MINISTRY OF EDUCATION AND SCIENCE OF UKRAINE
National aerospace university named after N. E. Zhukovsky
"Kharkov aviation institute"

A. A. Roienko

CIRCUITS THEORY AND MICROELECTRONICS

Tutorial

Kharkov «KhAI» 2014

UDC 621.382 + 621.3 (075.8)
LBC 31.211 + 32.844.1Я73
R73

Розглянуто ряд питань з теорії електричних кіл, а саме основні терміни і елементи електрики, основні закони теорії кіл та їх застосування, функціонування RC-ланцюгів. Другу частину посібника присвячено питанням мікроелектроніки, а саме достатньо глибоко висвітлено основні властивості напівпровідникових матеріалів, принципи дії p-n-переходу, напівпровідникових діодів, біполярних і польових транзисторів.

Для студентів технічних вузів, що навчаються за спеціальностями «Радіотехніка», «Комп'ютерні мережі та системи», «Телекомунікації».

Reviewers: doct. of tech. scien. I. M. Mycenko,
cand. of phys. and math. scien. V. B. Sinitskiy

Roienko, A. A.

R73 Circuits Theory and Microelectronics [Text]: tutorial / A. A. Roienko. – Kh.: Nat. aerospace un-ty named after N. E. Zhukovsky "Kharkov aviat. inst-t", 2014. – 100 p.

ISBN 978-966-662-358-7

The first part of tutorial deals with the questions of circuits theory such as main electricity concepts, basic electrical elements, main laws of circuits theory, RC-circuit charge and discharge processes. The second part is devoted to the microelectronic devices and gives rather deep overview of main properties of semiconductor materials, p-n junction operating principles, semiconductor diodes, bipolar junction and field-effect transistors.

Tutorial will be useful for the students of radioengineering, telecommunication, computer systems specialities.

Figs. 120. Tabls. 1. Bibliogr.: 7 titles

UDC 621.382 + 621.3 (075.8)
LBC 31.211 + 32.844.1Я73

ISBN 978-966-662-358-7

1. BASICS OF CIRCUITS THEORY

1.1. Electrical properties of materials

From the electrical point of view, all materials can be divided into three groups. But at first, remember that all materials are built of **atoms**. Each atom has **nucleus** and **electrons** rotated on the nucleus and bounded to it.

Definition. *Conductors* - materials that permit free motion of a large number of electrons. Most metals are good conductors. For example, *copper, silver, gold* etc.

Definition. *Insulators* (or nonconductors) are materials with molecules which haven't free electrons at all. Electrons are tightly bounded to their atoms. Large amount of energy is required to free the electrons from the influence of the nucleus. Examples of insulators are *rubber, plastics, glass and dry wood.*

Definition. *Semiconductors* - materials that conduct electricity, but offer opposition to current flow. These type of materials is neither good conductor nor good insulator. Examples of semiconductors are *carbon, silicon, germanium, tin, and lead.*

Examples. *Conductor* materials are widely used in different wires (for instance, Cat. E5 cable), in production of various printed circuit boards, etc. *Insulators* are usually applied for isolating wires of different polarities, in capacitors and other electrical devices. *Semiconductor* materials are the basis of the whole microelectronic field and, as a result, computer devices.

1.2. The concept of electricity

The basic concepts of electricity are closely bounded to the conductor materials. We start with the concept of voltage.

Electricity exists due to the movement of either positively charged particles (holes) or negatively charged particles (electrons) from point A to point B. Consider the movement of holes.

Definition. *Voltage* between two points of scheme is the energy released during the movement of unitary positive charge from the point with high potential (point A, for instance) to the point with lower potential (point B). Potential is defined as

$$\varphi = W_p / q, \tag{1.1}$$

where W_p denotes the potential energy of the particle, q is the particle's charge.

During its movement from one point to another, particle interacts with many atoms inside the conductor and step-by-step looses its energy. Then, taking into account all particles, voltage is defined as potential difference between two corresponding points of scheme:

$$V = \varphi_1 - \varphi_2. \tag{1.2}$$

The basic measurement unit for potential difference is the **Volt** (symbol V) and, that's why, potential difference is called **voltage**.

Term "*voltage*" has several synonyms: "**electromotive force**" (*emf*), "**potential difference**" and "**voltage drop**". In electrical formulas and equations voltage is usually symbolized with a capital **V** or with **E**.

Note. An object's electrical charge is determined by the number of electrons that the object has gained or lost.

Usually there are many free electrons (or, equivalently, positive charges) in the conductor. In neutral state, they drift from one atom to another in a random direction. But when potential difference is applied, the direction of free electrons movement is taken under control.

The strength of the applied voltage determines how many electrons change from a random motion to a more directional path through the wire.

Definition. The movement or flow of such electrons is called **electron current flow** or just **current**.

The symbol for current is "**I**" or "*i*". Quantitatively the current is measured by its *strength*. The current strength ***I*** equals to the amount of electricity (or charge) Δq transferred during time interval Δt through a conductor cross section:

$$I = \Delta q / \Delta t . \qquad (1.3)$$

The basic measurement for current is the **amperage** (A). The *direction of electron flow* is from a point of negative potential to a point of positive potential. The flow of electrons in one direction causes a flow of positive charges in the opposite direction. The flow of positive charges is known as **conventional current**.

Generally, electrical current can be classified as one of two general types: **Direct Current** (DC) or **Alternating Current** (AC). A *direct current* flows continuously in the direction. An *alternating current* periodically reverses direction. An example of DC is the current obtained from a *battery*. An example of AC is a *common household current*.

Definition. Resistance is defined as the opposition to the current flow. The amount of such an opposition mainly depends upon the amount of available free electrons in the material.

Resistance is measured in **Ohms** and is represented by the symbol "**R**". One Ohm is defined as the amount of resistance that will limit the current in conductor to 1 A when the potential difference applied to the conductor is 1 V. The short notation for Ohm is the Greek letter capital omega **Ω**.

Definition. Conductance is defined as the ability to conduct current and is opposite to resistance. If wire has a high conductance, then it has low resistance and vice-versa. The unit used to specify conductance is called "**Mho**" which is "Ohm" spelled backward. The symbol for "Mho" is the Greek letter omega inverted **℧**. The symbol for conductance used in formulas is **G**:

$$G = 1 / R . \qquad (1.4)$$

Electricity is generally used to do some sort of work, such as turning a motor or generating heat.

Definition. Specifically, **power** is the rate (or speed) at which work is done or the rate at which heat is generated. In equations, power is abbreviated with the capital letter **P**. Power is described as the current **I** in a circuit times the voltage **V** across the circuit:

$$P = V \cdot I. \tag{1.5}$$

The measurement unit for power is **Watt** and is abbreviated with the capital **W**.

In 1827, George Ohm discovered that there was a definite relationship between voltage, current and resistance in an electrical circuit. This relationship can be stated as follows.

Definition. Applied voltage equals circuit current times the circuit resistance:

$$V = I \cdot R. \tag{1.6}$$

1.3. Voltage, current and power in case of harmonic impact

In case of harmonic impact, current and voltage in some schemes vary in time accordingly to the sine or cosine law:

$$i(t) = I \cdot \cos(2\pi f \cdot t),\ v(t) = V \cdot \cos(2\pi f \cdot t),\ f = 1/T, \tag{1.7}$$

where $\omega = 2\pi f = 2\pi/T$ denotes a ***cycling frequency***, T is a ***signal cycle*** (or period).

Definition. ***Instantaneous power p(t)*** is denoted as

$$p(t) = i(t) \cdot v(t). \tag{1.8}$$

Power given off on the resistance R also depends on time and varies with the frequency doubled 2ω (Fig. 1.1):

$$p(t) = i(t) \cdot v(t) = I\cos(\omega t) \cdot V\cos(\omega t) = 0.5 \cdot I \cdot V \cdot [1 + \cos(2\omega t)], \tag{1.9}$$

Defi..'tion. **Average power P** is obtained by averaging the instantaneous power over the period:

$$P = \frac{1}{T}\int_0^T p(t)dt. \tag{1.10}$$

The concept of **average power** is used to estimate the *extracted power*. For our case, we get:

$$P = \frac{1}{T}\int_0^T p(t)dt = \frac{VI}{2} = \frac{V^2}{2R} = \frac{I^2}{2}R, \tag{1.11}$$

$$\int_0^T \cos(2\omega t)\,dt = \frac{1}{2\omega}\sin\left(2\frac{2\pi}{T}T\right) - 0 = 0. \qquad (1.12)$$

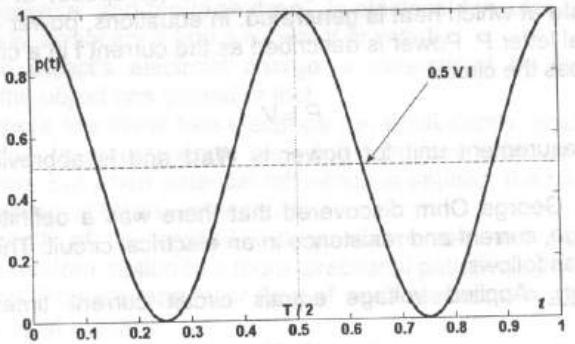


Fig. 1.1

1.4. Effective (active) voltage value

Let the harmonic voltage $v(t)$ impacts the resistance R. As a result, the average power P is extracted. Define the equivalent *direct voltage* $\boldsymbol{V_{eff}}$ such that when it applies to the resistance R the extracted power value P_0 will be the same as the *average power* P in case of alternating voltage.

The average power P (from Eq. (1.10) and Ohm's law (1.6)) can be written as:

$$P = \frac{1}{T}\int_0^T \frac{v^2(t)}{R}dt = \frac{V^2}{2R}. \qquad (1.13)$$

Definition. *Extracted power* P_0:

$$P_0 = V_{eff}^2 / R. \qquad (1.14)$$

Then setting P equal to P_0 we obtain:

$$\frac{V^2}{2R} = \frac{V_{eff}^2}{R} \text{ from which } V_{eff}^2 = \frac{V^2}{2} \text{ or } V_{eff} = \frac{V}{\sqrt{2}}. \qquad (1.15)$$

Thus:

1. Effective voltage V_{eff} in case of harmonic impact is $\sqrt{2}$=1.41 times smaller than the amplitude value V of signal $v(t)$.

2. Effective voltage value V_{eff} is also equal to the **root-mean-square voltage:**

$$P = \frac{1}{T}\int_0^T \frac{v^2(t)}{R}dt,\ P_0 = V_{eff}^2 / R,\ P = P_0, \text{ from which}$$

$$V_{eff} = \sqrt{\frac{1}{T}\int_0^T v^2(t)dt}\,. \qquad (1.16)$$

2. ELECTRICAL SCHEME AND ITS COMPONENTS

Definition. Electrical scheme is a set of *discrete* and *integral* components (Fig. 2.1). *Discrete components* are divided on *linear* and *nonlinear* elements.

Definition. Linear elements are the elements for which the dependence of current upon voltage is described by linear function. This group includes such real elements as *resistors*, *capacitors* and *inductances*.

Elements that have *nonlinear voltage-ampere characteristics* (VAC) belong to the class of *nonlinear elements*. This group consists of *diodes*, *transistors*, etc.

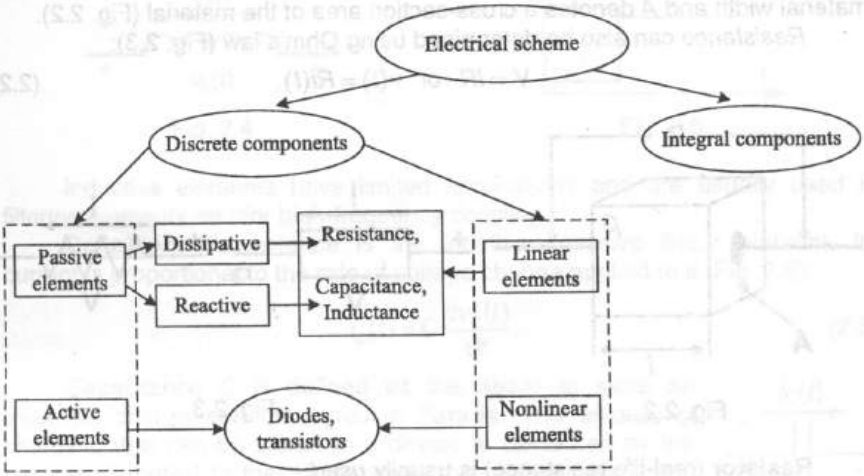


Fig. 2.1

Discrete elements are also divided on *active* and *passive* elements.

Definition. *Active elements* allow amplifying input alternating signals, whereas *passive elements* can only transform the energy of the signal from one form to another.

Passive discrete elements are divided into two classes according to the energy transformation mode: **dissipative** elements and **reactive** elements.

The example of *dissipative element* is a **resistance** that actively transforms energy from electro-magnetic kind to the thermal state. In opposite,

the *reactive elements* accumulate the energy of electrical or magnetic field and then return it to the circuit. Well-known examples of such elements are *capacitance* and *inductance*.

2.1. Idealized models of discrete passive elements

The mathematical analysis of real-life schemes is usually done using its **idealized model**. Then the set of idealized elements is necessary to replace the real elements. In this topic we will consider the main idealized passive elements and their properties.

Definition. *Resistance* is the idealized passive linear element and it can transform the current into voltage without lag effects. Parameter of the element is the *resistance* defined as:

$$R = \rho \frac{l}{A}, \tag{2.1}$$

where ρ denotes the *resistivity* of the material and is measured in Ω-*m*, l is a material width and A denotes a cross-section area of the material (Fig. 2.2).

Resistance can also be determined using Ohm's law (Fig. 2.3):

$$V = IR \quad \text{or} \quad v(t) = Ri(t). \tag{2.2}$$

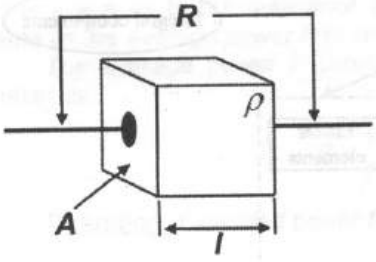


Fig. 2.2

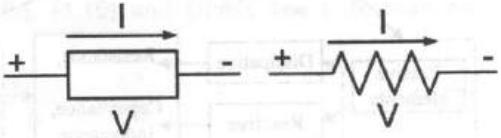


Fig. 2.3

Resistor (real-life resistance) is usually *used*:

1. As a load of active devices;
2. In filtering schemes;
3. As an additional timing element (time constant setting);
4. For the redistribution of the potential levels in a scheme etc.

Definition. *Inductance* is an idealized passive element (Fig. 2.4). Voltage drop on the inductance is proportional to the rate of current change flown through it:

The *main features* of the element:

1. The symbol used to indicate inductance in electrical formulas and equations is ***L***. The measurement units for L are called "**Henries**" and are

abbreviated by ***H***. Equation (2.3) is the mathematical representation for the voltage $v_L(t)$ induced in a coil with the inductance *L*.

$$v_L(t) = -L\frac{di(t)}{dt}. \qquad (2.3)$$

2. *Inductance* is naturally a *dynamic* element. Voltage drop on the element appears only when there are changes of current, flown through the inductance. In *static mode* (in case of direct current) the voltage drop on the element is *absent*.

3. Inductance is an *inertial element*, i. e. current through it cannot change instantly when potential difference is applied on it. That is, Fig. 2.5 cannot be observed because otherwise:

$$\frac{di}{dt} \to \infty, \quad v_L(t) \to \infty. \qquad (2.4)$$

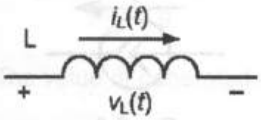


Fig. 2.4

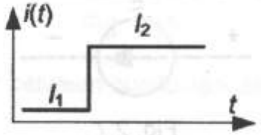


Fig. 2.5

Inductive elements have limited *applications* and are usually used in filtering elements and for high-frequency coupling.

Definition. *Capacitance* is the idealized passive linear element. Its current is proportional to the rate of voltage change applied to it (Fig. 2.6):

$$i_C(t) = C\frac{dv_C(t)}{dt}. \qquad (2.5)$$

Capacitance *C* is defined as the ability to store an electrical charge and measured in **Farads**. The amount of charge *q* that can be stored in a device is calculated by the voltage *V* applied to the device when the charge was stored times *C*:

$$q = CV_c. \qquad (2.6)$$

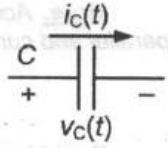


Fig. 2.6

2.2. Scheme active elements

Active elements are power sources which actuate currents or voltages in the scheme. Idealized active elements are divided into:

1) idealized independent voltage source (**emf generator**);
2) idealized independent current source (**current generator**).

Definition. *Emf generator* is the energy source with the terminal voltage that doesn't depend on the current flown through the source (Fig. 2.7). Numerical characteristic of the element is the terminal emf "***E***".

EMF action causes increasing of the potential, that's why it's directed from the low to the high potential inside the source.

Main features of the element:

1) Emf value doesn't depend on the load resistance;
2) The internal resistance of the source equals to 0, R_i=0;
3) There is no lack of energy in the source;
4) Emf generator is the source of the extremely high power because the current in the external circuit at any value of load resistance is $I=E/R_i$.

Definition. ***Current generator*** is the active element with the current that doesn't depend on the value of terminal voltage (Fig. 2.8). The main numerical parameter of the source is the generated current value *I*.

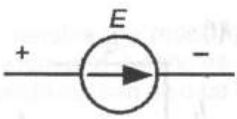


Fig. 2.7

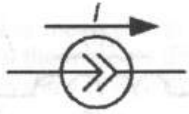


Fig. 2.8

Main features of the element:

1) Value of the current, generated by the element, doesn't depend on the load resistance;
2) Internal resistance of the source is extremely high (currents of other sources can't flow through the element);
3) Current generator is the source of the extremely high power; at any value of load resistance R_L its terminal voltage is $V=IR_L$ and power is determined as $P=VI$.

Note. *According to their features EMF generator can't be connected in parallel and current generators can't be connected in series.*

2.3. Real-life power supplies

Real-life sources are the *sources of finite power*.

Definition. *Real-life voltage source* has low internal resistance which is connected in series with the ideal emf generator (Fig. 2.9). When current through the load increases output terminal voltage decreases:

$$e = E - IR_i \qquad (2.7)$$

and emf *E* is divided between the input and the load resistances:

$$e = IR_L = \frac{E}{R_i + R_L} R_L, \quad e = \frac{E}{1 + R_i / R_L}. \tag{2.8}$$

Definition. *Real-life current supply* has high internal resistance (low conductance) and has parallel connection with the ideal current generator (Fig. 2.10).

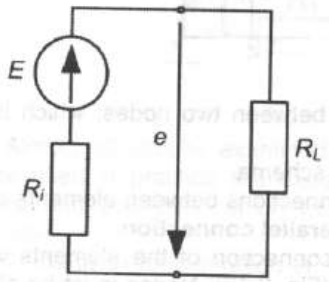


Fig. 2.9

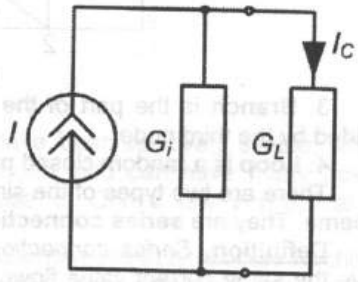


Fig. 2.10

The value of generated current is divided between conductances G_i and G_L:

$$I_C = VG_L = \frac{I}{G_i + G_L} G_L, \quad I_C = \frac{I}{1 + G_i / G_L}. \tag{2.9}$$

Current and voltage of the real-life sources are lower than the current and voltage of ideal supplies due to the internal resistances:

$$e = E - IR_i, \quad I_C = I - VG_i. \tag{2.10}$$

Real-life sources are close to the ideal ones if following ratios are fulfilled:

$\frac{R_i}{R_L} << 1$ - for the *emf* generator;

$\frac{G_i}{G_L} << 1$ - for the current generator.

2.4. Electrical circuit and its main laws

At first, let us introduce the following main concepts:

1. **Scheme** is a graphic representation of the device model, composed from idealized electrical elements (Fig. 2.11).
2. **Node** (or *junction*) is the connection point of three or more electrical elements.

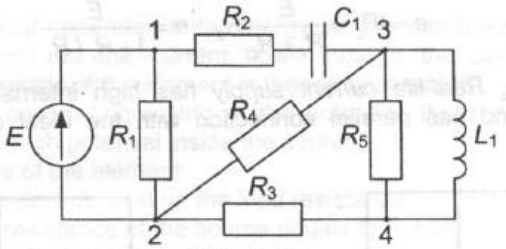


Fig. 2.11

3. **Branch** is the part of the scheme between two nodes, which is not divided by the third node.

4. **Loop** is a random closed path in the scheme.

There are two types of the simplest connections between elements of the scheme. They are **series connection** and **parallel connection**.

Definition. *Series connection* is the connection of the elements which have the same current value flown through (Fig. 2.12). Nodes must be absent between the elements connected in series.

In case of series connection the total resistance increases. Elements R_2, C_1 have series connection on Fig. 2.9 whereas R_2, R_3 don't have it.

Definition. *Parallel connection* is the connection of elements, which have the same applied voltage (Fig. 2.13). Element terminals connected in parallel are merged in pairs.

In case of parallel connection the overall resistance decreases. On Fig. 2.11 elements R_5, L_1 are connected in parallel as well as the elements R_2, R_1 on Fig. 2.13.

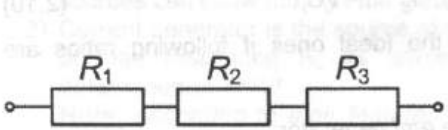


Fig. 2.12

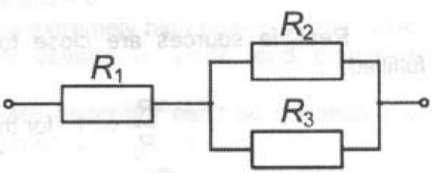


Fig. 2.13

Calculation example. On Fig. 2.14 the voltage value V_1 is known and equals to 10 V. The resistance value R_2=2 $k\Omega$ and current value I_2=1mA.

Then voltage drop on the resistance R_2 will be:

$$V_{R_2} = I_2 R_2 = 1 \cdot 10^{-3} \cdot 2 \cdot 10^{3} = 2\,\text{V}. \qquad (2.11)$$

Therefore, voltage value in the second node with respect to 0 potential will be

$$V_2 = V_1 - V_{R_2} = 10 - 2 = 8\,\text{V}. \qquad (2.12)$$

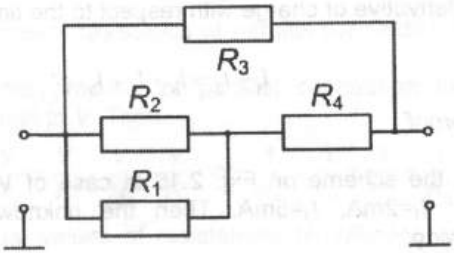


Fig. 2.14

Almost all circuits examined so far, have been relatively simple. Many circuits used in practice are extremely complex and can have many power sources and branches which make the use of Ohm's law impractical or impossible.

Through experimentation in 1857 the German physicist Gustav Kirchhoff developed methods to solve complex circuits. They are known today as Kirchhoff laws.

2.5. Kirchhoff laws. Law 1 - Kirchhoff current law

Definition. The current arriving at any node in a circuit is equal to the current leaving that node (Fig. 2.15):

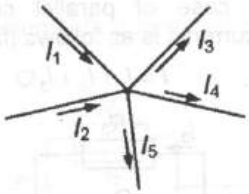


Fig. 2.15

$$I_1 + I_2 = I_3 + I_4 + I_5 \quad \text{or} \quad I_1 + I_2 - I_3 - I_4 - I_5 = 0 \tag{2.13}$$

or equivalently:

$$\sum_{k=1}^{5} I_k = 0 . \tag{2.14}$$

Proof. Kirchhoff current law is obtained from the **charge conservation law** which states of the impossibility to store charge in the node. In other words, number of charges entering the node (on Fig. 2.13, q_1 and q_2) is equal to number of charges leaving the node (on Fig. 2.13, q_3, q_4 and q_5), that is:

$$q_1 + q_2 = q_3 + q_4 + q_5 . \tag{2.15}$$

After taking derivative of charge with respect to the time we have currents equality:

$$I_1 + I_2 = I_3 + I_4 + I_5. \tag{2.16}$$

//end of the proof

Example. In the scheme on Fig. 2.16 in case of $V_1 > V_2$, there are the following currents: I_1=2mA, I_3=5mA. Then the unknown current can be calculated as following:

$$I_2 = I_3 - I_1 = 5 - 2 = 3mA. \tag{2.17}$$

Note. In case of AC under harmonic impact we should make geometric sum of the currents instead of algebraic one (Fig. 2.17).

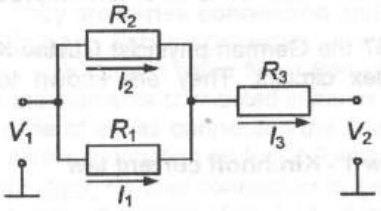


Fig. 2.16

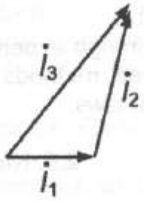


Fig. 2.17

Consequence 1. In case of parallel connection of the elements (Fig. 2.18) the equation for currents is as follows (from the 1st law):

$$I = I_1 + I_2 + I_3. \tag{2.18}$$

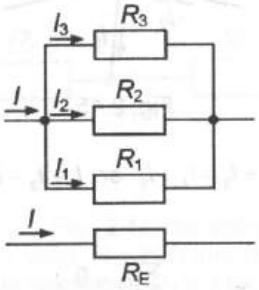


Fig. 2.18

Using Ohm's law we can write:

$$\frac{V}{R_E} = \frac{V_1}{R_1} + \frac{V_2}{R_2} + \frac{V_3}{R_3}, \tag{2.19}$$

where R_E is the equivalent resistance of parallel connection of resistors R_1, R_2 and R_3.

According to the property of parallel connection the voltage values $V_1=V_2=V_3$ and are equal to V. Then

$$\frac{V}{R_E}=\frac{V}{R_1}+\frac{V}{R_2}+\frac{V}{R_3} \quad \text{or} \quad \frac{1}{R_E}=\frac{1}{R_1}+\frac{1}{R_2}+\frac{1}{R_3}. \tag{2.20}$$

In case of equal values of resistances ($R_1=R_2=R_3$) we have $R=R_1/3$. If there are only two resistances in the scheme the formula (2.20) will be

$$\frac{1}{R_E}=\frac{1}{R_1}+\frac{1}{R_2}, \quad R_E=\frac{R_1R_2}{R_1+R_2}. \tag{2.21}$$

Note. *Equivalent resistance of two elements R_1 and R_2 connected in parallel is less than the value of the smallest resistance*, i. e.

$$\text{if } R_1<R_2, \text{ i. e. } \frac{R_1}{R_2}<1, \text{ then } R_E=\frac{R_1R_2}{R_1+R_2}=\frac{R_1}{1+\dfrac{R_1}{R_2}}<R_1. \tag{2.22}$$

If the difference in the resistance values R_1 and R_2 is significant, the equivalent resistance is almost equal to the smaller resistance:

$$\text{i. e., if } \frac{R_1}{R_2}=0.1, \text{ then } R_E=\frac{R_1}{1+0,1}\approx R_1. \tag{2.23}$$

Consequence 2. Two elements connected in parallel create a current divider (Fig. 2.19).

