signal analysis.

In fact, it turns out that total variables representing signals such as those input to an audio amplifier commonly consist of two components: a DC offset plus a time-varying component with a zero average. We will show that if the time-varying component is small, then the incremental amplification provided by the MOSFET amplifier to the time-varying component about the operating point defined by the input DC offset will be approximately linear.

Remember the notation of small signal analysis: vI = VI+vi, meaning the total variable vI is equal to the DC offset Vi and a small variable signal vi.

So, we add a small-time varying component to the DC offset as per the figure below:



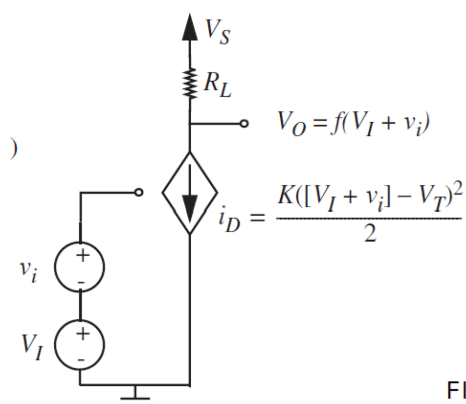


FIGURE 8.5

For the combined input signal shown in Figure 8.5 right picture, the response current iD through the MOSFET is the sum of two components: a bias current ID and a change id due to the incremental input signal vi as seen in Figure 8.5 left picture.

As depicted in Figure 8.5 (left graph), this combined current can be obtained by substituting for vGS as

$$iD = f(VI + vi) = ID + id = \frac{K([VI + vi] - VT)^2}{2} \quad (8.4)$$

Since we know that vi is small compared to VI, we can adopt the following linearization technique to obtain the combined response: Model the MOSFET characteristic curve accurately only in the vicinity of the bias point VI and disregard the rest of the curve. The Taylor series expansion is the natural tool for this task. If vi is small enough for us to ignore the second and higher order terms of the Taylor expansion equation, the following simplification results:

$$iD = ID + id \ = \frac{K(VI - VT)^2}{2} + K(VI - VT)vi \quad (8.7)\ (8.8)$$

$$\boldsymbol{ID = \frac{K(VI - VT)^2}{2} \quad (8.9)}$$

$$\boldsymbol{id = K(VI - VT)vi \quad (8.10)}$$

This result can be obtained quite simply by expanding equation 8.4, no need Taylor series (done by professor in YouTube lesson).

In equation 8.10, K(VI − VT) relates the input voltage to the current through the MOSFET. Notice that for a given DC bias, the K(VI −VT) term is a constant. Since the form of Equation 8.10 is similar to that for a conductance, the K(VI−VT) term is called the incremental transconductance gm of the MOSFET. Accordingly, we can write

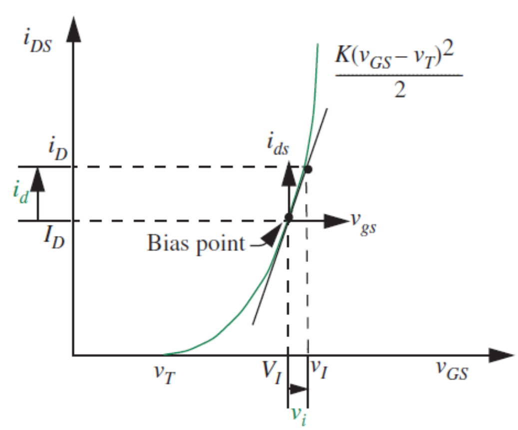
$$id \ = \ gm * vi \quad (8.12) \ \text{where} \ gm = K(VGS \ - \ VT) \ (8.13)$$

In our example VGS = VI, which is not always the case as I saw during the problems of the last chapter.

Graphical interpretation:

A graphical interpretation is shown in Figure 8.8. Equation 8.8 is a straight line passing through the DC operating point (VI, ID) and tangent to the curve at that point.

FIGURE 8.8 Incremental change in the output current for the MOSFET for a small change in the input voltage.



Small signal gain:

Returning to our amplifier, we can express the total output voltage vO as the sum of the output operating voltage VO and the incremental change vo as vO = VO + vo. From last chapter we know that vO = VS – iD*RL. Replacing vO and iD with their DC and incremental components we obtain:

$$VO = VS - ID * RL \quad (8.17)$$

$$vo = -id * RL = -gm * vi * RL \;\; (8.18)\,(8.19)$$

In other words,

$$\boldsymbol{Small\ signal\ gain} = \frac{\boldsymbol{vo}}{\boldsymbol{vi}} = \boldsymbol{-gm * RL = A} \quad (\mathbf{8.20})$$

Notice from Equation 8.20 that the small signal gain is a constant *−gm*RL*. Note, however, that *gm*, and therefore the gain, depends on the choice of bias point for the amplifier. Equation 8.19 demonstrates that for small excursions from a DC operating point, a linear amplifier results! This result forms the basis of the small-signal model.

<u>Direct path to final result using calculus:</u>

We can directly arrive at the small signal response be it voltage or current using basic calculus for circuit responses that are differentiable. Recall that the derivative of a function *y = f (x)* at the point *xo* is the slope of the function at that point, or *f'(xo)*. Given a small change *Δx* from a point *xo*, we can compute the response to the change as the product of the slope at that point and *Δx*. In other words,

$$f(xo + \Delta\mathrm{x}) = f(xo) + \frac{df(x)}{dx}\int_{xo}^{\cdot} \Delta\mathrm{x}$$

Thus, the incremental change in the output is given by

$$\Delta\mathrm{y} = \frac{\mathrm{dy}}{\mathrm{dx}}\int_{xo}^{\cdot} \Delta\mathrm{x} \quad (8.21)$$

In particular, we can obtain the incremental voltage gain directly from the voltage transfer function, without first determining the incremental output current. The input-output voltage relationship for the MOSFET amplifier is given by

$$vO = f(vI) = VS - K\frac{(vI - VT)^2}{2}RL$$

So,

$$vo = \frac{df(vI)}{dvI}\int_{vI=VI}^{\cdot} vi = -K(vI - VT)RL\int_{vI=VI}^{\cdot} vi = -gm * RL * vi$$

Not surprisingly, this result is the same as the one we obtained earlier.

8.2.1 SMALL-SIGNAL CIRCUIT REPRESENTATION

A model that involves only the small-signal variables of a circuit, and hence describes purely the small-signal behavior of that circuit, would greatly facilitate small-signal analysis. We can do that through the following steps:

1. Set each source to its operating-point value and determine the operating point branch voltages and currents for each component in the circuit. In simple terms: analyze the circuit at its operating point values.
2. Linearize the behavior of each circuit component about its operating point. That is, determine the linearized small-signal behavior of each component, and select a linear component to represent this behavior. The parameters of the small-signal components will commonly depend on the operating point voltages or currents. In simple words; here we work with total variables and we linearize the equations around the operating point.
3. Replace each original component in the circuit with its linearized equivalent and re-label the circuit with the small-signal branch variables. The resulting circuit is the desired small-signal model. In simple terms, we cancel the operating-point variables from the linearized equations so we can work with the small signal circuit alone.
4. Complete the small-signal analysis by combining the linearized equations to determine the desired small-signal variables in terms of the small-signal inputs at the sources.

The professor's explanation on this point might be useful, link here.

Small-signal circuit models for various devices are summarized in Figure 8.10.

- The small-signal equivalent model for an independent DC voltage source is a short circuit because its output voltage does not change for any perturbation of the current through it.
- The small-signal model for an independent DC current source is an open circuit.
- A resistor behaves identically for a large signal or a small-signal. Therefore, its two models are the same.
- For a MOSFET, the derivation resulting in equation 8.10 shows how to relate ids (the incremental drain to source current) to vgs (the incremental gate to source voltage). First element in Fig. 8.10.

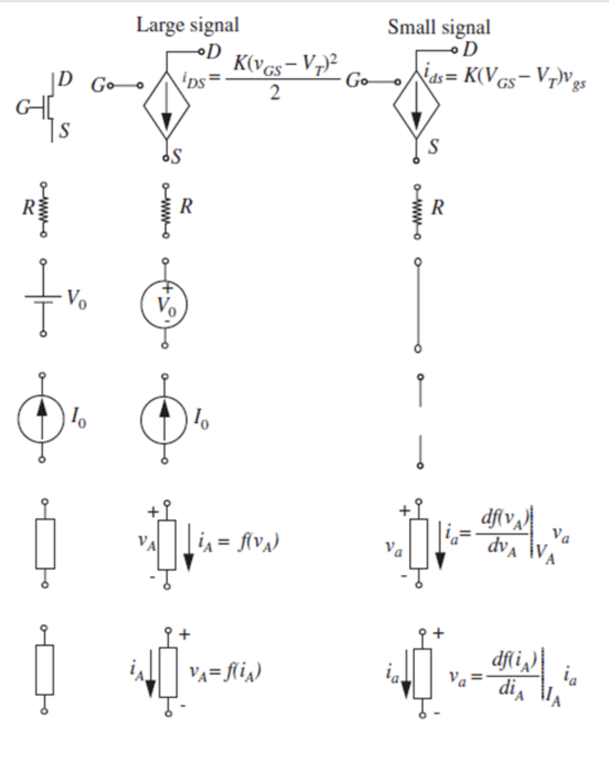


FIGURE 8.10 Small-signal equivalent models.

In general, if a device variable *xB* depends on some other variable *xA* as $xB = f(xA)$ then the incremental change in *xB* due to a small change in *xA* is given by

$$xb = \frac{df(xA)}{dxA}\int_{xA=XA}^{\cdot} xa \quad (8.22)$$

where *XA* is the operating point value of *xA*.

8.2.2 SMALL-SIGNAL CIRCUIT FOR THE MOSFET AMPLIFIER (METHOD APPLICATION EXAMPLE)

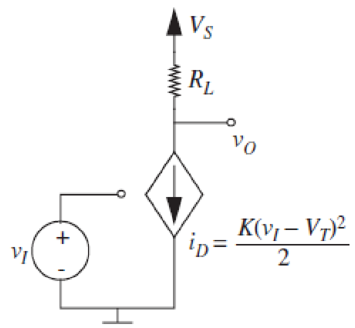


FIGURE 8.14 The MOSFET amplifier.

As an application of what we outlined just above, let us derive the small-signal equivalent circuit for the MOSFET amplifier in Figure 8.14.

The first step is to determine the operating point of the MOSFET amplifier for its bias voltages using the large-signal SCS circuit model. Assuming that the input bias voltage is VI, we can determine, as we did in the last chapter, the output operating current ID and the output operating voltage VO and we obtain:

$$ID = \frac{K}{2}(VI - VT)^2$$

$$VO = VS - ID * RL = VS - \frac{K}{2}(VI - VT)^2 RL$$

The second step is to determine the linearized small-signal models for each component. Referring to Figure 8.10, we see that the small-signal model for the DC power supply is a short, and for a resistor is the resistor itself. Finally, the linearized small-signal model for the MOSFET in saturation is a voltage-dependent current source whose small-signal current is linearly related to the small-signal gate-to-source voltage as:

$$ids = K(VGS - VT) * vgs$$

As the third step, we replace each original component in the circuit with its linearized equivalent and re-label the circuit with the small-signal branch variables vi, vo, and id as depicted in Figure 8.16.

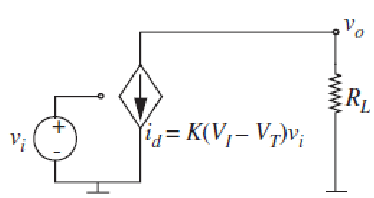


FIGURE 8.16 The small-signal SCS circuit model for the MOSFET amplifier.

The small-signal circuit model can be analyzed to determine the circuit response to small signals. For example, we can use Figure 8.16 to determine the small signal gain of the MOSFET amplifier. Applying KVL at the output, we get

$$vo = -id * RL = -K(VI - VT) * vi * RL$$

Thus, the small signal gain is given by

$$\frac{vo}{vi} = -K(VI - VT) * RL = -gm * RL$$

8.2.3 SELECTING AN OPERATING POINT

We have discussed the choice of operating point in the context of large signals before. When dealing with small signals, other criteria are often more important in selecting the operating point than just obtaining maximum dynamic range. One criterion is the small-signal gain of the amplifier. The small-signal gain of the amplifier is dependent on the input operating point voltage VI. The magnitude of the small-signal gain is given by:

$$\left|\frac{vo}{vi}\right| = K(VI - VT)RL \quad (8.30)$$

Figure 8.17 plots the magnitude of the gain for various values of VI. The graph indicates that the amplifier gain increases with increasing VI.

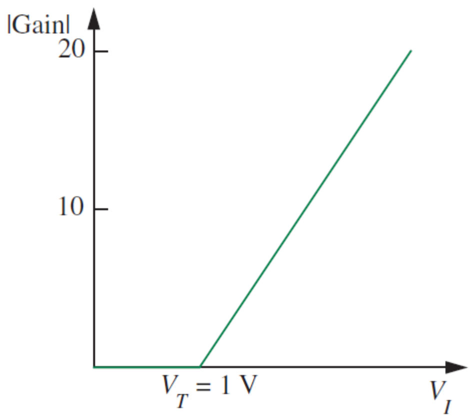


FIGURE 8.17 Magnitude of the small-signal gain of the amplifier for various values of the input operating point voltage V_I.

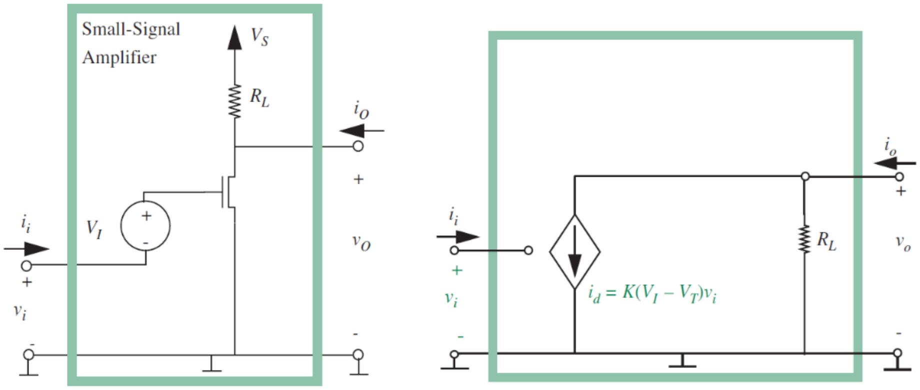
A second criterion is the output operating-point voltage. This is important when the amplifier must drive another circuit stage, and the output operating-point voltage of the amplifier determines the input operating-point voltage of the next stage.

Steps to set the operating point for both criteria:

1. Determine the voltage gain we want our small signal to achieve
2. Substitute the circuit values in equation 8.30 to find VI, our operating point. At this point the gain will be our target gain.
3. Use large signal analysis to verify the min and max point at which the MOSFET operates in saturation.
4. Make sure that our small signal vi swings within the limits found.

8.2.4 INPUT AND OUTPUT RESISTANCE, CURRENT AND POWER GAIN

The small-signal equivalent circuit also allows us to determine other important circuit parameters, such as the small-signal input resistance, output resistance, current gain, and power gain. Since the amplifier behaves as a linear network for small signals, it can be characterized by a Thévenin equivalent when viewed from any given port. The input and output resistance come in handy in this Thévenin characterization. Let us determine these values for the MOSFET amplifier using its small-signal circuit in Figure 8.16. Since these parameters are externally observed quantities, they are defined with respect to the external ports of the amplifier abstraction. Thus, it is important that we define precisely what constitutes the input and output ports of the small-signal amplifier. Figure 8.19 shows the relationship between the external ports of the amplifier circuit and the small-signal model. Notice that we have internalized the input bias voltage into the small-signal amplifier abstraction, so the user of the amplifier does not have to provide the appropriate input bias voltage. Instead, the user can simply provide a small input signal and observe the resulting signal output superimposed on the DC output offset.



(a) (b)

FIGURE 8.19 Amplifier input and output ports: (a) amplifier circuit; (b) small-signal model. As shown in the amplifier circuit, we have internalized the input bias voltage into the small-signal amplifier abstraction.

Input resistance *ri*

Incremental input resistance is the change in the input current for a small change in the input voltage.

