PREVIEW

Download the complete book

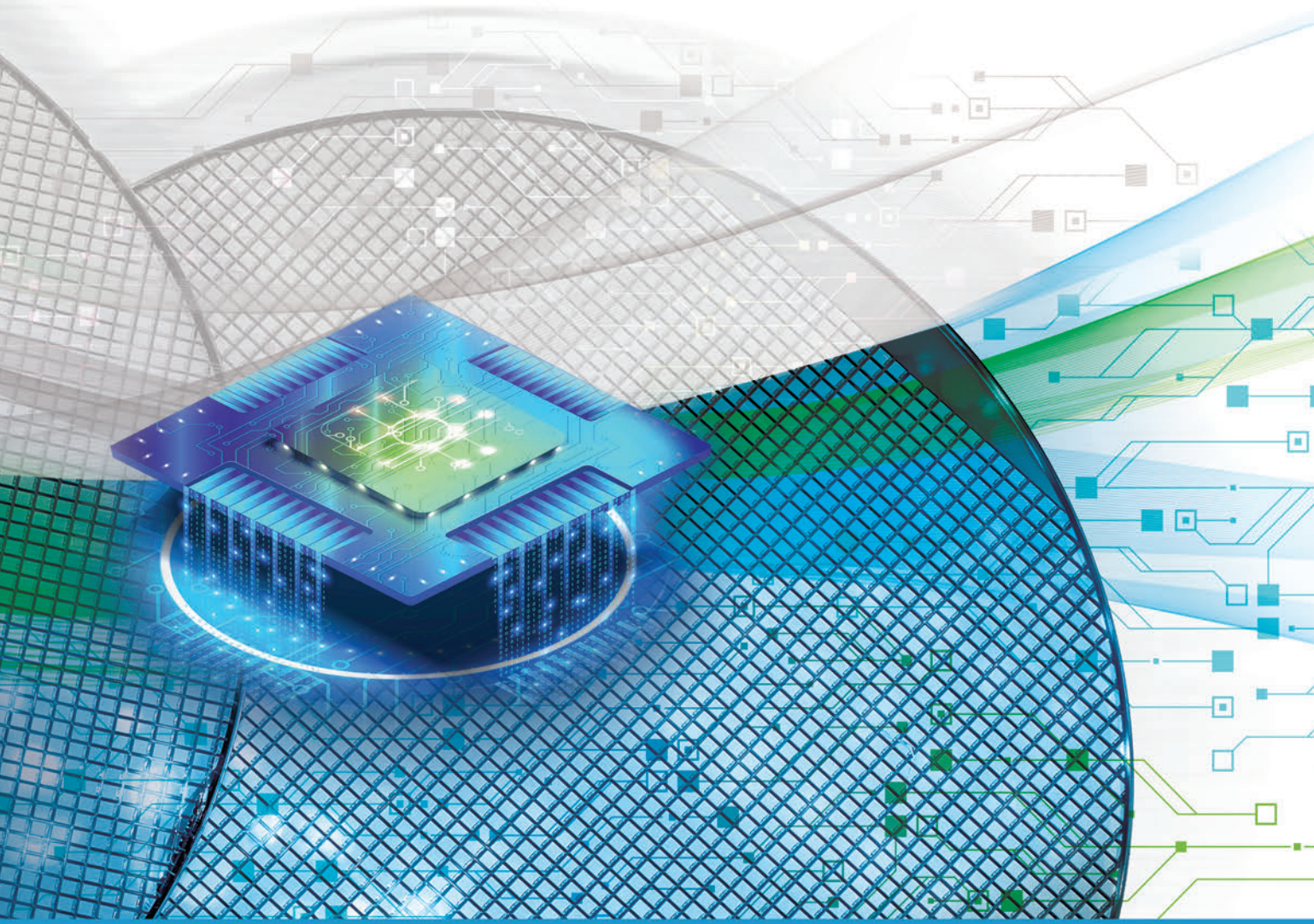
www.mks.com/mks-handbook

mks

MKS INSTRUMENTS HANDBOOK

Semiconductor Devices and Process Technology

2ND EDITION *by the Office of the CTO*



MKS Instruments Handbook 2nd Edition

Semiconductor Devices and Process Technology

by the Office of the CTO

The 2nd edition of "Semiconductor Devices and Process Technology" handbook is now available as **a free download at www.mks.com/mks-handbook.** As the Semiconductor industry continues to push the boundaries of possibility, new materials and processes are needed to support smaller and more intricate device structures. The 2nd edition Handbook includes new chapters on plasma technology and advanced temperature measurement, as well as updates and more in-depth information on thin films, vacuum technology, ion implantation, thermal processing, deposition, etch, chemical mechanical polishing (CMP), lithography, wafer inspection, and CMOS device structures and process steps. Initially created as an internal training source for our employees, the handbook has evolved into a go-to technical resource for our partners and customers.

Download the complete handbook **here.**

6/2023

MKS Instruments Handbook 2nd Edition

Semiconductor Devices and Process Technology

by the Office of the CTO

Contributing Authors

Larry Bourget
Gerardo Brucker, PhD
Mike Feaver
Gordon Hill
Yoshua Ichihashi
Keith Koai, PhD
Amanda Larson, PhD
Christiane Le Tiec, PhD
Linnell Martinez, PhD
Ilya Pokidov
Aaron Radomski
Guy Rosenzweig, PhD
Efi Rotem, PhD
Chiu-Ying Tai, PhD
Marc Tricard
Marijke Van Spyk, PhD

Reviewers

Anina Barth, PhD; Roberto Bosco, Remi Chazalmartin, Josh Ding, PhD; Dan Hazen, Eric Hill, Jim Kafka, PhD; Ann Kazakos, PhD; Alain Jeanne, PhD; Jaechul Jung, Dave Kelly, Moran Kuriel, John Liao, Jimmy Lui, PhD; Jean Louis Nicque, PhD; Joris Nunen, Andrei Rusu, Vyacheslav Ryaboy, PhD; Yunmei Shen, Roger Sonnemans, Guofeng Sun, Ike Ke, Sarah White, PhD; Chenglong Yang, PhD; Jim Ye, PhD; Hossein Zarrin, Oleg Zinoviev