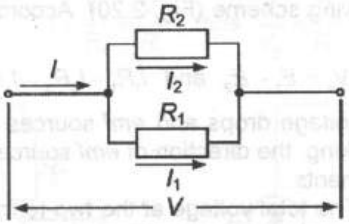


Fig. 2.19

Proof. Elements connected in parallel have the same applied voltage V:

$$I_1R_1=I_2R_2=V, \text{ then } \frac{I_1}{I_2}=\frac{R_2}{R_1}. \tag{2.24}$$

It's obvious that currents in branches are inversely proportional to their resistances. It allows to determine the current in the second branch if the current in the first branch is known:

$$I_2 = I_1 \frac{R_1}{R_2}. \quad (2.25)$$

In a similar way, we can determine currents in branches of the divider if total current and elements R_1 and R_2 of the divider are known:

$$V = IR_E = I\frac{R_1R_2}{R_1+R_2} = I_1R_1 = I_2R_2. \quad (2.26)$$

Then we have

$$I_1 = I\frac{R_2}{R_1+R_2} \text{ and } I_2 = I\frac{R_1}{R_1+R_2}. \quad (2.27)$$

Calculation example. the total current of the divider on Fig. 2.19 I=10 mA and elements $R_1 = 2k\Omega$, $R_2 = 3k\Omega$. Calculate current value in each branch of the scheme. Using formulas (2.27) we get

$$I_1 = I\frac{R_2}{R_1+R_2} = 10\frac{3}{2+3} = 6mA \text{ and } I_2 = I\frac{R_1}{R_1+R_2} = 10\frac{2}{2+3} = 4mA. \quad (2.28)$$

2.6. Kirchhoff laws. Law 2 - Kirchhoff voltage law

Definition. The sum of the voltage drops around a closed loop is equal to the sum of the voltage sources of that loop:

$$\sum_{k=1}^{N} V_k = \sum_{i=1}^{M} E_i. \quad (2.29)$$

Consider the following scheme (Fig. 2.20). According to Kirchhoff second law we can write:

$$V_1 - V_3 - V_4 = E_1 - E_2 \text{ and } I_1R_1 - I_3R_3 - I_4R_4 = E_1 - E_2. \quad (2.30)$$

The signs of the voltage drops and *emf* sources we define according to the direction of path tracking, the direction of *emf* source operations and current direction at the loop elements.

Consequence 1. The total voltage at the two-terminal network is equal to the sum of voltage drops at its elements (Fig. 2.21):

$$V = V_1 + V_2 + V_3, \quad IR_{SC} = IR_1 + IR_2 + IR_3. \quad (2.31)$$

Then

$$R_{SC} = R_1 + R_2 + R_3. \quad (2.32)$$

Thus, *the equivalent resistance of elements connected in series equals to the sum of resistances of these elements.*

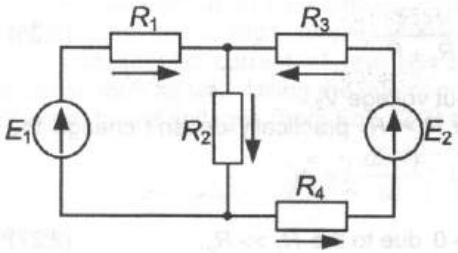


Fig. 2.20

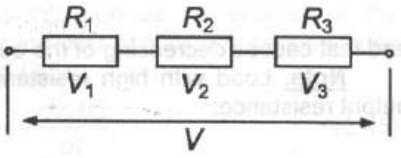


Fig. 2.21

Consequence 2. Voltage divider consists of two resistances, connected in series. The input voltage is applied to the elements R_1 and R_2, whereas the output voltage is taken only from one of them (R_2) (so called, **divider's arm**) (Fig. 2.22).

Determine the output voltage V_2 using the input voltage V_1:

$$V_2 = I_2 R_2 = IR_2 = \frac{V}{R_1 + R_2} R_2 = V \frac{R_2}{R_1 + R_2}, \quad V_2 = \frac{V}{1 + \frac{R_1}{R_2}}. \tag{2.33}$$

Application of voltage divider. If it is necessary to get voltage value V_2 as $V_2 = V/3$, then the divider resistances must satisfy the following equation:

$$R_1 = 2R_2. \tag{2.34}$$

Voltage drops at the divider resistances are proportional to the corresponding resistances:

$$I = \frac{V_1}{R_1} = \frac{V_2}{R_2}, \text{ thus } \frac{V_1}{V_2} = \frac{R_1}{R_2}. \tag{2.35}$$

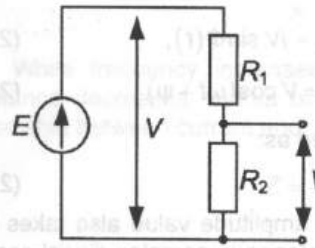


Fig. 2.22

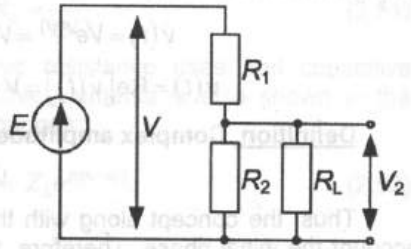


Fig. 2.23

Consequence 3. If we connect the load to the divider resistance (Fig. 2.23), then we generally will change the output voltage (decrease it). In fact, the equivalent resistance of divider lower arm will be decreased:

$$R_E^{\parallel} = \frac{R_2 R_L}{R_2 + R_L}, \tag{2.36}$$

and that causes decreasing of the output voltage V_2.

Note. Load with high resistance $R_L >> R_2$ practically doesn't change the output resistance:

$$R_E^{\parallel} = \frac{R_2}{1 + \frac{R_2}{R_L}}, \quad \frac{R_2}{R_L} \to 0 \text{ due to the } R_L >> R_2. \tag{2.37}$$

Then

$$R_E^{\parallel} \approx \frac{R_2}{1} \approx R_2. \tag{2.38}$$

2.7. Circuit calculation method. Complex amplitude method

Under harmonic impact currents and voltages at linear elements are described by cosine (or sine) law of variation:

$$v(t) = V\cos(\omega t + \psi), \tag{2.39}$$

where $v(t)$ is an instantaneous voltage, V denotes the amplitude, ψ is an initial phase, $\Phi(t)=\omega t+\psi$ denotes total instantaneous phase, $\omega=d\Phi/dt$.

Definition. Instantaneous voltage can be represented in a complex form:

$$\dot{v}(t) = Ve^{j(\omega t+\psi)} = Ve^{j\Phi(t)}. \tag{2.40}$$

Two considered forms of harmonic voltage representation are related as (based on Euler formula):

$$\dot{v}(t) = Ve^{j\Phi(t)} = V\cos\Phi(t) + jV\sin\Phi(t), \tag{2.41}$$

$$v(t) = \mathrm{Re}\left[\dot{v}(t)\right] = V\cos\Phi(t) = V\cos(\omega t + \psi). \tag{2.42}$$

Definition. Complex amplitude is defined as:

$$\dot{V} = Ve^{j\psi}. \tag{2.43}$$

Thus, the concept along with the actual amplitude value also takes into account the initial phase. Therefore, the instantaneous complex signal can be represented:

$$\dot{v}(t) = Ve^{j(\omega t+\psi)} = Ve^{j\psi}e^{j\omega t} = \dot{V}e^{j\omega t}. \tag{2.44}$$

All calculations in the complex amplitude method are made with respect to the *current and voltage complex amplitudes.*

Voltage and current of reactive elements can be represented in the complex form as well. Using the interconnection between instantaneous values of currents and voltages (from Eqs. (2.3) and (2.5))

$$v_L = L\frac{di_L(t)}{dt},\quad i_C = C\frac{dv_C(t)}{dt}, \tag{2.45}$$

we obtain

$$\dot{v}(t) = \dot{V}e^{j\omega t},\quad i(t) = \dot{I}e^{j\omega t}, \tag{2.46}$$

$$\dot{v}_L(t) = L\frac{d}{dt}\left(\dot{I}_L e^{j\omega t}\right) = Lj\omega\dot{I}e^{j\omega t} = \dot{V}_L e^{j\omega t}, \tag{2.47}$$

$$i_C(t) = C\frac{d}{dt}\left(\dot{V}_C e^{j\omega t}\right) = Cj\omega\dot{V}_C e^{j\omega t} = \dot{I}_C e^{j\omega t}. \tag{2.48}$$

Decreasing both parts of (2.47) and (2.48) by common multiplier $e^{j\omega t}$, we get *Ohm's law in a complex form*:

$$\dot{V}_L = \dot{I}_L(j\omega L) = \dot{I}_L Z_L,\quad \dot{I}_C = \dot{V}_C(j\omega C) = \dot{V}_C / Z_C. \tag{2.49}$$

Definitions. **Complex resistances of reactive elements** Z_L and Z_C are as follows:

$$Z_L = j\omega L = jX_L,\quad Z_C = \frac{1}{j\omega C} = -jX_C. \tag{2.50}$$

Fulfillment of Ohm's law (in complex form, Eq. (2.49)) *confirms the linear character* of the inductive and capacitive elements.

Definition. **Resistances of the reactive elements** are frequency dependent:

$$X_L = \omega L,\quad X_C = \frac{1}{\omega C}. \tag{2.51}$$

While frequency increases, inductive resistance rises and capacitive resistance decreases. Inertia of the reactive elements is also shown in the phase shift between current and voltage changes:

$$Z = \frac{\dot{V}}{\dot{I}} = \frac{Ve^{j\psi_V}}{Ie^{j\psi_I}} = Z_A e^{j(\psi_V - \psi_I)}, \tag{2.52}$$

$$Z_L = jX_L = X_L e^{j\frac{\pi}{2}},\quad Z_C = -jX_C = X_C e^{-j\frac{\pi}{2}}. \tag{2.53}$$

In the inductive element the voltage **leads** the current by 90^0: $\psi_U - \psi_I = \pi/2$.
And in capacitive element the voltage **lags** the current by 90^0: $\psi_U - \psi_I = -\pi/2$.

2.8. Main parameters of periodical signals

There are so-called *geometrical parameters of periodical signals*:

- signal cycle T;
- frequency $F=1/T$, $\omega=2\pi F=2\pi/T$;
- amplitude (maximal instantaneous signal value).

There are also *additional parameters* for periodic impulse sequence:

- pulse duration, τ;
- filling factor (or *off-duty factor*) $Q=T/\tau$;
- pulse duty factor (*fill factor*) $\gamma=1/Q$.

The following energy characteristics can be calculated for impulse signals:

- **instantaneous power** of the signal $p(t)=s^2(t)$, if $s(t)$ is voltage (current), then $s^2(t)$ is the power dissipated on the resistance $R=1\ \Omega$;
- **signal energy on the interval**:

$$E=\int_0^T p(t)dt=\int_0^T s^2(t)dt \tag{2.54}$$

because $P=E/T$, i. e. $E=PT$;

- **averaged signal power**:

$$P_{avg}=\frac{E}{T}=\frac{1}{T}\int_0^T s^2(t)dt\,; \tag{2.55}$$

- **mean value of impulse sequence** (constant component of the signal):

$$V_0=\frac{1}{T}\int_0^T v(t)dt\,. \tag{2.56}$$

The meaning of the constant component. Replace a sine pulse by rectangular with an equivalent area at the interval equaled to the signal period (Fig. 2.24). Then constant component is a height of the rectangle with an equivalent area and one side equaled to the signal period.

Area of sine impulse:

$$S_1=\int_0^{T/2} V_m\sin(\omega t)dt=-V_m\frac{1}{\omega}\cos(\omega t)\Big|_0^{T/2}=$$
$$=-V_m\frac{T}{2\pi}\left[\cos\left(\frac{2\pi}{T}\frac{T}{2}\right)-\cos 0\right]=-V_m\frac{T}{2\pi}(-2)=\frac{V_mT}{\pi}, \tag{2.57}$$

$$S_2=V_0T. \tag{2.58}$$

Then we have

$$S_1 = S_2,\ \frac{V_m T}{\pi} = V_0 T,\ V_0 = \frac{V_m}{\pi}. \qquad (2.59)$$

The value of constant component is determined for the periodical sequence of rectangular impulses in a similar way.

From a physical point of view constant component is such a constant voltage that is not passed by separating capacity. The value of constant component is determined for the periodical sequence of rectangular impulses in a similar way (Fig. 2.25):

$$V_m \tau = V_0 T,\ V_0 = V_m \frac{\tau}{T} = \frac{V_m}{Q}. \qquad (2.60)$$

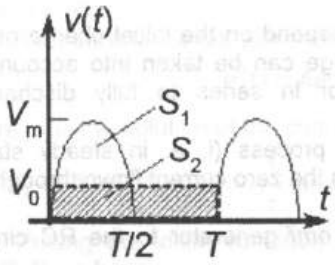


Fig. 2.24

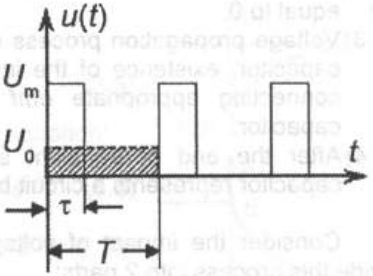


Fig. 2.25

3. RC-CIRCUITS

The simplest RC circuit consists of resistor and capacitor connected in series (Fig. 3.1). Due to capacitor's inertia the task of full description of such a circuit can be done only if current and voltage processes are observed during some period of time.

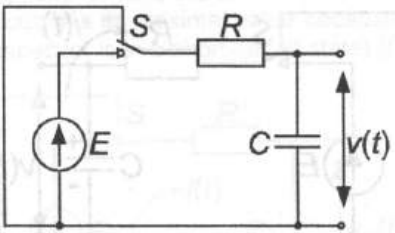


Fig. 3.1

It is known, that current through the capacity is determined by the velocity of voltage change:

$$i_C(t) = C\frac{dv_C(t)}{dt}. \qquad (3.1)$$

Therefore, voltage at the capacity can't be changed instantly, i. e. it keeps its previous value during switching:

$$v_C(t_0-) = v_C(t_0+). \qquad (3.2)$$

3.1. Main properties of RC circuit in switching mode

1) After circuit turning on, a capacitor should be considered as a short-circuit element for the current flown through it.
2) When capacitor is in a short-circuit mode, voltage on the capacitor is equal to 0.
3) Voltage propagation process doesn't depend on the initial charge of the capacitor; existence of the initial voltage can be taken into account by connecting appropriate *emf* generator in series to fully discharged capacitor.
4) After the end of transient switching process (i. e. in steady state), capacitor represents a circuit break with the zero current flown through it.

Consider the impact of voltage from *emf* generator to the RC circuit. Divide this process into 2 parts:

1) **Switching-off the *emf* generator** from RC circuit followed by capacitor discharge process;
2) **Switching-on the *emf* generator** that provides charging of the capacitor.

3.2. Capacitor discharge process

Scheme initial conditions: capacitor voltage $v(0)=E$; current in the circuit $I=0$; *emf* generator is connected to the circuit.

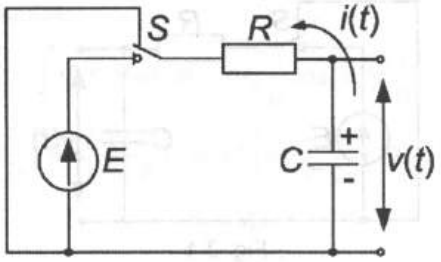


Fig. 3.2

At the moment $t=0$ *emf* generator is switched off from the circuit (Fig. 3.2). Capacity discharge process initiates the current in the circuit:

$$i_C(t) = C\frac{dv_C(t)}{dt}. \qquad (3.3)$$

Immediately after switching, voltage at the capacity remains the same $V_0=E$ and the current in the circuit has its maximal value $I_0=E/R$ (Fig. 3.3).

Transient process of capacity discharge is described by the following equation (from Kirchhoff voltage law):

$$C\frac{dv(t)}{dt} = -\frac{v(t)}{R}, \quad \tau\frac{dv(t)}{dt} + v(t) = 0, \qquad (3.4)$$

where $\tau = RC$.

After solving the equation we have

$$v(t) = Ae^{p_1 t}, \qquad (3.5)$$

where p_1 is the solution of the characteristic equation:

$$p\tau + 1 = 0, \text{ i. e. } p_1 = -\frac{1}{\tau} \text{ and } v(t) = Ae^{-t/\tau}. \qquad (3.6)$$

Putting the initial condition $t=0$, $v(0)=E$ into the Eq. (3.6) we have

$$v(t) = Ee^{-t/\tau}. \qquad (3.7)$$

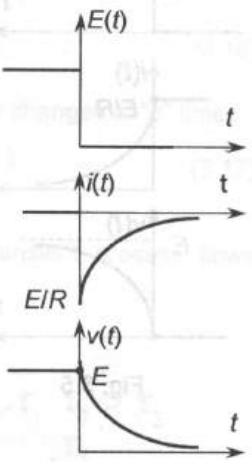


Fig. 3.3

3.3. Capacitor charge process

Scheme initial conditions: switching moment $t=0$; voltage at the capacity $v_C(0)=0$; current in the circuit $i(0)=0$; *emf* generator is switched off.

Scheme conditions immediately after switching:

- voltage at the capacity remains the same $v_C=0$;
- current in the circuit has its maximal value because *emf* E is applied only to the resistor (capacitor is in a short-circuit state) (Fig. 3.4).

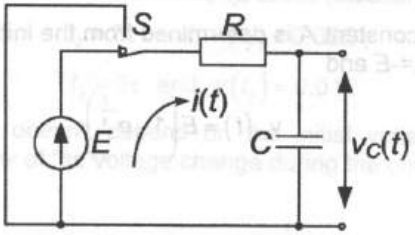


Fig. 3.4

During transient process while charging the capacitor, current in the circuit decreases and the voltage $v_C(t)$ increases. After completing the transient process the scheme steady state conditions are: $i(t)=0$, $v_C(t)=E$ (Fig. 3.5).

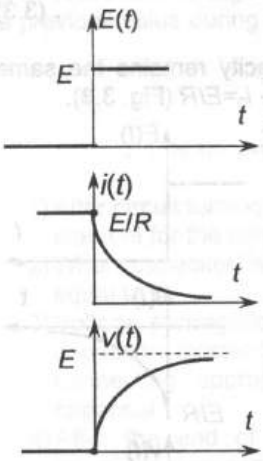


Fig. 3.5

Analytically transient process of the capacitor charge is described by the differential equation (from Kirchhoff voltage law):

$$E = v_C(t) + v_R(t) \quad \text{or} \quad E = v_C(t) + i(t)R. \tag{3.8}$$

Considering $i(t) = C\dfrac{dv_C(t)}{dt}$ we have

$$E = v_C(t) + \tau\frac{dv_C(t)}{dt}. \tag{3.9}$$

where $\tau = RC$.

Solution of the equation consists of the free component, which is determined only by the internal circuit parameters and of the stationary component:

$$v(t) = v_{free}(t) + v_{stationary}(t). \tag{3.10}$$

Free mode is determined while the external effect is absent E=0, i. e.

$$0 = v_C(t) + \tau\frac{dv_C(t)}{dt} \tag{3.11}$$

and

$$v_C(t) = v_{free}(t) = Ae^{-\frac{t}{\tau}}. \tag{3.12}$$

Steady state corresponds to the static case:

$$t \to \infty, \quad \frac{dv_C(t)}{dt} = 0, \text{ therefore } v_{st}(t) = E. \tag{3.13}$$

In general, switching process consists of two components:

$$v_C(t) = E + Ae^{-\frac{t}{\tau}}. \tag{3.14}$$

The unknown constant A is determined from the initial scheme conditions t=0, $v_C(0)$=0. Then A=-E and

$$v_C(t) = E\left(1 - e^{-\frac{t}{\tau}}\right). \tag{3.15}$$

3.4. Features of the transient processes in the RC-circuit

Transient process flows nonlinearly (irregularly) in time.

Discharge process. At the beginning, after switching the toggle, velocity of the voltage change is very high but then transient process is getting slower.

At the initial time *the velocity of the transient process is determined by the time parameter of the circuit - τ* (so-called, **time constant**):

$$v_C(t) = Ee^{-\frac{t}{\tau}}, \quad \left.\frac{dv_C(t)}{dt}\right|_{t=0} = -\frac{E}{\tau}. \tag{3.16}$$

During the time equaled to τ voltage at the capacity changes at "e" times:

$$v_C(\tau) = \frac{E}{e}, \tag{3.17}$$

i. e. at 63 % from the initial value (Fig. 3.6).

The greater the time constant is the slower transient process flows (Fig. 3.7).

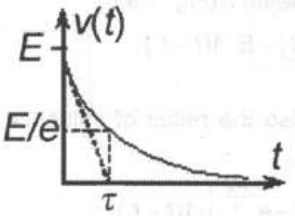


Fig. 3.6

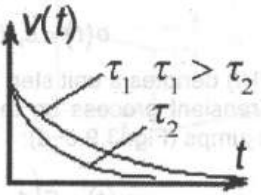


Fig. 3.7

Duration of the transient process is determined according to the threshold level of a transient voltage:

a) voltage is changed at 10 times:

$$v(t_1) = Ee^{-\frac{t_1}{\tau}} = \frac{E}{10} = 0.1E, \quad e^{\frac{t_1}{\tau}} = 10, \quad t_1 = \tau \ln 10, \quad t_1 = 2.3\tau; \tag{3.18}$$

b) practical duration of the transient process (discharge process):

$$t_2 = 3\tau \text{ and } v(t_2) = 0.05E, \tag{3.19}$$

$$t_3 = 5\tau \text{ and } v(t_3) = 0.01E. \tag{3.20}$$

Time constant doesn't depend on the initial voltage of the capacitor (Fig. 3.8). General law of the voltage change during the charging process of the capacitor:

$$v(t) = E + Ae^{-\frac{t}{\tau}}. \tag{3.21}$$

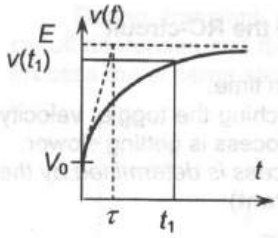


Fig. 3.8

If at $t=0$ initial voltage $v(0)=V_0$, then constant value is $V_0=E+A$, $A=V_0-E$. Then the capacitor charge in case of non-zero initial voltage is described by the following formula

$$v(t)=E+(V_0-E)e^{-\frac{t}{\tau}}. \qquad (3.22)$$

General formula for the determination of the transient process duration at the specified threshold level $v(t_1)$ (Fig. 3.8) (from Eq. (3.22)):

$$\frac{E-v(t_1)}{E-V_0}=e^{-\frac{t_1}{\tau}},\quad t_1=-\tau\ln\frac{E-v(t_1)}{E-V_0}\ \text{ or }\ t_1=\tau\ln\frac{E-V_0}{E-v(t_1)}. \qquad (3.23)$$

3.5. RC-circuit under impact of rectangular pulse

Rectangular pulse can be considered as a result of two polar voltage jumps shifted in time on the impulse duration value t_i (Fig. 3.9,b):

$$e(t)=s_1(t)-s_2(t)=E\cdot 1(t)-E\cdot 1(t-t_i). \qquad (3.24)$$

where $1(t)$ denotes a unit step (Fig. 3.9,a).

Transient process on the capacity is also the result of influence of two voltage jumps (Fig. 3.9,c–d):

$$v(t)=E\left(1-e^{-\frac{t}{\tau}}\right)\cdot 1(t)-E\left(1-e^{-\frac{t-t_i}{\tau}}\right)\cdot 1(t-t_i). \qquad (3.25)$$

Correspondingly, voltage changes on the resistor are determined by the law of current variation (Fig. 3.9,e):

$$v_R(t)=i(t)R=Ee^{-\frac{t}{\tau}}1(t)-Ee^{-\frac{t-t_i}{\tau}}1(t-t_i). \qquad (3.26)$$

Output pulse shape depends on the ratio of time constant of the circuit and impulse duration. The case of *relatively small time constant* $\tau<<t_i$ is described at stress diagrams (Fig. 3.9,d–e). The second extreme case corresponds to the *relatively high time constant* $\tau>>t_i$ (Fig. 3.10).

While impulse duration the capacitor charges up to the small voltage value V_m. Charge and discharge of the capacitor takes place at the initial part of the transient process curve, that's why voltage change is practically linear.

Pulse top drop at the resistor takes place according to the capacitor charge. Negative impulse tail is formed by the current of the discharge capacitor.

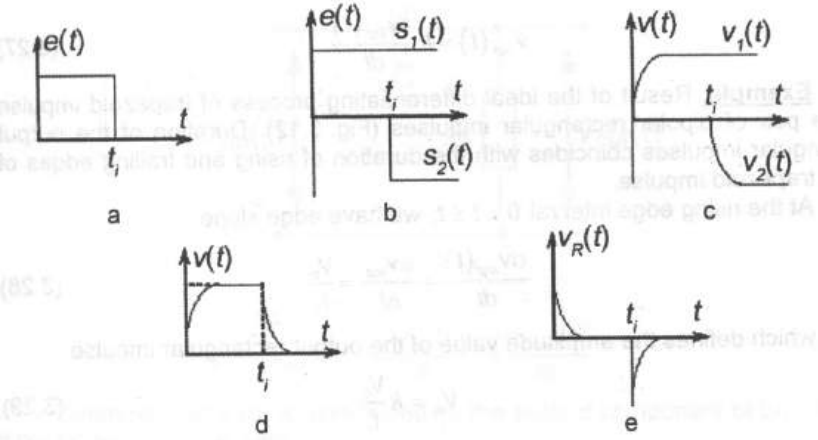


Fig. 3.9

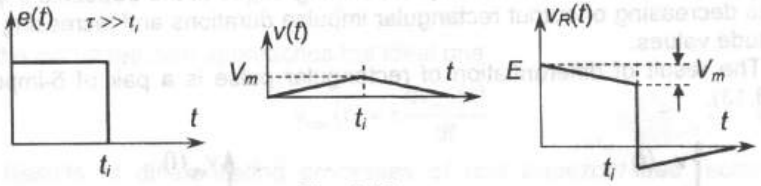


Fig. 3.10

For *undistorted impulse transmission* it is necessary to increase sufficiently time constant in comparison to the impulse duration.

Limiting mode corresponds to the impulse duration $t_i=3\tau$ when the voltage at the capacity reaches 95% from the steady-state value, i. e. 0.95E (Fig. 3.11).

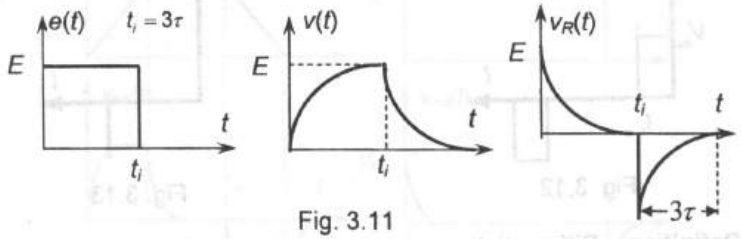


Fig. 3.11

3.6. Differentiating circuit

Definition. Differentiating circuit is a circuit where the output voltage is proportional to the velocity of input voltage change:

$$v_{out}(t) = k\frac{dv_{inp}(t)}{dt}. \qquad (3.27)$$

Example. Result of the ideal differentiating process of trapezoid impulse is the pair of bipolar rectangular impulses (Fig. 3.12). Duration of the output rectangular impulses coincides with the duration of rising and trailing edges of input trapezoid impulse.