Accordingly, as depicted in Figure 8.20 we compute *ri* by applying a small test-voltage *vtest* at the input and measuring the corresponding current *itest*. All other independent small-signal voltages or DC voltage sources are shorted. All other independent small-signal or DC current sources are turned into open circuits. The input resistance for the MOSFET amplifier is given by

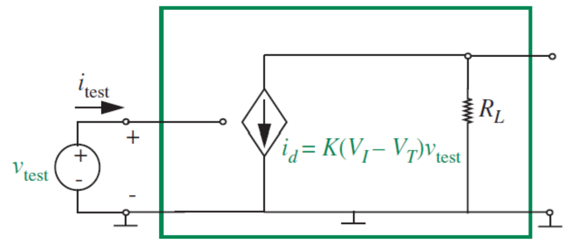


FIGURE 8.20 Input resistance measurement.

ri = vtest / itest = infinite (because itest = 0)

Output resistance *rout*

Incremental output resistance is the change in the output current for a small change in the output voltage. As depicted in Figure 8.21, we compute the *rout* by applying a small test voltage *vtest* at the output and measuring the corresponding current *itest*. As before, all other independent small-signal voltages or DC voltage sources are shorted. Thus, the small-signal input voltage vi is set to 0. Similarly, all other independent small-signal or DC current sources are turned into open circuits. The output resistance is given by

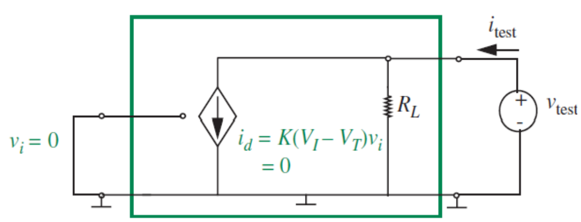


FIGURE 8.21 Output resistance measurement.

rout = vtest / itest = RL

Because the small signal voltage input *vi* is set to zero, the current through the MOSFET is 0. In other words, the MOSFET behaves like an open circuit. Thus, the output resistance for small signals is RL.

Current gain

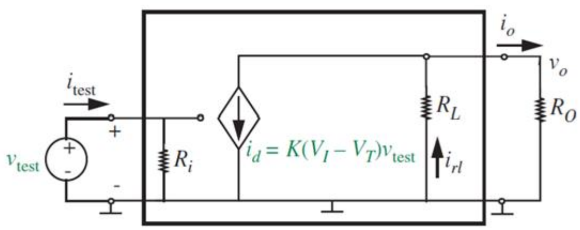


FIGURE 8.22 Current gain measurement. As an exercise, we place a resistance R_i between the input terminal and ground. For a MOSFET, $R_i = \infty$.

Current gain is the change in the output current divided by the change in the input current, for a given external load resistance. As depicted in Figure 8.22, we can compute the current gain by applying a small test voltage at the input and measuring both the input current itest and the output current io. The ratio io/itest is the current gain. Note that the output current is not the current that flows through the dependent current source, rather it is the current that is drawn by an external load resister RO. Because it is dependent on the value of the load resistor, the current gain is defined for a given load resistance. The introduction of an external load resistance also reduces the voltage gain of the amplifier because it appears in parallel with the internal load resistor RL. The current gain with an external load resistance RO is given by

Current gain = io / itest

Calculations are:

$$Current\ gain = \frac{\frac{vo}{RO}}{\frac{vtest}{Ri}} = \frac{vo}{vtest}\frac{Ri}{RO} = -K(VI - VT)(RL||RO)\frac{Ri}{RO}$$

Notice that the voltage gain of the amplifier with an external load is lower than an unloaded amplifier. For our idealized MOSFET, since Ri is infinite, the current gain is also infinite.

Important points to understand this paragraph (ChatGPT)
Adding Ri (input resistance) and RO (output resistance) in the current gain analysis is a practical step to make the model more realistic and reflect real-world conditions. Purpose of Ri: - In a perfect MOSFET, the gate does not draw any current, so the Ri is infinite. - However, in real-world amplifiers, the input will often interface with other circuits that have finite input resistance. - Adding Ri simulates a practical scenario where the amplifier's input draws some (small) current, and you can see how it affects the current gain. Purpose of RO: - Similarly, the output of an amplifier often connects to an external load, represented by RO. - RO impacts the current that flows at the output and the overall performance of the amplifier. - Including RO allows us to calculate a current gain that depends on the interaction between the internal resistance of the circuit and the external load. Why Do This? The exercise is meant to illustrate: - How real-world resistances (finite Ri and RO) impact the amplifier's behavior.

- How the current gain depends not just on the intrinsic properties of the MOSFET but also on the external circuit it's embedded in.

Without Ri and RO, the theoretical current gain would always appear infinite (because i=0), which doesn't reflect practical situations.

In summary: This exercise bridges the gap between the idealized MOSFET amplifier and a realistic, practical circuit that interfaces with real-world loads and sources.

Power gain

Incremental power gain is the ratio of the power supplied by the amplifier to an external load to that supplied to the amplifier by the input source.

Since we already have the voltage gain and the current gain calculated before, the power gain is given by

$$Power\ gain = \frac{vo * io}{vtest * itest} = \frac{vo}{vtest}\frac{io}{itest} = [K(VI - VT)(RL||RO)]^2 \frac{Ri}{RO}$$

As before, since for our MOSFET amplifier Ri is infinite, the power gain is also infinite. In practical circuits, however there is always some input resistance, so the power gain is finite.

Page 427, example 8.2 for a simple example of what we have discussed so far.

Below follow some examples of analysis of different types of amplifiers (differential, BJT, operational etc.) Basically we apply the general method studied above to some specific amplifiers circuits.

PAGE 429 DIFFERENTIAL AMPLIFIER

When a signal is noisy, straightforward use of an amplifier would amplify both the signal and the noise. But a differential amplifier can be used to amplify the signal by a much larger gain relative to the noise. The differential amplifier is a building block for high-quality amplification and is useful for processing small signals.

How does it work? (The paragraph following here has been mostly written be me cause the book I believe is not clear)

There are two wires carrying two input signals carrying both the same amount of noise. Usually, these two wires come out of a sensor which measures, for example, external temperature. One of the two wires has an inverter signal so for example one has +1.5V and the other -0.5V with a noise of 0.5V (so the original clean measurement is 1V and -1V). Then from these two input wires we calculate two components, the difference-mode component (we subtract one from the other, in our example = 2V) and the common mode component (in our example = 0.5V). The difference mode component is passed into a MOSFET with high gain AD and the common mode component into a MOSFET with low gain AC. As long as the noise is the same in both wires, the difference mode component will stay the same.

The wires are often twisted together to ensure that when there is noise, the same amount of noise infects both wires. In other situations, the two wires might both carry a common DC bias, and also in this case the differential amplifier is of help.

The differential amplifier abstraction is shown in Figure 8.26. It is a two-port device with one differential input port and one single-ended output port. The input port has two input terminals. The + input is called the non-inverting input and the – input is called the inverting input. It has an output port across which vO appears.

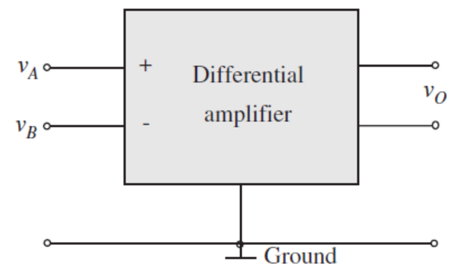


FIGURE 8.26 Difference amplifier black box representation.

The behavior of the difference amplifier is best explained by considering its effect on the following signals related to the two components, vA and vB:

1. A difference-mode component signal, VD = VA - VB
2. A common-mode component signal, VC = (VA + VB) / 2

The output of the difference amplifier is a function of these two components of the input, VO = AD*VD + AC*VC. AD is the difference-mode gain and AC is the common-mode gain.

The key in using a difference amplifier is to encode the useful signal in the difference-mode component and the noise in the common-mode component. Then if we make AD large and AC small, we achieve our goal of noise reduction. Usually we use the common-mode rejection ratio (CMRR) to describe the ability of the amplifier to reject the common-mode noise:

CMRR = AD / AC

For an example and extensive practical exercise of the MOSFET implementation of the Difference Amplifier refer to book page 431. Figure 8.28 and 8.29, the circuit for the example, are kept as a reference for what explained above.

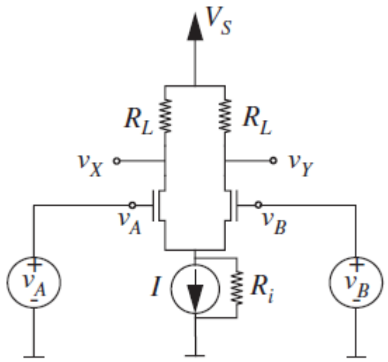


FIGURE 8.28 Source-coupled difference amplifier. All voltages are measured with respect to ground.

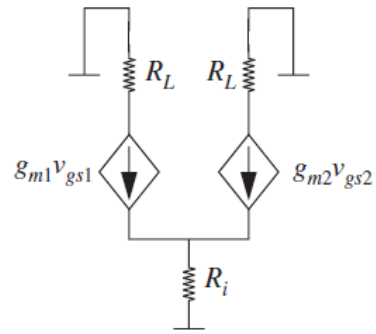


FIGURE 8.29 Source-coupled difference amplifier — small-signal model.

PAGE 436 - SOURCE FOLLOWER

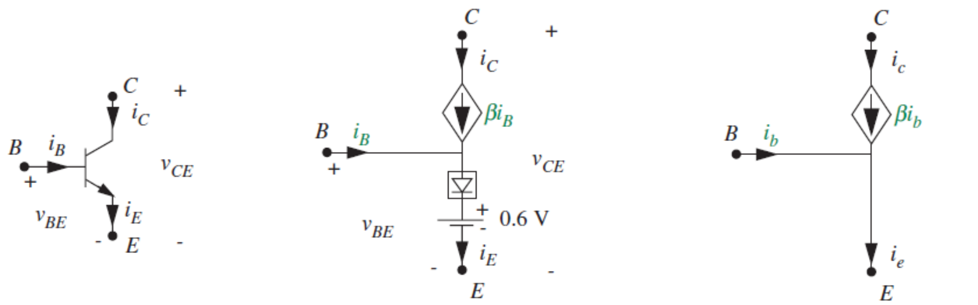
This circuit we have seen before. Full example at page 436. The source-follower output resistance can be made low. The low output resistance makes the source follower very useful as a buffer device, which can provide a large amount of current gain. If I am not wrong means that the input voltage is not distorted and I get a boost in the current output.

PAGE 438 - SMALL SIGNAL MODEL FOR THE BJT

Considering the BJT in its active region (as we use it as an amplifier), we can ignore the base-to-collector diode as it behaves like an open circuit. Figure 8.42c depicts the small signal circuit of the piecewise linear model of figure 8.42b. We can easily see that $i_C = \beta i_B$ and thus for incremental variables $i_c = \beta i_b$.

The full example for the BJT on page 438 onwards. We will find that the gain of the BJT amplifier is independent of the operating point, provided the BJT operates in the active region. For a given BJT device (that is, a fixed value for ß) the gain can be increased by increasing RL or decreasing RI.

FIGURE 8.42 Small-signal model for the BJT.



(a) BJT symbol

(b) BJT large-signal model assuming BJT is in active region

(c) BJT small-signal model

PAGE 443 – OPEREATIONAL AMPLIFIER

--- skipped --

CHAPTER 9: ENERGY STORAGE ELEMENTS

Until now, time has not been important:

- The analyses we have performed so far have been static.
- All circuit responses at a given time have depended only on the circuit inputs at that time.

Therefore, our circuits have so far responded to input changes infinitely fast. But in reality circuits do take time to respond to their inputs, and this delay is often of significant importance.

As an example of circuit delays, and the importance of time in describing the response of a circuit, consider an inverter. Its ideal response, based on our analysis of electronic circuits to this point, is shown in Figure 9.2. In reality though, the output shown in Figure 9.3 is more likely to occur.

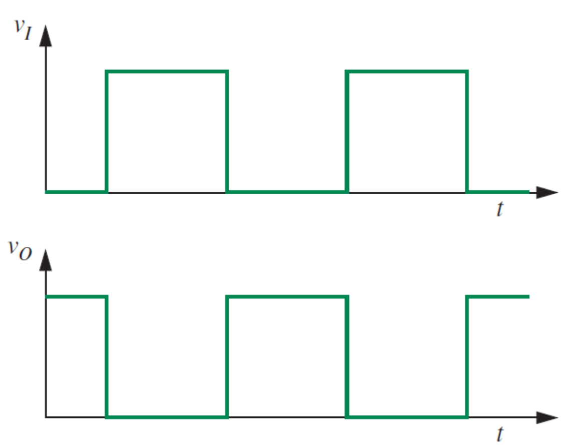


FIGURE 9.2 Ideal response of the first inverter to a square-wave input.

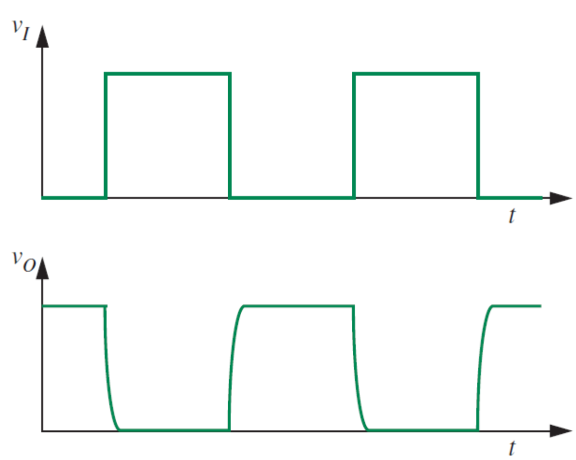


FIGURE 9.3 Observed response of the first inverter to a square-wave input.

To explain the temporal behavior of circuit responses we must introduce two new elements: capacitors and inductors.

9.1 CONSTITUTIVE LAWS

Here we introduce the capacitor and inductor in the abstract and develop the constitutive laws that relate their branch variables.

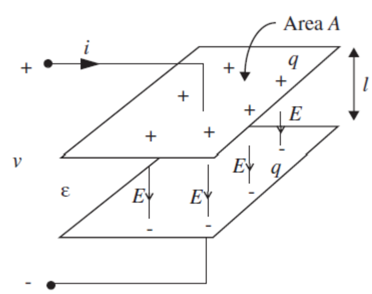
Some vocabulary:

- Dielectric: A material that doesn't conduct electricity but can support an electric field. Think of it as an insulator that can store electrical energy by polarizing when placed in an electric field.
- Permittivity: A measure of how much a material can "permit" or allow the formation of an electric field within it. It determines how well the material can store electrical energy when exposed to an electric field.
- Magnetic permeability (symbol: μ) is a property of a material that describes how well it can support the formation of a magnetic field within itself. It measures the ability of a material to conduct magnetic flux, or in simpler terms, how easily a magnetic field can pass through the material.

9.1.1 CAPACITORS

An idealized two terminal linear capacitor is shown in Figure 9.11. In this capacitor each terminal is connected to a conducting plate. The two plates are parallel and are separated by a **gap of length l**. Their **area of overlap is A**. Note that these dimensions will be functions of time if the geometry of the capacitor varies. The gap is filled with an insulating linear **dielectric having permittivity ϵ**. As current enters the positive terminal of the capacitor, it transports the **electric charge q** onto the corresponding plate; the unit of charge is the Coulomb [C]. Simultaneously, an identical current exits the negative terminal and transports an equal charge off the other plate. Thus, although charge is separated within the capacitor, no net charge accumulates within it, as is required for lumped circuit elements by the lumped matter discipline discussed in Chapter 1. The charge q on the positive plate and its image charge −q on the negative plate produce an electric field within the dielectric. It follows from Maxwell's Equations and the properties of linear dielectrics that the strength E of this field is

FIGURE 9.11 An idealized parallel-plate capacitor.



$$E(t) = \frac{q(t)}{\epsilon A(t)} \quad (9.1)$$

and its direction points from the positive plate to the negative plate. The electric field can then be integrated across the dielectric from the positive plate to the negative plate to yield

$$v(t) = l(t)E(t) \quad (9.2)$$

Combining the two equations results in

$$q(t) = \frac{\epsilon A(t)}{l(t)} v(t) \quad (9.3)$$

And we define

$$C(t) = \frac{\epsilon A(t)}{l(t)} \quad (9.4)$$

Where C is the capacitance of the capacitor having the units of Coulombs/Volt, or Farads[F].

By substitution in equation 9.3, we have Equation 9.5: $q(t) = C(t)v(t)$. **The capacitor exhibits an algebraic relation between its branch voltage and its stored charge**. The rate at which charge is transported onto the positive plate of the capacitor is

$$\frac{dq(t)}{dt} = i(t) \quad (9.6)$$

From Equation 9.6 we see that the Ampere is equivalent to a Coulomb/second. Equation 9.6 can be combined with Equation 9.5 to yield

$$i(t) = \frac{d\big(C(t)v(t)\big)}{dt} \quad (9.7)$$

which is the element law for an ideal linear capacitor. Unless stated otherwise, we will assume in this text that capacitors are both linear and time-invariant. For linear, time-invariant capacitors, Equations 9.5 and 9.7 reduce respectively to

$$\boldsymbol{q(t) = Cv(t) \quad (9.8)}$$

$$\boldsymbol{i(t) = C\frac{dv(t)}{dt} \quad (9.9)}$$

The symbol for capacitors is shown in Figure 9.12, along with the graph of the relation between the branch voltage and stored charge of the capacitor.

