TRANSISTORS

Bruno Riccò
University of Bologna, Italy

Keywords: Transistors, BJT, MOSFET, Semiconductors, Electron Devices, Silicon Technology, Transistor Models, pn Junction, Microelectronics

Contents

Summary

Electronic systems are formed by circuits, in turn constituted by elementary components, connected to one another to perform specific functions. Among such components, by far the most important ones are semiconductor "transistors" that, consequently, represent the foundation of the "information society" of our time. Therefore, the importance of the subject cannot be overestimated, and any treatment of "circuits and systems" must assume at least basic knowledge about transistors. These, however, can be treated at different level of depth and details, from semiconductor physics to mathematic modeling. Hence, when tackling the subject a choice is always in order and this introductory chapter has decided to provide the "minimum" knowledge necessary to understand the rest of this volume, dedicated to circuits and systems realized with a number, often extremely large, of transistors. Such a choice involves a number of consequences: a) only the most important type of transistors (MOSFETs and BJTs) are considered, while others (JFETs, MESFETs, HEMTs,…) have been

sacrificed because of the available space; b) the treatment is focused on what is needed to “use” transistors, namely their external current vs. voltage characteristics, rather than to the physical reasons for their behavior (although a brief phenomenological description of transistor physics is given “for completeness” at the end of the chapter); c) no attempt is made to discuss transistor characteristics and performance from a quantitative point of view, as these aspects vary considerably to meet applications requirements and are critically dependent on technological details; d) only the “classic” transistor structures are considered, although to-day devices incorporate a number of structural variations. In spite of these limitations, the description given in this chapter should be sufficient to follow the rest of this volume and represent a useful entry for more complete and specialized analysis.

1. Introduction

Transistors are certainly one of the most important inventions of the XX Century and the foundation of all electronic circuits, hence of the whole Information and Communication Technologies (ICT) characterizing our modern society, hence it is appropriate that they open a volume dedicated to “Circuit and Systems”.

Transistors are based on semiconductor physics, are fabricated by means of a sophisticated technology, have been used in many different applications (for analog, digital and power applications, low power consumption, high frequencies,…) and have been developed in a variety of versions (BJT, MOSFET, JFET, MESFETS, IGFET,…).

As a whole, all these aspects of transistors represent a huge body of knowledge, that would require much more than the space allowed in this volume, primarily dedicated to their use and applications.
Thus, in planning this chapter a number of very drastic decisions have been necessary and it is important to state them clearly before entering into the subject.

At the beginning, the presentation that follows will be kept at a general, hence also somewhat abstract, level, for the sake of conciseness, but also in order to enlighten the essential commonality of transistors, regardless of specific implementations.

Then, when a choice becomes necessary, only two of the many types of transistors will be considered, namely the (so called) npn BJT and the n-channel, enhancement- mode MOSFET, by far the most important of all transistors. The reason for such a choice is that, starting from these two noticeable examples, the extrapolation to all other type of transistors should be (relatively) simple, at least as far as concepts are concerned.

The considered devices will be first treated “at their terminals”, i.e. from the point of view of their use, rather than starting from the physics they are based on. This priority seems most appropriate for a volume dedicated to “circuits and systems” (and not to electron devices). Furthermore, it also helps in emphasizing ideas and approaches common to all type of transistors.

In particular, the models representing the transistor in circuit design and analysis are briefly discussed.

Finally, a brief description of the basic physics of transistor is given. Here too, the discussion has been limited to the cases of the npn BJT and the n-channel, enhancement- mode MOSFET, and kept at a absolute (though hopefully reasonable) minimum from the point of view of the physical insight.

On the whole, the target will not be completeness and rigor, but rather an insight of transistors to people primarily interested in their motivated use.

2. Elementary Bricks of Electronics

All electronic circuits, digital, analog, or mixed, are made of large numbers of a few elementary components, put together to realize "architectures" that can feature outstanding complexity and performance. In general, such elementary components can be classified into two categories: passive (resistors, capacitors, connectors,…) and active (mainly transistors). These latter are much more important than the others, and in many cases are the only ones used to realize electronic circuits and systems. In particular, for example, microprocessors are essentially made only of transistors (nowadays billions of them) and of a (complex) systems of interconnects.