At the rising edge interval $0 < t \le t_1$ we have edge slope:

$$\frac{dv_{inp}(t)}{dt} = \frac{\Delta v_{inp}}{\Delta t} = \frac{V_0}{t_1} \qquad (3.28)$$

which defines the amplitude value of the output rectangular impulse:

$$V_m = k\frac{V_0}{t_1} \qquad (3.29)$$

where k has dimension of time and denotes some constant value.

Decreasing durations of rising and trailing edges of the trapezoid impulse causes decreasing of output rectangular impulse durations and increasing their amplitude values.

The result of differentiation of rectangular pulse is a pair of δ-impulses (Fig. 3.13).

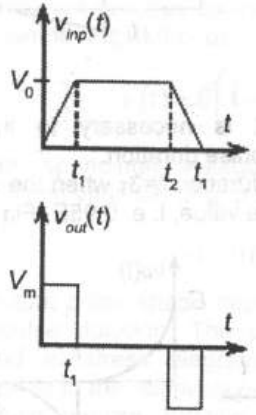


Fig. 3.12

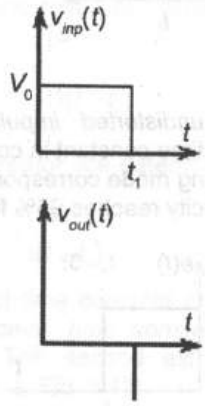


Fig. 3.13

Definition. Differentiating circuit based on RC-scheme (Fig. 3.14) differentiates output signal *with some error*:

$$v_{out}(t) = Ri(t) = RC\frac{dv_C(t)}{dt} = \tau\frac{d}{dt}\left[v_{inp}(t) - v_{out}(t)\right]; \qquad (3.30)$$

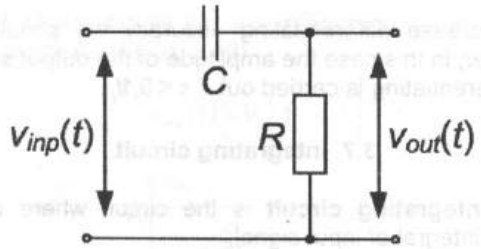


Fig. 3.14

$$v_{out}(t) = \tau \frac{dv_{inp}(t)}{dt} - \tau \frac{dv_{out}(t)}{dt}. \tag{3.31}$$

Differentiating error is determined by the second component of Eq. (3.31). If the following ratio is held

$$\frac{dv_{out}(t)}{dt} << \frac{dv_{inp}(t)}{dt} \tag{3.32}$$

then the circuit reaction approaches the ideal one:

$$v_{out}(t) \approx \tau \frac{dv_{inp}(t)}{dt}. \tag{3.33}$$

Results of differentiating processes of *real trapezoid* and rectangular impulses are depicted in Fig. 3.15. *Ideal differentiating* is carried out on the impulse parts with constant or linear-varying voltage, except periods of capacitor charging and discharging.

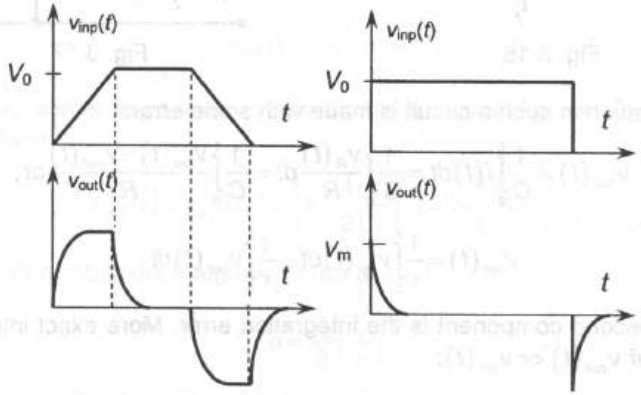


Fig. 3.15

Note. To increase differentiating accuracy we should decrease time constant τ. However, in this case the amplitude of the output signal decreases.

Satisfied differentiating is carried out if $\tau \le 0,1t_i$.

3.7. Integrating circuit

Definition. Integrating circuit is the circuit where output voltage is proportional to the integral of input signal:

$$v_{out}(t) = k\int_0^t v_{inp}(t)dt. \qquad (3.34)$$

As a result of ideal integration of rectangular impulse, linearly-increased voltage is generated:

$$v_{out}(t) = k\int_0^t V_0 dt = V_0 kt, \quad t_i \ge t > 0, \qquad (3.35)$$

where *k* has the dimension inversed to time (Fig. 3.16).

Important. The simplest integrating circuit is implemented on the basis of RC-circuit where output voltage is taken from the capacitor (Fig. 3.17).

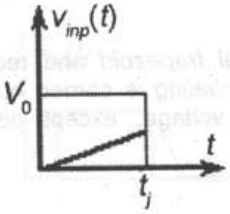


Fig. 3.16

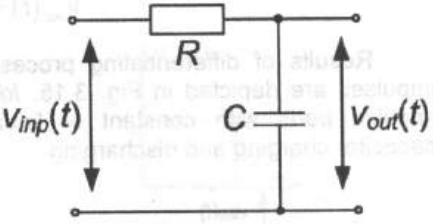


Fig. 3.17

Integration in such a circuit is made with some error:

$$v_{out}(t) = \frac{1}{C}\int_0^t i(t)dt = \frac{1}{C}\int_0^t \frac{v_R(t)}{R}dt = \frac{1}{C}\int_0^t \frac{v_{inp}(t) - v_{out}(t)}{R}dt; \qquad (3.36)$$

$$v_{out}(t) = \frac{1}{\tau}\int_0^t v_{inp}(t)dt - \frac{1}{\tau}\int_0^t v_{out}(t)dt. \qquad (3.37)$$

The second component is the integration error. More exact integration is carried out if $v_{out}(t) << v_{inp}(t)$:

$$v_{out}(t) \approx \frac{1}{\tau}\int_0^t v_{inp}(t)dt. \qquad (3.38)$$

Note. In case of rectangular impulse integration output signal differs from ideal linearly-increased signal (Fig. 3.18):

$$v_{out}(t) = V_0\left(1 - e^{-\frac{t}{\tau}}\right). \tag{3.39}$$

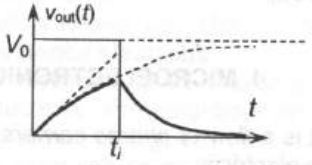


Fig. 3.18

Initial part corresponds to the exact integration. The biggest error corresponds to the end of impulse.

Important. To implement more exact integration it's necessary to increase time constant of the circuit $\tau >> t_{imp}$ to work at the initial part of the charging curve. That causes the small amplitude of the output impulse.

Note. After impulse ending the ideal integrator keeps voltage level, whereas the real integrator has the exponential decay due to capacity discharge.

Integration error can be determined from the comparison of output signal from the ideal integrator (*under the impulse impact*):

$$v_{out}(t) = \frac{t}{\tau}V_0. \tag{3.40}$$

and real one

$$v_{out}(t) = V_0\left(1 - e^{-\frac{t}{\tau}}\right). \tag{3.41}$$

Making series expansion of Eq. (3.41) and limiting with 3 expansion terms, we have:

$$v_{out}(t) \approx V_0\left[1 - 1 + \frac{1}{\tau}t - \frac{1}{2}\left(\frac{t}{\tau}\right)^2\right] = V_0\frac{t}{\tau}\left(1 - \frac{t}{2\tau}\right). \tag{3.42}$$

Definition. Absolute integration error at $t = t_i$:

$$\delta = \frac{V_0}{2}\left(\frac{t_i}{\tau}\right)^2. \tag{3.43}$$

Definition. Relative error of integration process:

$$\gamma = \frac{\delta}{V_0 t_i / \tau} = \frac{t_i}{2\tau}. \qquad (3.44)$$

Example. If $\gamma = 0.01$ and $t_i/\tau = 0.02$, we have $\tau = 50t_i$.

Note. If there are no strict requirements to the integrating circuit, then time constant $\tau = (5...10)t_i$.

4. MICROELECTRONICS

Electrical current is a *flow of charge carriers*. There are two requirements for the existence of the electricity:

- There must be a supply of charge carriers;
- The charge carriers must be free to move.

One of the best conductors is the metals because their atomic structure is such that there are many free electrons wandering within the material. At the other extreme are the non-conductors or insulators, including most plastics and ceramics, which have no free electrons. Insulators offer very low level of conductivity under the apllied voltage source.

Materials which have a conductivity level somewhere between the extremes of insulators and conductors are called **semiconductors**. Such a name was given because *semiconductor materials are insulators at low temperatures but are conductors at room temperature and above.*

4.1. Main properties of semiconductor materials

Two of the most widely-used semiconductors are the **silicon** (Si) and **germanium** (Ge). Important property of the semiconductor materials is that their characteristics can be altered significantly through the **addition of impurity** or **impact of heat and light**.

Definition. Addition of impurity to the semiconductor material can change its properties from relatively poor conductor to a good conductor of electricity; such a process is known as "**doping**".

Definition. Intrinsic semiconductor material is such a material that has been carefully refined to reduce the impurities to a very low level – essentially as pure as can be made by modern technology.

Some of the unique qualities of the Ge and Si noticed above are due to their atomic structure. The atoms of both materials form a very definite *pattern* that is periodic in nature (i. e., continually repeats itself).

Definition. One complete pattern is called **crystal** and the periodic arrangement of the atoms is called **lattice**. For Ge and Si the crystal has 3D-diamond structure of Fig. 4.1.

Examine the structure of the typical semiconductor atom. It is known that the atom is composed of three basic particles: the *electron*, the *proton* and the

neutron. In the atomic structure the neutrons and protons form the *nucleus*, while the electrons revolve around the nucleus in fixed *orbits*.

The Born models of the Ge and Si are shown in Fig. 4.2. As indicated, the Ge atom has 32 orbiting electrons, while Si has 14 orbiting electrons. In each case, there are 4 electrons in the outer shell.

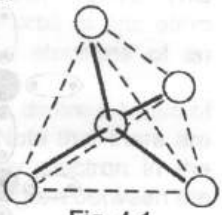


Fig. 4.1

Definition. Outer shell is known as **valence shell** and its electrons are called **valence electrons**.

The ideal number of electrons to fill this shell is eight. Sharing four outer electrons with neighbour atoms leads to the crystal structure shown on Fig. 4.1.

At different time each atom should have full outer shell, i. e. *eight outer electrons*.

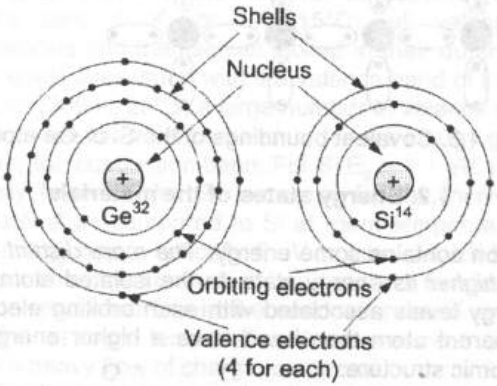


Fig. 4.2. The atomic structure of Ge (*left*) and Si (*right*)

Definition. This creates a bond between the adjacent atoms, holding them together in the lattice (Fig. 4.3). Such kind of atom bonding is called **covalent bonding**.

Although the covalent bond will result in a stronger bond between the valence electrons and their parent atoms, it is still possible for the valence electrons to absorb sufficient kinetic energy from natural sources to break the covalent bond and obtain "**free**" state.

Definition. The term "**free**" denotes that the motion of such electrons is quite sensitive to applied electric field established by outer voltage source or any potential difference.

Definition. Free electrons in the material are called as **intrinsic charge carriers** if their appearance is due to the natural causes only. Increasing of the semiconductor material temperature results in a substantial increasing of the number of free electrons.

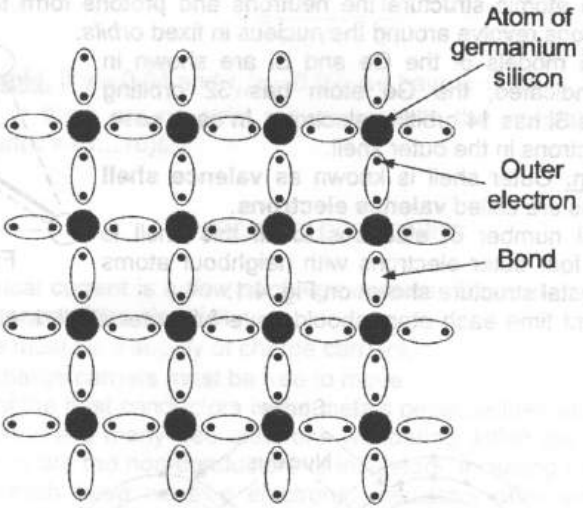


Fig. 4.3. Covalent boundings of the Si or Ge atoms

4.2. Energy states of the materials

Each electron contains some energy. *The more distant the electron from the nucleus, the higher its energy state.* In the isolated atomic structure there are discrete energy levels associated with each orbiting electron (Fig. 4.4). If electron left its parent atom then it will have a higher energy state than any electron in the atomic structure.

Between the discrete energy levels there are **gaps** in which no electrons in the isolated atomic structure can appear.

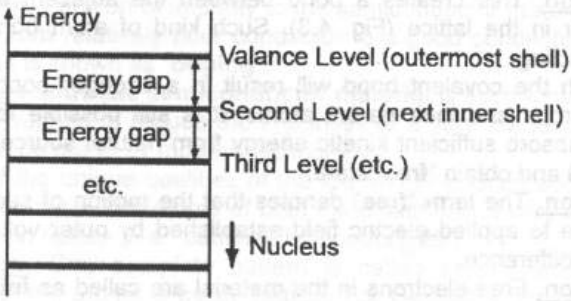


Fig. 4.4. Discrete levels in isolated atomic structures

When the atoms of a material are brought closer to each other in order to form the lattice structure, there is an interaction between atoms. This interaction results in the effect that electrons in a particular orbit of one atom will have slightly different energy levels comparing to the electrons of an adjoining atom.

The net result of such a process is an expansion of the discrete levels of possible energy states to the **bands** as shown in Fig. 4.5. Note that there are **boundary levels** and maximal energy states in which any electron in the atomic lattice can find itself and there remains a **forbidden region** between the **valence band** and the **ionization level (or conduction band)**.

Definition. Ionization is the mechanism when an electron absorbs sufficient energy to break away from the atomic structure and enter the conduction band.

The energy associated with each electron is measured in **electron-volts** (**eV**). At absolute zero, 0 K (or -273.15°C), all valence electrons of semiconductor materials find themselves locked in their outermost shell of the atom with energy levels associated with the valence band of Fig. 4.5. However, at room temperature (300 K, 25°C) a large number of valence electrons acquire sufficient energy to leave the valence band, cross the energy gap defined by E_g in Fig. 4.5 and enter the conduction band. For Si E_g is 1.1 eV, for Ge - 0.67 eV.

The obviously lower E_g for Ge accounts for the increased number of carriers in that material as compared to Si at room temperature. Note that *for the insulator* the energy gap is typically *5 eV* or more, which limits the number of electrons that can enter the conduction band at room temperature.

The conductor has electrons in the conduction band even at 0 K. Quite obviously, therefore, at room temperature there are more than enough free carriers to sustain a heavy flow of charge, or current.

Next consider the situation when certain impurities are added to the intrinsic semiconductor material.

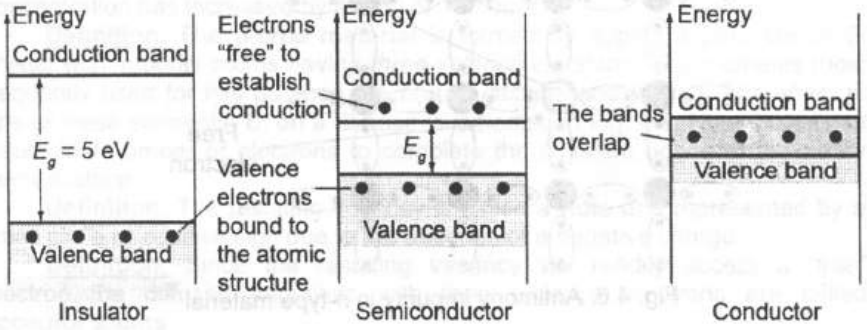


Fig. 4.5. Conduction and valence bands of an insulator, semiconductor and conductor

4.3. Impurity in semiconductor materials

The characteristics of semiconductor materials (namely, conduction) can be altered significantly by the addition of certain impurity atoms (doping process) into the relatively pure semiconductor material. In practice a very small percentage of dopant atoms is introduced into the lattice - for example, 1 part in 10 million.

Definition. A semiconductor material that has been subjected to the doping process is called an **extrinsic material.**

There are two extrinsic materials of immeasurable importance to semiconductor device fabrication: **n-type** and **p-type**. Consider each of these materials.

Definition. The **n-type material** is created by introducing those impurity elements that have *five valence electrons*, such as *antimony*, *arsenic* and *phosphorus*. The effect of such impurity elements is indicated in Fig. 4.6 (using antimony as the impurity in a Si base). Note that the four covalent bonds are still present.

There is, however, an additional *fifth electron* due to the impurity atom, which is unassociated with any particular covalent bond. This remaining electron is *loosely bound* to its parent (antimony) atom and is relatively free to move within the newly formed *n*-type material.

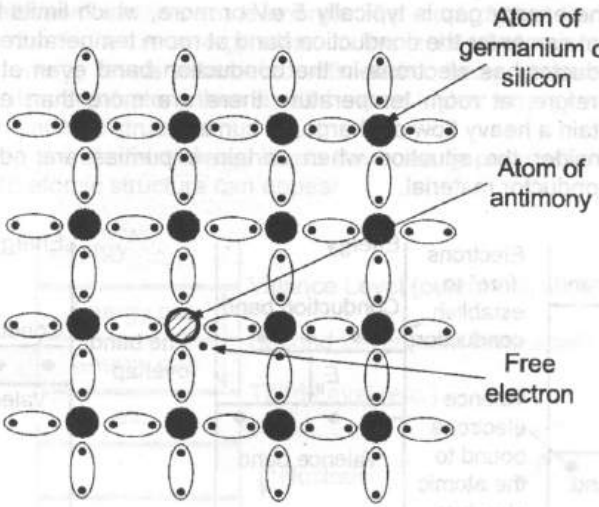


Fig. 4.6. Antimony impurity in *n*-type material

Definition: Since the inserted impurity atom *has donated* a relatively "free" electron to the structure, *diffused impurities* with five valence electrons are called **donor atoms**.

The effect of this doping process on the relative conductivity can be described in best way using the energy-band diagram (Fig. 4.7). Note that a discrete energy level (called the **donor level**) appears in the *forbidden band* with an E_g significantly less than that of the intrinsic material.

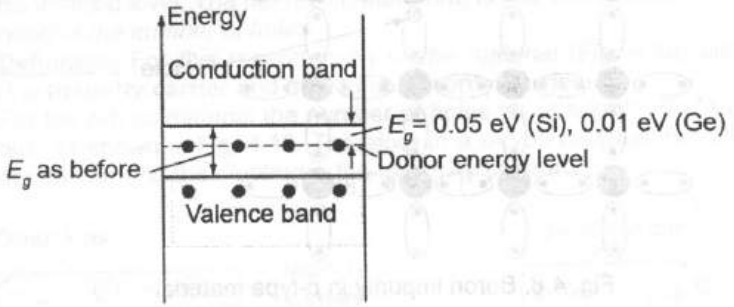


Fig. 4.7. Effect of donor impurities on the energy band structure

Those "free" electrons due to the added impurity sit at this energy level and have less difficulty for absorbing a sufficient measure of thermal energy to move into the conduction band at room temperature. The result is that at room temperature, there are a large number of carriers (electrons) in the conduction band and the conductivity of the material increases significantly.

At room temperature in an intrinsic Si material there is about one free electron for every 10^{12} atoms (1 to 10^9 for Ge). If our dosage level were 1 in 10 million (10^7), the ratio ($10^{12}/10^7=10^5$) would indicate that the carrier concentration has increased by a ratio of 100 000:1.

Definition. The **p-type material** is formed by doping a pure Ge or Si crystal with impurity atoms having *three valence* electrons. The elements most frequently used for this purpose are *boron*, *gallium*, and *indium*. The effect of one of these elements, B, on a Si base is indicated in Fig. 4.8. Now there is an insufficient number of electrons to complete the covalent bonds of the newly formed lattice.

Definition. The resulting vacancy is called a **hole** and represented by a small circle or *positive sign* due to the absence of a negative charge.

Definition. Since the resulting vacancy will readily accept a "free" electron the diffused impurities with three valence electrons are called **acceptor atoms**.

The resulting *p*-type material is *electrically neutral*. The same is correct for the *n*-type material.

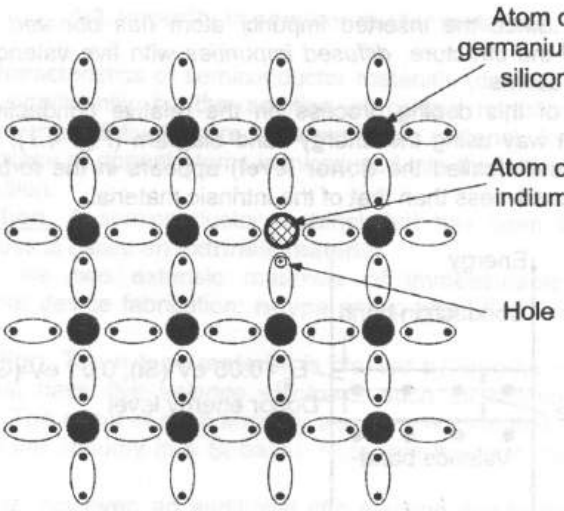


Fig. 4.8. Boron impurity in *p*-type material

4.4. The current direction in the semiconductor materials

The effect of the hole on conduction is shown in Fig. 4.9. If a valence electron acquires sufficient kinetic energy to break its covalent bond and fills the void created by a hole, then a vacancy, or hole, will be created in the covalent bond that released the electron. There is, therefore, a transfer of holes to the left and electrons to the right, as shown in Fig. 4.9. The direction to be used further is that of **conventional flow**, which is indicated by the direction of **hole flow**.

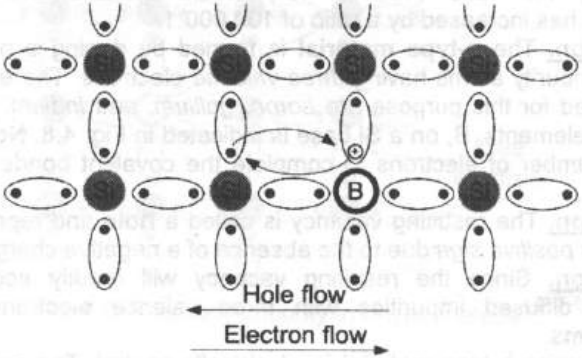


Fig. 4.9. Electron versus hole flow

4.5. Majority and Minority charge carriers

In the intrinsic state, the number of free electrons in Ge or Si is due only to those few electrons in the valence band, that have acquired sufficient energy from thermal or light sources to break the covalent bond or to the *few impurities* that could not be removed. The vacancies left behind in the covalent bonding structure represent very limited supply of holes.

In an *n*-type material, the number of holes has not changed significantly from this intrinsic level. The net result, therefore, is that the *number of electrons far outweighs the number of holes.*

Definition. For this reason in an *n-type material* (Fig. 4.10) *electron* is called the **majority carrier** and *hole* - the **minority carrier**.

For the *p*-type material the number of holes far outweighs the number of electrons, as shown in Fig. 4.10. Therefore in a *p-type material* the *hole* is the **majority carrier** and the *electron* is the **minority carrier**.

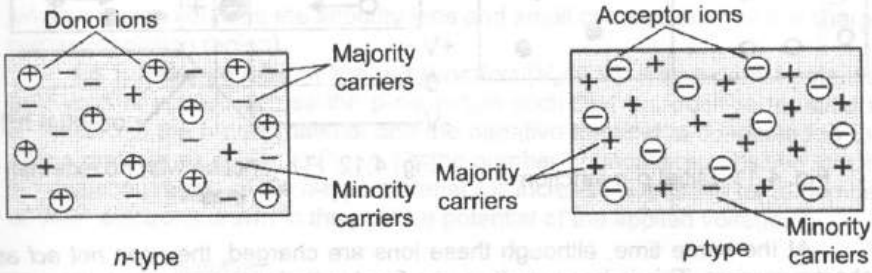


Fig. 4.10. Two semiconductor material types

When the fifth electron of a *donor atom* leaves the parent atom, the atom remaining acquires a *net positive charge*: hence the positive sign in the donor-ion representation. For similar reasons, the *negative sign* appears in the **acceptor ion**.

The *n*- and *p*-type materials represent the basic building blocks of semiconductor devices. We will find in the next section that the "*joining*" of a single *n*-type material with a *p*-type material will result in a semiconductor element of considerable importance in electronic systems.

4.6. *P-n*-junction

Consider now a bar of silicon, doped so that half is *p*-type and half is *n*-type (Fig. 4.11). The bar is not connected to a circuit, so no external electric field is applied to it. Electrons are free to wander at random direction.

Ion appearance. Some of the electrons from the *n*-type material (*major charge carries*) wander across the junction, attracted by the holes in the *p*-type

material. The holes in the region of the junction become filled. In the *p*-type region (to the left of the junction) the filling of holes means that the atoms have, on average, more electrons than they should. The atoms have become **negative ions** (Fig. 4.12). Similarly, in the *n*-type region the atoms have, on average, lost electrons leaving holes and so have become **positive ions** (Fig. 4.12).

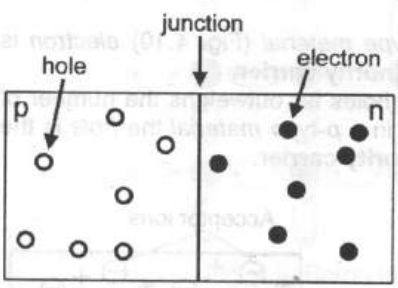


Fig. 4.11. Initial *p-n*-junction

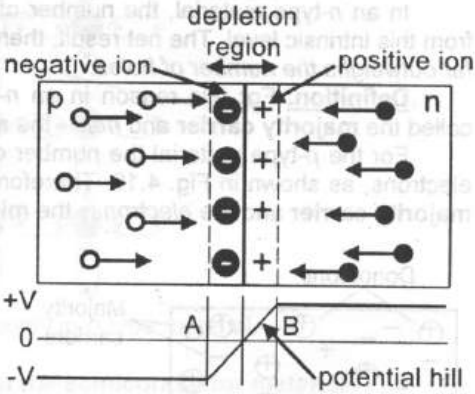


Fig. 4.12. *P-n*-junction with no external bias

At the same time, although these ions are charged, they *can not act as charge carriers*. This is because they are fixed in their places in the lattice. *The result is that the region on side A of the junction has overall negative charge and the region on side B has overall positive charge.*

There is a **potential difference** (p.d.) across the junction. The effect of this p.d. is the same as if a cell was connected across the junction. This is not a real cell but we refer to its effects as a virtual cell. At a silicon *p-n*-junction, the p.d. is about 0.6 V. At a gemlanium junction, it is about 0.2 V.