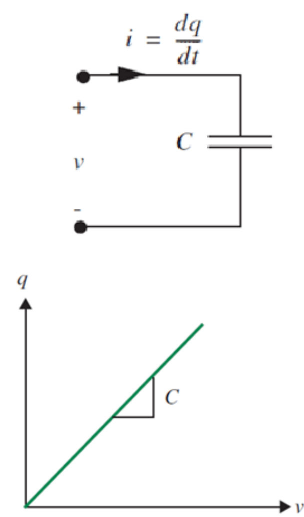


FIGURE 9.12 The symbol and voltage-charge relation for the ideal linear capacitor. The element law for the capacitor is $i = Cdv/dt$.

One of the important properties of a capacitor is its **memory property**. In fact, it is this property that allows the capacitor to be the primary memory element in all integrated circuits. To see this property we integrate 9.6 to produce 9.11 and we integrate 9.8 to produce 9.12

$$q(t) = \int_{-\infty}^{t} i(t)dt \quad (9.11)$$

$$v(t) = \frac{1}{C}\int_{-\infty}^{t} i(t)dt \quad (9.12)$$

Equation 9.12 shows that the branch voltage of a capacitor depends on the entire past history of its branch current, which is the essence of memory.

From equation 9.13 below it follows that q(t1) perfectly summarizes, or memorizes, the entire accumulated history of i(t) for t ≤ t1. Thus if q(t1) is known, it is necessary and sufficient to know i only over the interval t1 ≤ t ≤ t2 in order to determine q(t2). For this reason, **q is referred to as the state of the capacitor**.

$$q(t2) = \int_{-\infty}^{t2} i(t)dt = \int_{t1}^{t2} i(t)dt + \int_{-\infty}^{t1} i(t)dt = \int_{t1}^{t2} i(t)dt + q(t1)\ (9.13)$$

For linear time-invariant capacitors, v can also easily serve as a state because v is proportionally related to q through the constant C. Accordingly, we can rewrite Equation 9.12 as

$$v(t2) = \frac{1}{C}\int_{-\infty}^{t2} i(t)dt = \frac{1}{C}\int_{t1}^{t2} i(t)dt + v(t1)\ (9.14)$$

Thus v(t1) also memorizes the entire accumulated history of i(t) for t < t1 and can serve as the state of the capacitor.

Associated with the ability to exhibit memory is the **property of energy storage**, which is often exploited by circuits that process energy. To determine the electric energy wE stored in a capacitor, we recognize that the power iv is the rate at which energy is delivered to the capacitor through its port. Thus,

$$\frac{dwE(t)}{dt} = i(t)v(t)\ \ (9.15)$$

Next, substitute for i using Equation 9.6, cancel the time differentials, and omit the parametric time dependence to obtain

$$dwE = dq * v\ \ \ (9.16)$$

Equation 9.16 is a statement of incremental energy storage within the capacitor. It states that the transport of the incremental charge dq from the negative plate of the capacitor to the top plate across the electric potential difference v stores the incremental energy dwE within the capacitor. To obtain the total stored electric energy, we must integrate Equation 9.16 with v treated as a function of q. This yields

$$wE = \int_0^q v(x)dx\ \ \ (9.17)$$

where x is a dummy variable of integration. Finally, substitution of Equation 9.8 and integration yields

$$\boldsymbol{Stored\ energy = wE(t) = \frac{q^2(t)}{2C} = \frac{Cv(t)^2}{2}\ \ \ (9.18)}$$

as the electric energy stored in a capacitor. The units of energy is the Joule [J], or Watt-second.

Recommended video here.

9.1.2 INDUCTORS

The inductor models the effect of magnetic fields. Figure 9.13 shows a two-terminal linear inductor where a coil with a terminal on each end is wound with **N turns** around a toroidal core made from an insulator having **magnetic permeability μ**. The **length around the core is l** and its **cross-sectional area is A**.

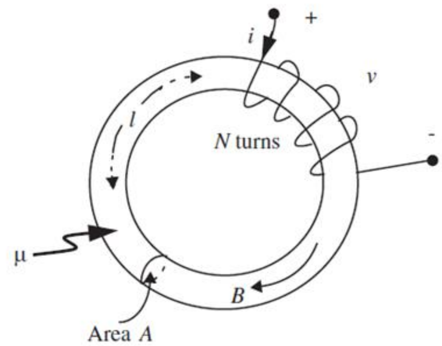


FIGURE 9.13 An idealized toroidal inductor.

The current in the coil produces a magnetic flux in the inductor. Ideally, this magnetic flux does not stray significantly from the core, so that the flux outside the element is negligible. Thus the inductor can be treated as a lumped circuit element that satisfies the lumped matter discipline discussed in Chapter 1. From Maxwell's Equations and the properties of permeable materials, the density B of the flux is

$$B(t) = \frac{\mu N i(t)}{l(t)} \quad (9.19)$$

and its direction is around the core. The magnetic flux density can be integrated across the core to yield

$$\Phi(t) = A(t)B(t) \quad (9.20)$$

where Φ is the total flux passing through the core, and hence through one turn of the coil. Since the flux Φ is linked N times by the N-turn coil, the total flux λ linked by the coil is

$$\lambda(t) = N\Phi(t) = NA(t)B(t) \quad (9.21)$$

The units of flux linkage is the Weber [Wb]. Combining Equations 9.19 and 9.21 results in

$$\boldsymbol{\lambda(t) = \frac{\mu N^2 A(t)}{l(t)} i(t) \quad (9.22)}$$

We define L, the inductance of the inductor, as

$$\boldsymbol{L(t) = \frac{\mu N^2 A(t)}{l(t)} \quad (9.23)}$$

L has the units of Webers/Ampere, or Henrys [H]. So (9.24): $\lambda(t) = L(t)i(t)$.

The inductor exhibits an algebraic relation between its branch current and its flux linkage. Again, from Maxwell's Equations, the rate at which flux linkage builds up in the inductor is

$$\frac{d\lambda(t)}{dt} = v(t) \quad (9.25)$$

From the equation we see that the Volt is equivalent to a Weber/second. By combining 9.25 and 9.24 we yield the element law for an ideal linear inductor:

$$v(t) = \frac{d\big(L(t)i(t)\big)}{dt} \quad (9.26)$$

For time-invariant inductors, 9.24 and 9.26 reduce to:

$$\boldsymbol{\lambda(t) = Li(t)} \quad \mathbf{(9.27)}$$

$$\boldsymbol{v(t) = L\frac{di(t)}{dt}} \quad \mathbf{(9.28)}$$

The symbol for an ideal linear inductor, together with the graph of the relation between branch current and flux linkage of the inductor is shown in Figure 9.14.

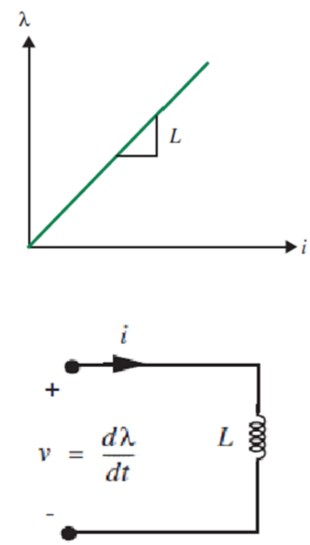


FIGURE 9.14 The symbol and current-flux-linkage relation for an ideal linear inductor. The element law for an inductor is $v = L di/dt$.

Same as capacitor, a property of an inductor is its **memory property**. We integrate Equation 9.25:

$$\lambda(t) = \int_{-\infty}^{t} v(t)dt \quad (9.29)$$

Or with the substitution of equation 9.27:

$$i(t) = \frac{1}{L}\int_{-\infty}^{t} v(t)dt \quad (9.30)$$

Equation 9.30 shows that the branch current of an inductor depends on the entire past history of its branch voltage, which is the essence of memory. In the same manner as with the capacitor, as long as we know λ(t1), which summarizes the entire history of v(t) for t ≤ t1, it is necessary and sufficient to know v only over the interval t1 ≤ t ≤ t2 in order to determine λ(t2). For this reason, **λ, the total flux linked by the coil, is referred to as the state of the inductor**.

$$\lambda(t2) = \int_{-\infty}^{t2} v(t)dt = \int_{t1}^{t2} v(t)dt + \lambda \quad (9.31)$$

For linear time-invariant inductors, i can also serve as a state because it is proportionally related to λ through the constant L. So we can rewrite 9.30:

$$i(t2) = \frac{1}{L}\int_{-\infty}^{t2} v(t)dt = \frac{1}{L}\int_{t1}^{t2} v(t)dt + i(t1) \quad (9.32)$$

Inductors have also the **property of energy storage**. To determine the magnetic energy wM stored in an inductor, we recognize that the power iv is the rate at which energy is delivered to the inductor through its port. Thus,

$$\frac{dwM(t)}{dt} = i(t)v(t) \quad (9.33)$$

Next, substitute for v using Equation 9.25, cancel the time differentials, and omit the parametric time dependence to obtain

$$dwM = i * d\lambda \quad (9.34)$$

To obtain the total stored magnetic energy, we must integrate Equation 9.34 with i treated as a function of λ. This yields

$$wM = \int_0^{\lambda} i(x)dx \quad (9.35)$$

where x is a dummy variable of integration. Finally, substitution of 9.27 and integration yields

$$\boldsymbol{Stored\ energy = wM(t) = \frac{\lambda^2 t(t)}{2L} = \frac{Li(t)^2}{2} \quad (9.36)}$$

as the magnetic energy stored in an inductor.

9.2 SERIES AND PARALLEL CONNECTIONS

9.2.1 CAPACITORS

For capacitors in series, as in figure 9.15, we have the following rules/assumptions:

- The capacitors were not charged when connected.
- Since the capacitors share a common current, they also store a common charge q, being q(t) = C1v1(t) = C2v2(t)
- Using KVL we see v(t) = v1(t) + v2(t)

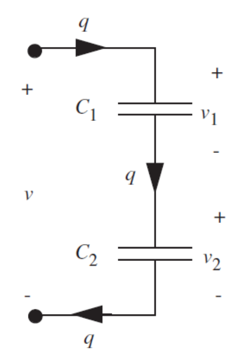


FIGURE 9.15 Two capacitors in series.

Finally, since the effective capacitance C of the two series capacitors is q/v, it follows that $\frac{1}{C} = \frac{v(t)}{q(t)} = \frac{1}{C1} + \frac{1}{C2}$ which is the same as equation 9.39, from which we see that the reciprocal capacitances of capacitors in series add.

$$C = \frac{C1C2}{C1 + C2} \quad (9.39)$$

For capacitors in parallel, as in figure 9.16, we have the following rules/assumptions:

- Since the capacitors share a common voltage v, it follows from 9.8 that $v(t) = \frac{q1(t)}{C1} + \frac{q2(t)}{C2}$
- Using KCL and equation 9.11 we see that q(t) = q1(t) + q2(t)

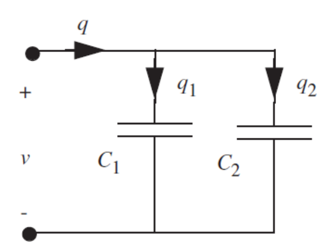


FIGURE 9.16 Two capacitors in parallel.

Finally, since the effective capacitance C of the two parallel capacitors is q/v, it follows equation 9.42, from which we see that the capacitances of capacitors in parallel add.

$$C = \frac{q(t)}{v(t)} = C1 + C2 \quad (9.42)$$

9.2.2 INDUCTORS

For inductors in series as in Figure 9.18:

- We assume that neither inductor carried a current at the time of their connection.
- Since the two inductors share a common current i, it follows from 9.27 that $i(t) = \frac{\lambda 1(t)}{L1} = \frac{\lambda 2(t)}{L2}$
- Using KVL and 9.29 we see that $\lambda(t) = \lambda 1(t) + \lambda 2(t)$

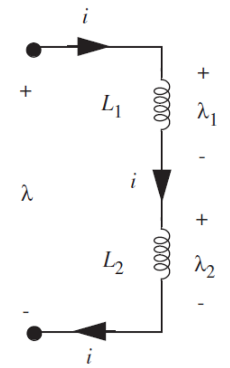


FIGURE 9.18 Two inductors in series.

Finally, since the effective inductance L of the two series inductors is λ/i, it follows 9.45 from which we see that the inductances of inductors in series add.

$$L = \frac{\lambda(t)}{i(t)} = L1 + L2 \quad (9.45)$$

For inductors in parallel as in Figure 9.19:

- Since the two inductors share a common voltage, it follows from Equation 9.29 that they share a common flux linkage λ, as shown in Figure 9.19.
- Thus, following Equation 9.27, $\lambda(t) = L1i1(t) + L2i2(t)$
- Using KCL we see that i(t) = i1(t) + i2(t)

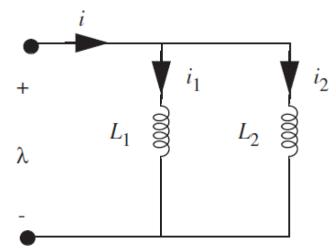


FIGURE 9.19 Two inductors in parallel.

Finally since the inductance L of the two parallel inductors is λ/i it follows that $\frac{1}{L} = \frac{i(t)}{\lambda(t)} = \frac{1}{L1} + \frac{1}{L2}$, which is the same as equation 9.48, from which we see that the reciprocal inductances of inductors in parallel add.

$$L = \frac{L1L2}{L1 + L2} \quad (9.48)$$

9.3 SPECIAL EXAMPLES (Not sure is summarized everything ok)

We examine here several parasitic capacitances and inductances that are commonly encountered inside integrated circuits, and in external wiring connections to them and other circuit elements.

9.3.1 MOSFET GATE CAPACITANCE

In Figure 9.22 an N-channel MOSFET with source and substrate grounded, and positive voltages applied to its gate and drain. Notice the silicon dioxide dielectric that separates its gate and channel.

As the positive gate voltage is applied, electrons flow from the source into the channel and accumulate beneath the gate. When the gate voltage exceeds the threshold voltage of the MOSFET, the electron density beneath the gate becomes sufficiently high to invert the channel from p-type silicon to n-type silicon. Thus, a continuous n-type channel forms between the source and drain, thereby allowing electrons to flow from the source to the drain, and hence current to flow from the drain to the source, in the response to the positive drain voltage.

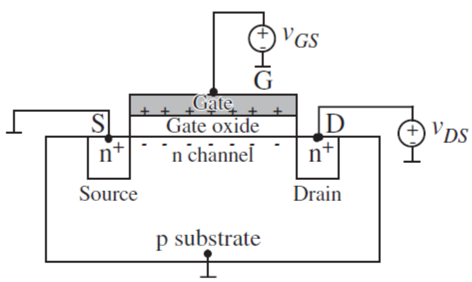


FIGURE 9.22 MOSFET with a positive voltage applied at the gate relative to the source and substrate.

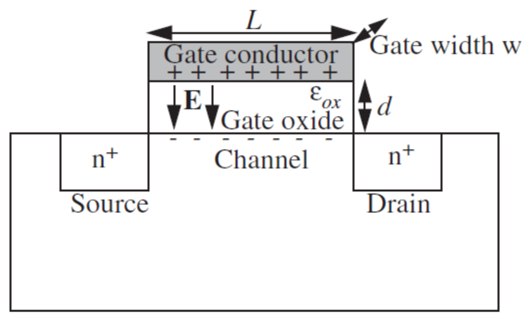


FIGURE 9.23 Electric charge and field within the MOSFET with a positive voltage applied at the gate relative to the source and substrate.

The important observation here is that in the process of inverting its channel, and turning itself on, the MOSFET actually forms a parallel-plate capacitor between its gate and channel. This is emphasized in Figure 9.23, which shows the electric field E in the silicon dioxide emanating from the positive charge on the gate and terminating on the negative charge in the channel. The gate-to-channel capacitance is approximately $\frac{\epsilon OXLW}{d}$ where ϵOX ≈ 3.9ϵo is the permittivity of the silicon dioxide, d is its thickness, L is the channel length and W its width. LW is the gate area. This capacitance is referred to as the gate-to-source capacitance of the MOSFET, CGS:

$$CGS = \frac{\epsilon OXLW}{d} \quad (9.49) \quad usually\ COX = \frac{\epsilon OX}{d}$$

This realization also leads to the augmented switch-resistor-capacitor (SRC) model of the MOSFET shown in Figure 9.24. Here, a lumped capacitor is added to the SR model to account for the charge that must be supplied to the gate conductor and channel in order to turn on the MOSFET. Because the SRC model contains a capacitor between the gate and source terminals of the MOSFET, a current will flow into the gate terminal and out from the source terminal of that model as the gate-to-source voltage of the MOSFET varies. This current transports the charge that accumulates within the MOSFET as seen in Figures 9.22 and 9.23. Following Equation 9.9, the current is given by

$$IG = CGS\frac{dvGS}{dt} \quad (9.50) \quad where\ CGS = COX * LW$$

From Equation 9.50 we can now begin to see the reason for the inverter behavior observed in Figure 9.3. It will take time for the gate current to transport charge onto the gate, and hence it will take time for the gate voltage to rise. Thus, it will take time for the inverter to pass a signal from its input to its output.