In brief, transistors represent the basic and essential building blocks of all electronics systems, and their outstanding properties (very small dimensions, high performance, extremely low power consumption, very high reliability,…) are the foundation of the whole Information and Communication Technology (ICT) characterizing the present human society. In turn, such properties are the result of a fabrication technology (overwhelmingly, though not exclusively, based on silicon) that is one of the most sophisticated ones ever developed by man, that has made it possible to fabricate extremely small, reliable and low-power devices, in the form of integrated circuits (ICs), and more recently also of "integrated systems" (ISs) of huge complexity. These are at the heart of very important "specific" products (computers, telecom systems, robots,…), but are also "embedded" in an ever growing number of everyday objects to make them "intelligent" and "communicative". For this reasons, microelectronics, i.e. the entire activity leading to the production of ICs and ISs (including design, testing, packaging,…) is considered an "enabling" and pervasive technology entering in most products of our society.

Therefore, the importance of transistors for modern economy and society, can only be underestimated.

From a physical point of view, transistors are small devices (in general this word indicates single objects, and its meaning changes with the context, so that, for example, it can be used to indicate single transistors, as it will be done in this chapter, but also entire ISs, such as as microprocessors). They exploit the electrical properties of semiconductors (in practice, overwhelmingly, though not exclusively, Silicon) and are fabricated with a long sequence of technological steps leading to the creation of elementary structures (junctions, interfaces, contacts,…).

The variety of possibilities in this field has produced many types of transistors, denoted with different names and acronyms (not unique even for the same device), such as, in particular: Bipolar Junction Transistor (BJT); Metal- Oxide- Semiconductor Transistor (MOST), also called Field Effect Transistor (FET) or MOSFET; Junction Field Effect Transistor (JFET); Metal – Semiconductor Field Effect Transistor (MESFET); Insulated Gate Field Effect Transistor (IGFET); Hetero-junction Bipolar Transistor (HBT);High Electron Mobility Transistor (HEMT)….

Of all these devices, the MOSFET and the BJT are certainly the most important and the only ones that will be treated here.

Regardless of their specific type, however, transistors implement (with different values of such relevant parameters as dimension, speed, power consumption…) the abstract concept of "electronic valve", on which all electronics is based.

Therefore, it is convenient to start from such a concept, keeping the discussion as general as possible. Furthermore, we will first focus on what is needed to use transistors in the design of electronic circuits. Only after this, we will (briefly) consider the physical basis of transistors, in order to provide an idea of how they work and are fabricated.

3. The Basic Electronic Device

The electronic valve is a device featuring "two ports", each formed by a couple of terminals, one used as input and the other as output, as illustrated in Fig. 1.

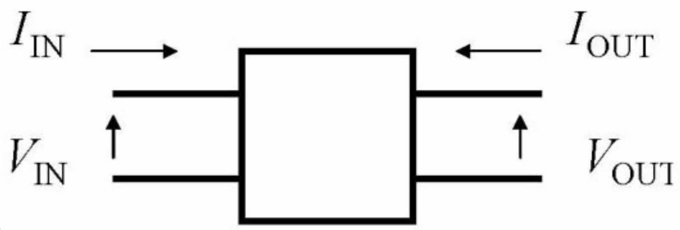

Figure 1. Schematic ("black-box") representation of an electronic valve

The input port, characterized by the voltage V_{IN} and the current I_{IN}, is used to "control" the electrical state of the output one, described by the voltage V_{OUT} and current I_{OUT} . The physics of the device within the "black" box separating the two ports, of course, determines the relationship between the couples of variables ($V_{IN,}I_{IN}$) and (V_{OUT}, I_{OUT}). Furthermore, in general, such a relationship depends also on the operating conditions (static, dynamic,…) as well on a number of external parameters (such as, for instance, temperature).

The control implicit in the concept of "valve" means that the electrical status of the output can be determined acting on the input (and, at first approximation, not vice versa) Furthermore, in practice, the power needed at the input port is much lower than that at the output one.

Thus, in simple terms, "it is possible to control large effects with little effort".

In particular, changing V_{IN} and/or I_{IN} the output port can be switched off (i.e. making $I_{\mathrm{OUT}} \approx 0$) or on ($I_{\mathrm{OUT}}$ substantially $\neq 0$). When this is the only effect of interest, the valve is used as an electronic (controlled) switch, foundation of all digital circuits. In this case, the most important features of the operation are the current difference between on and off state and the speed with which the device can be switched from one state to the other.