Definition. As more and more electrons wander across the junction and fill the holes on the other side, the charged regions A and B become wider. Eventually, the negative charge in region A prevents more electrons from being attracted through to the *p*-type material, and the charged region does not become any wider. It contains no charge carriers and, for this reason, it is called the **depletion region**.

We think of the p.d. across the depletion region as a **potential hill**, the potential rising as we pass from the *p*-type side to the *n*-type side. The hill is so steep that electrons can no longer pass down it.

Physical interpretation. There is also more detailed explanation concerning the processes through the *p-n*-junction. At the initial state, there is a

big difference in concentrations of the similar major charge carries (MajCCs) at the both sides of the junction. Due to this fact, MajCCs become flowing through the junction under the **diffusion process**.

Definition. This results in the certain value of **diffusion current** j_d which is directed towards a hole motion.

Due to the motion of MajCCs from the both sides of the junction, positive and negative ions appear. These ions create a **diffusion electrical field** directed from positive to negative ions. This field accelerates minor charge carries (MinCCs) from depletion region and becomes decelerating force for MajCCs.

Definition. The flow of the MinCCs through the junction under the diffusion field creates a **drift current** j_{dr}. The structure comes to the equlibrium state when the value of drift current becomes equal to the value of diffusion current.

Summary. Thus, in the absence of an applied bias voltage, the net flow of charge in any one direction for a semiconductor diode is zero. Then *p-n*-junction region contains the impurity ions and small concentration of the charge carries (depletion region).

4.6.1. Reverse bias of the *p-n*-junction (V_d<0 V). If an external potential of V volts is applied across the *p-n*-junction such that the positive terminal is connected to the *n*-type material and the negative terminal is connected to the *p*-type material as shown in Fig. 4.13, the number of uncovered positive ions in the depletion region of the *n*-type material will increase due to the large number of "free" electrons drawn to the positive potential of the applied voltage.

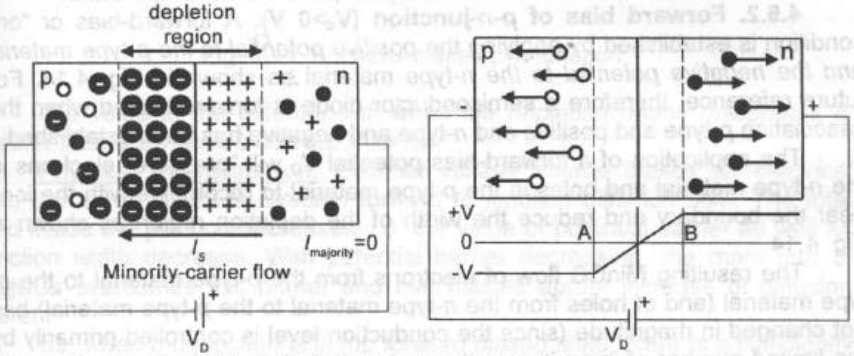


Fig. 4.13. Reverse-biased *p-n*-junction

For similar reasons, the number of uncovered negative ions will increase in the *p*-type material. *The net effect, therefore, is a widening of the depletion region.* This widening of the depletion region will establish too great a barrier for the MajCCs to overcome, effectively reducing the majority carrier flow to zero as shown in Fig. 4.13.

The number of MinCCs, however, that find themselves entering the depletion region will not change, resulting in *minority-carrier flow vectors* of the same magnitude as in the case with no applied voltage.

Definition. The current that exists under reverse-bias conditions is called the **reverse saturation current** and is represented by I_s. The term "*saturation*" comes from the fact that it reaches its maximum level quickly and does not change significantly with increase in the reverse-bias potential.

The reverse saturation current is seldom more than a few microamperes except for high-power devices. In fact, in recent years its level is typically in the nanoampere range for silicon devices and in the low-microampere range for germanium.

Physical interpretation. Reverse bias of *p-n*-junction creates conditions for MinCCs drift, whereas MajCCs *diffusion current is practically absent.*

Such a situation explains by the fact that the field, created by the external source E_{ex} coincides with the direction of *p-n*-junction field E_i created by ions in depletion region. This causes the increasing of potential barrier height for MajCCs, that's why *diffusion current tends to 0.*

At the same time, the net field in the structure equaled to the sum of the E_{ex} and E_i, accelerates the MinCCs and creates a reverse current of *p-n*-junction j_{dr}, namely, the saturation current. It is directed as earlier from high potential to the low one. *Small value of the reverse saturation current is caused by the little concentration of MinCCs.*

Hence, *reverse current of the junction has a drift current character and it is formed by the MinCC motion* and is considered to be a heat flow since either MinCC concentration and saturation current depend on the temperature.

4.6.2. Forward bias of *p-n*-junction (V_d>0 V). A forward-bias or "on" condition is established by applying the *positive potential to the p-type material and the negative potential to the n-type* material as shown in Fig. 4.14. For future reference, therefore a semiconductor diode is forward-biased when the association *p*-type and positive and *n*-type and negative has been established.

The application of a forward-bias potential V_D will "pressure" electrons in the *n*-type material and holes in the *p*-type material to recombine with the ions near the boundary and reduce the width of the depletion region as shown in Fig. 4.14.

The resulting MinCC flow of electrons from the *p*-type material to the *n*-type material (and of holes from the *n*-type material to the *p*-type material) has not changed in magnitude (since the conduction level is controlled primarily by the limited number of impurities in the material), but the reduction in the width of the depletion region has resulted in a heavy majority flow across the junction.

An electron of the *n*-type material now "sees" a reduced barrier at the junction due to the reduced depletion region and a strong attraction for the positive potential applied to the *p*-type material. As the applied bias increases in magnitude the depletion region will continue to decrease in width until a flood of electrons can pass through the junction, resulting in a rise in current.

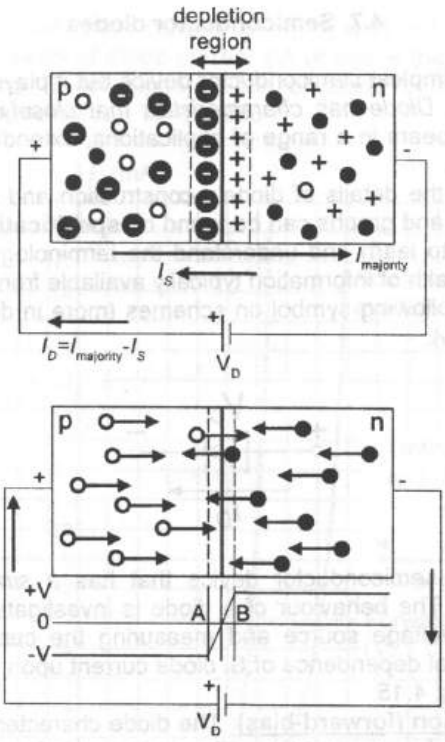


Fig. 4.14. Forward-biased *p-n*-junction

The value of the external p.d. when the depletion region disappears is greater than 0.2 V for Ge and 0.6 V for Si.

Physical interpretation. *External voltage source makes an electrical field inside the structure directed forward to junction diffusion field.* The total field inside the junction decreases, i.e. the value of potential barrier as well as junction width decrease. With potential barrier decreasing, the main part of MajCCs overcomes this barrier and that causes abrupt increase of diffusion current.

So, *when p-n-junction is in the forward biased state, there is a sufficiently high direct current, which has diffusion character and is created by the MajCCs.*

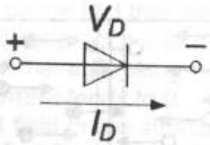
Summary. Thus, the property of unidirectional conduction of *p-n*-junction is caused by the existence of charges of two types: MajCCs (*high concentration*) and MinCCs (*low concentration*), and also the factor which controls the flows of these charges (*ion diffusion field*). *Diffusion field accelerates MinCCs and creates breaking effect on MajCCs.* The certain polarity of external bias source increases or decreases this control factor.

4.7. Semiconductor diodes

Diode is the simplest semiconductor device but it plays a very vital role in electronic systems. *Diode has characteristics that closely match those of a simple switch.* It appears in a range of applications, extending from the simple to the very complex.

In addition to the details of diode's construction and characteristics, the very important data and graphs can be found on **specification sheets**. We will see some of them to learn and understand the terminology employed and to demonstrate the wealth of information typically available from manufacturers.

Diode has a following symbol on schemes (more in details this point will be highlighted below):



Definition. A semiconductor device that has a *single p-n-junction* is known as a **diode**. The behaviour of a diode is investigated by connecting it across a variable voltage source and measuring the current which passes through it. The typical dependence of Si diode current upon the applied voltage is depicted in the Fig. 4.15.

Plot explanation (forward-bias). The diode characteristics coincide with the characteristics of a *p-n*-junction. The right part of the plot shows what happens when the diode is in forward-biased state. When the applied p.d. is in range from zero to 0.6 V (for *Si*), its value is insufficient for electrons to overcome the depletion region of the junction. No current flows.

As the bias is further increased, the depletion region will continue to decrease in width until a flood of electrons can pass through the junction. *Current increases exponentially* (not a straight line) as shown in the forward-bias region of the characteristics, i. e. **the diode does not obey Ohm law**.

Note that the *vertical scale* of the plot is measured *in mA* (although some semiconductor diodes can have a vertical scale measured in amperes) and the *horizontal scale* in the forward-bias region has a *maximum of 1 V*.

Definition. The potential at which the rise of the curve occurs is commonly referred to as the **offset, threshold**, or **firing potential** (the notation V_T).

Approximating formula. It can be demonstrated through the use of solid-state physics that the *general characteristic of a semiconductor diode* can be defined by the following equation for the forward- and reverse-bias regions:

$$I_D = I_S\left(e^{kV_D/T_K} - 1\right), \qquad (4.1)$$

where I_S is a *reverse saturation current*; k=11.6/η with η=1 for Ge and η=2 for Si for relatively low levels of diode current (at or below the knee of the curve) and η=1 for Ge and Si for higher levels of diode current (in the rapidly increasing section of the curve); T_K=T_C+273°.

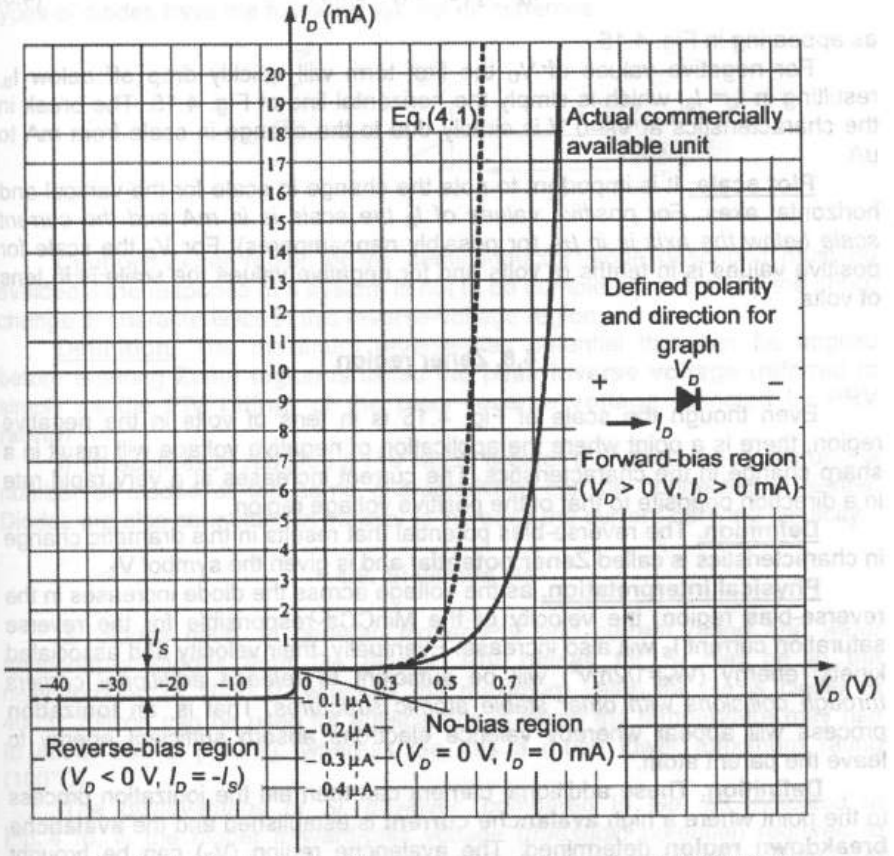


Fig. 4.15. The typical voltage-amperage dependence of Si diode

A plot of Eq. (4.1) is provided in Fig. 4.15. If we expand Eq. (4.1) into the following form,

$$I_D = I_S e^{kV_D/T_K} - I_S \qquad (4.2)$$

the contributing component for each region of Fig. 4.15 can easily be described.

For positive values of V_D the first term of the equation above will grow very quickly and overpower the effect of the second term. The result is that for positive values of V_D, I_D will be positive and grow as the function $y=e^x$.

At V_D=0 V, Eq. (4.1) becomes

$$I_D = I_S\left(e^0 - 1\right) = 0\ mA \tag{4.3}$$

as appearing in Fig. 4.15.

For negative values of V_D the first term will quickly drop off below I_S, resulting in I_D=-I_S, which is simply the horizontal line of Fig. 4.15. The break in the characteristics at V_D=0 V is simply due to the change in scale from mA to μA.

Plot scale. It is important to note the change in scale for the vertical and horizontal axes. *For positive values of I_D the scale is in mA and the current scale below the axis is in μA* (or possibly nanoamperes). For V_D the scale for positive values is in tenths of volts and for negative values the scale is in tens of volts.

4.8. Zener region

Even though the scale of Fig. 4.15 is in tens of volts in the negative region, there is a point where the application of negative voltage will result in a sharp change in the characteristics. The current increases at a very rapid rate in a direction opposite to that of the positive voltage region.

Definition. The reverse-bias potential that results in this dramatic change in characteristics is called **Zener potential** and is given the symbol V_Z.

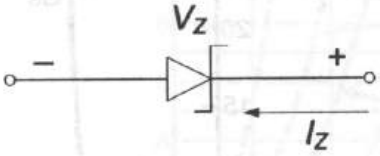
Physical interpretation. as the voltage across the diode increases in the reverse-bias region, the velocity of the MinCCs responsible for the reverse saturation current I_S will also increase. Eventually, their velocity and associated kinetic energy (W_K=1/2mv^2) will be sufficient to *release additional carriers through collisions with other stable atomic structures*. That is, an **ionization** process will appear whereby valence electrons absorb sufficient energy to leave the parent atom.

Definition. These additional carriers can then aid the ionization process to the point where a high **avalanche current** is established and the **avalanche breakdown region** determined. The avalanche region (V_Z) can be brought closer to the vertical axis by increasing the doping levels in the *p*- and *n*-type materials.

Definition. However, as V_Z decreases to very low levels, such as -5 V, another mechanism, called **Zener breakdown**, will contribute to the sharp change in the characteristic. It occurs because there is a strong electric field in the region of the junction that can disrupt the bonding forces within the atom and "generate" carriers.

Although Zener breakdown mechanism is a significant contributor only at lower levels of V_Z, this sharp change in the characteristic at any level is called Zener region.

Definition. Diodes employing this unique portion of the characteristic of a *p-n*-junction are called **Zener diodes or voltage-reference diodes**. These types of diodes have the following symbol on schemes:



Caution. Zener region of the semiconductor diode described must be avoided if the response of a system is not to be completely altered by the sharp change in characteristics in this reverse-voltage region.

Definition. The maximum reverse-bias potential that can be applied before entering Zener region is called the **peak inverse voltage** (referred to simply as the **PIV rating**) or the **peak reverse voltage** (denoted by **PRV rating**).

If an application requires a PIV rating greater than that of a single unit, a number of diodes of the same characteristics can be connected in series. Diodes are also connected in parallel to increase the current-carrying capacity.

4.9. Si versus Ge

Si diodes have, in general, higher PIV and current rating and wider temperature ranges than Ge diodes. PIV ratings for *Si* can be in the neighborhood of 1000 V, whereas the maximum value for *Ge* is closer to 400 V. Silicon can be used for applications in which the temperature may rise to about 200°C (400°F), whereas *Ge* has a much lower maximum rating (100°C).

The *disadvantage of Si*, however, compared to *Ge*, as indicated in Fig. 4.16, is the higher forward-bias voltage (*threshold potential*) required to reach the region of upward swing. It is typically of the order of magnitude of 0.7 V for commercially available *Si* diodes and 0.3 V for *Ge* diodes.

Generally, the whole characteristics of *Si* as compared to *Ge* still make it the choice in the majority of commercially available units.

Temperature effects. One more reason for choosing the *Si* material is the temperature influence on the diode characteristics. The common saturation current for a *Ge* diode is in the order of 1 or 2 μA at 25°C (see Fig. 4.16) and about 100 μA at 100°C. The last value is high enough for considering open-circuit state in the reverse-bias region.

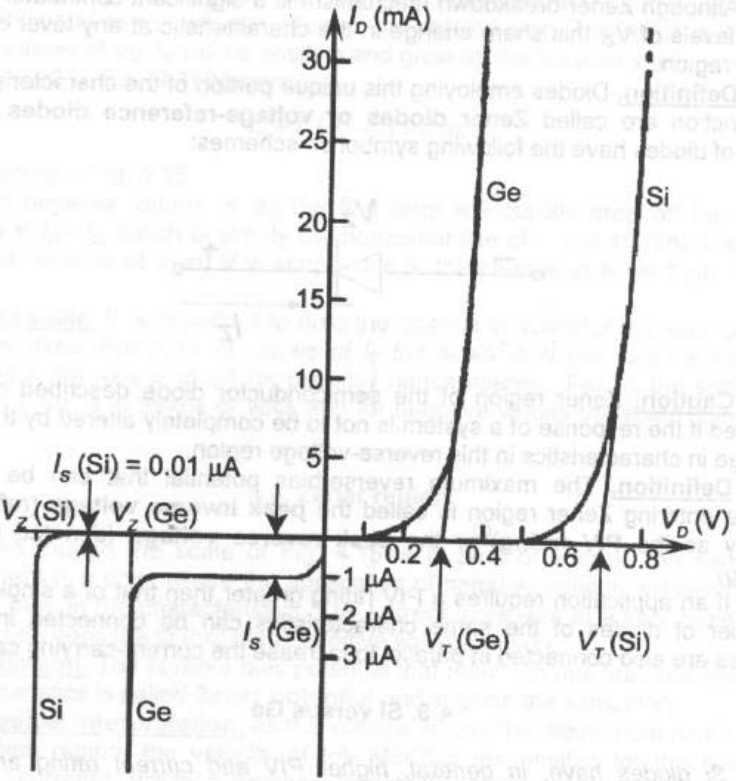


Fig. 4.16. Comparison of *Si* and *Ge* semiconductor diodes

At the same time, typical values of reverse current for *Si* diodes are much lower than that of *Ge* for similar power and voltage levels (see Fig. 4.16). Moreover, even at high temperatures the level of saturation current for *Si* diodes do not reach the same levels obtained for *Ge* (compare 1-2 μA for *Si* and 100μA for *Ge* diodes).

The breakdown voltage is also increasing with temperature (Fig. 4.17). It has been found experimentally that the reverse saturation current will just about double in magnitude for every 10°C increase in temperature.

As to the forward-bias region, the forward characteristics are actually becoming more "ideal," but generally *the temperatures beyond the normal operating range can have a very detrimental effect on the diode's maximum power and current levels.*

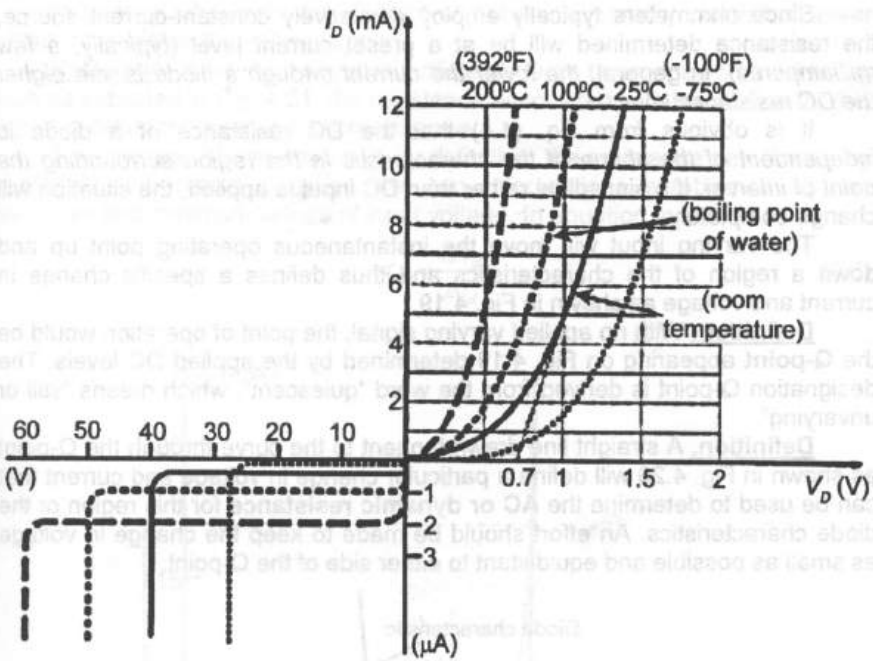


Fig. 4.17. Variation in *Si* diode characteristics with temperature change

4.10. Main diode characteristics. Resistance types

As the operating point of a diode moves from one region to another the resistance of the diode will also change due to the nonlinear shape of the characteristic curve. The type of applied voltage or signal defines the resistance level of interest.

There are three different resistance levels. Consider the first one – **DC or Static Resistance**.

The application of a DC voltage to a circuit containing a semiconductor diode will result in an operating point on the characteristic curve that will not change with time. The resistance of the diode at the operating point can be found simply by finding the corresponding levels of V_D and I_D (Fig. 4.18):

$$R_D = \frac{V_D}{I_D}. \tag{4.4}$$

The DC resistance levels at the knee and below will be *greater* than the resistance levels obtained for the vertical rise section of the characteristics. The resistance levels in the reverse-bias region will naturally be quite high.

Since ohmmeters typically employ a relatively constant-current source, the resistance determined will be at a preset current level (*typically, a few milliamperes*). In general, *the lower the current through a diode is, the higher the DC resistance level.*

It is obvious from Eq. (4.4) that the DC resistance of a diode is *independent of the shape of the characteristic in the region surrounding the point of interest.* If a sinusoidal rather than DC input is applied, the situation will change completely.

The varying input will move the instantaneous operating point up and down a region of the characteristics and thus defines a specific change in current and voltage as shown in Fig. 4.19.

Definition. With no applied varying signal, the point of operation would be the **Q-point** appearing on Fig. 4.19 determined by the applied DC levels. The designation Q-point is derived from the word "quiescent", which means "still or unvarying".

Definition. A straight line drawn tangent to the curve through the Q-point as shown in Fig. 4.20 will define a particular change in voltage and current that can be used to determine the **AC or dynamic resistance** for this region of the diode characteristics. An effort should be made to keep the change in voltage as small as possible and equidistant to either side of the Q-point.

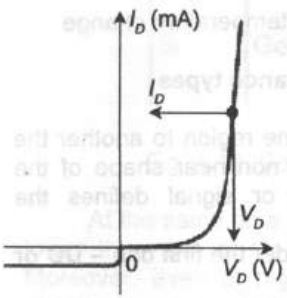


Fig. 4.18

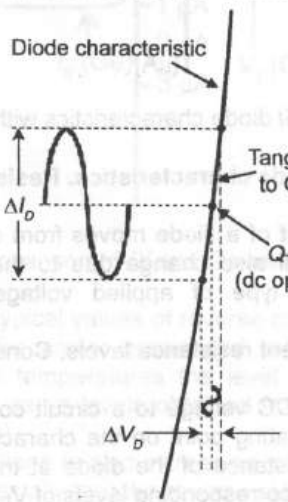


Fig. 4.19

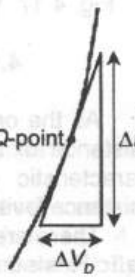


Fig. 4.20

In equation form:

$$r_d = \frac{\Delta V_d}{\Delta I_d} \qquad (4.5)$$

In general, *the lower the Q-point of operation (smaller current or lower voltage) the higher the AC resistance.*

Definition. If the input signal is sufficiently large to produce a broad swing such as indicated in Fig. 4.21, the resistance associated with the device for this region is called the **average AC resistance**.

The average AC resistance is, by definition, the resistance determined by a straight line drawn between the two intersections established by the maximum and minimum values of input voltage. In equation form:

$$r_{av} = \frac{\Delta V_d}{\Delta I_d}. \qquad (4.6)$$

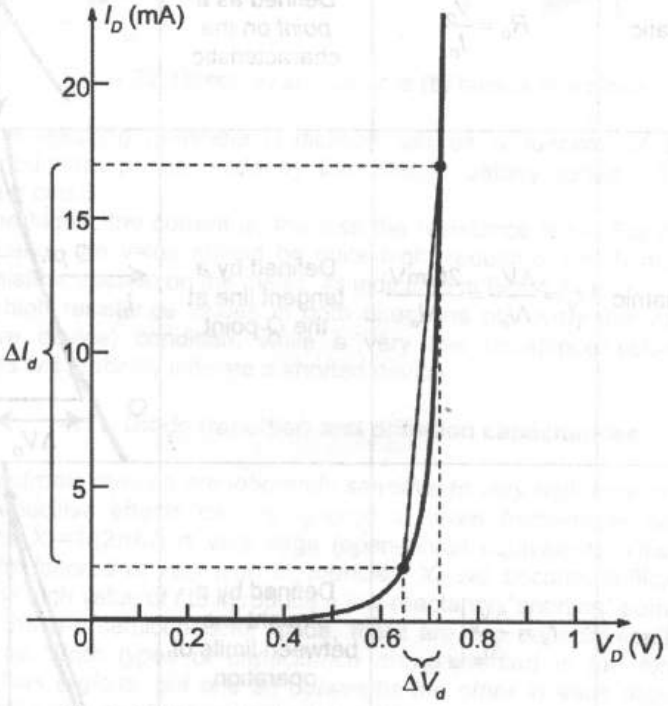


Fig. 4.21. Determining the average AC resistance between indicated limits

For situation indicated in Fig. 4.21:

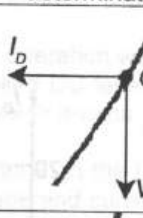
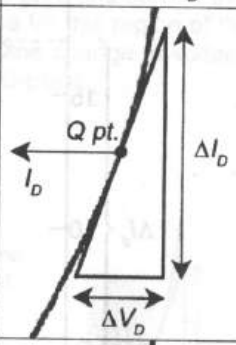
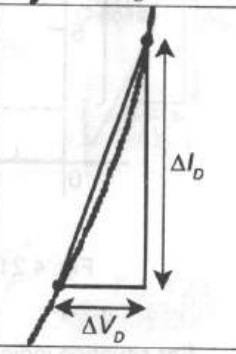
$$\Delta I_d = 15\ mA,\ \Delta V_d = 0.075\ V$$

$$r_{av} = 0.075\ V / 15\ \text{mA} = 5\ \Omega$$

Equation (4.6) defines a value that is considered the average of the AC values from 2 to 17 mA.

The fact that one resistance level can be used for such a wide range of the characteristics is quite useful in the implementation of diode equivalent circuits. Thus, *the lower the level of currents used to determine the average resistance the higher the resistance level* (see also Table 4.1).