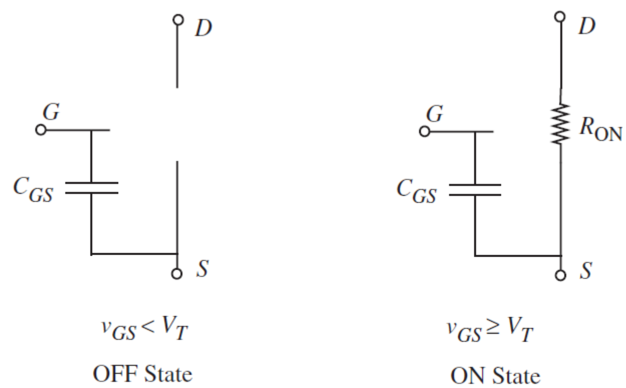


FIGURE 9.24 The switch-resistor-capacitor (SRC) model of the MOSFET.

9.3.2 WIRING LOOP INDUCTANCE

To estimate the inductance of a wiring loop, let's imagine a wire forms a circle loop. The circle loop has a radius R and the wire has a radius A. Its inductance L is given approximately by

$$L = \mu o R \left(\ln \left(\frac{8R}{A} \right) - 2 \right) \quad (9.52)$$

This expression can also be used to successfully approximate the inductance of many noncircular wiring loops. µ0 represents the permeability of free space (also known as the magnetic constant). Its value is approximately $4\pi \times 10^{-7}$ H/m

9.3.3 IC WIRING CAPACITANCE AND INDUCTANCE

-to read-

9.3.4 TRANSFORMERS

A transformer is a two-port device made by winding a second coil around the inductor. Let the first coil have N1 turns and the second coil have N2 turns. The symbol for an ideal transformer having this construction is shown in Figure 9.29. The two dots indicate the ends of the two coils that are wound in the same direction. In an ideal transformer, the coils are wound so tightly against each other that each of their turns links the same flux Φ (t). It then follows from Equations 9.25 and 9.21 that

$$v1 = N1\frac{d\Phi(t)}{dt}$$

$$v2 = N2\frac{d\Phi(t)}{dt}$$

so that

$$\frac{v1(t)}{N1} = \frac{v2(t)}{N2} \quad (9.59)$$

In an ideal transformer the core permeability is infinite. For a single-coil inductor carrying a finite flux Φ(t) = λ(t)/N, Equation 9.22 shows that the total ampere-turns Ni(t) flowing around the core through the coil must vanish as μ becomes infinite. In an ideal transformer, the total ampere turns must similarly vanish, and so

$$N1i1(t) + N2i2(t) = 0 \quad (9.60)$$

$$N1i1(t) = -N2i2(t) \quad (9.61)$$

Equations 9.59 and 9.61 are the constitutive equations for an ideal transformer. By combining 9.59 and 9.61, it can be observed that

$$v1(t)i1(t) = -v2(t)i2(t) \quad (9.62)$$

Thus, the power flowing into one port of an ideal transformer must instantaneously flow out from the second port. Said differently, an ideal transformer cannot store energy. This is consistent with having an infinitely permeable core.

9.4 SIMPLE CIRCUIT EXAMPLES

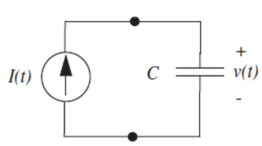


FIGURE 9.31 A current source driving a capacitor.

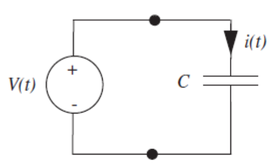


FIGURE 9.32 A voltage source driving a capacitor.

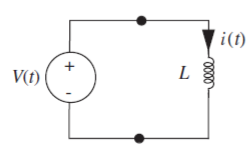


FIGURE 9.33 A voltage source driving an inductor.

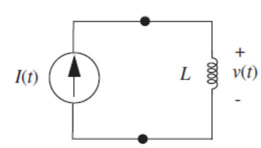


FIGURE 9.34 A current source driving an inductor.

Figure 9.31. The current I from the source must circulate through the capacitor. If we know the current through the capacitor, we know also the voltage v(t) from equations studied in 9.1.

Figure 9.32. The voltage from the source must also appear across the capacitor. So we can find the current circulating through both elements using equations studied in 9.1.

Figure 9.33. The voltage from the source must also appear across the inductor. So we can find the current across both elements. Same reasoning for Figure 9.34.

9.4.2 STEP INPUTS

<u>A current source driving a capacitor: Figure 9.31 and 9.37</u>

Let's consider the source step function:

$$I(t) = \begin{cases} 0 & t \leq 0 \\ Io & t > 0 \end{cases} \quad (9.75)$$

Current steps to Io at t = 0. Substituting 9. 75 into 9.12, the voltage across the capacitor is:

$$v(t) = \begin{cases} 0 & t \leq 0 \\ \dfrac{Io * t}{C} & t > 0 \end{cases} \quad (9.76)$$

***********Why do I multiply by t in 9.76? Voltage across capacitor is $v(t) = \frac{1}{C}\int I(t)dt$. Since the current I(t) is a step input, meaning I(t)=Io for t>0, substituting this constant current into the integral gives: $v(t) = \frac{Io}{C}\int_0^t 1dt = \frac{Io}{C} * t$. This is why the result includes t — it reflects the fact that the voltage increases linearly over time due to the constant current charging the capacitor. Physical interpretation is that for constant Io, voltage across the capacitor increases linearly with time as it accumulates charge**********

Once the current source steps to Io, it starts deliver charge to the capacitor at that constant rate. The charge then accumulates linearly in the capacitor, much like water would accumulate in a glass from a faucet set to deliver that water at a constant rate. Since charge and voltage are proportional through the constant capacitance C of the linear time-invariant capacitor, the voltage across the capacitor also increases linearly.

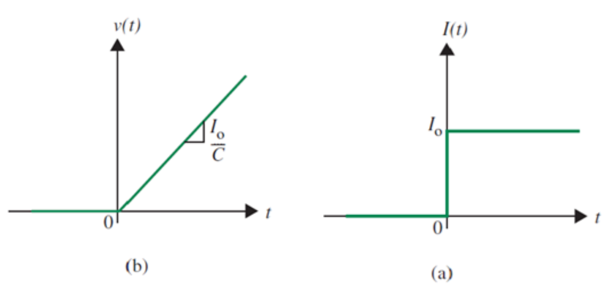


FIGURE 9.37 The current and voltage in the circuit shown in Figure 9.36.

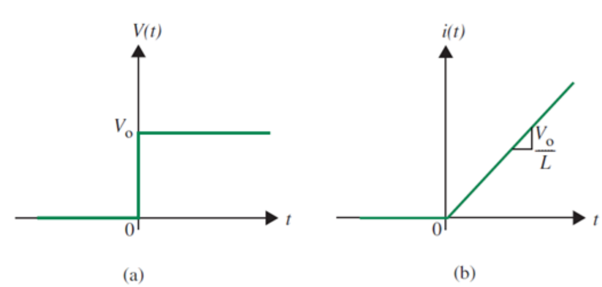


FIGURE 9.39 The current and voltage in the circuit shown in Figure 9.38.

<u>A voltage source driving an inductor: Figure 9.33 and 9.39</u>

Similar to the above. In summary:

$$V(t) = \begin{cases} 0 & t \leq 0 \\ Vo & t > 0 \end{cases} \quad (9.77)$$

Voltage steps to Vo at t = 0. Current through the inductor:

$$i(t) = \begin{cases} 0 & t \leq 0 \\ \dfrac{Vo * t}{L} & t > 0 \end{cases} \quad (9.77)$$

A voltage source driving a capacitor: Figure 9.32

Let’s consider Figure 9.32 and analyze its operation with a voltage source that takes a discontinuous step as:

$$V(t) = \begin{cases} 0 & t \leq 0 \\ Vo & t > 0 \end{cases} \quad (9.79)$$

Handling of the derivation of the function at t = 0:

To analyze its operation with a source that takes a discontinuous step, we refer to Equation 9.64, $i(t) = C\frac{dV(t)}{dt}$, and notice that we must contend with the differentiation of the step at t = 0. We can develop an understanding of this differentiation with the help of the ramping unit step function u(t;T) defined in Figure 9.41. Here, u(t;T) is a function of time t, having the ramp duration T as a parameter. Note that the ramp in u(t;T), which occurs over the period 0 ≤ t ≤ T, becomes increasingly steeper as the ramp width T approaches 0. In fact, it is u(t;T) in the limit T → 0, or simply u(t), as illustrated in Figure 9.42. Notice that the ideal unit step is the function at work in Equation 9.79. Recognizing this limiting behavior, our approach to dealing with the differentiation of a step will be to take a more roundabout, but easier route: We will compute the response of the circuit to a ramping unit step function, and then take the limit as T → 0.

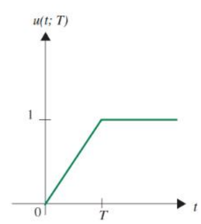


FIGURE 9.41 The ramping unit step function $u(t; T)$.

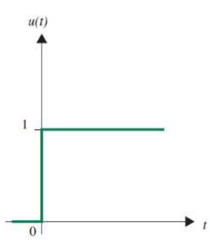


FIGURE 9.42 The unit step function $u(t)$.

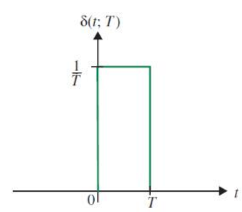


FIGURE 9.43 The unit-area pulse function $\delta(t; T)$ obtained by differentiating the ramping unit step function $u(t; T)$.

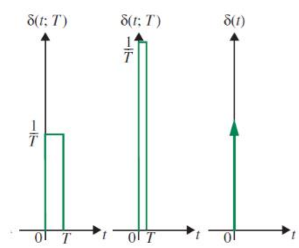


FIGURE 9.44 The unit-area pulse function becomes the unit impulse in the limit as $T \to 0$.

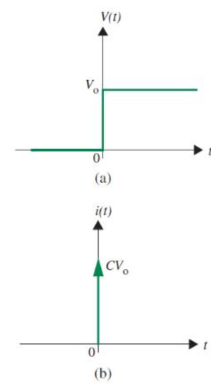


FIGURE 9.45 The voltage and current in the circuit shown in Figure 9.32 for a step voltage input.

Thus, we can rewrite Equation 9.79 in terms of the unit step function as:

$$V(t) = Vo * \lim_{T \to 0} u(t;T) = Vo * u(t) \quad (9.80)$$

The ramping unit step function u(t;T) can be differentiated to yield the unit-area pulse function δ(t;T) shown in Figure 9.43. This function becomes increasingly narrow and tall as T approaches 0, but in doing so it maintains unit area, as depicted in Figure 9.44. In the limit T → 0, δ(t;T) becomes the unit impulse (see the right-most graph in Figure 9.44), which we will simply denote by δ(t).

The unit impulse has several important properties for our purposes. These properties are:

$$\delta(t) = 0 \;\; for\; t \neq 0$$

$$\int_{-\infty}^{t} \delta(t)dt = u(t) \;\; => \;\; \delta(t) = \frac{du(t)}{dt}$$

$$\int_{-\infty}^{\infty} \delta(t)dt = 1$$

Note that in accordance with Figure 9.43, the units of δ(t) are reciprocal time.

Back to the circuit:

Suppose that the voltage source in the figure produces the ramping voltage step given by

$$V(t) = Vo * u(t;T) \ \ (9.84)$$

Substitution of Equation 9.84 into Equation 9.64 $i(t) = C\frac{dV(t)}{dt}$, then yields

$$i(t) = C * Vo * \delta(t;T) \ (9.85)$$

Equation 9.85 shows that the voltage source supplies the current C*Vo/T during the period 0 ≤ t ≤ T as it ramps up the capacitor voltage from 0 to Vo; note that δ(t;T) = 1/T during that period. In ramping up the capacitor voltage to Vo, the source delivers the charge C*Vo, in accordance with Equation 9.8. This can be verified by integrating i over 0 < t < T.

Now consider the circuit behavior described by Equations 9.84 and 9.85 in the limit T → 0. In this case, V becomes the discontinuous voltage step described by Equations 9.79 and 9.80, and i becomes

$$i(t) = C * Vo * \delta(t) \ (9.86)$$

This response can also be obtained directly by substituting Equation 9.80 into Equation 9.64, and then making use of Equation 9.82.

From our limiting interpretation of the impulse, we see that i in Equation 9.86 is a current that instantaneously delivers the charge C*Vo to the capacitor at t = 0. Thus, the charge stored in the capacitor takes a step at t = 0, and so the voltage steps too as driven by the source. This illustrates an important point made earlier, namely that it takes an infinite current to cause the charge stored by, and hence the voltage appearing across, a capacitor to take a discontinuous step. Thus, except under unusual circumstances involving infinite currents, the state of a capacitor is a continuous function of time.

<u>A current source driving an inductor: Figure 9.34</u>

We consider a current step input of the following form (Figure 9.47):

$$I(t) = \begin{cases} 0 & t \leq 0 \\ Io & t > 0 \end{cases} \ \ (9.87)$$

Following the steps of the previous example, using the step function, we find:

$$v(t) = L * Io * \delta(t,T) \ (9.89)$$

Equation 9.89 shows that the current source supplies the voltage L*Io/T during the period 0 < t < T as it ramps up the inductor current from 0 to Io. Note again that δ(t,T) = 1/T during that period. In ramping up the inductor current to Io, the source delivers the flux linkage L*Io in accordance with Equation 9.27. Consider the circuit behavior described by Equation 9.89 in the limit T ⟶ 0. The result is:

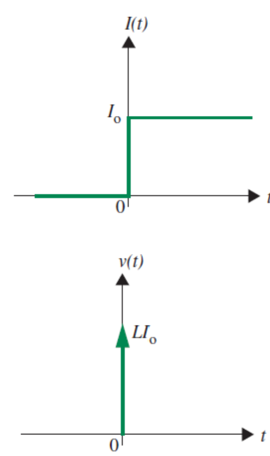


FIGURE 9.47 The current and voltage in the circuit shown in Figure 9.46 for a step current input.

$$v(t) = L * Io * \delta(t)\ (9.91)$$

The forms of the input current step and the voltage impulse response for the circuit in Figure 9.34 are depicted in Figure 9.47.

From our limiting interpretation of the impulse we now see that v in Equation 9.91 is a voltage that instantaneously delivers the flux linkage LIo to the inductor at t = 0. Thus, the flux linked by the inductor takes a step at t = 0, and so the current steps too as driven by the source.

9.4.3 IMPULSE INPUTS

Here the example in figure 9.31 (a current source driving a capacitor) is reviewed, using the impulse function we have seen in the previous two examples. Assuming we have a current impulse with an area under the impulse Q, or:

$$\int_{-\infty}^{\infty} i(t)dt = \int_{-\infty}^{\infty} Q\delta(t)dt = Q$$

Let us analyze the circuit in figure 9.31 for an impulse input current with strength Q. In other words, an input current of the form I(t) = Qδ(t). So we need to find the voltage across the capacitor, by integrating the current through the capacitor:

$$v(t) = \frac{1}{C}\int_{-\infty}^{t} I(t)dt = \frac{1}{C}Qu(t)$$

Thus, a current impulse of strength Q that occurs at time t deposits a charge Q on the capacitor. This charge results in the capacitor voltage jumping to (1/C)*Q at time t.

“TBH not sure why we use the impulse function to solve this circuit when we already solved it before. Also impulse function as I understand are mathematical constructions used to simplify our lives so they have to be taken with a grain of salt”

9.4/B ChatGPT - CONSIDERATIONS ON STEP FUNCTION u(t;T), PULSES $\frac{d}{dt}u(t;T)$, IMPULSES δ(t)

The need for step function, pulses and impulse functions come from the need to analyze the circuit operation with a source that takes a discontinuous step. If voltage increases to 5V for example, we need to analyze the moment of the switch from 0 to 5. If we want to find the current in a circuit with a capacitor, we need to use i(t)=C*dV(t)/dt and thus use differentiation at t = 0. The functions we discuss come to help in this process.

Step-by-Step explanation on how to go from step function to impulse:

1. Ramping Unit Step Function u(t;T):
    a. u(t;T) represents a ramp that starts at 0 and linearly increases to 1 over a duration T, after which it remains constant at 1. See figure 9.41.

b. Mathematically: $u(t;T) = \begin{cases} \frac{t}{T} & 0 \leq t < T \\ 1 & t \geq T \end{cases}$

c. This means that for 0 ⩽ t < T the function increases linearly with a slope of 1/T, and after T it stays constant at 1.