On the contrary, in analog circuits transistors are used to process “signals”, i.e. continuous variations of voltages and/or currents. In this case, the device is biased in steady state, i.e. both ports are on and feature static currents and voltages to which the signals are superimposed, while the most important operation is signal amplification, namely reproduction, but with larger amplitude, at the output port of the signal applied at the input. The main parameters of interest are the “gain”, i.e. the ratio of the output over input signal amplitude, and the bandwidth, i.e. the maximum signal frequency at which such a gain is satisfactorily high.

At a first, intuitive but rough, approximation the behavior described above can also be interpreted as if at the output port the device presented a resistance whose value can be controlled by acting on the input. From such a vision, the idea of a “transfer resistor”, hence the name “transistor”.

Without entering in more details, it is important to realize that the input/ output relationship of the valve is the only information required to use the device, since it describes the way in which it interacts with the external world (in practice with other devices). In particular, for this purpose it is not necessary to understand the physics determining such a relationship, although this is of course always useful (for instance to support motivated choices).

4. Device Characteristics

The relationship between ($V_{\mathrm{IN},} I_{I\mathrm{N}}$) and ($V_{\mathrm{OUT}}, I_{\mathrm{OUT}}$) can be determined experimentally, and be represented by means of mathematical (or numerical) models to be used in circuit design. In general, it depends not only on the device itself, but also from the circuit connected to the output port and this, of course, makes it extremely difficult to dominate the situation.

For this reason, it is common to divide the problem into two sub-parts: one representing the output- input relationship under static conditions, i.e. with $V_{\mathrm{IN},} I_{I\mathrm{N}}, V_{\mathrm{OUT}}$ and I_{OUT} fixed in time (or changing slowly); the other describing the dynamic effects due to (rapid) variation of the electrical situation.

The first part gives rise to the so called “static characteristics”, important because they are due only to the device itself and are essential to determine the “operating point” of the device in analog applications. These characteristics are also called “large signal”, because they consider the actual values of currents and voltages.

As for dynamic modeling, the problem is tackled adding (lumped) capacitors at the device terminal. Such elements, approximately account for the changes of the electric charges within the device depending on applied voltages. For each of such charges (Q_{N}) and terminal voltage $\left(V_{\mathrm{LM}}\right)$ from which it depends, i.e. for each (generic) relationship $Q_{\mathrm{N}} = f\left(V_{\mathrm{LM}}\right)$ when the voltage changes, so does the charge, thus a dynamic current (i_{LM}) flows between the L and M terminals. Such a current can then be described defining the differential capacitance (capacitor)

$$C_{LM} = dQ_N / dV_{LM} \, , \tag{1}$$

to be used in equations such as: .

$$i_{\mathrm{LM}} = C_{\mathrm{LM}} dV_{\mathrm{LM}} / dt \, . \tag{2}$$

Each voltage dependent charge within the device gives rise to differential capacitances (one for each voltage it depends on). Such capacitances, however, depend on the physics of the device and cannot be described in abstract form.

As for the static characteristics linking V_{OUT} and I_{OUT} to V_{IN} and I_{IN}, the most common (and useful) are the (family of) curves:

- I_{IN} vs. V_{IN} (input characteristic);
- I_{OUT} vs. V_{IN} , with V_{OUT} as a (fixed) parameter (trans-characteristic) ;
- I_{OUT} vs. V_{OUT} with V_{IN} as a (fixed) parameter (output characteristic).

The second of these curves is called “transfer characteristics”, as it links the current at the output port to the voltage applied at input terminals.

These characteristics looks as schematically illustrated in Figure 2, and, as can be seen, a value of (the module of) V_{IN}, called “threshold voltage (V_{T}) exists, below which the transistors is switched off (i.e. I_{OUT} is negligible for all practical purposes).

The transistor static characteristics can also be modeled by means of either a set of analytical expression, or by different numerical methods (for instance look-up tables).