Table 4.1. Summary table on resistance parameters

Type	Equation	Special characteristics	Graphical determination
DC or static	$R_D = \frac{V_D}{I_D}$	Defined as a *point* on the characteristic	
AC or dynamic	$r_d = \frac{\Delta V_d}{\Delta I_d} = \frac{26\text{ mV}}{\Delta I_d}$	Defined by a tangent line at the *Q*-point	
Average ac	$r_d = \frac{\Delta V_d}{\Delta I_d}\Big\|_{\text{pt. to pt.}}$	Defined by a straight line between limits of operation	

Diode testing. Based on the described resistance properties of the diode, there is a way for checking the diode condition with the **ohmmeter**.

We found that the forward-bias resistance of a semiconductor diode is quite low compared to the reverse-bias level. Therefore, if we measure the resistance of a diode using the connections indicated in Fig. 4.22,a, we can expect a relatively low level.

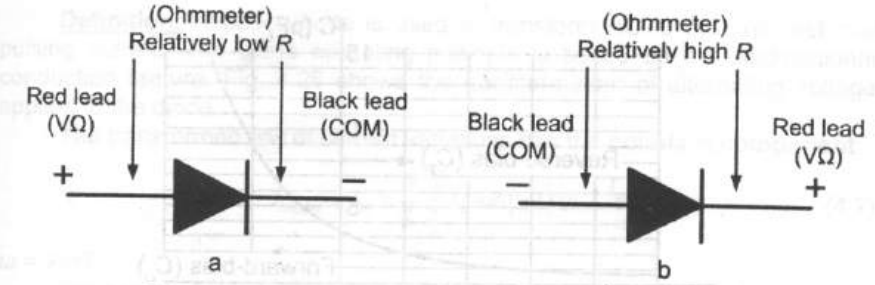


Fig. 4.22. Direct (a) and reverse (b) biases of a diode

The resulting ohmmeter indication will be a function of the current established through the diode by the internal battery (often 1.5 V) of the ohmmeter circuit.

The higher the current is, the less the resistance level. For the reverse-bias situation the value should be quite high, requiring a high measurement limit (resistance scale) on the meter, as indicated in Fig. 4.22,b.

A high resistance values in both directions obviously indicate an open (defective device) condition, while a very low resistance values in both directions will probably indicate a shorted device.

4.11. Diode transition and diffusion capacitances

Electronic devices are inherently sensitive to very high frequencies. Most shunt capacitive effects can be ignored at lower frequencies because the reactance $X_C=1/(2\pi fC)$ is very large (open-circuit equivalent). *This, however, cannot be ignored at very high frequencies.* X_C will become sufficiently small due to the high value of f to introduce a low-reactance "shorting" path.

In the *p-n* semiconductor diode, there are *two capacitive effects* to be considered. *Both types of capacitance are presented in the forward- and reverse-bias regions, but one so outweighs the other in each region that we consider the effects of only one in each region.*

Definition. In the reverse-bias region there is a **transition- or depletion-region capacitance** (C_T), while in the forward-bias region we have the **diffusion (C_D) or storage capacitance**.

Recall that the basic equation for the capacitance of a parallel-plate capacitor is defined by $C=\varepsilon A/d$, where ε is the permittivity of the dielectric (insulator) between the plates of area A separated by a distance d.

In the reverse-bias region there is a depletion region (free of carriers) that behaves essentially like an insulator between the layers of opposite charge. Since the depletion width (*d*) will increase with increased reverse-bias potential, the resulting transition capacitance will decrease, as shown in Fig. 4.23.

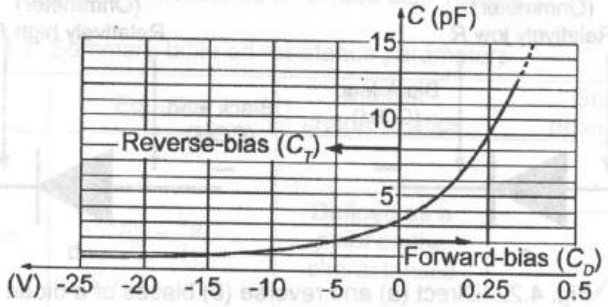


Fig. 4.23. Transition and diffusion capacitance versus applied bias for a Si diode

The fact that the capacitance is dependent on the applied reverse-bias potential has application in a number of electronic systems.

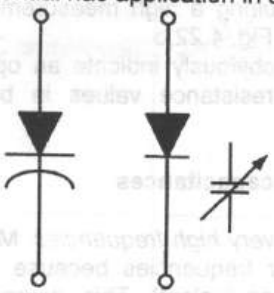


Fig. 4.24

Definition. A diode whose operation is wholly dependent on this phenomenon is called ***varicap*** (also called ***varactor, VVC*** - voltage-variable capacitor or *tuning*) that means semiconductor, voltage-dependent, variable capacitors. This type of diodes is depicted on the schemes as in Fig. 4.24.

Although the effect described above will also be present in the forward-bias region, it is overshadowed by a capacitance effect directly dependent on the rate at which charge is injected into the regions just outside the depletion region. The result is that increased levels of current will result in increased levels of *diffusion capacitance*.

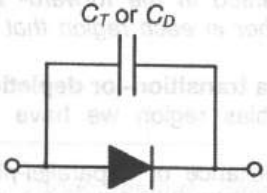


Fig. 4.25

However, increased levels of current result in reduced levels of associated resistance, and the resulting time constant (τ=RC), which is very important in high-speed applications, does not become excessive.

The capacitive effects described above are represented by a capacitor in parallel with the ideal diode, as shown in Fig. 4.25. For

low- or mid-frequency applications (except in the power area), however, the capacitor is normally not included in the diode symbol.

4.12. Rectifier diodes

Definition. Rectifier diode is used to transform AC to the unidirectional pulsing current. The diode operating principle is based on the unidirectional conduction feature. Fig. 4.26 shows the transformation of alternating voltage applied to the diode.

The transformed law of current variations has the **constant component**:

$$I_{MEAN} = I_M = \frac{1}{T}\int_0^{T/2} I_m \sin(\omega t)\,dt = \frac{I_m}{\pi}, \qquad (4.7)$$

$\omega = 2\pi/T$.

Because of this the aforementioned transformation is known as **rectification**.

Half-wave rectifier scheme. The scheme has a rectifier diode VD, load resistor R and filtering capacitor C (Fig. 4.27).

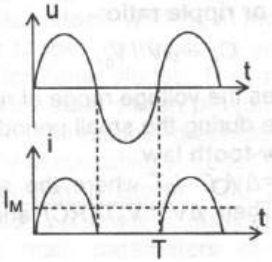


Fig. 4.26

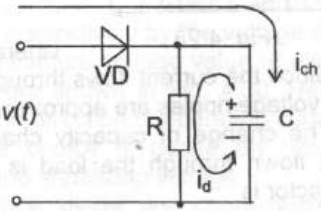


Fig. 4.27

Diode transforms input alternating voltage to the impulse voltage V_R. In case of capacitor absence rectified voltage can be selected on the load resistance. *The main effect of capacitor application is the smoothing of output voltage V_{RC}.* In other words, this means that capacitor filters or selects the constant component V_0. During the time interval t voltage applied to the capacitor exceeds the instant value of input voltage ($V_C > V(t)$), therefore the diode is closed.

Capacitor is slowly discharged through the resistor. As soon as input voltage exceeds the capacitor voltage ($V(t) > V_C$), the diode is being opened and capacitor charge process starts and lasts during the time t_{CH}. In this case diode lets through short impulses of charge current i_{CH} of duration t_{CH}. Note that charge impulse current i_{CH} is much greater than the average value of rectified

current I_{AV} since $i_{CH}t_{CH} = I_{AV}T$ according to the charge equality condition (Fig. 4.28).

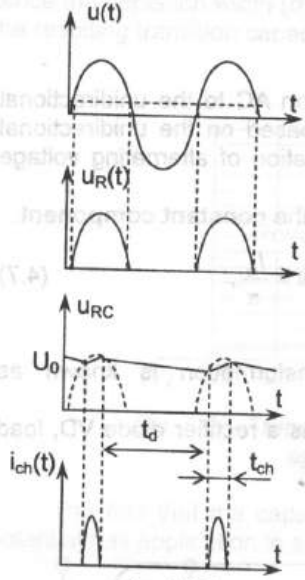


Fig. 4.28

In order to increase the effectiveness of rectified voltage smoothing, the load resistor should be short-circuited by capacitor in respect to the variable component. That is the ratio between capacitive and active resistances should satisfy the following inequality:

$$X = \frac{1}{\omega C} << R. \qquad (4.8)$$

Taking in mind $\omega = 2\pi/T$ and suppose that the 10 times difference in resistance values is enough, we get the equivalent parity:

$$\tau_p = RC = 10T / 2\pi > T. \qquad (4.9)$$

Hence, the capacitor discharge time constant should enlarge the input voltage cycle.

Definition. The main parameter for defining the quality of rectifier scheme is known as **ripple factor or ripple ratio**:

$$C_R = \Delta V / V_0, \qquad (4.10)$$

where ΔV denotes the voltage range of ripples.

Since the current flows through the diode during the small period of time, output voltage ripples are approximated by **saw-tooth law**.

The change of capacity charge is $\Delta q = \Delta V\ C = I_0 T$ where the averaged current flown through the load is $I_0 = V_0/R$. Then $\Delta V = V_0 T/(RC)$ and hence ripple factor is

$$C_R = \frac{T}{\tau_p} = \frac{1}{fRC}. \qquad (4.11)$$

4.13. Voltage-reference diodes

Definition. **Voltage-reference diodes** (VRD) - semiconductor devices which use high slope sections of voltage-ampere characteristic (VAC), i. e. ones with weak dependence of voltage on the current variations.

Diode purpose. VRD implements a voltage stabilization process, i. e. maintaining substantially constant voltage value in a wide range of current variations. The diode VAC has two sections with a high slope for both forward and reverse bias (Fig. 4.29).

The diode operates in the stabilization area if:

1. The direct voltage V_D exceeds a certain threshold value $V_{FWD} > V_{THR}$.

2. The reverse voltage V_{REV} is smaller than the breakdown voltage $V_{REV}<V_{BDN}$.

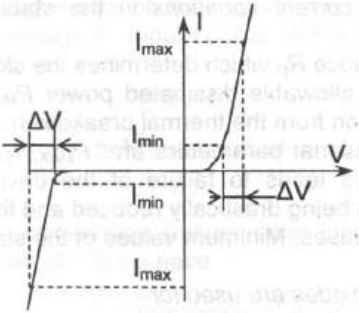


Fig. 4.29

There are two kinds of reference diodes:

1. **Zener diode** (or **stabilitron**) which is using a section with reverse-biased voltages smaller than the electrical breakdown value (reverse bias $V_{REV}<V_{BDN}$; V_{BDN} equals -3 V, for instance).

2. **Stabistors**, working in the area of the forward branch diode.

The ranges of stabilized voltage are V_{ST}=0.7 ... 2 V for stabistors and V_{ST}<-3 V for Zener diodes. They are defined respectively by the voltage drop in the forward-biased silicon *p-n*-junction and minimum voltage value of electrical breakdown. Conditional graphical notations are shown in Fig. 4.30. Here are the polarity of operating voltages as well as the direction of operating currents.

Selecting a diode. The initial diode choice is based on two main classification parameters: *dissipated power* and *stabilizing voltage*.

The main parameters of the reference diode are characterized, in particular, by the "geometric" VAC features (Fig. 4.31).

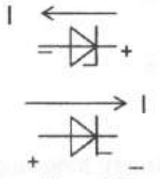


Fig. 4.30

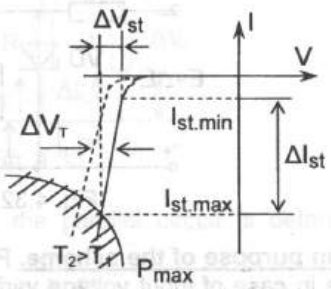


Fig. 4.31

These parameters include:

1. *Nominal (mean) value of the stabilization voltage* or a range of its changes ($V_{ST.MIN}$ ... $V_{ST.MAX}$).

2. *The range of current variations* in the stabilization area ($I_{ST.MIN}$ ... ... $I_{ST.MAX}$).

3. *Dynamic resistance* R_D which determines the slope of characteristic.

4. The maximum allowable dissipated power $P_{MAX}=I_{ST.MAX}\ V_{ST.MAX}$ which protects the diode junction from the thermal breakdown.

The limiting operational parameters are: P_{MAX}, $I_{ST.MAX}$, $I_{ST.MIN}$. Exceeding the first two parameters leads to failure of the device. When $I<I_{ST.MIN}$ the steepness of the VAC is being drastically reduced and the level of noise on pre-breakdown section increases. Minimum values of the stabilization current $I_{ST.MIN}$ are within 0.5 ... 3 mA.

General-purpose diodes are used for:

1. Getting a fixed (reference) voltage.

2. DC stabilization in a certain range of current changes.

3. In AC circuits - to limit the instantaneous voltage values, i. e. to obtain constant voltage in a small time interval.

4. In pulse circuits as high-speed cut-off items, as these diodes have no injection accumulation processes.

5. Pulse reference diodes are used to limit (commit) the amplitude of the pulse signal. In such diodes junction capacitance is minimized and this fact allows receiving short duration of transient processes.

4.14. Parametric voltage stabilizer

The voltage stabilization principle is disclosed on the basis of the simple circuit (Fig. 4.32). It has two main elements: a stabilitron VD and ballast resistor R_B. Load resistor R_L is a part of the circuit for stabilizing the voltage V_L.

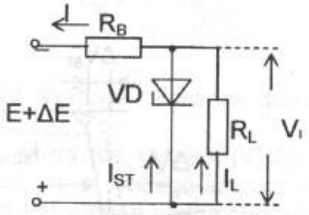


Fig. 4.32

The main purpose of the scheme. Provides practical immutability of the output voltage in case of input voltage variations ΔE with respect to the rated voltage E. It should be noted that the diode is connected in parallel with the load resistor, as in this case voltage values on the two elements coincide.

The ballast resistance performs a dual role:

1. In the **static mode** it limits the current value running through the diode.
2. In the **dynamic mode** it provides the effect of stabilizing.

Static mode. Static mode of the circuit operation is analyzed at a constant nominal input voltage E. Input source emf is distributed between the load and ballast resistors:

$$E = V_L + V_B \tag{4.12}$$

and the total current I is divided between the diode and load resistor:

$$I = I_{ST} + I_L. \tag{4.13}$$

Let's establish the relationship between the input and output voltages. Then, based on (4.12) with (4.13) we have:

$$E = V_L + (I_{ST} + I_L)R_B. \tag{4.14}$$

Expressing the currents through the output voltage, we get:

$$E = V_L + \left(\frac{V_L}{R_{ST}} + \frac{V_L}{R_L}\right)R_B \tag{4.15}$$

or finally

$$E = U_L\left(1 + \frac{R_B}{R_{ST}} + \frac{R_B}{R_L}\right), \tag{4.16}$$

where R_{ST} is the diode's static resistance.

Dynamic mode. let's replace the diode with dynamic resistance R_D. Then the equivalent circuit with respect to the variations of input voltage is shown in Fig. 4.33.

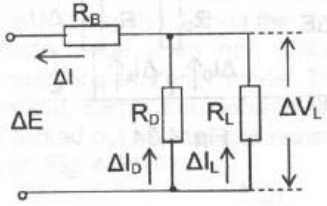


Fig. 4.33

The equivalent resistance of the parallel circuit is determined by the formula

$$R_E = \frac{R_D R_L}{R_D + R_L} \tag{4.17}$$

and since $R_D << R_L$, the approximate equality is right:

$$R_E = \frac{R_D}{\frac{R_D}{R_L}+1} \approx R_D. \qquad (4.18)$$

In this scheme the principle of voltage divider with respect to the input voltage changes is implemented, i. e., ΔE is distributed between R_B and R_D. Since $R_D << R_B$, the main voltage increment falls on the ballast resistor R_B and the output voltage changes slightly:

$$\Delta V_L = \Delta E \frac{R_D}{R_B + R_D} \approx \Delta E \frac{R_D}{R_B}, \qquad (4.19)$$

i. e. $\Delta V_L << \Delta E$.

Hence, the ballast resistor's purpose and name become obvious - it plays the role of ballast on which the main voltage change is dropping, so only a small part of changing voltage is released on the load resistor.

In this circuit the current increments divider is also implemented as the auxiliary effect. Total current change, which arose because of the instability of the input voltage, is divided between the parallel elements R_D and R_L. Since $R_D << R_L$, the main change in the current passes through the diode $\Delta I_D >> \Delta I_L$ and therefore $\Delta V_L = I_L R_L$ will be also small. It follows from the equivalent circuit (Fig. 4.34) that

$$\Delta V_L = \Delta I \frac{R_D R_L}{R_D + R_L}. \qquad (4.20)$$

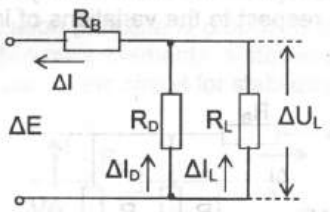


Fig. 4.34

Then

$$\Delta I_L = \Delta I \frac{R_D}{R_D + R_L} \approx \Delta I \frac{R_D}{R_L}, \text{ i. e. } \Delta I_L < \Delta I. \qquad (4.21)$$

In general, the following ratio according to the second Kirchhoff's law can be written for the dynamic mode:

$$\Delta E = \Delta V_L + (\Delta I_{ST} + \Delta I_L) R_B \qquad (4.22)$$

or by analogy with (4.16)

$$\Delta E = \Delta V_L \left(1 + \frac{R_B}{R_L} + \frac{R_B}{R_D}\right). \qquad (4.23)$$

Definition. The stabilization quality of the whole circuit is determined by the **stabilization coefficient**:

$$C_{ST} = \frac{\Delta E/E}{\Delta V_L/V_L} \qquad (4.24)$$

which shows how many times the relative changes in the input voltage are greater than the output variations.

Using the relationship between input and output voltages in static (4.16) and dynamic (4.23) modes, we get

$$C_{ST} = \frac{1 + \frac{R_B}{R_D} + \frac{R_B}{R_L}}{1 + \frac{R_B}{R_{ST}} + \frac{R_B}{R_L}}. \qquad (4.25)$$

Thus, the realization of the circuit implemented the principle of voltage stabilization on the basis of nonlinear voltage divider is made. The diode connected in one arm of the divider, as a non-linear element, has considerable differences in two parameters: static and dynamic resistances ($R_{ST} >> R_D$). Because of this, it keeps a relatively high static voltage at a small value of the voltage drop with respect to its variations.

4.15. Impulse limiters

Circuit 1. *Provides a pulse amplitude selection* (Fig. 4.35). If the pulse amplitude is less than the limit threshold, then the circuit is in open state. In this case, the pulse amplitude value does not sufficient to provide electrical breakdown of a diode operating in Zener mode. The pulses with amplitudes exceeding the limit threshold, switch the diode to breakdown mode. Then the "tops" of the pulses are picked out on the load resistor (excluding voltage drop on the diode) (see plots on Fig. 4.35).

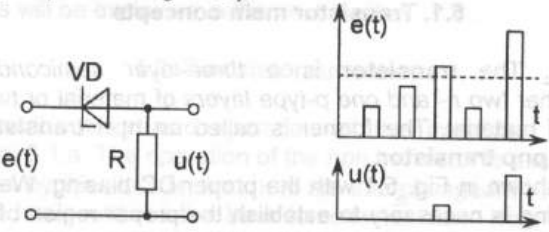


Fig. 4.35

Circuit 2. *Fixes the pulse amplitude* (Fig. 4.36). If stabilitron's limit threshold is less than any of pulses' amplitude, the diode operates in the breakdown mode. Then the reference pulse amplitude is supported on the diode (see plots on Fig. 4.36).

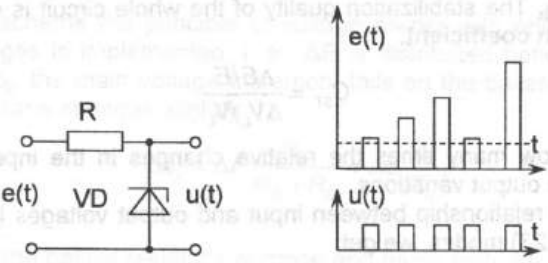


Fig. 4.36

5. BIPOLAR JUNCTION TRANSISTOR

During the period 1904–1947, the vacuum tube was undoubtedly the electronic device of interest and developm Fig. 4.36 r 23, 1947, however, the electronics industry was to experi a completely new direction of interest and development. It was on the afternoon of this day that **Walter H. Brattain** and **John Bardeen** demonstrated the amplifying action of the first transistor at the Bell Telephone Laboratories.

The advantages of this three-terminal solid-state device over the tube were immediately obvious:

- it was smaller and lightweight;
- had no heater requirement or heater loss;
- had rugged construction;
- and was more efficient since less power was absorbed by the device itself;
- it was instantly available for use, requiring no warm-up period;
- and lower operating voltages were possible.

5.1. Transistor main concepts

Definition. The **transistor** is *a three-layer semiconductor device* consisting of either *two n- and one p-type layers* of material *or two p- and one n-type layers* of material. The former is called an **npn transistor**, while the latter is called a **pnp transistor**.

Both are shown in Fig. 5.1 with the proper DC biasing. We will find later that the dc biasing is necessary to establish the proper region of operation for ac amplification.

Definition. The three layers of the transistor are known as the **collector**, the **base** and the **emitter**.

Accordingly the terminals in Fig. 5.1 have been indicated by the capital letters E for emitter, C for collector, and B for base. The *emitter* layer is *heavily doped*, the *base lightly doped*, and the *collector* only *lightly doped*.

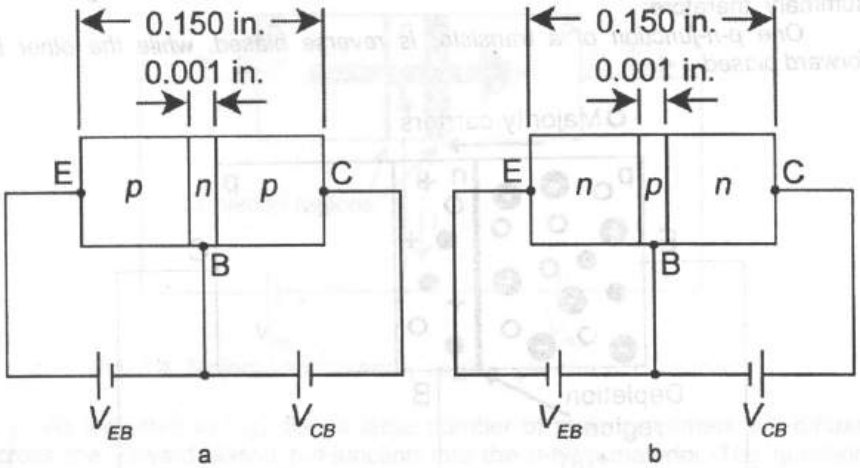


Fig. 5.1. Types of transistors: *pnp* (a); *npn* (b)

The outer layers have widths much greater than the sandwiched p- or n-type material. For the transistors shown in Fig. 5.1 the ratio of the total width to that of the center layer is 0.150/0.001=150:1.

Note. The doping of the sandwiched layer is also considerably less than that of the outer layers (typically, 10:1 or less). This lower doping level decreases the conductivity (increases the resistance) of this material by limiting the number of "free" carriers.

The fact that *conduction occurs through all three layers, which means that it involves both electrons (negative) and holes (positive) as charge carriers*, is why these are sometimes called *bipolar transistors*. Their full name is **bipolar junction transistors** (BJT) because their action depends on the properties of a *p-n*-junction, as will be explained in the next section.

5.2. Transistor operation

Let's describe the basic operation of the transistor using the *pnp* transistor of Fig. 5.1,a. The operation of the *npn* transistor is exactly the same if the roles played by the electron and hole are interchanged. In Fig. 5.2 the *pnp* transistor is redrawn without the base-to-collector bias. Note the similarities between this situation and that of the forward-biased diode. The depletion

region has been reduced in width due to the applied bias, resulting in a heavy flow of majority carriers from the *p*- to the *n*-type material.

Let us now remove the base-to-emitter bias of the *pnp* transistor of Fig. 5.1,a as shown in Fig. 5.3. Consider the similarities between this situation and that of the reverse-biased diode. Recall that the flow of majority carriers is zero, resulting in only a minority-carrier flow, as indicated in Fig. 5.3. In summary, therefore:

One p-n-junction of a transistor is reverse biased, while the other is forward biased.

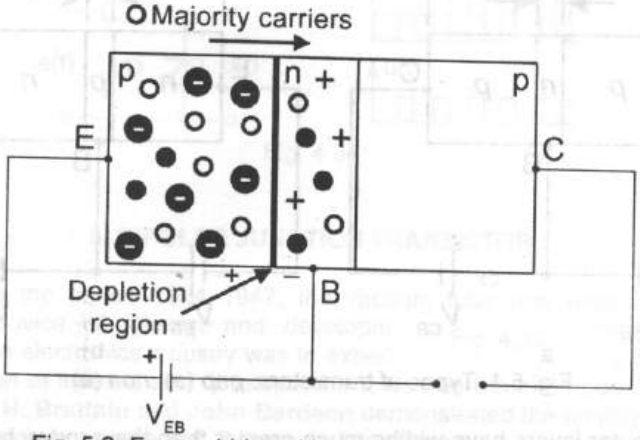


Fig. 5.2. Forward-biased junction of a *pnp* transistor

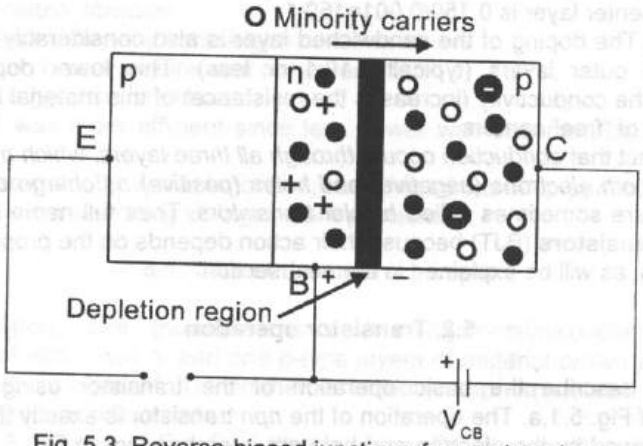


Fig. 5.3. Reverse-biased junction of a *pnp* transistor

In Fig. 5.4 both biasing potentials have been applied to a *pnp* transistor, with the resulting majority- and minority-carrier flow indicated. *Note in Fig. 5.4 the widths of the depletion regions, indicating clearly which junction is forward-biased and which is reverse-biased.*

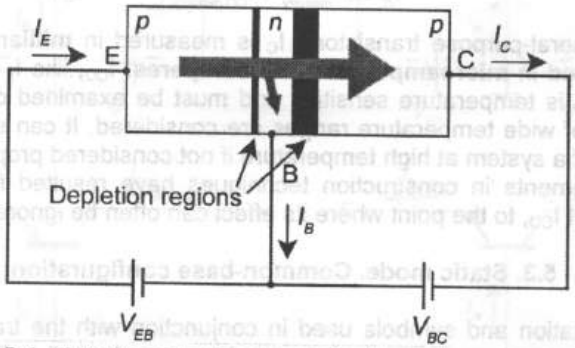


Fig. 5.4. Majority and minority carrier flow of a pnp transistor

As indicated in Fig. 5.4, a large number of majority carriers will diffuse across the forward-biased *p-n*-junction into the *n*-type material. The question then is whether these carriers will contribute directly to the base current I_B or pass directly into the *p*-type material?