2. Differentiating u(t;T) with respect to t:
   a. The derivative of u(t;T) with respect to time t results in a pulse of constant height 1/T over the interval 0 ⩽ t < T.
   b. Mathematically: $\frac{d}{dt}u(t;T) = \begin{cases} \frac{1}{T} & 0 \leq t < T \\ 0 & t \geq T \end{cases}$
   c. This results in a rectangular pulse with height 1/T and duration T.
   d. As seen just above, the differentiation is done in the two different regions of the curve. The derivative of t/T is 1/T. The derivative of 1 is 0.
3. Unit Area Condition:
   a. The key property of this pulse is that it maintains unit area, meaning the integral of the pulse over time is always 1, regardless of the value of T.
   b. The area of the pulse is given by: Area = 1/T * T = 1
   c. As T becomes smaller, the pulse becomes narrower and taller while keeping the total area equal to 1.
   d. The pulse has a constant height of 1/T over the time interval [0,T], as in figure 9.43.
4. Limit as T ⟶ 0:
   a. Why does this result in an impulse as T → 0?
   b. When T ⟶ 0, the pulse becomes infinitely narrow (T → 0) and infinitely tall (1/T → ∞). Despite this, the area of the pulse remains constant at 1. This results in the unit impulse function δ(t).
   c. The impulse function δ(t) has zero width and infinite height, but its integral over all time is equal to 1.

So, if you want to graph these functions:

Function	Name	Graph	Figure
u(t;T)	Ramping unit step function	You draw a line that starts at 0, increases linearly to 1 over time T, and stays at 1.	9.41
δ(t;T)	Unit area pulse function	You draw a rectangle starting at t = 0 with height 1/T and width T.	9.43
δ(t) (when T⟶0)	Unit impulse	You draw an arrow at t=0 with an area of 1, representing the ideal impulse.	9.44 right
u(t)	Unit step function		9.42

In the graphs, the Y-axis represents the amplitude of the mathematical function itself, which is used to model how voltage or current changes over time. Once you multiply these functions by a voltage or current value, you get the actual voltage or current input to the circuit.

If it is not clear, here the difference between a pulse δ(t;T) and an impulse δ(t). A pulse is simply a signal that lasts for a short period of time and has a certain height. It's used to represent events where a quantity (like voltage or current) changes quickly, but not instantly. As T ⟶ 0, this pulse becomes narrower and taller, approaching an ideal impulse, which models an instantaneous event with zero duration but finite effect (area = 1).

<u>Practical Example:</u>

We apply a step voltage of 5V at t=0, so we have: Vinput(t)=5·u(t)

Here, u(t) is a step function (ranging from 0 to 1), and multiplying by 5 scales it represents 5 volts. A step voltage represents a sudden change in voltage from 0V to a constant value at t=0. Once the voltage steps up, it stays constant indefinitely.

If you want to apply an impulse of 5V·s at t=0, you use: Vinput(t)=5·δ(t)

Here, multiplying δ(t) by 5 represents an impulse with a total area of 5V·s (volts multiplied by seconds). An impulse voltage represents an extremely short, high-magnitude burst of voltage at t=0, with a total area (integral) equal to the desired impulse strength. This can be used when we have a circuit with capacitors for example in that instant the capacitors will charge and achieve the steady-state (see Poblem 9.1 in the notes as an example).

9.5 ENERGY, CHARGE, AND FLUX CONSERVATION

Here we consider connections in the presence of initial charge and flux linkage.

Consider the parallel connection of the two initially-charged capacitors shown in Figure 9.51; the connection occurs when the switch closes. We wish to determine the state of the capacitors after the switch is closed. KCL applied to the bottom node of the circuit dictates that

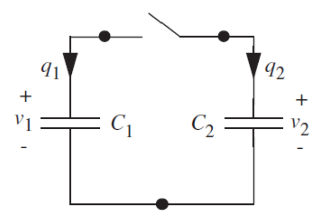


FIGURE 9.51 Two capacitors connected in parallel through a switch.

$$\frac{dq1(t)}{dt} + \frac{dq2(t)}{dt} = \frac{d}{dt}\big(q1(t) + q2(t)\big) = 0 \quad (9.98)$$

Equation 9.98 suggests that the total charge q1+q2 is constant, hence conserved for all time, even when switch closes.

Let the switch close at t = 0. For t > 0, KVL applied to the loop in figure 9.51 dictates:

$$v1(t) = v2(t) \ (9.99)$$

$$\frac{q1(t)}{C1} = \frac{q2(t)}{C2} \ (9.100)$$

Before the switch is closed, from 9.98 we have:

$$q1(t) + q2(t) = Q1 + Q2 \quad (9.101)$$

$$with\ Q1\ and\ Q2\ the\ charges\ prior\ to\ switch\ closure$$

Next, 9.100 and 9.101 can be jointly solved:

$$\boldsymbol{q1(t) = \frac{C1}{C1 + C2}(Q1 + Q2)\ (9.102)}$$

$$\boldsymbol{q2(t) = \frac{C2}{C1 + C2}(Q1 + Q2)\ (9.103)}$$

for t > 0. Substituting 9.102 and 9.103 in 9.8:

$$v1(t) = \frac{q1(t)}{C1} = \frac{Q1 + Q2}{C1 + C2}\ \ (9.104)$$

$$v2(t) = \frac{q2(t)}{C2} = \frac{Q1 + Q2}{C1 + C2}\ \ (9.105)$$

for t > 0. According to 9.102 and 9.103 the capacitors share the total charge in proportion to their capacitance. While charge is conserved during the closure of the switch, it is interesting to note that energy is not. Using 9.18, total energy stored before the switch closes is:

$$we(t < 0) = \frac{Q^2 1}{2C1} + \frac{Q^2 2}{2C2}\ \ (9.106)$$

Instead the total energy stored after the switch is closed is:

$$wE(t > 0) = \frac{(Q1 + Q2)^2}{2(C1 + C2)}\ \ (9.107)$$

Subtracting 9.107 from 9.106 the energy lost is:

$$\frac{1}{2} * \frac{C1C2}{C1 + C2}\left(\frac{Q1}{C1} - \frac{Q2}{C2}\right)^2$$

CHAPTER 10: FIRST-ORDER TRANSIENTS IN LINEAR ELECTRICAL NETWORKS

This chapter will augment our digital abstraction with the concept of delay to include the effects of capacitors and inductors. Because they can store energy, capacitors and inductors display the memory property and offer signal-processing possibilities not available in circuits containing only resistors. Apply a square-wave voltage to a multi-resistor linear circuit, and all the voltages and currents in the network will have the same square-wave shape. But include one capacitor in the circuit and very different waveforms will appear. The linear analysis techniques already developed - node equations, superposition, etc. - are adequate for finding appropriate network equations to analyze these kinds of circuits. However, the formulations turn out to be differential equations.

This chapter will discuss systems containing a single storage element, namely, a single capacitor or a single inductor. Such systems are described by simple, first-order differential equations.

10.1 ANALYSIS OF RC CIRCUITS

10.1.1 PARALLEL RC CIRCUIT, STEP INPUT

Shown in Figure 10.2a is a simple source-resistor-capacitor circuit. It can be seen as a Norton transformation applied to a more complicated circuit (with always only one capacitor though). Let us assume we wish to find the capacitor voltage vC. Using the node method we obtain (note top node has the same voltage vC):

$$i(t) = \frac{vC}{R} + C\frac{dvC}{dt} \quad (10.1)$$

Rewriting:

$$\frac{dvC}{dt} + \frac{vC}{RC} = \frac{i(t)}{C} \quad (10.2)$$

To find vC(t), we must solve a nonhomogeneous, linear first-order ordinary differential equation with constant coefficients. We will use the method of homogeneous and particular solutions explained briefly in 10.1.A below.

Skipping the mathematical details in 10.1.A, assuming that the current source i(t) is a step function i(t) = Io for t > 0 as per Figure 10.2b, assuming that voltage on the capacitor was zero before the current step was applied, we get the complete solution for t>0, as plotted in Figure 10.2c and as per equation below:

$$\boldsymbol{vC = IoR\left(1 - e^{-\frac{t}{RC}}\right)} \quad (\mathbf{10.20})$$

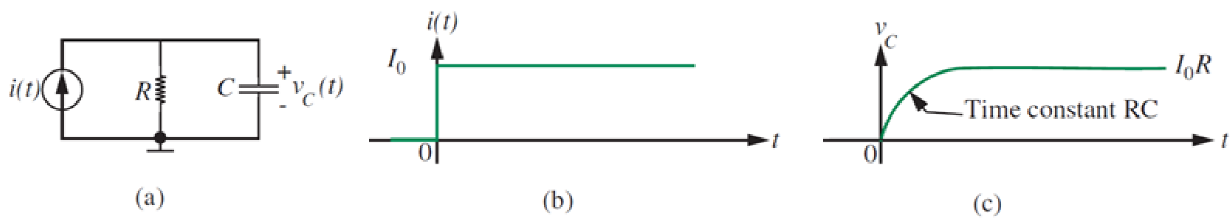


FIGURE 10.2 Capacitor charging transient.

Some observations. First, notice that capacitor voltage vC = 0 at t = 0 and reaches its final value of IoR for large t. The increase from 0 to IoR has a time constant RC. The final value of IoR for the capacitor voltage implies that all of the current from the current source flows through the resistor, and the capacitor behaves like an open circuit (for large t).

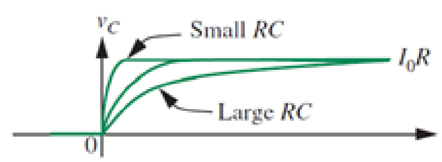


FIGURE 10.4 Significance of the RC time constant.

Second, the initial value of 0 for the capacitor voltage implies that at t = 0 all the current from the current source must be flowing through the capacitor, and none through the resistor. Thus, the capacitor behaves like an instantaneous short circuit at t = 0.

Third, the physical significance of the time constant RC can now be seen. Illustrated in Figure 10.4, it is the temporal scale factor that determines how rapidly the transient goes to completion.

10.1.A MATH: METHOD OF HOMOGENEOUS AND PARTICULAR SOLUTIONS

Let's say we need to solve $\frac{dvC}{dt}+\frac{vC}{RC}=\frac{i(t)}{C}$ for Vc (first example in 10.1.1). Assume that the current source i(t) is a step function i(t) = Io for t > 0 as per Figure 10.2b. We assume now that the voltage on the capacitor was zero before the current step was applied.

The method of homogeneous and particular solutions proceeds in three steps:

1. Find the homogeneous solution vCH. We set the driving function in the original differential equation to zero: $\frac{dvCH}{dt}+\frac{vCH}{RC}=0$. We assume a solution of the form $vCH = Ae^{st}$. By substitution we have $Ase^{st}+\frac{Ae^{st}}{RC}=0$. Discarding the trivial solution A = 0, we find $s=-\frac{1}{RC}$ which is the characteristic equation of the system. So, the homogeneous solution is of the form $vCH = Ae^{-\frac{t}{RC}}$. The product RC has the dimensions of time and is called the time constant of the circuit.
2. Find the particular solution vCP. We are looking for any solution to $Io=\frac{vCP}{R}+C\frac{dvCP}{dt}$. Since Io is constant in time for t>0, one solution is also a constant vCP = K. Substituting K in the equation we find the particular solution $vCP = IoR$.
3. The total solution is then the sum of the homogeneous solution and the particular solution. Use the initial conditions to solve for the remaining constants. So in our case the total solution is $vC = Ae^{-\frac{t}{RC}}+IoR$. We need to evaluate the constant A. For t < 0 we know vC = 0 and for t > 0 our found solution is valid. These two parts of the solution are patched together by a continuity condition derived from Equation 9.9: An instantaneous jump in capacitor voltage requires an infinite spike in current, so for finite current, the capacitor voltage must be continuous. This circuit cannot support infinite capacitor current (because i(t) is finite, the infinite current would have to come from the resistor, and this is impossible). Thus we are justified in assuming continuity of vC, hence can equate the solutions for negative time and positive time by solving at t = 0. So $0 = A + IoR$, and so $A=-IoR$. And thus, the complete solution is: $vC=-IoRe^{-\frac{t}{RC}}+IoR$, plotted in Figure 10.2c

10.1.2 RC DISCHARGE TRANSIENT

With the capacitor now charged, assume that the current source is suddenly set to zero as suggested in Figure 10.5a. The relevant circuit to analyze the RC turn-off or discharge transient is in Figure 10.5c. The voltage on the capacitor vC = IoR at t < 0.

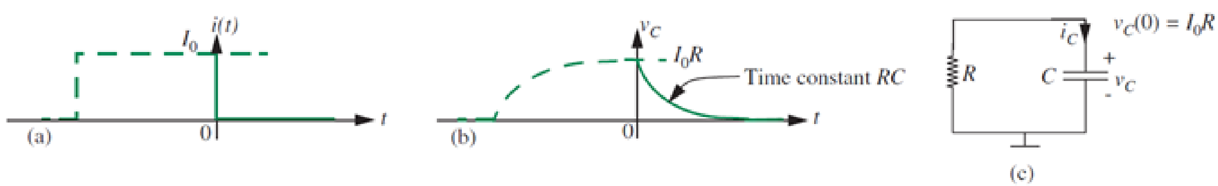


FIGURE 10.5 RC discharge transient.

Because the drive current is zero, the differential equation for t > 0 is now

$$0=\frac{vC}{R}+C\frac{dvC}{dt}$$

The homogenous solution is the same as before, $vCH = Ae^{-\frac{t}{RC}}$ but now the particular solution is zero, so the homogenous solution is the total solution. Since at t = 0, vC = IoR, it follows that A = IoR at t = 0. So the capacitor voltage waveform for t > 0 and sketched in Figure 10.5b, is

$$\boldsymbol{vC = IoRe^{-\frac{t}{RC}}} \quad \mathbf{(10.25)}$$

In general, for a resistor and capacitor circuit with an initial voltage vC(0) on the capacitor, the capacitor voltage waveform for t > 0 is

$$\boldsymbol{vC = vC(0)e^{-\frac{t}{RC}}} \quad \mathbf{(10.26)}$$

10.1.3 SERIES RC CIRCUIT, STEP INPUT

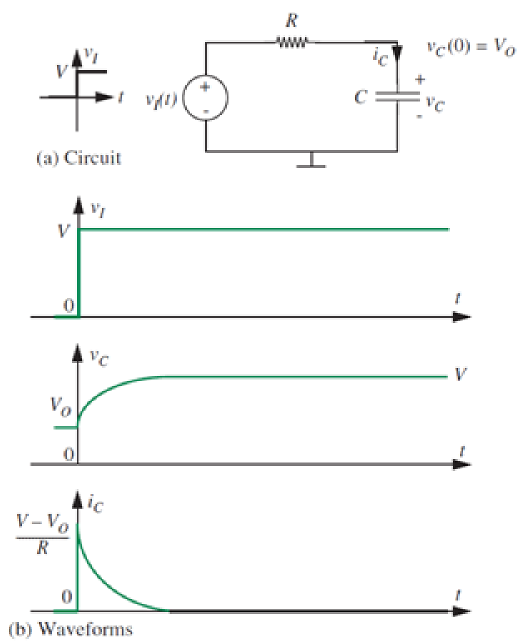


FIGURE 10.7 Series RC circuit with step input.

Let's now work with a Thevenin circuit, as in Figure 10.7 and determine the capacitor voltage as a function of time. The input waveform vS is assumed to be a voltage step of magnitude V applied at t = 0, but this time we assume the capacitor voltage is VO just before the step: vC = VO at t < 0.

Node method to find the differential equation: we apply KCL at the node with voltage vC, we get $\frac{vC - vI}{R} + C\frac{dvC}{dt} = 0$, which is the same as $\frac{dvC}{dt} + \frac{vC}{RC} = \frac{vI}{RC}$. We solve the differential equation as usual (page 514), and find the complete solution for the capacitor voltage for t > 0 as

$$\begin{aligned} \boldsymbol{vC} &= \boldsymbol{V + (VO - V)e^{-\frac{t}{RC}}} \\ &= \boldsymbol{VOe^{-\frac{t}{RC}} + V\left(1 - e^{-\frac{t}{RC}}\right)} \quad \mathbf{10.36} \end{aligned}$$

Finally, from equation 9.9, the current through the capacitor is

$$\boldsymbol{iC = C\frac{dvC}{dt} = \frac{V - VO}{R}e^{-\frac{t}{RC}}} \quad \mathbf{(10.38)}$$

This expression for iC also matches our expectations since iC must be 0 when t ≫ 0, and since the capacitor behaves like a voltage source with voltage VO during the step transition at t = 0, the current at t = 0 must equal (V - VO)/R. These waveforms are shown in Figure 10.7b.

In case we want the voltage vR across the resistor, by KVL vR = vI – vC, or vR = iC*R.