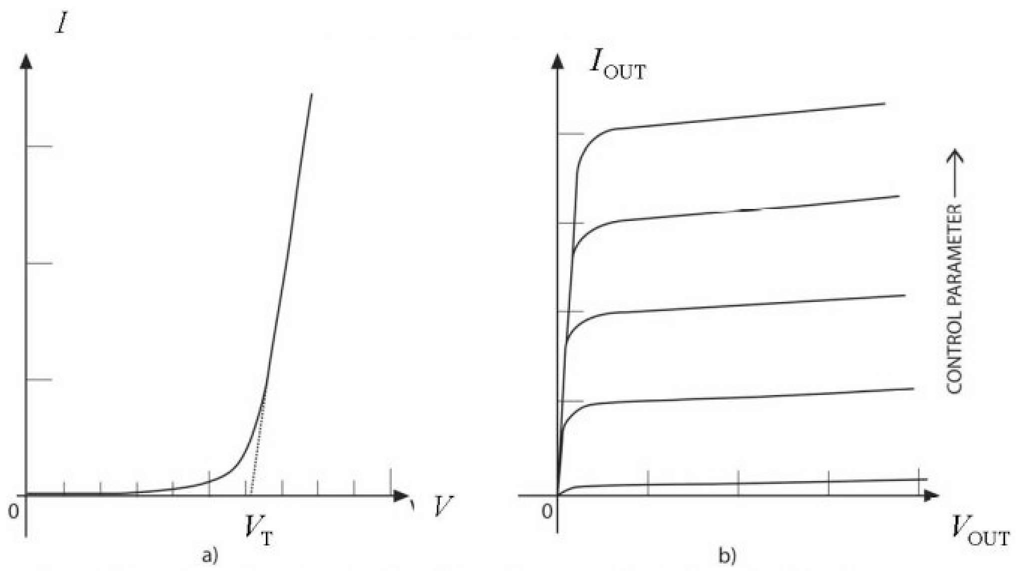

Figure 2. Typical transistor characteristics. a) the voltage on the horizontal axis is V_{IN} ; the curve describes the "input characteristic" or the "trans- characteristics" if the current on the vertical axis is I_{IN} or I_{OUT} , respectively. In both cases, the existence of a "threshold voltage" for the transistor turn-on is illustrated. b) these curves represent the transistor "output characteristic" linking I_{OUT} and V_{OUT} . The control parameter of each curve can be either a current (in particular $I_{IN} = I_B$ in the BJT), or a voltage ($V_{IN} = V_{GS}$ in MOSFETs).

The analytical models, in principle derived from the device physics, are preferable, because they provide more insight and should be independent of geometry, dimensions,… However, the physics of modern transistors is very complex and the attempt to capture it all with analytical equations normally leads to inaccurate description of the current-voltage relationships.

As already mentioned, in analog application the device is biased with fixed values of $V_{IN,} I_{IN}, V_{OUT},$, and I_{OUT} to which small signals ($\delta V_{IN,} \delta I_{IN}, \delta V_{OUT}$, and δI_{OUT}) are superimposed, In this case, the interest is more on the relationship between the signals rather than on the real (i.e. entire) voltages and currents. To deal with analog circuits, models are then required that consider only the signals and get rid of the bias voltages and currents. Normally, this model is derived assuming that the signal are slow, so that the device behavior can still be (approximately) described by its static model, whose equations are then linearized essentially as in the following example:

$$I_{OUT} = f\left(V_{IN}, V_{OUT}\right), \text{ hence, } I_{OUT0} + \delta I_{OUT} = f\left(V_{IN0} + \delta V_{IN}, V_{OUT0} + \delta I_{OUT}\right) \quad (3)$$

Here, the sub 0 indicates bias (i.e. steady state) values of voltages and currents. Expanding the last term of this equation up to first order (around the bias point), it is

$$I_{OUT0} + \delta I_{OUT} = f\left(V_{IN0}, V_{OUT0}\right) + \left(dI_{OUT}/dV_{IN}\right)_0 \delta V_{IN} + \left(dI_{OUT}/dV_{OUT}\right)_0 \delta V_{OUT} \quad (4)$$

where,

$$f\left(V_{\mathrm{IN0,}}V_{\mathrm{OUT0}}\right)=I_{\mathrm{OUT0}}.$$

Then, defining:

$$\left(dI_{\mathrm{OUT}}/dV_{\mathrm{IN}}\right)_0=g_{\mathrm{m}} \text{ and } \left(dI_{\mathrm{OUT}}/dV_{\mathrm{OUT}}\right)_0=g_{\mathrm{OUT}}\ , \qquad (5)$$

from Eq. (4),

$$dI_{\mathrm{OUT}}=g_{\mathrm{m}}\,\delta V_{\mathrm{IN}}+g_{\mathrm{OUT}}\delta V_{\mathrm{OUT}} \qquad (6)$$

Equation (6) describes a relationship among signals and can be represented in terms of a “small signal equivalent circuit” (in this particular case formed by two controlled current generators whose parameters, g_{m} and g_{OUT} are derivatives taken at the device bias point).