Since the sandwiched *n*-type material is very thin and has a low conductivity, a very small number of these carriers will take this path of high resistance to the base terminal.

The magnitude of the base current is typically on the order of microamperes as compared to milliamperes for the emitter and collector currents. The larger number of these majority carriers will diffuse across the reverse-biased junction into the *p*-type material connected to the collector terminal as indicated in Fig. 5.4.

The reason for the relative ease with which the majority carriers can cross the reverse-biased junction is easily understood if we consider that for the reverse-biased diode the injected majority carriers will appear as minority carriers in the *n*-type material. In other words, there has been an injection of minority carriers into the *n*-type base region material. Combining this with the fact that all the minority carriers in the depletion region will cross the reverse-biased junction of a diode accounts for the flow indicated in Fig. 5.4.

Applying Kirchhoff current law to the transistor of Fig. 5.4 as if it were a single node, we obtain:

$$I_E = I_C + I_B \tag{5.1}$$

and find that *the emitter current is the sum of the collector and base currents.*

Definition. The **collector current**, however, is comprised of two components - the majority and minority carriers as indicated in Fig. 5.4. The minority-current component is called the **leakage current** and is given the symbol I_{CO} (I_C current with emitter terminal Open). The collector current, therefore, is determined in total by the equation:

$$I_C = I_{C\ majority} + I_{CO\ minority}. \tag{5.2}$$

For general-purpose transistors, I_C is measured in **milliamperes**, while I_{CO} is measured in **microamperes** or **nanoamperes**. I_{CO}, like I_s for a reverse-biased diode, is temperature sensitive and must be examined carefully when applications of wide temperature ranges are considered. It can severely affect the stability of a system at high temperature if not considered properly.

Improvements in construction techniques have resulted in significantly lower levels of I_{CO}, to the point where its effect can often be ignored.

5.3. Static mode. Common-base configuration

The notation and symbols used in conjunction with the transistor in the majority of texts and manuals published today are indicated in Fig. 5.5 for the common-base configuration with *pnp* and *npn* transistors.

Definition. The **common-base terminology** is derived from the fact that the *base is common to both the input and output sides of the configuration.* In addition, *the base is usually the terminal closest to, or at, ground potential.* Remember that we refer to conventional (hole) flow rather than electron flow.

For the transistor the arrow in the graphic symbol defines the direction of emitter current (conventional flow) through the device.

All the current directions appearing in Fig. 5.5 are the actual directions as defined by the choice of conventional flow. Note in each case that $I_E=I_C+I_B$. Note also that the *applied biasing (voltage sources) are such as to establish current in the direction indicated for each branch.*

To fully describe the behavior of a three-terminal device such as the common-base amplifiers of Fig. 5.5 *requires two sets of characteristics - one for the driving point or input parameters and the other for the output side.*

The *input set* for the common-base amplifier as shown in Fig. 5.6 will relate an input current (I_E) to input voltage (V_{BE}) for various levels of output voltage (V_{CB}).

The *output set* will relate an output current (I_C) to an output voltage (V_{CB}) for various levels of input current (I_E) as shown in Fig. 5.7. The output or collector set of characteristics has three basic regions of interest, as indicated in Fig. 5.7: **the active, cutoff, and saturation regions.**

Definition. The **active region** is the region normally employed for linear (undistorted) amplifiers. In particular, *in the active region the collector-base junction is reverse-biased, while the base-emitter junction is forward-biased.*

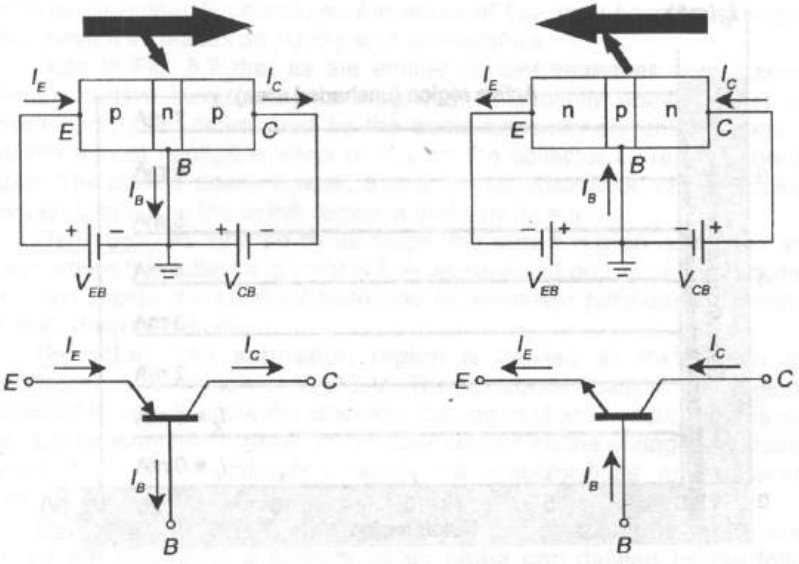


Fig. 5.5. Notation and symbols used with the common-base configuration: *pnp* transistor (a); *npn* transistor (b)

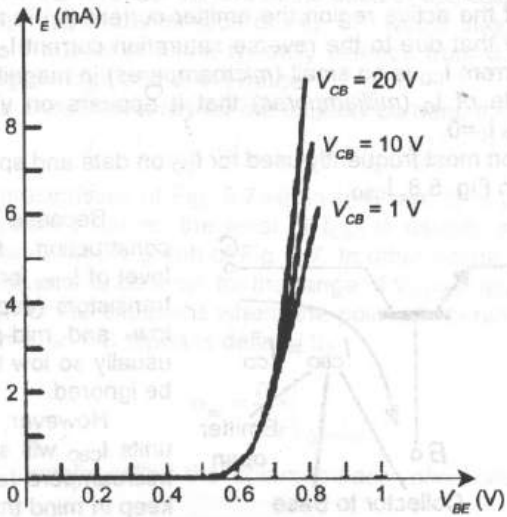


Fig. 5.6. Input or driving point characteristics for a common-base silicon transistor amplifier

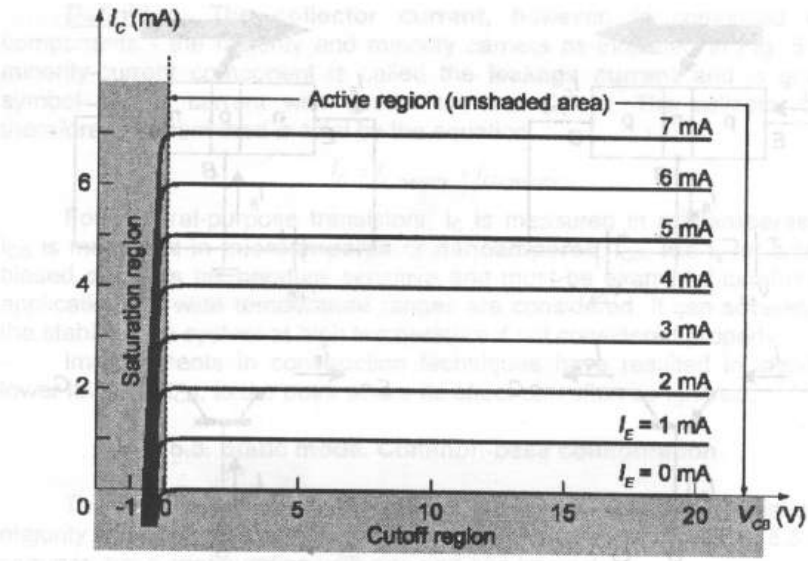


Fig. 5.7. Output or collector characteristics for a common-base transistor amplifier

The active region is defined by the biasing arrangements of Fig. 5.5. At the lower end of the active region the emitter current (I_E) is zero, the collector current is simply that due to the reverse saturation current I_{CO}, as indicated in Fig. 5.7. The current I_{CO} is so small (*microamperes*) in magnitude compared to the vertical scale of I_C (*milliamperes*) that it appears on virtually the same horizontal line as I_C=0.

The notation most frequently used for I_{CO} on data and specification sheets is, as indicated in Fig. 5.8, I_{CBO}.

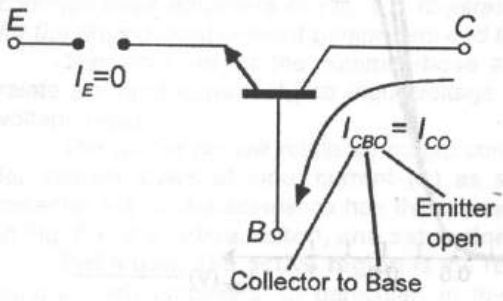


Fig. 5.8. Reverse saturation current

Because of improved construction techniques, the level of I_{CBO} for general-purpose transistors (especially Si) in the low- and mid-power ranges is usually so low that its effect can be ignored.

However, for higher power units I_{CBO} will still appear in the microampere range. In addition, keep in mind that I_{CBO}, like I_s, for the diode (both reverse leakage currents) is temperature

sensitive. *At higher temperatures the effect of I_{CBO} may become an important factor since it increases so rapidly with temperature.*

Note in Fig. 5.7 that as the emitter current increases above zero, the collector current increases to a magnitude essentially equal to that of the emitter current as determined by the basic transistor-current relations. Note also the almost negligible effect of V_{CB} on the collector current for the active region. The curves clearly indicate that a first approximation to the relationship between I_E and I_C in the active region is given by $I_C \approx I_E$.

Definition. As inferred by its name, the **cutoff region** is defined as that region where the collector current is 0 A, as revealed on Fig. 5.7. In addition, *in the cutoff region the collector-base and base-emitter junctions of a transistor are both reverse-biased.*

Definition. The **saturation region** is defined as that region of the characteristics to the left of V_{CB}=0 V. The horizontal scale in this region was expanded to clearly show the dramatic change in characteristics in this region. Note the exponential increase in collector current as the voltage V_{CB} increases toward 0 V. *In the saturation region the collector-base and base-emitter junctions are forward-biased.*

Definition. In the DC mode the levels of I_C and I_E due to the majority carriers are related by a quantity called **alpha** and defined by the following equation:

$$\alpha_{dc} = I_C / I_E, \tag{5.3}$$

where I_C and I_E are the levels of current at the point of operation.

Even though the characteristics of Fig. 5.7 would suggest that α=1, for practical devices the level of alpha typically extends from 0.90 to 0.998, with most approaching the high end of the range.

Since *alpha is defined solely for the majority carriers*, then:

$$I_C = I_{C\,majority} + I_{CO\,minority} = \alpha I_E + I_{CBO}. \tag{5.4}$$

For the characteristics of Fig. 5.7 when I_E=0 mA, I_C is therefore equal to I_{CBO}, but as mentioned earlier, the level of I_{CBO} is usually so small that it is virtually undetectable on the graph of Fig. 5.7. In other words, when I_E=0 mA in Fig. 5.7, I_C also appears to be 0 mA for the range of V_{CB} values.

Definition. For AC situations where the point of operation moves on the characteristic curve, an **AC alpha** is defined by

$$\alpha_{ac} = \left.\frac{\Delta I_C}{\Delta I_E}\right|_{V_{CB}=const}. \tag{5.5}$$

The α_{ac} is formally called the *common-base, short-circuit, amplification factor.*

Note. For most situations the magnitudes of α_{ac} and α_{dc} are quite close, permitting the use of the magnitude of one for the other.

Biasing. The proper biasing of the common-base configuration in the active region can be determined quickly using the approximation $I_C \approx I_E$ and assuming for the moment that $I_B \approx 0$ A.

The result is the configuration of Fig. 5.9 for the *pnp* transistor. The arrow of the symbol defines the direction of conventional flow for $I_E \approx I_C$. The DC supplies are then inserted with a polarity that will support the resulting current direction. For the *npn* transistor the polarities will be reversed.

Transistor amplifying action. Let us introduce the basic amplifying action of the transistor on a surface level using the network of Fig. 5.10. The **DC biasing** does not appear in the figure since our interest will be limited to the **AC response**.

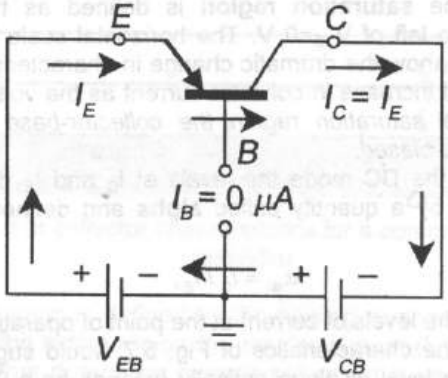


Fig. 5.9. Majority and minority carrier flow of a *pnp* transistor

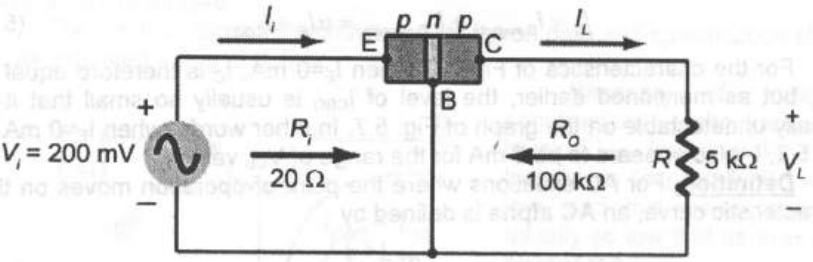


Fig. 5.10. Basic voltage amplification action of the common-base configuration

For the common-base configuration the *AC input resistance* determined by the characteristics of Fig. 5.6 is quite small and typically varies from 10 to 100 Ω.

The *output resistance* as determined by the curves of Fig. 5.7 is quite high (*the more horizontal the curves the higher the resistance*) and typically

varies from 50 kΩ to 1 MΩ (100 kΩ for the transistor of Fig. 5.8). *The difference in resistance is due to the forward-biased junction at the input (base to emitter) and the reverse-biased junction at the output (base to collector).*

Using a common value of 20 Ω for the input resistance, we can find that

$$I_i = V_i / R_i = 200\ mV / 20\ \Omega = 10\ mA. \qquad (5.6)$$

If we assume for the moment that $\alpha_{ac}=1$ ($I_c=I_e$), $I_L=I_i=10$ mA, then

$$V_L = I_L R_0 = 10\ mA \cdot 5\ k\Omega = 50\ V. \qquad (5.7)$$

Hence, the voltage amplification is $A_v=V_L/V_i=250$.

Typical values of voltage amplification for the common-base configuration vary from 50 to 300. The current amplification (I_C/I_E) is always less than 1 for the common-base configuration. This latter inequality should be obvious since α is always less than 1.

Note. The basic amplifying action was produced by transferring a current *I* from a low- to a high-resistance circuit. The combination of the two terms in italics results in the label transistor; that is,

*tran*sfer + re*sistor* → ***transistor.***

Calculation examples (for data on Figs. 5.6 and 5.7).

(a) Using the depicted characteristics, determine the resulting collector current if I_E=3 mA and V_{CB}=10 V.

(b) Using the characteristics, determine the resulting collector current if I_E remains at 3 mA but V_{CB} is reduced to 2 V.

(c) Using the characteristics, determine V_{BE} if I_C=4 mA and V_{CB} = 20 V.

Solutions.

(a) The characteristics clearly indicate that $I_C = I_E$ = 3mA.

(b) The effect of changing V_{CB} is negligible and I_C continues to be 3mA.

(c) From Fig. 5.7, $I_E = I_C$ = 4 mA. In Fig. 5.6 the resulting level of V_{BE} is about 0.74 V.

5.4. Static mode. Common-emitter configuration

The most frequently encountered transistor configuration appears in Fig. 5.11 for the *pnp* and *npn* transistors.

Definition. It is called the **common-emitter configuration** since the emitter is common or reference to both the input and output terminals (in this case common to both the base and collector terminals).

Two sets of characteristics are again necessary to describe fully the behavior of the common-emitter configuration: one for the input or base-emitter circuit and one for the output or collector-emitter circuit. Both are shown in Fig. 5.12, 5.13.

The emitter, collector, and base currents are shown in their actual conventional current direction. *Even though the transistor configuration has*

changed, the current relations developed earlier for the common-base configuration are still applicable. That is, $I_E = I_C + I_B$ and $I_C = \alpha I_E$.

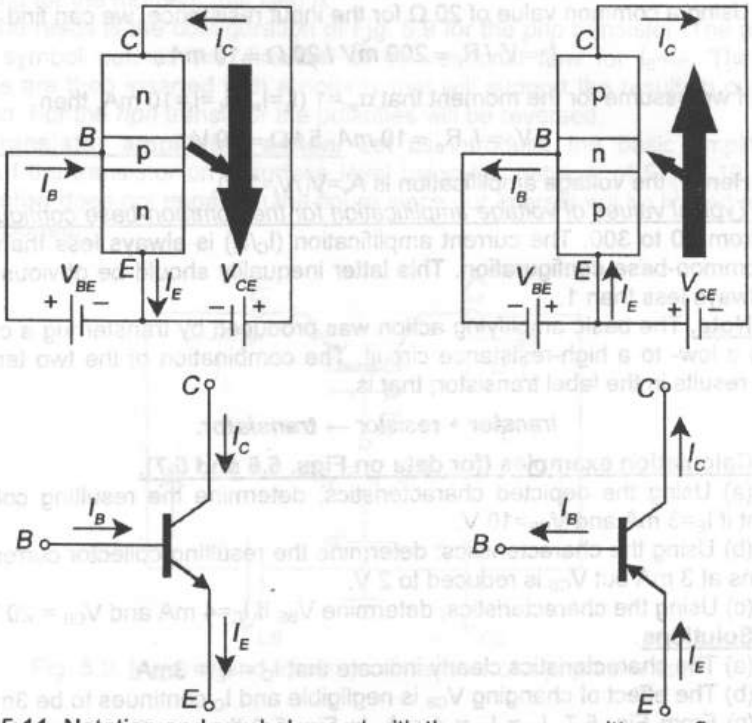


Fig. 5.11. Notation and symbols used with the common-emitter configuration: *npn* transistor (left column); *pnp* transistor (right column)

Note. On the characteristics of Fig. 5.13 the magnitude of I_B is in microamperes, compared to milliamperes of I_C. Consider also that the curves of I_B are not as horizontal as those obtained for I_E in the common-base configuration, indicating that the collector-to-emitter voltage will influence the magnitude of the collector current.

In Fig. 5.13 the *active region* exists to the right of the vertical dashed line at V_{CEsat} and above the curve for I_B equal to zero. The region to the left of V_{CEsat} is called the *saturation region*.

Active region. *In the active region of a common-emitter amplifier the collector-base junction is reverse-biased, while the base-emitter junction is forward-biased.* These are the same conditions that existed in the active region of the common-base configuration. The active region of the common-emitter configuration can be employed for voltage, current, or power amplification.

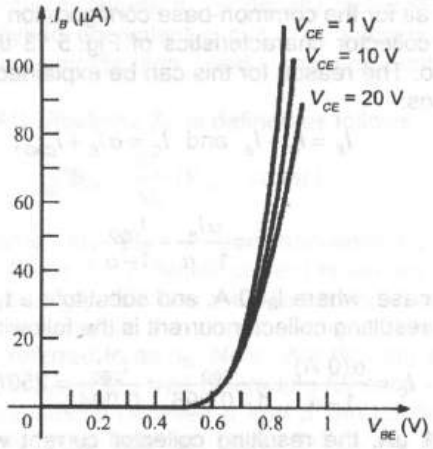


Fig. 5.12. Input or driving point characteristics for a common-emitter silicon transistor amplifier

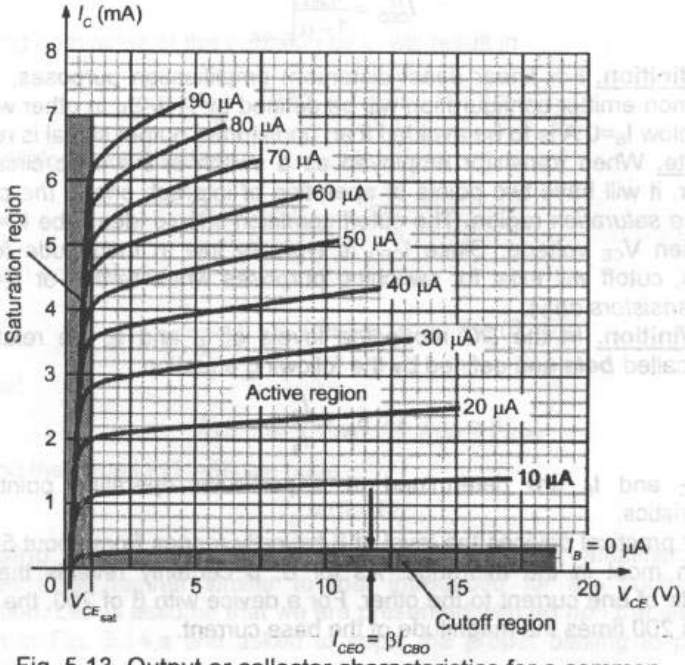


Fig. 5.13. Output or collector characteristics for a common-emitter transistor amplifier

Cutoff region. The *cutoff region* for the common-emitter configuration is not as well defined as for the common-base configuration.

Note on the collector characteristics of Fig. 5.13 that I_C is not equal to zero when I_B is zero. The reason for this can be explained in such a way. From the main expressions:

$$I_E = I_C + I_B \quad \text{and} \quad I_C = \alpha I_E + I_{CBO}, \tag{5.8}$$

we obtain:

$$I_C = \frac{\alpha I_B}{1-\alpha} + \frac{I_{CBO}}{1-\alpha}. \tag{5.9}$$

Consider the case, where I_B=0 A, and substitute a typical value of α such as 0.996. Then the resulting collector current is the following:

$$I_C = \frac{\alpha(0\ A)}{1-\alpha} + \frac{I_{CBO}}{1-0.996} = \frac{I_{CBO}}{0.004} = 250 I_{CBO}. \tag{5.10}$$

If I_{CBO} were 1 μA, the resulting collector current with I_B=0 A would be 250(1μA)=0.25 mA, as reflected in the characteristics of Fig. 5.13. For future reference, the collector current defined by the condition I_B=0 A will be as:

$$I_{CEO} = \left.\frac{I_{CBO}}{1-\alpha}\right|_{I_B=0A}. \tag{5.11}$$

Definition. For linear (least distortion) amplification purposes, cutoff for the common-emitter configuration will be defined by I_C=I_{CEO}. In other words, the region below I_B=0 A is to be avoided if an undistorted output signal is required.

Note. When transistor employed as a switch in the logic circuitry of a computer, it will have two points of operation of interest: one *in the cutoff* and one *in the saturation* region. The cutoff condition should ideally be I_C=0 mA for the chosen V_{CE} voltage. Since I_{CEO} is typically low in magnitude for silicon materials, cutoff will exist for switching purposes when I_B=0 A or I_C=I_{CEO} (*for silicon transistors only*).

Definition. In the *DC mode* the levels of I_C and I_B are related by a quantity called ***beta*** and defined by the following equation:

$$\beta_{dc} = \frac{I_C}{I_B} \tag{5.12}$$

where I_C and I_B are determined at a particular operating point on the characteristics.

For practical devices the level of β typically ranges from about 50 to over 400, with most in the midrange. As for α, β certainly reveals the relative magnitude of one current to the other. For a device with β of 200, the collector current is 200 times the magnitude of the base current.

Note. On specification sheets β_{dc} is usually included as h_{FE} with the h derived from an AC *h*ybrid equivalent circuit. The subscripts FE are derived from *Forward*-current amplification and common-*emitter* configuration, respectively.

Definition. For AC situations β_{ac} is defined as follows:

$$\beta_{ac} = \frac{\Delta I_C}{\Delta I_B}\,(V_{CE} = const). \tag{5.13}$$

The formal name for β_{ac} is **common-emitter, forward-current, amplification factor**. Since the collector current is usually the output current for a common-emitter configuration and the base current the input current, the term *amplification* is included in the nomenclature above. On specification sheets β_{ac} is normally referred to as h_{fe}. Note, *that the only difference between the notation used for the β_{DC} is the type of lettering for each subscript quantity.*

Definition. A *relationship between α and β* can be developed using the basic relationships introduced thus far. Using $\beta=I_C/I_B$ we have $I_B = I_C/\beta$, and from $\alpha=I_C/I_E$ we have $I_E=I_C/\alpha$. Substituting into $I_E=I_C+I_B$ we obtain

$$\frac{I_C}{\alpha} = I_C + \frac{I_C}{\beta} \tag{5.14}$$

and dividing both sides of the equation by I_C will result in

$$\frac{1}{\alpha} = 1 + \frac{1}{\beta}. \text{ Then } \alpha = \frac{\beta}{\beta+1} \text{ or } \beta = \frac{\alpha}{1-\alpha}. \tag{5.15}$$

In addition, recall that

$$I_{CEO} = \frac{I_{CBO}}{1-\alpha} \tag{5.16}$$

but using an equivalence of

$$\frac{1}{1-\alpha} = \beta + 1 \text{ (derived from the (5.15)),} \tag{5.17}$$

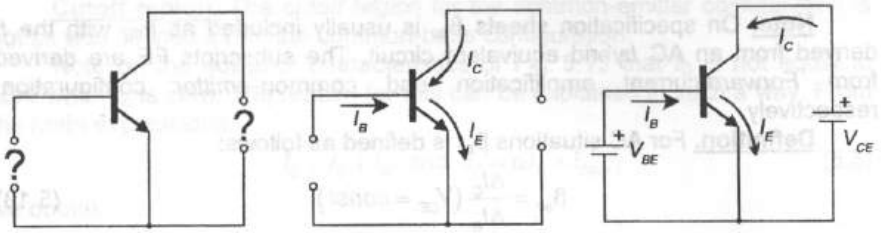
we find that

$$I_{CEO} = (\beta+1)I_{CBO} \text{ or } I_{CEO} \cong \beta I_{CBO}. \tag{5.18}$$

Using the equation $I_C=\beta I_B$ we have

$$I_E = (\beta+1)I_B. \tag{5.19}$$

Biasing. The proper biasing of a common-emitter amplifier can be determined in a manner similar to that introduced for the common-base configuration. Let us assume that we are presented with an *npn* transistor such as shown in Fig. 5.14,a and asked to apply the proper biasing to place the device in the active region.



a b c

Fig. 5.14. Determining the proper biasing arrangement for a common-emitter *npn* transistor configuration

The first step is to indicate the direction of I_E as established by the arrow in the transistor symbol as shown in Fig. 5.14,b. Next, the other currents are introduced as shown, keeping in mind the Kirchhoff's current law relationship: $I_C+I_B=I_E$. Finally, the supplies are introduced with polarities that will support the resulting directions of I_B and I_C as shown in Fig. 5.14,c to complete the picture.

The same approach can be applied to *pnp* transistors. If the transistor of Fig. 5.14 was a *pnp* transistor all the currents and polarities of Fig. 5.14,c would be reversed.

5.5. Static mode. Common-collector configuration

The third and final transistor configuration is the common-collector configuration, shown in Fig. 5.15 with the proper current directions and voltage notation.