As one final point of interest, notice that Equation 10.36 was derived assuming both an initial nonzero state (VO) and a nonzero input (a step of voltage V). Substituting V = 0 in the complete solution, we obtain the so called zero input response (ZIR - the response for nonzero initial conditions, but where the input drive is zero):

$$vC = VOe^{-\frac{t}{RC}} \quad (10.39)$$

and substituting VO = 0 in Equation 10.36 we obtain the zero-state response (ZSR):

$$vC = V - Ve^{-\frac{t}{RC}} = V(1 - e^{-\frac{t}{RC}}) \quad (10.40)$$

Notice also that the total response is the sum of the ZIR and the ZSR.

10.1.4 SERIES RC CIRCUIT, SQUARE-WAVE INPUT (just FYI)

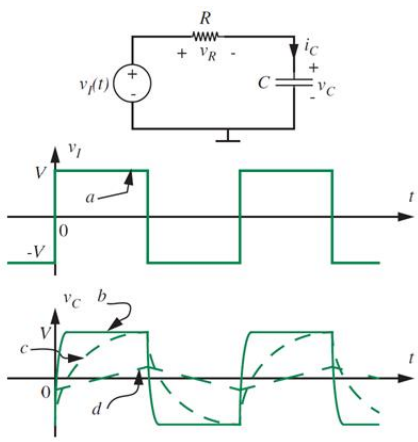


FIGURE 10.8 Response to square wave.

Briefly, in the case we apply a square wave input to a series RC circuit, the transient will not go to completion if the circuit time constant is much longer than the square wave period, as seen for curve d in figure 10.8. So, for large RC, the capacitor voltage is approximately the integral of the input voltage wave. We can see it from the differential equations describing the circuit: by applying KVL we have vI = iC*R+vC, and substituting for the capacitor element law we get $vI = RC\frac{dvC}{dt} + vC$. It can be seen that as the circuit time constant becomes bigger, the capacitor voltage vC must become smaller. Integrating both sides in the equation just written, and ignoring vC since it is small for big RC, we get $vC \approx \frac{1}{RC}\int vI\, dt + K$ where K = 0. So for large RC, the capacitor voltage is roughly the integral of the input voltage.

This behavior is crucial in signal processing. When RC is large, the circuit acts like an integrator. This means the output voltage is proportional to the integral of the input. Instead of copying the square input, it smooths it out, which is handy for filtering noise or creating different wave shapes (like turning sharp signals into smooth ones).

10.2 ANALYSIS OF RL CIRCUITS

10.2.1 SERIES RL CIRCUIT, STEP INPUT

Figure 10.9 will serve as a simple illustration of a transient involving an inductor. The input waveform vS is assumed to be a voltage step applied at t = 0 (see Figure 10.9a), and the inductor current is assumed to be zero just before the step: iL = 0 for t < 0. Suppose we want to solve for the current iL. We can follow the exact same approach as before (page 518).

The complete solution for our circuit for t > 0 is:

$$\boldsymbol{iL = \frac{V}{R}\left(1 - e^{-\frac{R}{L}*t}\right)} \quad (\mathbf{10.59})$$

And from equation 9.28, the voltage across the inductor is:

$$\boldsymbol{vL = L\frac{diL}{dt} = Ve^{-\frac{R}{L}*t}} \quad (\mathbf{10.60})$$

These waveforms are shown in Figure 10.9b.

Notice that the inductor current has an initial value of 0 and a final value of V/R. Thus, the inductor behaves like an instantaneous open circuit at t = 0 and a short circuit for t ≫ 0.

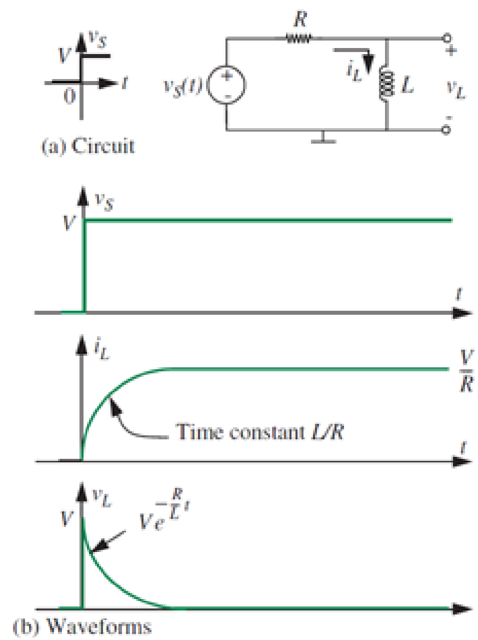


FIGURE 10.9 Inductor current buildup.

10.3 INTUITIVE ANALYSIS

The several examples with step-function drive that we worked previously suggest that such circuits have a very limited range of solutions. It turns out that for simple excitations, such as the step and the impulse, the response of first-order systems can be sketched easily using some intuition.

Let us illustrate using the step response of a series RC circuit in Figure 10.11a as an example. We will address the most general case, namely one in which there is both a nonzero initial state and a nonzero input. Initial voltage on the capacitor is VO. At t = 0 there is an input step voltage of magnitude V. For the purposes of sketching our result, we assume V > VO. As in Figure 10.11a, switch S1 is initially closed and S2 is open, resulting in the voltage VO being applied directly across the capacitor. Just before t = 0, that is, at $t = 0^-$, S1 is opened (S2 remains open). Then, at t = 0, S2 is closed (S1 remains open). The closing of S2 and opening of S1 results in a series RC circuit with a step voltage V applied at t = 0.

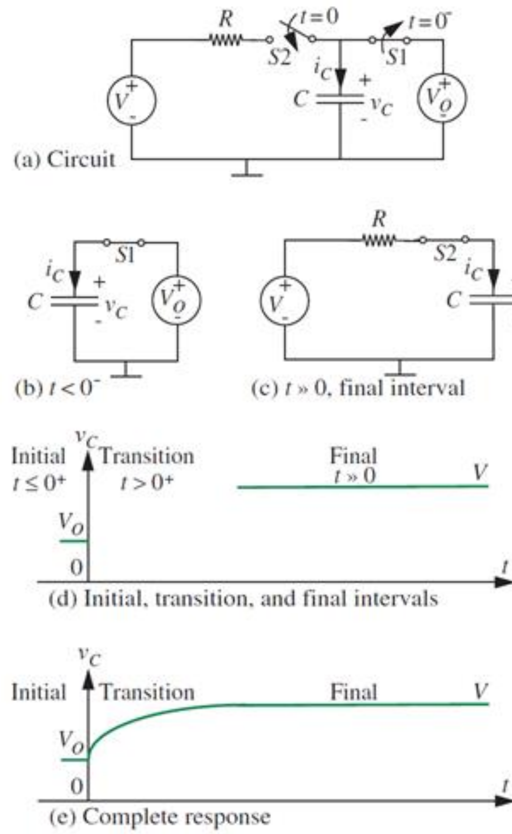


FIGURE 10.11 Step response of series RC circuit. The arrangement of switches provides for the initial voltage V_O on the capacitor, and an input step voltage of magnitude V at $t = 0$.

Suppose we are interested in sketching the voltage vC as a function of time. The form of the response can be sketched intuitively by identifying three intervals of operation as indicated in Figure 10.11d. The overall response can be quickly sketched through inspection by first determining the initial and final interval values of the voltage on the capacitor.

1. Initial Interval ($t \leq 0^+$). The capacitor voltage is VO during $t = 0^-$. Next, notice that in the short period of time between $t = 0^-$ and t = 0, and still within the initial interval, the capacitor is not connected to any other circuit (S1 is opened at $t = 0^-$ and S2 is closed immediately at t = 0). Assuming the capacitor is ideal, it holds its charge and so its voltage remains at VO until the switch S1 is closed. At t = 0, S1 is closed, resulting in a finite step of magnitude V being applied to a series RC circuit in which the capacitor has a voltage VO across it. Let us now determine the capacitor voltage at $t = 0^+$ just after the step. From the element law of the capacitor (Equation 9.7), we know that an instantaneous jump in capacitor voltage requires an infinite spike (that is, an impulse) in current. Since a finite step voltage applied across a resistor cannot support an infinite spike in current, we conclude that the capacitor voltage cannot change instantaneously, rather it must be continuous. Thus, the voltage across the capacitor at $t = 0^+$ must also be VO. The voltage across the capacitor during the initial interval $t \leq 0^+$ is sketched in Figure 10.11d.
2. Final Interval ($t \gg 0$). Observe here our situation is identical to the circuit in Figure 10.11c. The capacitor, when all transients have died out, has reached its final state, and hence the capacitor

current is zero (behaving like an open circuit). Thus, to satisfy KVL, the capacitor voltage must be equal to V. This value is sketched in figure 10.11d.

3. Transition Interval $t > 0^+$. During this interval, observe that the capacitor voltage cannot jump instantaneously from VO to V due to the continuity condition. Specifically, we know from the solution to the homogeneous equation for the RC circuit that the transient follows an exponential form, either rising $(1 - e^{-\frac{t}{RC}})$ or falling $\left(e^{-\frac{t}{RC}}\right)$ with time constant RC. (For the corresponding inductor-resistor circuit the time constant will be L/R). In our case, since V > VO, the transient will be a rising exponential.
4. Complete response. This is sketched in Figure 10.11e. The corresponding equation
$$vC = V + (VO - V)e^{-\frac{t}{RC}} \ (same\ as\ 10.36)$$
In other words, for t ≥ 0
$$\boldsymbol{vC = final\ value + (inital\ value - final\ value)e^{-\frac{t}{time\ costant}} \ \ (10.61)}$$
Re-arranging terms:
$$vC = VOe^{-\frac{t}{RC}} + V\left(1 - e^{-\frac{t}{RC}}\right) \ (10.62)$$

Sometimes we want the response related to the capacitor current. The response can be determined from the voltage response and the element law of the capacitor. Or the current can be directly obtained with the same process just used above to get the voltage response (see page 524). An interesting fact about this example with current as the unknown to find is that the current is not necessarily continuous. So the current would jump. I wonder then, a step function is not suitable to model this, a impulse is required. (elaborate)

Inductors can be treated in a similar manner, with the difference that the state variable for an inductor is its current. So the inductor current is continuous, instead the voltage is not (recall, from Equation 9.26, an instantaneous jump in inductor current requires an infinite spike, that is, an impulse, in the voltage). More on page 254.

Intuitive analysis for impulses is discussed in 10.6.4.

From ChatGPT (about impulses and step functions):

For capacitors, the voltage does not make any sudden jumps, because doing so would demand an infinitely large current. Meanwhile, the current can hop around right away if the circuit conditions change abruptly.

- Voltage is continuous on a capacitor: no sudden leaps without infinite current.
- Current can change suddenly: at t=0, it jumps to (V−V0)/R, then settles down to zero over time.

Because for a capacitor, the voltage across it is tied to the charge on its plates, Q=C·V. If the voltage tried to jump instantly, you'd need an infinite current to change the charge that fast - impossible in a real circuit. So, the capacitor's voltage just can't make a sudden leap; it stays continuous. However, the current can hop around if the voltage changes abruptly.

For an inductor, it's the opposite because it stores energy in its current:

How does this relate with step functions and impulses?

Steps become impulses when you differentiate:

- A step in voltage across a capacitor is like a vertical jump in V(t). The capacitor current is $iC = C\frac{dV}{dt}$
- The derivative of a step is an impulse (a super-spike). Meaning if the capacitor's voltage literally stepped up instantly, the math implies an infinite current spike—impractical in the real world.

For a capacitor, the current is tied to the change (derivative) of voltage. So a sudden voltage step would be an impulse current. Opposite for inductors.

In short, step changes in one quantity (voltage for capacitors, current for inductors) imply an impulse in the other. Physically, we can't really supply infinite current or infinite voltage, so these theoretical "impulses" remind us that capacitor voltage (or inductor current) can't instantly jump in a real circuit.

Now the question is why we choose voltage and current to have these roles for capacitors and inductors?

Device Physics Dictates Which Quantity "Must" Be Continuous

- A capacitor stores energy in its electric field; mathematically, Q=C*V. That makes the voltage its "state"—the physical storage variable—so the voltage can't jump without requiring an infinite current

So if I understand correctly is all about the element law for the devices.

10.4 PROPAGATION DELAY AND THE DIGITAL ABSTRACTION

The RC effects we have seen thus far are the source of delays in digital circuits.

Consider the two-inverter digital circuit shown in Figure 10.14-10.15 in which inverter A drives inverter B.

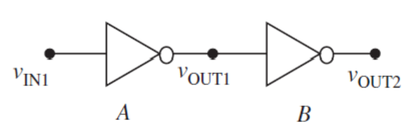


FIGURE 10.14 Inverters connected in series.

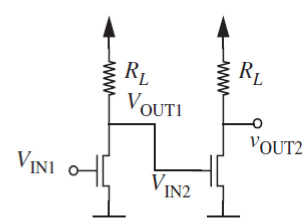


FIGURE 10.15 Internal circuits of the inverters.

Ideally, the input vIN1 (corresponding to a sequence of 1's and 0's of the form shown in Figure 10.16) should produce the ideal output vOUT1 (ideal). However, in practice, if we were to observe the output vOUT1 on an oscilloscope, we would notice the output changes from one valid voltage level (i.e. a logical 0) to another valid voltage level (i.e. a logical 1) more slowly over a small period of time as suggested by the signal marked vOUT1 (actual) in Figure 10.16.

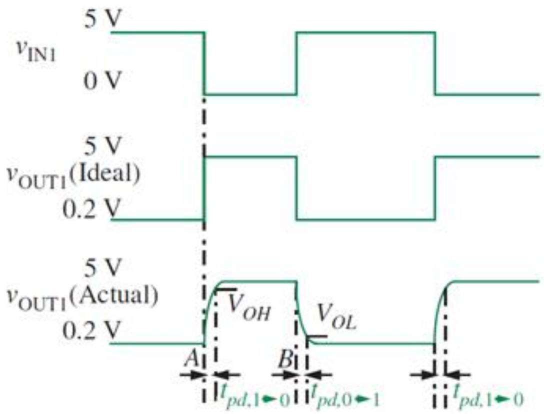


FIGURE 10.16 Characteristics of ideal and actual inverters.

How does this slow transition affect the behavior of the digital circuit? Recall that the vOUT1 signal represents a digital signal, so it must reach VOH so that the gate that produced it adheres to the static discipline and we obtain a nonzero noise margin. As suggested in the lowermost signal in Figure 10.16, notice that the VOH crossing happens at a time interval $t(pd), 1 \rightarrow 0$, after the input changes from 1 to 0. Thus, effectively, there is a delay of $t(pd), 1 \rightarrow 0$ between the moment that the input changes to 0 to the moment that the output changes to a valid 1.

This period of time is called the propagation delay through inverter A for a 1 to 0 transition at the input and is denoted as $t(pd), 1 \rightarrow 0$. The inverter is also later characterized by a 0 → 1 propagation delay denoted as $t(pd), 0 \rightarrow 1$. The two delays are not necessarily equal. By simplicity we characterize digital gates by one single propagation delay, *t(pd)*, being the max of the two.

10.4.2 COMPUTING *t(pd)* FROM THE SRC MOSFET MODEL

Let's calculate the propagation delay for the SRC MOSFET model in figure 10.18.

FIGURE 10.18 Internal circuits of the inverters.

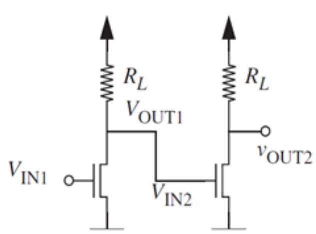


FIGURE 10.19 SRC circuit model of inverters connected in series when the input is high.

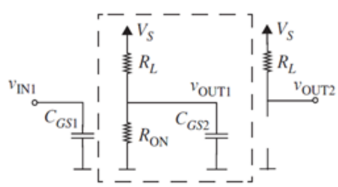
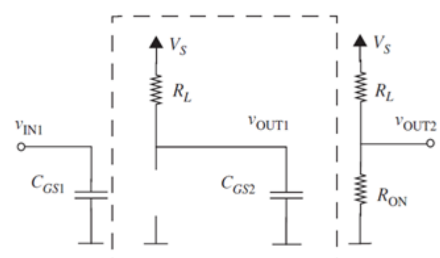


FIGURE 10.20 SRC circuit model of inverters connected in series when the input is low.



When alternating logical 1's and 0's are applied to the input to the inverter pair, and the inverters are allowed to reach steady state after each transition, the equivalent circuit model alternates between the two circuits in Figures 10.19 and 10.20.

Imagine vIN1 has been high for a long time; focus on the dashed part of the circuit in figure 10.19. Since the circuit is in its steady state, the capacitor behaves like an open circuit, and the voltage across the capacitor will be established by the voltage divider subcircuit. Assuming RL ≫ RON, the capacitor's voltage will be close to zero volts.