As for the definitions of Eq. (3), g_{m} is called “trans-conductance”, to recall that controlled current and controlling voltage refer to different device terminals; g_{OUT}, instead, is called “output conductance”, as it links current and voltage at the same (output) port.

-
-
-

Bibliography

Since transistors represent the foundation of modern microelectronics, all books dedicated to electronic “Circuits” and/or “Systems” contain chapters on transistors, normally treated at the “external terminals” (I-V characteristics and simple models), just as in this chapter. Examples of this kind of books are:

1. A. S. Sedra and K. C. Smith, (1990) “Microelectronic Circuits”, Saunders College Publishing, 3rd edition

2. J. M. Rabaey, (1996) “Digital Integrated Circuits: A Design Perspective”, Prentice Hall Electronics and VLSI Series

3. H. Haznedar, (1991) “Digital Microelectronics”, The Benjamin/ Cummings Publishing Company,

Furthermore, many books are specifically dedicated to semiconductor devices, and deal with device physics and technology in much more details than it is done here. Among these books, for instance:

4. A. Barlev, (1992) "Semiconductor and Electronic Devices", Prentice Hall, Third edition.

5. H. C. Casey, jr., (1999) "Devices for Integrated Circuits", John Wiley & Sons.

6. D. A. Neamen, (1997) "Semiconductor Physics and Devices”, Second Edition, Irwin.

7. R. F. Pierret, (1996) "Semiconductor Device Fundamentals", Addison-Wesley.

8. D. L. Pulfrey and N. G. Tarr, (1989) "Introduction of Microelectronic Devices", Prentice Hall.

9. D. J. Roulston, (1999) "An Introduction to the Physics of Semiconductor Devices", Oxford.

10. M. Shur, (1990) "Physics of Semiconductor Devices", Prentice Hall.

11. J. Singh, (2001) "Semiconductor Devices", John Wiley & Sons.

12. G. Streetman and S. Banerjee, (2000) "Solid State Electronic Devices, Fifth Edition", Prentice Hall.

13. S. M. Sze, (1981) "Physics of Semiconductor Devices, Second Edition", John Wiley & Sons.

14. S. Wang, (1989) "Fundamentals of Semiconductor Theory and Device Physics", Prentice Hall.

Biographical Sketch

Bruno Riccò graduated in Electronics at the University of Bologna (Italy) in 1971; in 1975 he received a Ph.D. in Physics from the University of Cambridge (U.K.) where he worked at the Cavendish Laboratory; in 1980 he became Full Professor of Electronics at the University of Padua (Italy) and in 1983 at the University of Bologna (Italy); in the period 1981 – 1986 he was Visiting Professor at the University of Stanford, at the IBM Thomas J. Watson Research Center (Yorktown Heights) and at the University of Washington; from 1986 to 1996 he was European Editor of the IEEE Transaction on Electron Devices; in 1995 he received the G. Marconi Award for research by the Italian Association of Electrical and Electronics Engineers (AEI); 1998 he became President of the Italian Group of Electronics Engineers; in 1999 he was appointed European representative for the International Electron Device Meeting (IEDM); in 1999 he founded the first university spin-off in Italy; in 2002 he was elected Chairman of the IEEE North Italy Section; in 2003 he was nominated Fellow of the IEEE; since 2004 he is the President of the company incubator “I Tech Off; since 1999 he is the President of the T3Lab, an initiative of the University of Bologna and Confindustria Bologna for Technology Transfer. In his career, Prof. Riccò has held several courses on Electron Devices, Digital ICs, Microelectronics, IC Testing, Self-Checking Architectures and Reliability. As for research, he worked in the field of microelectronics and is (co-) author of over 350 publications, more than half published on major international Journals. He also published 3 textbooks and holds several international patents.