Note. The common-collector configuration is used primarily for **impedance-matching purposes** since *it has a high input impedance and low output impedance, opposite to that of the common-base and common-emitter configurations.*

A common-collector circuit configuration is provided in Fig. 5.16 with the load resistor connected from emitter to ground. Note that the collector is tied to ground even though the transistor is connected in a manner similar to the common-emitter configuration. *From a design viewpoint, there is no need for a set of common-collector characteristics to choose the parameters of the circuit of Fig. 5.16.* It can be designed using the common-emitter characteristics of Subsection 5.4.

For all practical purposes, the output characteristics of the common-collector configuration are the same as for the common-emitter configuration. For the common-collector configuration the output characteristics are a plot of I_E versus V_{EC} for a range of values of I_B.

The input current, therefore, is the same for both the common-emitter and common-collector characteristics.

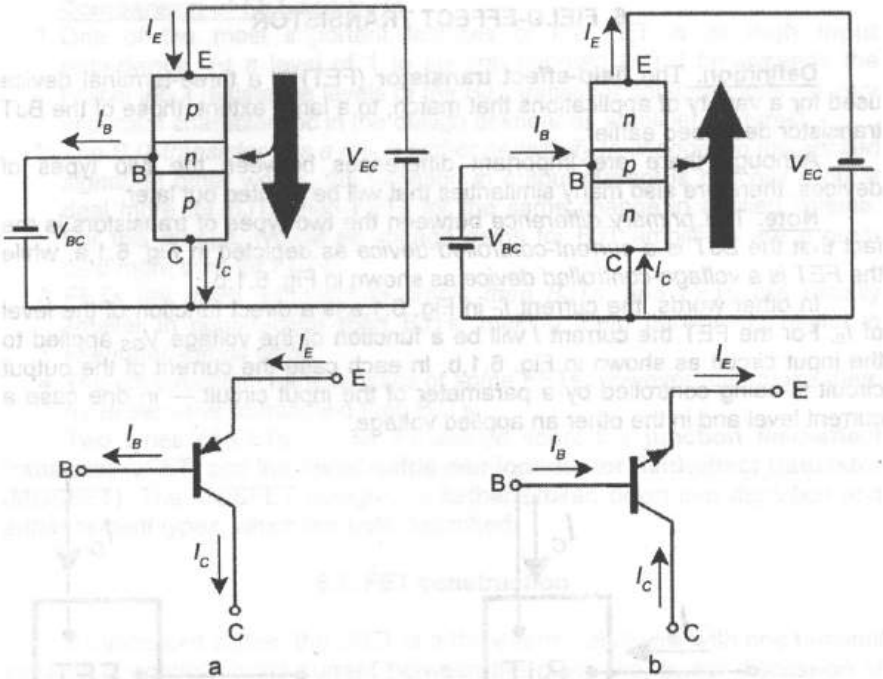


Fig. 5.15. Notation and symbols used with the common-collector configuration: *pnp* transistor (a); *npn* transistor (b)

The horizontal voltage axis for the common-collector configuration is obtained by simply changing the sign of the collector-to-emitter voltage of the common-emitter characteristics.

Finally, there is an almost unnoticeable change in the vertical scale of I_C of the common-emitter characteristics if I_C is replaced by I_E for the common-collector characteristics (since $\alpha \approx 1$).

For the input circuit of the common-collector configuration the common-emitter base characteristics are sufficient for obtaining the required information.

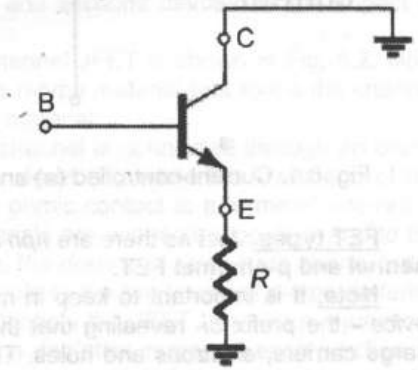


Fig. 5.16. Common-collector configuration used for impedance-matching purposes

6. FIELD-EFFECT TRANSISTOR

Definition. The **field-effect transistor** (FET) is a three-terminal device used for a variety of applications that match, to a large extent, those of the BJT transistor described earlier.

Although there are important differences between the two types of devices, there are also many similarities that will be pointed out later.

Note. The *primary difference* between the two types of transistors is the fact that the *BJT is a current-controlled device* as depicted in Fig. 6.1,a, while the *FET is a voltage-controlled device* as shown in Fig. 6.1,b.

In other words, the current I_C in Fig. 6.1,a is a direct function of the level of I_B. For the FET the current I will be a function of the voltage V_{GS} applied to the input circuit as shown in Fig. 6.1,b. In each case the current of the output circuit is being controlled by a parameter of the input circuit — in one case a current level and in the other an applied voltage.

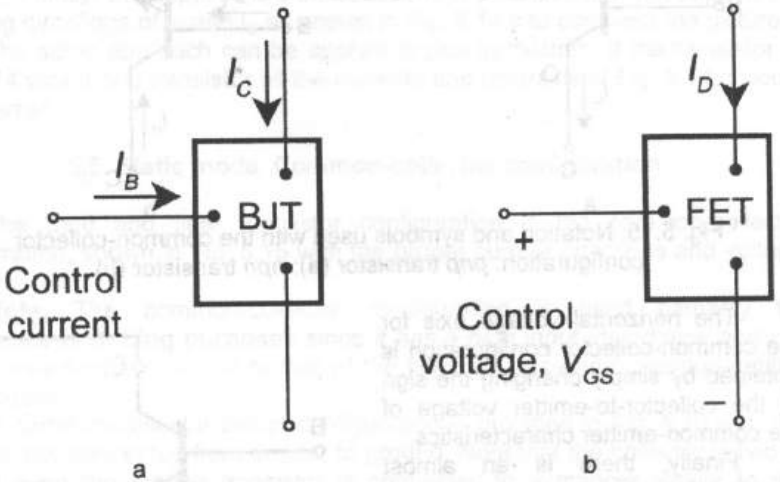


Fig. 6.1. Current-controlled (a) and voltage-controlled (b) amplifiers

FET types. Just as there are *npn* and *pnp* bipolar transistors, there are **n-channel** and **p-channel FET**.

Note. It is important to keep in mind that the *BJT transistor is a bipolar device* - the prefix *bi-* revealing that the conduction level is a function of two charge carriers, electrons and holes. The *FET is a unipolar device* depending solely on either electron (*n*-channel) or hole (*p*-channel) conduction.

Definition. In the FET an *electric field* is established by the charges present that will control the conduction path of the output circuit without the need for direct contact between the controlling and controlled quantities.

Comparison of FET and BJT:

1. One of the most important features of the FET is its **high input impedance**. At a level of 1 to several hundred MΩ it far exceeds the typical input resistance levels of the BJT transistor configurations - a very important characteristic in the design of linear ac amplifier systems.
2. *The BJT transistor has a much higher sensitivity to changes in the applied signal.* In other words, the variation in output current is typically a great deal more for BJTs than FETs for the same change in applied voltage. For this reason, typical ac voltage gains for BJT amplifiers are a great deal more than for FETs.
3. *FETs are more temperature stable than BJTs*, and FETs are usually smaller in construction than BJTs, making them particularly useful in *integrated-circuit chips.*
4. The construction characteristics of some FETs, however, can make them more sensitive to handling than BJTs.

Two types of FETs will be introduced soon: the **junction field-effect transistor** (JFET) and the **metal-oxide-semiconductor field-effect transistor** (MOSFET). The MOSFET category is further broken down into *depletion* and *enhancement types*, which are both described.

6.1. FET construction

As indicated earlier, the JFET is a three-terminal device with one terminal capable of controlling the current between the other two. In our discussion of the BJT transistor the *npn* transistor was employed through the major part of the analysis and design sections, with a section devoted to the impact of using a *pnp* transistor. For the JFET transistor the *n*-channel device will appear as the prominent device, with paragraphs and sections devoted to the impact of using a *p*-channel JFET.

The basic construction of the *n*-channel JFET is shown in Fig. 6.2. Note that the major part of the structure is the *n*-type material that forms the channel between the embedded layers of *p*-type material.

Definition. The top of the *n*-type channel is connected through an ohmic contact to a terminal referred to as the **drain** (D), while the lower end of the same material is connected through an ohmic contact to a terminal referred to as the **source** (S). The two *p*-type materials are connected together and to the **gate** (G) terminal. In essence, therefore, *the drain and source are connected to the ends of the n-type channel and the gate to the two layers of p-type material.*

In the absence of any applied potentials the JFET has two *p-n*-junctions under no-bias conditions. The result is a depletion region at each junction as shown in Fig. 6.2 that resembles the same region of a diode under no-bias conditions.

Recall also that a *depletion region* is that region void of free carriers and therefore unable to support conduction through the region.

The water analogy of Fig. 6.3 does provide a sense for the JFET control at the gate terminal and the appropriateness of the terminology applied to the terminals of the device. The source of water pressure can be likened to the applied voltage from drain to source that will establish a flow of water (electrons) from the spigot (source). The "gate," through an applied signal (potential), controls the flow of water (charge) to the "drain." The drain and source terminals are at opposite ends of the *n*-channel as introduced in Fig. 6.2 because the terminology is defined for electron flow.

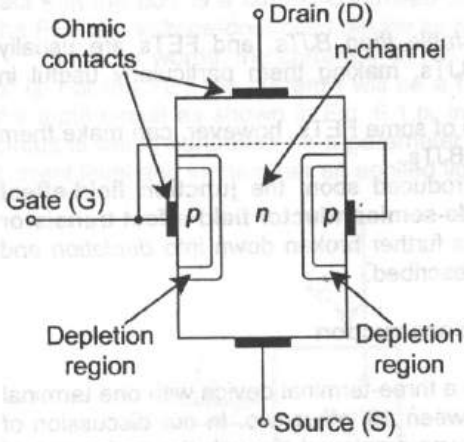


Fig. 6.2. FET construction

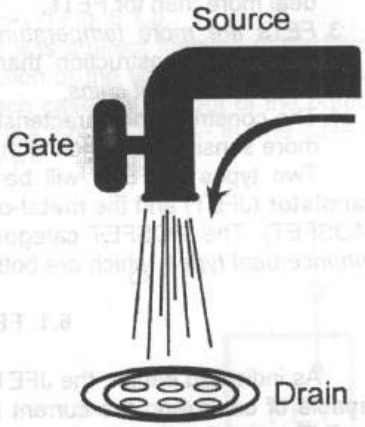


Fig. 6.3. Water analogy for the JFET control mechanism

6.2. V_{GS}=0 V, V_{DS} Some Positive Value

In Fig. 6.4 a *positive voltage V_{DS} has been applied across the channel* and the *gate has been connected directly to the source* to establish the condition V_{GS} = 0 V. The result is a *gate and source terminal at the same potential* and a depletion region in the low end of each *p*-material similar to the distribution of the no-bias conditions of Fig. 6.2.

The instant voltage V_{DS} is applied, the electrons will be drawn to the drain terminal, establishing the conventional current I_D with the defined direction of Fig. 6.4. The path of charge flow clearly reveals that the drain and source currents are equivalent (I_D=I_S). Under the conditions appearing in Fig. 6.4, the flow of charges is relatively uninhibited and limited solely by the resistance of the *n*-channel between drain and source.

As the voltage V_{DS} is increased from 0 to a few volts, the current will increase as determined by Ohm's law and the plot of I_D versus V_{DS} will appear as shown in Fig. 6.5. The relative straightness of the plot reveals that for the region of low values of V_{DS}, the resistance is essentially constant.

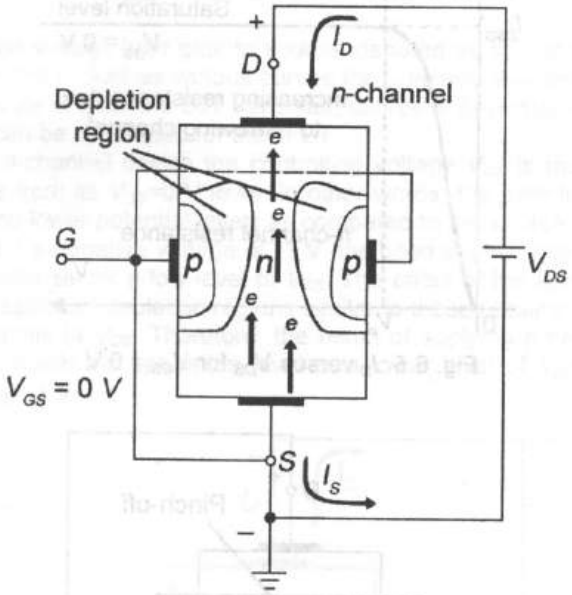


Fig. 6.4. JFET in the V_{GS} = 0 V and V_{DS} > 0 V

As V_{DS} increases and approaches a level referred to as V_P in Fig. 6.5, the depletion regions of Fig. 6.4 will widen, causing a noticeable reduction in the channel width. The reduced path of conduction causes the resistance to increase and the curve in the graph of Fig. 6.5 to occur. *The more horizontal the curve, the higher the resistance, suggesting that the resistance is approaching "infinite" ohms in the horizontal region.*

Definition. If V_{DS} is increased to a level where it appears that the two depletion regions would "touch" as shown in Fig. 6.6, a condition referred to as **pinch-off** will result. The level of V_{DS} that establishes this condition is referred to as the **pinch-off voltage** and is denoted by V_P as shown in Fig. 6.5.

In actuality, the term *pinch-off* is a misnomer in that it suggests the current I_D is pinched off and drops to 0 A. As shown in Fig. 6.5, however, this is hardly the case — I_D maintains a saturation level defined as I_{DSS} in Fig. 6.5. *In reality a very small channel still exists, with a current of very high density.*

As V_{DS} is increased beyond V_P, the region of close encounter between the two depletion regions will increase in length along the channel, but the level of I_D remains essentially the same. In essence, therefore, once $V_{DS}>V_P$ the JFET has the characteristics of a current source. The choice of notation I_{DSS} is derived from the fact that it is *the Drain-to-Source current with a Short-circuit connection from gate to source.*

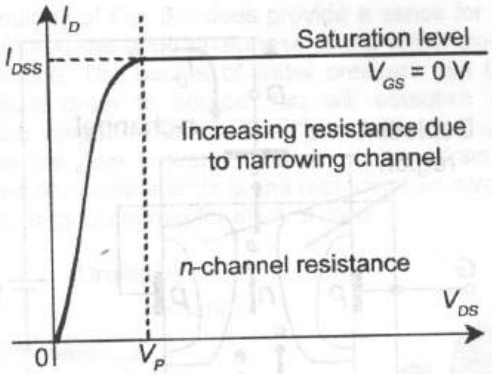


Fig. 6.5. I_D versus V_{DS} for $V_{GS} = 0$ V

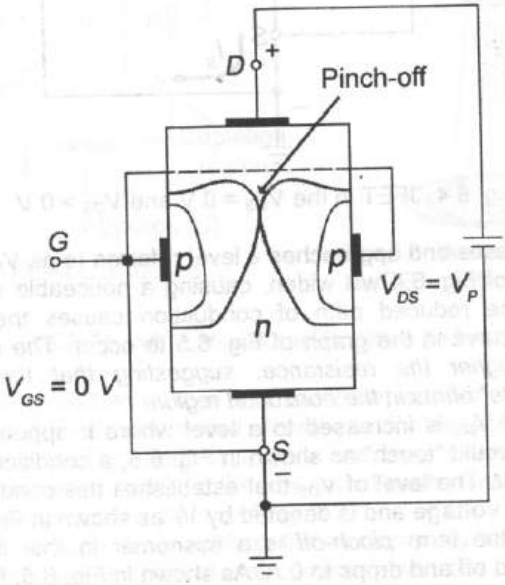


Fig. 6.6. Pinch-off ($V_{GS} = 0$ V, $V_{DS} = V_P$)

As we continue to investigate the characteristics of the device we will find that: *I_{DSS} is the maximum drain current for a JFET and is defined by the conditions V_{GS}=0 V and $V_{DS} > |V_P|$*. **Note** in Fig. 6.5 that $V_{GS} = 0$ V for the entire length of the curve. Next we will describe how the characteristics of Fig. 6.5 are affected by changes in the level of V_{GS}.

6.3. $V_{GS} < 0$ V, V_{DS} Some Positive Value

Note. The voltage from gate to source, denoted as V_{GS}, is the controlling voltage of the JFET. Just as various curves for I_C versus V_{CE} were established for different levels of I_B for the BJT transistor, curves of I_D versus V_{DS} for various levels of V_{GS} can be developed for the JFET.

For the *n*-channel device the controlling voltage V_{GS} is made more and more negative from its V_{GS}=0 V level. In other words, the gate terminal will be set at lower and lower potential levels as compared to the source.

In Fig. 6.7 a negative voltage of -1 V has been applied between the gate and source terminals for a low level of V_{DS}. The effect of the applied negative-bias V_{GS} is to establish depletion regions similar to those obtained with V_{GS}=0 V but at lower levels of V_{DS}. Therefore, the result of applying a negative bias to the gate is to reach the saturation level at a lower level of V_{DS} as shown in Fig. 6.8 for V_{GS}=-1 V.

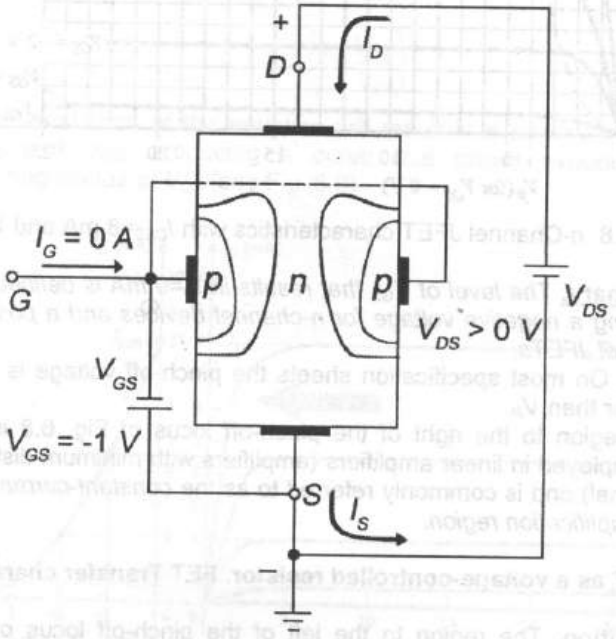


Fig. 6.7. Application of a negative voltage to the gate of a JFET

The resulting saturation level for I_D has been reduced and in fact will continue to decrease as V_{GS} is made more and more negative. Note also on Fig. 6.8 how *the pinch-off voltage continues to drop in a parabolic manner as*

V_{GS} *becomes more and more negative*. Eventually, V_{GS} when V_{GS}=-V_P will be sufficiently negative to establish a saturation level that is essentially 0 mA, and for all practical purposes the device has been "turned off."

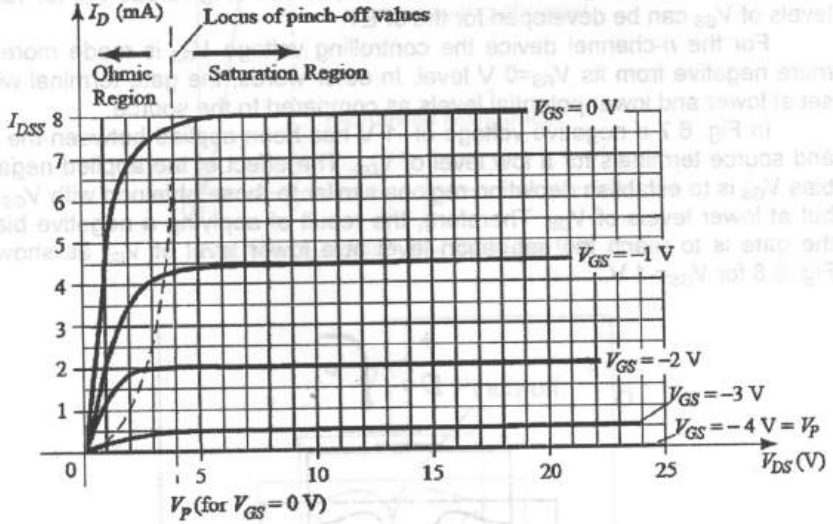


Fig. 6.8. *n*-Channel JFET characteristics with I_{DSS}=8 mA and V_P=4 V

Summary. *The level of V_{GS} that results in I_D=0 mA is defined by V_{GS}=V_P, with V_P being a negative voltage for n-channel devices and a positive voltage for p-channel JFETs.*

Note. On most specification sheets the pinch-off voltage is specified as $V_{GS(off)}$ rather than V_P.

The region to the right of the pinch-off locus of Fig. 6.8 is the region typically employed in linear amplifiers (amplifiers with minimum distortion of the applied signal) and is commonly referred to as the *constant-current, saturation,* or *linear amplification region.*

6.4. FET as a voltage-controlled resistor. FET Transfer characteristic

Definition. The region to the left of the pinch-off locus of Fig. 6.8 is referred to as the **ohmic or voltage-controlled resistance region**. In this region the JFET can actually be employed as a variable resistor (possibly for an automatic gain control system) whose resistance is controlled by the applied gate-to-source voltage.

Note in Fig. 6.8 that the slope of each curve and therefore the resistance of the device between drain and source for $V_{DS}=V_P$ is a function of the applied voltage V_{GS}. As V_{GS} becomes more and more negative, the slope of each curve becomes more and more horizontal, corresponding with an increasing resistance level.

The following equation will provide a good first approximation to the resistance level in terms of the applied voltage V_{GS}:

$$r_d = \frac{r_0}{\left(1 - V_{GS}/V_P\right)^2}, \tag{6.1}$$

where r_0 is the resistance with $V_{GS}=0$ V and r_d the resistance at a particular level of V_{GS}.

For an *n*-channel JFET with r_0 equal to 10 kΩ (V_{GS} = 0 V, V_P=-6 V), Eq. (6.1) will result in 40 kΩ at V_{GS}=-3 V.

Definition. The relationship between I_D and V_{GS} is defined by **Shockley's equation**:

$$I_D = I_{DSS}\left(1 - \frac{V_{GS}}{V_P}\right)^2. \tag{6.2}$$

The squared term of the equation will result in a nonlinear relationship between I_D and V_{GS}, producing a curve that grows exponentially with decreasing magnitudes of V_{GS} (see Fig. 6.9).

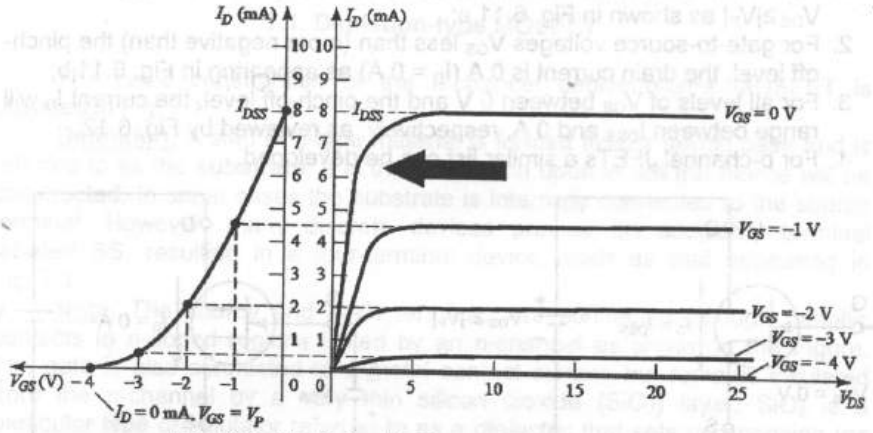


Fig. 6.9. Obtaining FET transfer characteristic

Note. *The transfer characteristics defined by Shockley's equation are unaffected by the network in which the device is employed.*

6.5. FET Symbols

The graphic symbols for the *n*-channel and *p*-channel JFETs are provided in Fig. 6.10. Note that the arrow is pointing in for the *n*-channel device of Fig. 6.10,a to represent the direction in which I_G would flow if the *p*-*n*-junction were forward-biased. For the *p*-channel device (Fig. 6.10,b) the only difference in the symbol is the direction of the arrow.

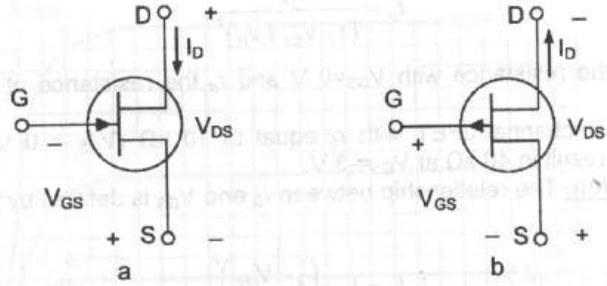


Fig. 6.10. JFET symbols: *n*-channel (a); *p*-channel (b)

Summary. A number of important parameters and relationships were introduced in this section. A few that will surface frequently in the analysis to follow in this chapter and the next for *n*-channel JFETs include the following:

1. The maximum current is defined as I_{DSS} and occurs when $V_{GS} = 0$ V and $V_{DS} \geq |V_P|$ as shown in Fig. 6.11,a;
2. For gate-to-source voltages V_{GS} less than (more negative than) the pinch-off level, the drain current is 0 A ($I_D = 0$ A) as appearing in Fig. 6.11,b;
3. For all levels of V_{GS} between 0 V and the pinch-off level, the current I_D will range between I_{DSS} and 0 A, respectively, as reviewed by Fig. 6.12.
4. For p-channel JFETs a similar list can be developed.

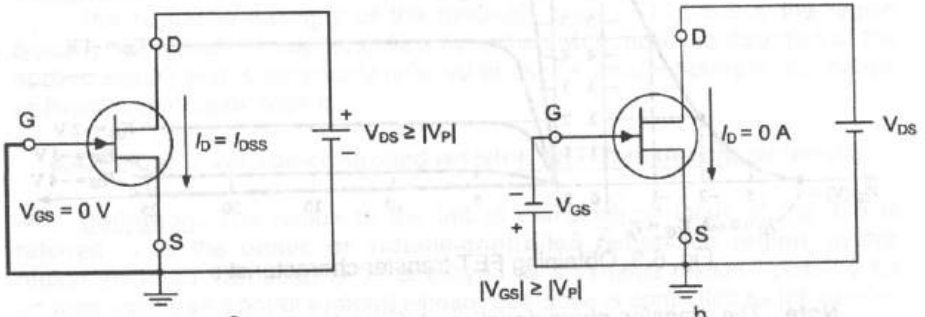


Fig. 6.11. $V_{GS} = 0$ V, $I_D = I_{DSS}$ (a); cutoff ($I_D = 0$ A) V_{GS} less than the pinch-off level (b)

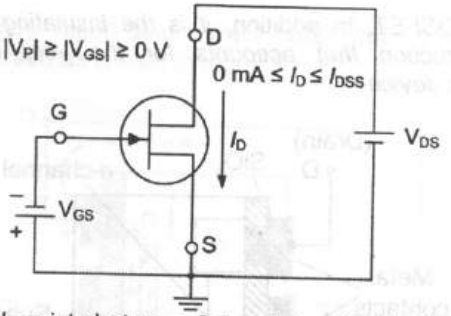


Fig. 6.12. I_D exists between 0 A and I_{DSS} for V_{GS} less than or equal to 0 V and greater than the pinch-off level

7. METAL-OXIDE-SEMICONDUCTOR-FIELD-EFFECT TRANSISTORS

As noted earlier, there are two types of FETs: **JFETs** and **MOSFETs**. MOSFETs are further broken down into *depletion type* and *enhancement type*.