Next, focus on the time instant when the input voltage vIN1 switches from a high to a low value, turning the first MOSFET off. At this transition instant, as we said, the voltage across CGS2, will be initially close to 0 V. This is depicted as the time instant A in Figure 10.16. After the first MOSFET turns off, Figure 10.20 applies. Focus again on the part of the circuit bound by the dashed box. It is easy to see that the circuit inside the dashed box is a first-order RC circuit. Remember, the voltage across CGS2 is low initially. Now, VS begins to charge CGS2 through the resistor RL. The equivalent RC circuit for the devices in the box is shown in Figure 10.21. As the capacitor charges up, the output voltage of inverter A rises. This voltage must rise above the valid logical output high threshold, namely VOH, to satisfy the static discipline. Notice that although the second MOSFET will turn on when vOUT1 crosses its VT threshold, we require vOUT1 to reach VOH to achieve a modest noise margin. Notice that the presence of the capacitor CGS2 makes vOUT1 take a finite amount of time to rise to the required VOH level. As we saw before, this interval of time is called the propagation delay for the inverter for a high to low transition at the input and is denoted by $t(pd), 1 \rightarrow 0$.

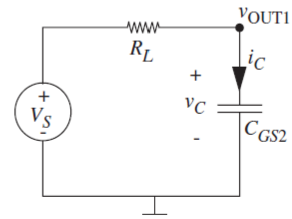


FIGURE 10.21 Equivalent circuit when C_{GS2} is charging.

Next, let us consider the time instant when the input voltage switches from 0 to 5 volts, turning the first MOSFET on. Let us assume that the 0 to 5 V transition happens after a sufficiently long period of time so that CGS2 is initially charged up to its steady state value of 5 V. When the first gate is turned on, CGS2 begins to discharge. The RC circuit and its Thévenin equivalent for the discharge is shown in Figure 10.22. For the logical 0 to logical 1 transition at the input to be reflected at the output of inverter A, the voltage across CGS2 needs to go below the valid logical output low threshold, VOL. As before, although MOSFET B will turn off when vOUT1 drops below 1 volt, we require the output to go below VOL to provide for an adequate noise margin. The time interval corresponding to the output capacitor discharge for an inverter is denoted by $t(pd), 1 \rightarrow 0$.

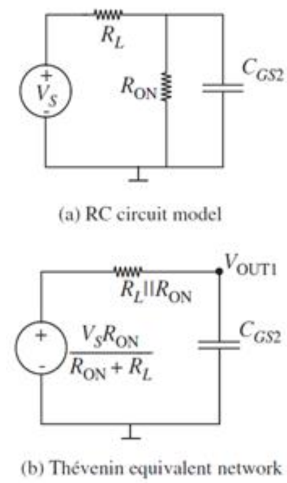


FIGURE 10.22 Equivalent circuit when C_{GS2} is discharging.

For convenience, we usually characterize devices with a single $t(pd)$ without defining their surrounding environment, so this simple device model can be used to obtain quick estimates of digital circuit delays when multiple gates are connected together. In reality though, $t(pd)$ will change based on some conditions, for example the size of the capacitance.

<u>Computing $t(pd), 0 \rightarrow 1$</u>

Let's apply what is above to compute the propagation delay for a low to high transition at the input of the inverter. Assumptions: V(OL)=1V, V(OH)=4V, R(ON)=1KΩ, VT=1V, RL=10KΩ. We need to calculate the time taken for the capacitor CGS2's voltage to drop from 5 to V(OL)=1V. Figure 10.22 shows the equivalent circuit when input is high. Using the node method, we find the differential equation that we need to solve. To find the time t it takes for the voltage across the capacitor to drop below 1V, we must solve for the value of t that satisfies:

$$Vth + (Vs - Vth)e^{-\frac{t}{Rth * Cgs2}} < 1$$

Rearranging:

$$t > -Rth * Cgs2 * \ln\left(\frac{1 - Vth}{Vs - Vth}\right)$$

with Rth = RL||RON and $Vth = Vs\frac{R(ON)}{R(ON)+RL}$. In our case substituting values, we get 0.2 nanoseconds. See page 534 for full example.

<u>Computing $t(pd), 1 \rightarrow 0$</u>

When the input vIN1 goes low, the circuit model that applies is shown in Figure 10.21. In this case, we know that initial voltage vC0 on the capacitor is determined by the voltage-divider relationship just seen above. Our goal is to solve for the time it takes for the capacitor to charge up to V(OH) = 4 volts. We follow the procedure outlined above, use the node method to find the differential equation, solve it and then to determine the delay we solve for t, which must be:

$$t > -RL * CGS2 * \ln\left(\frac{11}{50}\right) \quad (10.73)$$

Resulting in 1.5 nanoseconds for our example. Page 534 for full example.

Notice the RL*CGS2 factor in Equation 10.73. In typical circuits, a ballpark estimate of the delay can be obtained by simply taking the product of the capacitance and the effective resistance through which it charges.

<u>Computing $t(pd)$</u>

We use the greater of the two calculated above.

Worth a look at example 10.1 about how wire length affects the propagation delay on a set of inverters. Page 535.

10.5 STATE AND STATE VARIABLES

10.5.1 THE CONCEPT OF STATE

As we said in chapter 9, if for a capacitor we have $q(t2) = \int_{t1}^{t2} i(t)dt + q(t1)$, all the relevant past history of the circuit prior to t1 is summarized in q(t1). Variables that have this property are called state variables. Thus, if we know the value of the state variable at a certain time, and the value of the input variable thereafter, we can find the value of the state variable for any subsequent time. For linear time-invariant capacitors, the capacitor voltage is also a state variable, because q=Cv.

The first-order differential equations for the RC and RL circuits can all be written as state equations

$$\frac{d}{dt}(state\ variable) = f(state\ variable, input\ variable) \quad (10.80)$$

For the linear case, f is a linear function, so Equation 10.80 becomes

$$\frac{d}{dt}(state\ variable) = K1(state\ variable\ present\ value) + K2(input\ variable)\ \ (10.81)$$

Which means that for example, $\frac{dvC}{dt} + \frac{vC}{RC} = \frac{i(t)}{C}$ can be written in the canonic state equation form of Equation 10.81 as $\frac{dvC}{dt} = -\frac{vC}{RC} + \frac{i(t)}{C}$ (10.82) where the only state variable is vC. Making a parallel with equation 10.81, $K1 = -\frac{1}{RC}$ and $K2 = \frac{1}{C}$.

10.5.2 COMPUTER ANALYSIS USING THE STATE EQUATION

Quick intro: if I know the initial state and the future input signal, then I can know the entire waveform of vC even for complex (non-linear) equations. Here a method to estimate this waveform is introduced that can be used on a computer, since it implies calculating each small value for vC and thus finding the entirety of the waveform. Not sure which equations require this method though. Anyway, just take the method in 10.83 as correct.

One advantage of the state equation formulation is that even in the nonlinear case, the equations can be readily solved on a computer (just a note, somehow should be all the equations we have seen so far are linear, indeed they are called first order linear differential equations). If the input signal and the initial value of the state variable are known, then the slope of the state variable, that is,

$$\frac{d}{dt}(state\ variable)$$

can be found from Equation 10.81. The value of the state variable at time t+Δt can now be estimated by standard numerical methods (Euler's method, Runge-Kutta, etc.). The process can now be repeated until the entire waveform is found.

Continuing with our example in Equation 10.82, suppose that the input signal i(t) is known for all time. Also, suppose that the value of the state variable at time t = t0, namely vC(t0), is known. Then, Euler's method approximates the value of the state variable at time t = t0 + Δt as

$$vC(t0 + \Delta t) = vC(t0) - \frac{vC(t0)}{RC}\Delta t + \frac{i(t0)}{C}\Delta t\ \ \ (10.83)$$

Subsequent values of vC can be determined using the same process.

By choosing small enough values of Δt, a computer can determine the waveform for vC(t) to an arbitrary degree of accuracy. This process illustrates the fact that the initial state contains all the information that is necessary to determine the entire future behavior of the system from the initial state and the subsequent input.

This procedure works even for circuits with many capacitors and inductors, linear or nonlinear, because these higher-order circuits can be formulated in terms of a set of first-order state equations like Equation 10.80, one for each energy-storage element (with an independent state variable) in the network.

Chapter 12 discusses such an example in Section 12.10.1.

10.5.3 ZERO-INPUT AND ZERO-STATE RESPONSE

Another advantage of the state variable point of view is that it allows us to solve transient problems by **superposition**.

1. We find first the zero input response, the response for the true initial conditions, with the input drive zero.
2. Then we find the zero-state response, the response of the circuit when the initial state is zero, that is, all capacitor voltages and inductor currents are initially zero.
3. The total response is the sum of the zero-input response (ZIR) and the zero-state response (ZSR).

Relating these ideas to Equation 10.81 and following the three steps just outlines:

1. Finding the zero-input-response involves solving the equation:
$$\frac{d}{dt}(state\ variable) = K1(state\ variable\ present\ value)\ (10.84)$$
using the true initial conditions for the state variable.
2. Finding the zero-state response involves solving the equation:
$$\frac{d}{dt}(state\ variable) = K1(state\ variable\ present\ value) + K2(input\ variable)\ (10.85)$$
with the initial value of the state variable set to 0.

I strongly suggest to check example on page 541 in yellow where a problem is solved first by using the normal method we have seen at the beginning of this chapter, and then the same solution is found using the method explained here by superposition.

Basically, I divide the problem in two, I first calculate ZIR considering that there is a voltage supply V1 for t<0, but for t≥0 the input is zero. For the ZIR, the particular solution will by definition be zero. Hence all ZIRs will be homogeneous solutions. Then I calculate ZSR considering that for t<0 the input is zero, and for t≥0 the input is some value V2. This method still requires solving differential equations.

A major advantage to the state variable point of view is that any arbitrary ZIR can be added to any ZSR solution we have already worked out.

10.6 ADDITIONAL EXAMPLES

10.6.1 EFFECT OF WIRE INDUCTANCE IN DIGITAL CIRCUITS

See book, page 545

10.6.2 RAMP INPUTS AND LINEARITY

See book, page 545

10.6.3 RESPONSE OF AN RC CIRCUIT TO SHORT PULSES AND THE IMPULSE RESPONSE

See book

10.6.4 INTUITIVE METHOD FOR THE IMPULSE RESPONSE

See book

10.6.5 CLOCK SIGNALS AND CLOCK FANOUT

See book

10.6.6 SERIES RL CIRCUIT WITH SINE WAVE INPUT

See book

10.7 DIGITAL MEMORY

This chapter demonstrated previously that the memory aspect of capacitors and inductors formalized using the notion of state variables provided many uses in the analog domain. The same memory property can also be utilized in the digital domain to implement digital memory using the analogous concept of digital state.

10.7.1 THE CONCEPT OF DIGITAL STATE

Memory enables short live external inputs to be available to system circuitry for longer period of times. Basically, I can store a variable in memory and read it or use it any time I desire.

Future results depend only on the value stored in memory and future inputs. Future results do not depend on how the memory was time sequenced, just its final state (for example if I store 6 in memory I do not care if it came from 3+3 or 3*2, I care only about the number 6). This observation stems from the concept of a ‘‘state variable’’ that we saw earlier. The value stored in memory is simply a digital state variable in a manner analogous to an analog state variable value stored on a capacitor.

10.7.2 AN ABSTRACT DIGITAL MEMORY ELEMENT

Figure 10.50 shows an abstract memory element that can store one bit of data. The input dIN is copied into memory when the store signal is high. The value stored in the memory is available to be read as the output dOUT. If no new value is written into the memory, the last written value is stored indefinitely.

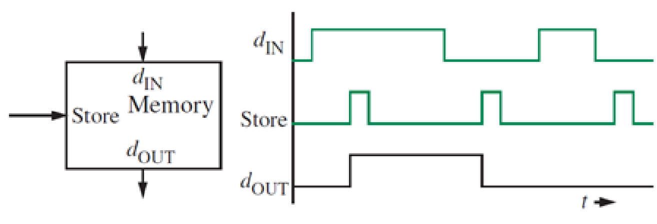


FIGURE 10.50 An abstract one-bit memory element.

10.7.3 DESIGN OF THE DIGITAL MEMORY ELEMENT

The memory element must be designed so that it stores indefinitely any value that has been written into it. Recall that a capacitor has the same property. Provided its discharge path has a high time constant (RC), a capacitor can store a charge for a long period of time. Furthermore, we can use a switch to enable charging the capacitor from a given input.

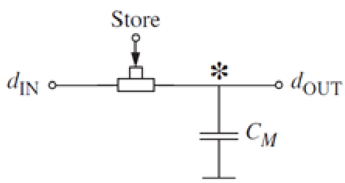


FIGURE 10.52 Circuit implementation of a memory element.

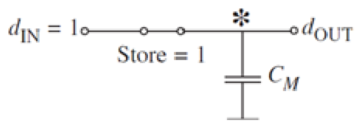


FIGURE 10.54 Charging up the memory capacitor, when the store signal is high.

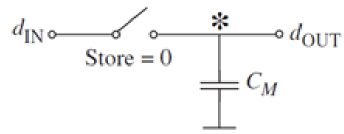


FIGURE 10.56 Charge storage in the memory capacitor, when the store signal is low.

Based on this intuition, consider the simple memory element circuit comprising a capacitor and an ideal switch shown in Figure 10.52. Analyzing the circuit:

- When the switch is ON (circuit in Figure 10.54) the capacitor charges (or discharges) to the value of the input voltage at the dIN terminal.
- When the switch is OFF (circuit in Figure 10.56), the dOUT terminal of the capacitor begins to float and the charge previously deposited on the capacitor is held in place. In the ideal case, if the resistance between the dOUT terminal and GND is infinite, the capacitor will hold the charge forever.

The waveforms shown in Figure 10.50 will apply to the memory element circuit shown in Figure 10.52 under the following idealized assumptions:

- When the store signal is high, the RC time constant associated with the capacitor circuit is negligible, and when the store signal is low, the RC time constant associated with the capacitor circuit is infinite. What this means is that when the store signal is high, the capacitor will charge/discharge instantaneously, instead when it is low, will hold the charge forever. The RC time constant of the circuit when the store signal is high is given by the product of CM and the sum of the on resistances of the switch and the driving element. Similarly, the RC time constant of the circuit when the store signal is low is given by the product of CM and the resistance seen by the capacitor.

The last problem with our memory element circuit is that the static discipline requires that in order to obtain positive noise margins, the voltage threshold requirements on the outputs of gates must be more stringent than those on the inputs (VIH had to be restored to VOH, where VOH > VIH). To inter-operate digital memory elements with our digital gates, we require that our digital memory elements satisfy the

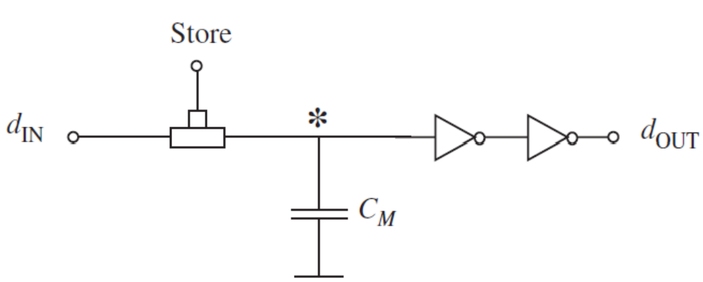


FIGURE 10.57 Circuit implementation of a signal restoring memory element.

same set of voltage thresholds. But our digital memory circuit as described in Figure 10.52 is non-restoring. As per Figure 10.57, a simple modification of our memory circuit can make it restoring. This design adds a pair of series connected inverters (a buffer) to the output of our previous memory element circuit. The buffer will restore a VIH voltage on the capacitor terminal to a VOH voltage at the dOUT output. Interestingly, when a buffer is included in the memory element circuit, we do not need to implement a special capacitor to hold charge. Rather, the gate capacitance of the buffer, CGS, forms the memory capacitance CM.

By isolating the capacitor from the circuit that reads the stored value, the buffer offers added advantages. Devices that read the value stored on the capacitor might have relatively low resistances associated with them, thereby discharging the capacitor in our original unbuffered memory circuit. In contrast, the buffered design of the memory element circuit protects the capacitor's charge from the external circuit. By careful design of the memory element, the input resistance of the buffer can be made to be very large, thereby ensuring a large discharge time constant.