Definition. The terms **depletion** and **enhancement** define their basic mode of operation, while the label **MOSFET** stands for **metal-oxide-semiconductor-field-effect transistor**.

Since there are differences in the characteristics and operation of each type of MOSFET, they are covered in separate sections.

7.1. Depletion-type MOSFET

The basic construction of the *n*-channel depletion-type MOSFET is provided in Fig. 7.1.

Definition. A slab of *p*-type material is formed from a silicon base and is referred to as the **substrate**. It is the foundation upon which the device will be constructed. In some cases the substrate is internally connected to the source terminal. However, many discrete devices provide an additional terminal **labeled SS**, resulting in a four-terminal device, such as that appearing in Fig. 7.1.

Note. The *source* and *drain terminals* are connected through metallic contacts to *n*-doped regions linked by an *n*-channel as shown in the Figure. The gate is also connected to a metal contact surface but remains insulated from the *n*-channel by a very thin silicon dioxide (SiO_2) layer. SiO_2 is a particular type of insulator referred to as a *dielectric* that sets up opposing (as revealed by the prefix *di-*) electric fields within the dielectric when exposed to an externally applied field.

The fact that the SiO_2 layer is an insulating layer reveals the following fact: *there is no direct electrical connection between the gate terminal and the*

channel of a MOSFET. In addition, *it is the insulating layer of SiO_2 in the MOSFET construction that accounts for the very desirable high input impedance of the device.*

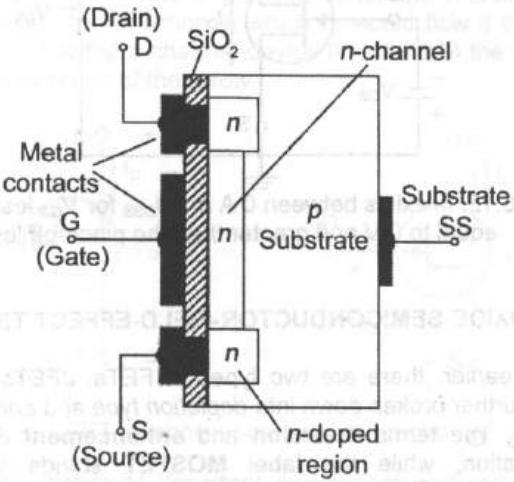


Fig. 7.1. *n*-type depletion type MOSFET

In fact, the input resistance of a MOSFET is often that of the typical JFET, even though the input impedance of most JFETs is sufficiently high for most applications. The very high input impedance continues to fully support the fact that the gate current (I_G) is essentially zero amperes for dc-biased configurations.

Note. The reason for the label metal-oxide-semiconductor FET is now fairly obvious: *metal* for the drain, source, and gate connections to the proper surface - in particular, the gate terminal and the control to be offered by the surface area of the contact, the *oxide* for the silicon dioxide insulating layer, and the *semiconductor* for the basic structure on which the *n*- and *p*-type regions are diffused. The insulating layer between the gate and channel has resulted in another name for the device: *insulated gate FET* or *IGFET,* although this label is used less and less in current literature.

In Fig. 7.2 the gate-to-source voltage is set to zero volts by the direct connection from one terminal to the other, and a voltage V_{DS} is applied across the drain-to-source terminals. The result is an attraction for the positive potential at the drain by the *free* electrons of the *n*-channel and a current similar to that established through the channel of the JFET. In fact, the resulting current with $V_{GS} = 0$ V continues to be labeled I_{DSS}, as shown in Fig. 7.2.

In Fig. 7.3 V_{GS} has been set at a negative voltage such as -1 V. The negative potential at the gate will tend to pressure electrons toward the *p*-type

substrate (like charges repel) and attract holes from the *p*-type substrate (opposite charges attract) as shown in Fig. 7.3.

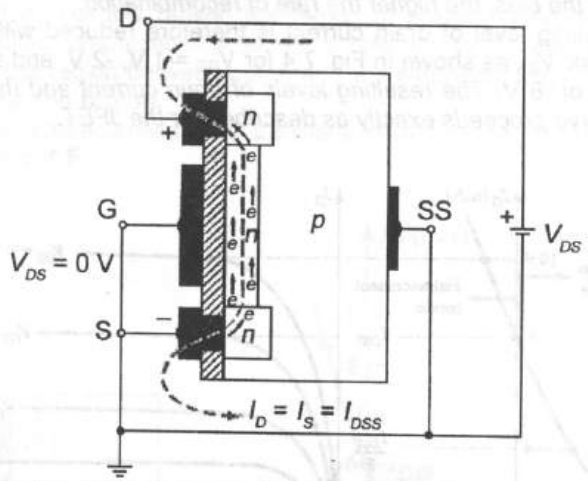


Fig. 7.2. *n*-Channel depletion-type MOSFET with V_{GS} = 0 V and an applied voltage V_{DD}

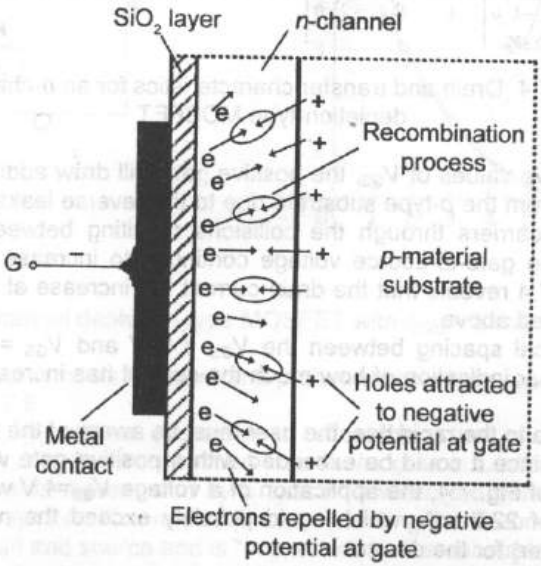


Fig. 7.3. Reduction in free carriers in channel due to a negative potential at the gate terminal

Depending on the magnitude of the negative bias established by V_{GS}, a level of recombination between electrons and holes will occur that will reduce the number of free electrons in the *n*-channel available for conduction. *The more negative the bias, the higher the rate of recombination.*

The resulting level of drain current is therefore reduced with increasing negative bias for V_{GS} as shown in Fig. 7.4 for V_{GS} =-1 V, -2 V, and so on, to the pinch-off level of -6 V. *The resulting levels of drain current and the plotting of the transfer curve proceeds exactly as described for the JFET.*

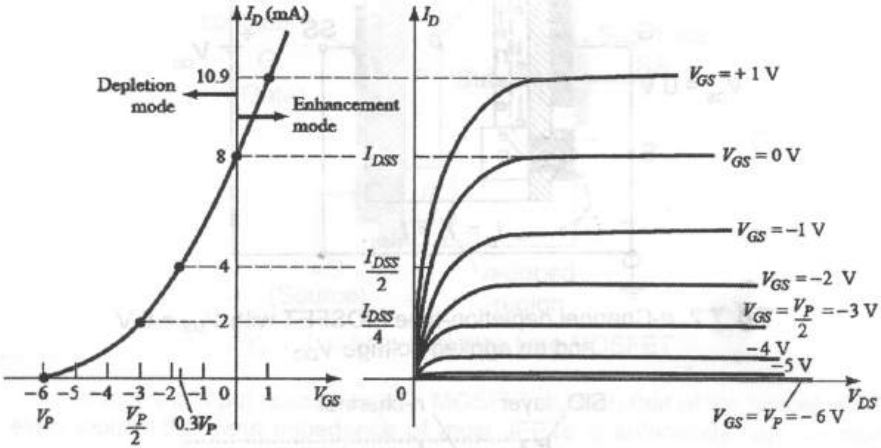


Fig. 7.4. Drain and transfer characteristics for an *n*-channel depletion-type MOSFET

For positive values of V_{GS}, the positive gate will draw additional electrons (free carriers) from the *p*-type substrate due to the reverse leakage current and establish new carriers through the collisions resulting between accelerating particles. As the gate-to-source voltage continues to increase in the positive direction, Fig. 7.4 reveals that the drain current will increase at a rapid rate for the reasons listed above.

The vertical spacing between the V_{GS} = 0 V and V_{GS} =-1 V curves of Fig. 7.4 is a clear indication of how much the current has increased for the 1 V change in V_{GS}.

Note. Due to the rapid rise, the user must be aware of the maximum drain current rating since it could be exceeded with a positive gate voltage. That is, for the device of Fig. 7.4, the application of a voltage V_{GS}=4 V would result in a drain current of 22.2 mA, which could possibly exceed the maximum rating (current or power) for the device.

As revealed above, the application of a positive gate-to-source voltage has "**enhanced**" the level of free carriers in the channel compared to that encountered with V_{GS}= 0 V.

Definition. For this reason the region of positive gate voltages on the drain or transfer characteristics is often referred to as the **enhancement region**, with the region between cutoff and the saturation level of I_{DSS} referred to as the **depletion region**.

Transfer characteristic and transistor structure for *p*-type device are depicted on Fig. 7.5.

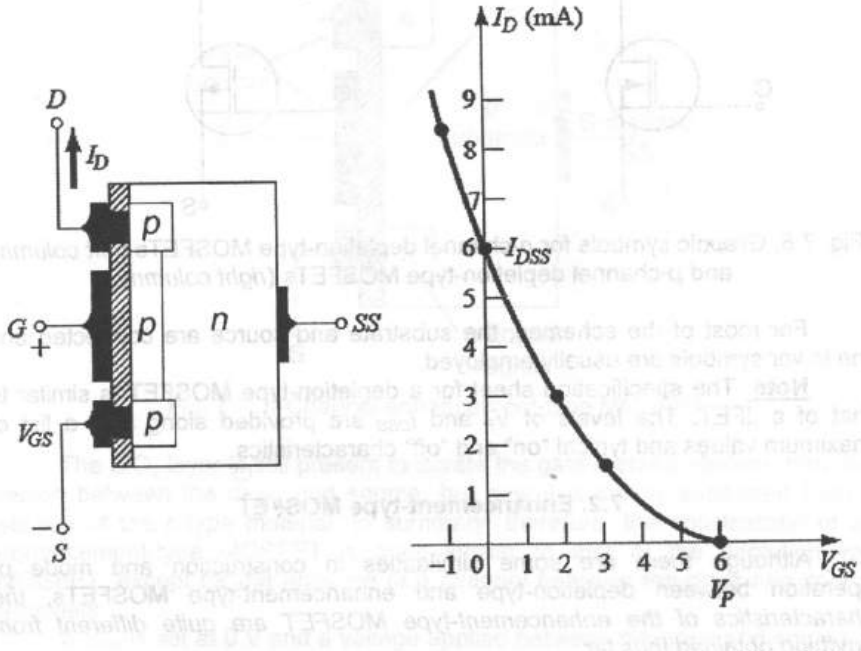


Fig. 7.5. *p*-Channel depletion-type MOSFET with I_{DSS} = 6 mA and V_P=-6 V

The graphic symbols for an *n*- and *p*-channel depletion-type MOSFET are provided in Fig. 7.6.

Note how the symbols chosen try to reflect the actual construction of the device. *The lack of a direct connection (due to the gate insulation) between the gate and channel is represented by a space between the gate and the other terminals of the symbol.* The vertical line representing the channel is connected between the drain and source and is "supported" by the substrate.

Two symbols are provided for each type of channel to reflect the fact that in some cases the substrate is externally available while in others it is not.

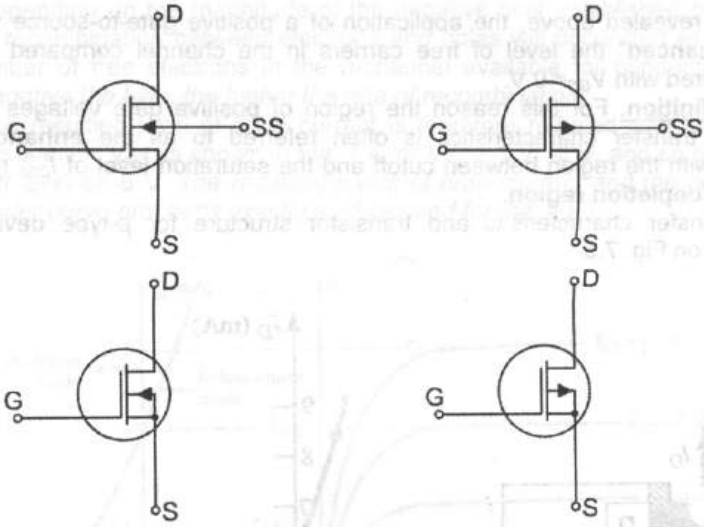


Fig. 7.6. Graphic symbols for *n*-channel depletion-type MOSFETs (*left column*) and *p*-channel depletion-type MOSFETs (*right column*)

For most of the schemes, the substrate and source are connected and the lower symbols are usually employed.

Note. The specification sheet for a depletion-type MOSFET is similar to that of a JFET. The levels of V_P and I_{DSS} are provided along with a list of maximum values and typical "on" and "off" characteristics.

7.2. Enhancement-type MOSFET

Although there are some similarities in construction and mode of operation between depletion-type and enhancement-type MOSFETs, *the characteristics of the enhancement-type MOSFET are quite different from anything obtained thus far.*

The transfer curve is not defined by Shockley's equation, and the drain current is now cut off until the gate-to-source voltage reaches a specific magnitude. In particular, current control in an *n*-channel device is now affected by a positive gate-to-source voltage rather than the range of negative voltages encountered for *n*-channel JFETs and *n*-channel depletion-type MOSFETs.

The basic construction of the *n*-channel enhancement-type MOSFET is provided in Fig. 7.7. A slab of *p*-type material is formed from a silicon base and is again referred to as the substrate. As with the depletion-type MOSFET, the substrate is sometimes internally connected to the source terminal, while in other cases a fourth lead is made available for external control of its potential

level. *The source and drain terminals are again connected through metallic contacts to n-doped regions, but note in Fig. 7.7 the absence of a channel between the two n-doped regions.*

Important. *The absence of a channel as a constructed component of the device is the primary difference between the construction of depletion-type and enhancement-type MOSFETs.*

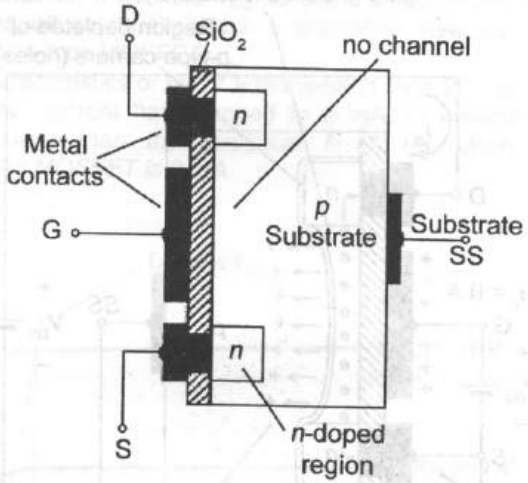


Fig. 7.7. *n*-Channel enhancement-type MOSFET

The SiO_2 layer is still present to isolate the gate metallic platform from the region between the drain and source, but now it is simply separated from a section of the *p*-type material. In summary, therefore, the construction of an enhancement-type MOSFET is quite similar to that of the depletion-type MOSFET, except for *the absence of a channel between the drain and source terminals.*

If V_{GS} is set at 0 V and a voltage applied between the drain and source of the device of Fig. 7.7, the absence of an *n*-channel (with its generous number of free carriers) will result in a current of effectively zero amperes - quite different from the depletion-type MOSFET and JFET where $I_D = I_{DSS}$.

It is not sufficient to have a large accumulation of carriers (electrons) at the drain and source (due to the *n*-doped regions) if a path fails to exist between the two. With V_{DS} some positive voltage, V_{GS} at 0 V, and terminal *SS* directly connected to the source, there are in fact two reverse-biased *p-n*-junctions between the *n*-doped regions and the *p*-substrate to oppose any significant flow between drain and source.

In Fig. 7.8 both V_{DS} and V_{GS} have been set at some positive voltage greater than 0 V, establishing the drain and gate at a positive potential with

respect to the source. The positive potential at the gate will pressure the holes (since like charges repel) in the *p*-substrate along the edge of the SiO_2 layer to leave the area and enter deeper regions of the *p*-substrate, as shown in the figure.

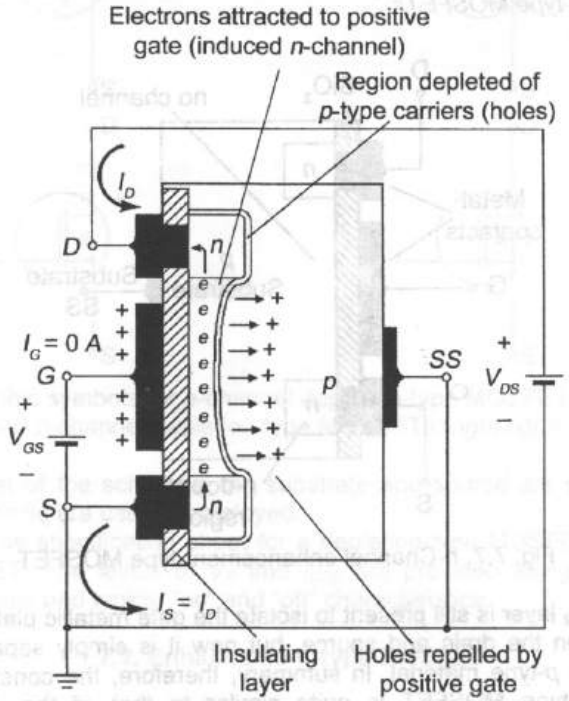


Fig. 7.8. Channel formation in the *n*-channel enhancement-type MOSFET

The result is a depletion region near the SiO_2 insulating layer void of holes. However, the electrons in the *p*-substrate (the minority carriers of the material) will be attracted to the positive gate and accumulate in the region near the surface of the SiO_2 layer. The SiO_2 layer and its insulating qualities will prevent the negative carriers from being absorbed at the gate terminal.

As V_{GS} increases in magnitude, the concentration of electrons near the SiO_2 surface increases until eventually the induced *n*-type region can support a measurable flow between drain and source.

Definition. The level of V_{GS} that results in the significant increase in drain current is called the **threshold voltage** and is given the symbol V_T. On specification sheets it is referred to as $V_{GS(Th)}$, although V_T is less unwieldy and will be used in the analysis to follow.

Definition. Since the channel is nonexistent with V_{GS} = 0 V and "enhanced" by the application of a positive gate-to-source voltage, this type of MOSFET is called an **enhancement-type MOSFET**. Both depletion- and enhancement-type MOSFETs have enhancement-type regions, but the label was applied to the latter since it is its only mode of operation.

As V_{GS} is increased beyond the threshold level, the density of free carriers in the induced channel will increase, resulting in an increased level of drain current. However, if we hold V_{GS} constant and increase the level of V_{DS}, the drain current will eventually reach a saturation level as occurred for the JFET and depletion-type MOSFET.

For the characteristics of Fig. 7.9 the level of V_T is 2 V, as revealed by the fact that the drain current has dropped to 0 mA. In general, therefore: *For values of V_{GS} less than the threshold level, the drain current of an enhancement-type MOSFET is 0 mA.*

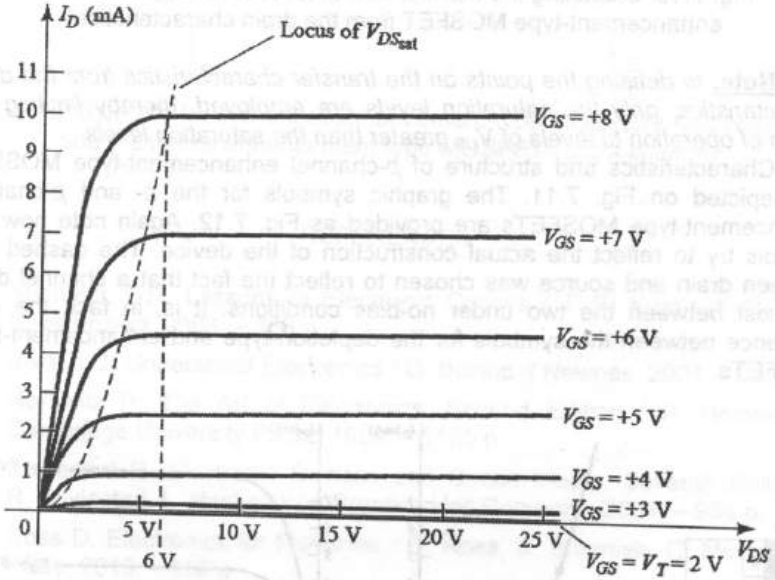


Fig. 7.9. Drain characteristics of an *n*-channel enhancement-type MOSFET with V_T = 2 V

In Fig. 7.10 the drain and transfer characteristics have been set side by side to describe the transfer process from one to the other. Essentially, it proceeds as introduced earlier for the JFET and depletion-type MOSFETs. In this case, however, it must be remembered that the drain current is 0 mA for $V_{GS} < V_T$. At this point a measurable current will result for I_D and will increase.

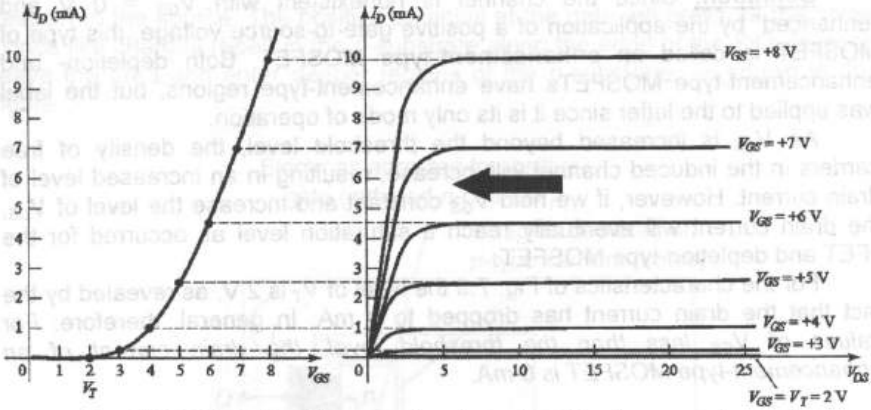


Fig. 7.10. Sketching the transfer characteristics for an *n*-channel enhancement-type MOSFET from the drain characteristics

Note. *In defining the points on the transfer characteristics from the drain characteristics, only the saturation levels are employed, thereby limiting the region of operation to levels of V_{DS} greater than the saturation levels.*

Characteristics and structure of *p*-channel enhancement-type MOSFET are depicted on Fig. 7.11. The graphic symbols for the *n*- and *p*-channel enhancement-type MOSFETs are provided as Fig. 7.12. Again note how the symbols try to reflect the actual construction of the device. The dashed line between drain and source was chosen to reflect the fact that a channel does not exist between the two under no-bias conditions. It is, in fact, the only difference between the symbols for the depletion-type and enhancement-type MOSFETs.

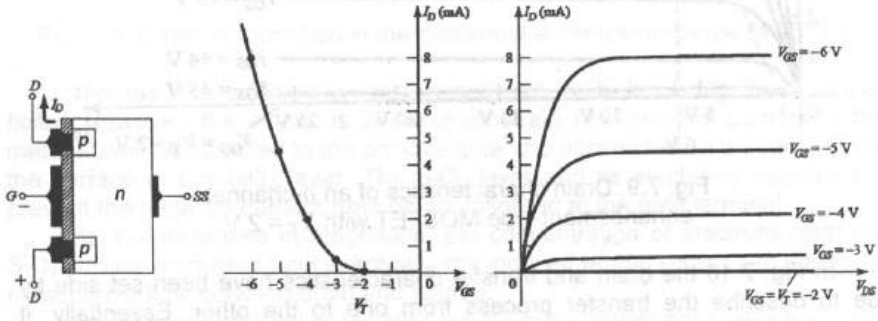


Fig. 7.11. *p*-channel enhancement-type MOSFET

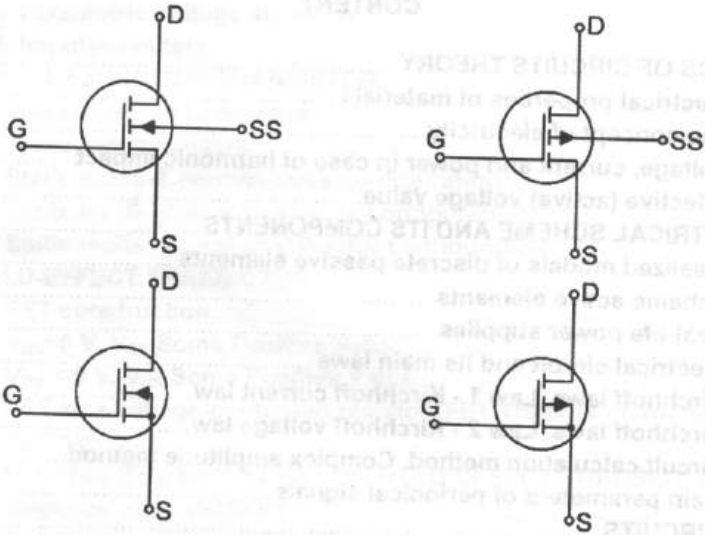


Fig. 7.12. Symbols for *n*-channel enhancement-type MOSFETs (*left column*) and *p*-channel enhancement-type MOSFETs (*right column*)

REFERENCES

1. Kuphaldt T. R. Lessons in Electronic Circuits / T. R. Kuphaldt // Open Book Project, 2003. – 366 p.
2. Bishop O. Understand Electronics / O. Bishop // Newnes, 2001. – 330 p.
3. Horowitz P. The Art of Electronics. Second Edition / P. Horowitz // Cambridge University Press, 1989. – 1125 p.
4. Boylestad R. Electronic Devices and Circuit theory. Seventh Edition / R. Boylestad, L. Nashelsky // Prentice Hall Company, 2001. – 934 p.
5. Ross D. Electronics for Dummies / D. Ross, C. Shamieh, G. McComb // Wiley, 2010. – 362 p.
6. Filipovic M. Understanding Electronics Components / M. Filipovic // in html-format.
7. Pierce J. F. Transistor Circuit Theory and Design / J. F. Pierce // Charles E. Merrill Books Inc., 1963. – 71 p.

CONTENT

Навчальне видання

Роєнко Олексій Олександрович

ТЕОРІЯ ЕЛЕКТРИЧНИХ КІЛ І МІКРОЕЛЕКТРОНІКА

(Англійською мовою)

Редактор Є. В. Пизіна
Технічний редактор Л. О. Кузьменко

Зв. план, 2014
Підписано до друку 30.12.2014
Формат 60x84 1/16. Папір офс. № 2. Офс. друк
Ум. друк. арк. 5,5. Обл.-вид. арк. 6,25. Наклад 100 пр.
Замовлення 404. Ціна вільна

Видавець і виготовлювач
Національний аерокосмічний університет ім. М. Є. Жуковського
«Харківський авіаційний інститут»
61070, Харків-70, вул. Чкалова, 17
http://www.khai.edu
Видавничий центр «ХАІ»
61070, Харків-70, вул. Чкалова, 17
izdat@khai.edu

Свідоцтво про внесення суб'єкта видавничої справи
до Державного реєстру видавців, виготовлювачів і розповсюджувачів
видавничої продукції сер. ДК № 391 від 30.03.2001