In practice, capacitors will leak their charge over time due to parasitic resistances (suppose it is RP). In this situation, for how long will the value stored in the capacitor remain valid after the store signal is de-asserted? There are two cases to consider:

1. If a 0 is stored on the capacitor, then the 0 value will be held indefinitely even with a low parasitic resistance.
2. If a 1 is written on the capacitor (let's assume that the voltage corresponding to a 1 is VS), the value stored on the capacitor will be read as a valid 1 by the buffer until it reaches the voltage threshold V(IH). Thus, the period over which the memory element will store a valid 1 is the interval over which the voltage drops from VS to VIH. We can compute this duration from capacitor discharge dynamics (for example, see equation 10.26). When a capacitor CM charged to an initial voltage VS discharges through a resistor RP, its voltage vC as a function of time is given by the following equation:

$$vC = VSe^{-\frac{t}{RpCm}}$$

The time taken for vC to drop from VS to VIH is given by:

$$tVS \to VIH = -RpCm * ln\frac{VIH}{VS}$$

10.7.4 A STATIC MEMORY ELEMENT

The one-bit memory element that we have discussed thus far is called a dynamic one-bit memory element or a dynamic D-latch. It is dynamic in the sense that it stores a value written into it only for a finite amount

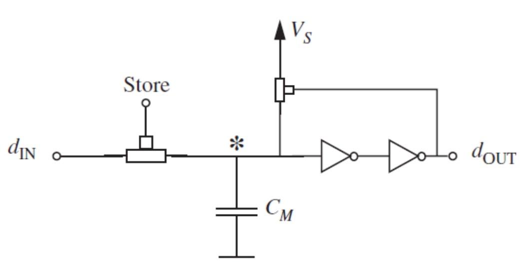


FIGURE 10.61 Circuit implementation of a static memory element using a trickle switch.

of time (due to nonzero parasitic resistances in practical implementations). The static one-bit memory element or a static D-latch is another type of memory element that has the same logic properties as the dynamic D-latch, but can store a value written into it indefinitely. Figure 10.61 shows one possible circuit for a static memory element.

YOUTUBE MIT COURSE

The Youtube course lesson one at this link.

LESSON 8 – 9 – AMPLIFIERS AND MOSFET

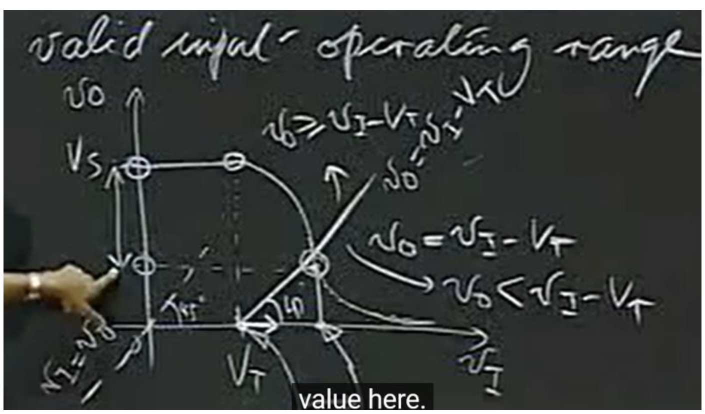
Lecture 8 focuses on the amplifier. After the first 5 minutes of recap of previous lessons, at min 5:30 we think about why we need amplifiers, which is noise immunity for analog circuits and provision of power gain for digital circuits. From min 19 to 27 we talk about about dependent sources. From minute 27 to the end we talk about the amplifiers.

Lesson 9 reviews previous lessons up until min 10, then till min 21 introduces MOSFET gradually going through the models like the switch model, the SR model. For the remaining part of the video we talk about the SCS model.

LESSON 9B – LARGE SIGNAL ANALYSIS

Introduction of lesson up until min 3:30. Min 3.30 to 8:15 review of the Mosfet and saturation state. From 8:45 introduction of large signal analysis, which involves two steps:

1. Write down the transfer function of our circuit (VO vs VI) – This is easy we have done before.
2. Find out the valid input operating range to keep saturation mode and the corresponding out range.

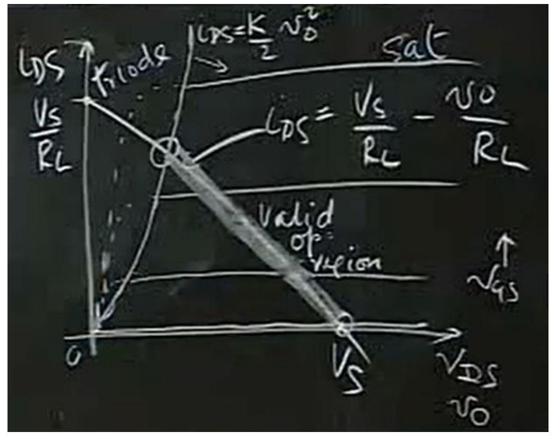


For step 2, to build intuition, the graph represents the transfer function $VO = f(VI)$. We need to find the two points where the MOSFET exits saturation.

The point where the MOSFET exits cutoff and goes in saturation is easy to find, being VI = VT and VO = VS. The other point between saturation and the triode region, is the point in the graph where the two lines, having equations VO = VI - VT and $VO = VS - \frac{K}{2}(VI - VT)^2 RL$ intersect. We need to mathematically find the intersection point.

So, the valid input range is VI: VT → $VT + \frac{-1+\sqrt{1+2K*RL*VS}}{K*RL}$ and the corresponding output range is VO: VS → $\frac{-1+\sqrt{1+2K*RL*VS}}{K*RL}$.

Another process to get the same result is by using the IDS vs VDS relation instead of the VI vs VO transfer function is shown in the graph below:



The conditions to be in saturations are VI ≥ VT and VO ≥ VI – VT. This last constraint could be related to the current as $IDS \ \leq \frac{\mathrm{K}}{2} VO$.

The two equations to solve together are $IDS = \frac{K}{2} VO^2$ and $IDS = \frac{VS}{RL} - \frac{VO}{RL}$ which represent the two lines in the graph.

Large signal analysis applies to many different circuits and devices (for example BJT) and the range can be found with a similar analysis as the one just done with the MOSFET.

From min 43 to the end, the DC voltage offset and the small signal analysis that follows are introduced.

LESSON 10 – AMPLIFIER: SMALL SIGNAL ANALYSIS

Review of previous lessons up until 7:20. From 7:30 starts explaining the small signal analysis and through math and a simplification we get to the equation vo = - RL*K(VI-VT)*vi, meaning the output voltage for a small perturbation vi on top of a voltage offset VI.

Another way to get the same result is to take the derivative (so the slope) of the function $VO = VS - \frac{K*RL}{2}(VI - VT)^2$, and multiply it by the small perturbation vi.

LESSON 11 – SMALL SIGNAL CIRCUIT METHOD

Review up until 9:30. From 9:30 to 23:00 we talk on how to choose the operating point; the considerations are basically two, the gain of the amplification and the input swing range. From 23:30 we start small signal analysis. He starts saying that he wants to do the small signal analysis in a simple way avoiding complicated math. Basically he said that we need to re-draw the circuit in a linear way, and he calls this new circuit the small signal circuit.

We can express our circuit with KVL and KCL and we get something like this (first KVL for input loop, second KVL for output loop, and third KCL):

VI + …. ; VO + …. ; … + iDS + ….

If we add a small perturbation around the operating point, we can write the same KVL and KCL rules as:

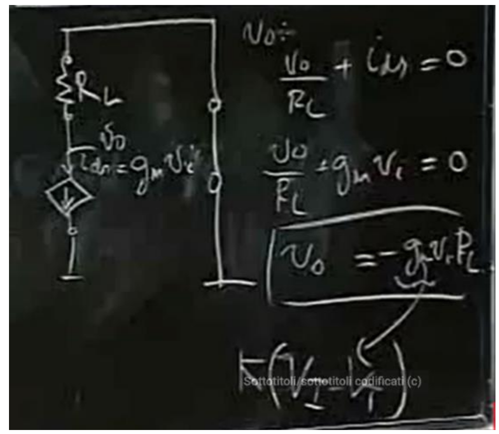
VI + vi …. ; VO + vo …. ; … + iDS + ids ….

Then we can cancel the large signal values and leave the small signal values. These will satisfy KVL and KCL:

vi + …. ; vo + …. ; … + ids + ….

The small signal circuit method (32:00) prescribes the following steps:

1. Find the operating point using large signal analysis
2. Develop small signal model of elements (what we have done just above)
3. Replace original elements with small signal elements (what we have done just above) and analyze resulting circuit (linear circuit)



Some examples:

A) MOSFET
B) DC Supply VS
C) Resistor

Following this procedure with our MOSFET circuit, we can simplify the circuit to the small signal circuit below and solve it as in figure.

LESSON 12 – CAPACITORS

With the new memory devices we are going to study, the output of a circuit will not only depend on the inputs alone of a circuit, but it is going to depend also on the background of the circuit, where it "has been" in the past.

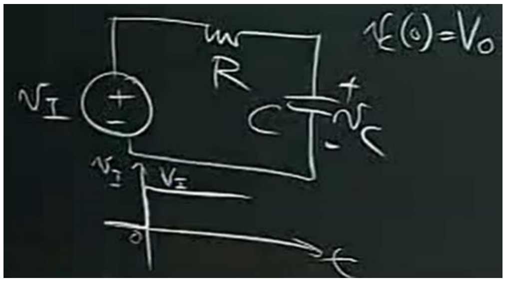
From 2:00 to 11:30 he makes the example of the two inverters that the second is going to be a bit delayed as it is shown in the book at the beginning of chapter 9.

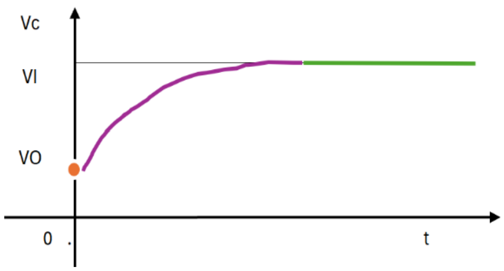
At 11:47 he talks about the capacitor. He starts with a claim, that the delay we saw in the inverters come from the presence of a capacitor somewhere in those inverters. He goes on explaining where the capacitor in a MOSFET is as done in the book in chapter 9 paragraph 3 I think. Then he makes the new MOSFET model with a capacitor inside to account for the delay (also done in book).

At 17:45 it explains the capacitor directly and quickly writes down the main equations. At 28:16 he puts the capacitor already in some simple circuits. He goes on like the last 15/20 minutes solving the circuit with a voltage source, a resistor and a capacitor in series. Most of this time is spent mathematically solving the differential equation for the circuit with the method of homogenous and particular solution (basically in the whole course this is the method used):

- Find the particular solution. Guess the value and check if left and right part match.
- Find the homogenous solution. Solve equation with the drive set to 0.
- Sum both above

LESSON 13 – DIGITAL CIRCUIT SPEED

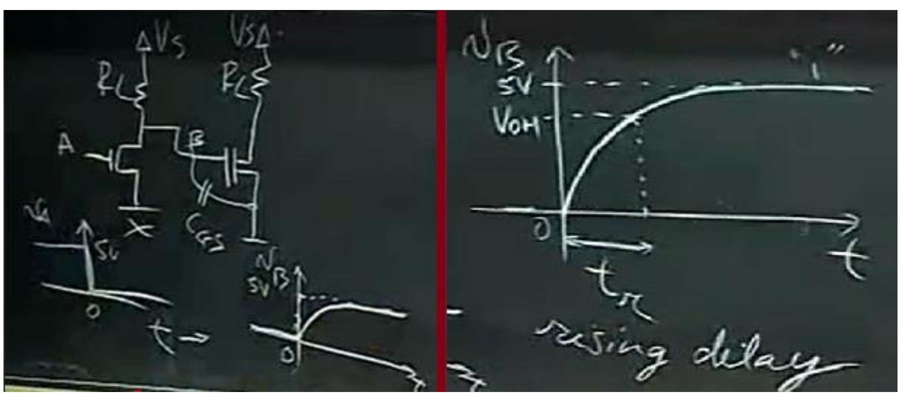




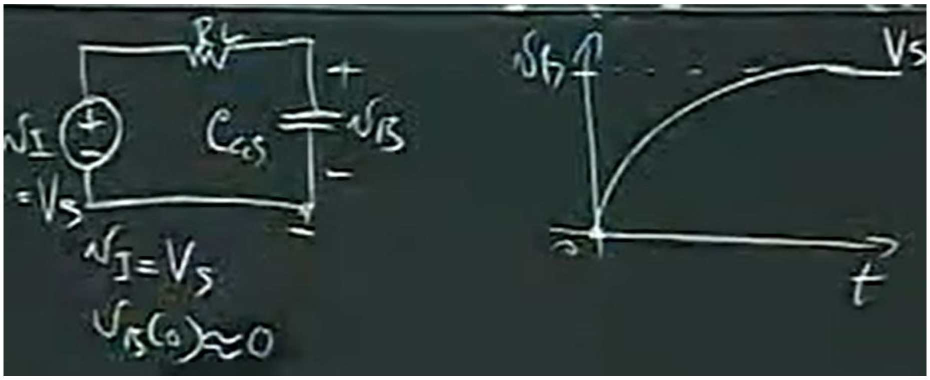
Review of previous lesson till 7:17. A more intuitive method that does not require solving differential equations is presented. Suppose we have the circuit and a step input VI as in figure, as well as a initial condition Vc(0)=VO. The solution is found in steps. First, we draw VO in orange, we know that the capacitor starts there. Then for t≫0 we know that vC = VI so we draw the green part of the curve. And we saw before that there are only two ways the capacitor can charge, so we can draw the purple part of the curve. Regarding the expression of the curve it is: $vC = VO + (VI - VO)(1 - e^{-\frac{t}{RC}})$.

If VI < VO we could work the solution out the same way, with an equation $vC = VO + (VO - VI)e^{-\frac{t}{RC}}$

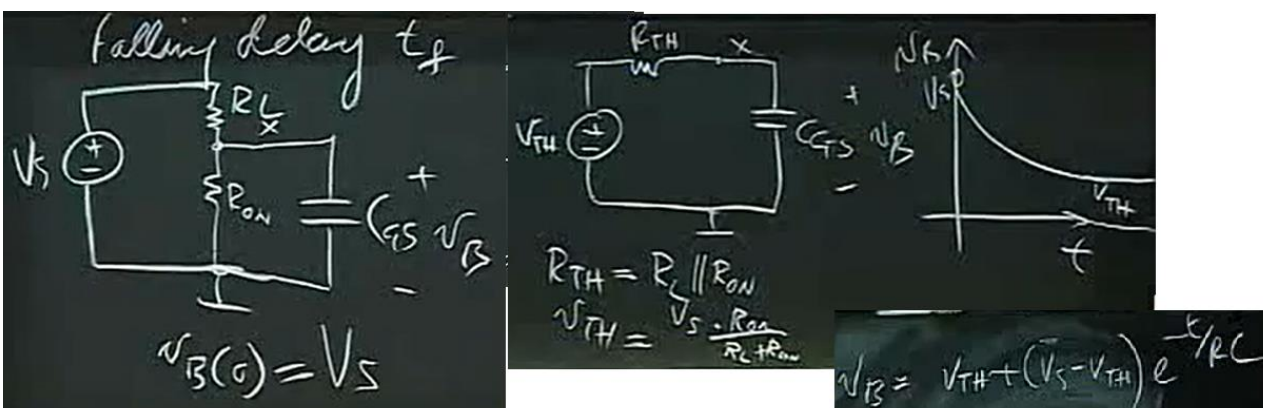
At 17:10 he says he wants to compute a rising and falling delays for the two inverters circuit. So, in figure below we have the inverters and across the gate B a capacitor CGS will form. We assume that the input VA is at 5V and will step down to 0V at t=0. Due to this we expect the input VB to step from 0 to 5V. Because of the capacitance we know is not a step change, but a rising exponential as per figure below. So when we talk about delay we want to know how long does it take to VB to go from a logical low to a logical high (from 0 volt to VOH). The graph for this on the right.



Let's compute the number of the delay. We first draw the equivalent circuit as per figure below. The circuit is basically the equivalent circuit when the first MOSFET is OFF. To solve this circuit, we can draw the graph on the right. Very simple. From this we can find the equation $VB = 0 + VS\left(1 - e^{-\frac{t}{RC}}\right)$. What is left is to calculate Tr (from the graph above which is actually the same redrawn below). Basically, I need to find for which t the output is VOH. So $VOH = VS\left(1 - e^{-\frac{t}{RC}}\right)$, which gives me the result $tr = -RL * Cgs * \ln\left(1 - \frac{VOH}{VS}\right)$.



At 32:52 we start calculating the falling delay for the inverters circuit. In this case the equivalent circuit is the one in figure below on the left. We can simplify this circuit using Thevenin so we get the circuit in the middle for which I can calculate VTH and RTH. Third step is to calculate the voltage across the capacitor



(figure on the right). The reasoning is the same as before, we know the voltage at t=0 is VS, we know the voltage for t≫0 is VTH, and boom! The equation of the curve is written under the graph. What is left is not to calculate the delay tf. Same as before, $VOL = VTH + (VS - VTH)e^{-\frac{tf}{RC}}$ and solving it I get $tf = -RTH * Cgs * ln\frac{\mathrm{VOL-VTH}}{VS-VTH}$.

For a quick approximation of the two delays, we have calculated we can use the following two quick formulas:

- For rising delay: RL*Cgs
- For falling delay: RTH*Cgs

At 47:13 he brough a EMS problem on a chip which will be the base for next lesson.

LESSON 14 –

Recap until 3:25.