Chief Editor

Gordon Hill

Contributing Editor

Tom K. Whidden, PhD, TechLink Writers Group

Graphics & Production

Crystal Heroux
Jane Lawlor

Foreword

Semiconductor devices have become essential and ubiquitous parts of our everyday lives, enabled by an industry that has pursued a relentless reduction in cost per function for over 60 years. Today's devices contain billions of transistors interconnected by kilometers of "wires," yet all fitting in a fingernail-sized integrated circuit. These devices are only possible because of manufacturing processes that continue to increase in complexity and require unprecedented innovation in physical structure, new materials, process control, and throughput. And looking ahead, the new challenges presented by the inexorable progression toward even more miniaturization and complexity only raise the demands for innovation and creativity.

Some of the most important unit processes necessary to sustain this pace of innovation include Etching, Deposition, Lithography, Metrology, Inspection, and Wet Cleaning. Improvements in these unit processes, in turn, are dependent on critical subsystem technologies and instruments that measure, control, and sustain pristine vacuum systems, deliver critical gases, reactive species, and power to these systems, and provide state-of-the-art laser, optics, and photonics capabilities to define and measure sub-nanometer features. MKS Instruments uniquely serves this market with the broadest portfolio of such solutions in the industry, or what we call our "Surround the Wafer℠" strategy.

Realizing that the trends of miniaturization and complexity are extending beyond the semiconductor chip to the other critical building blocks of advanced electronic devices – such as Advanced PCBs and Package Substrates, MKS expanded its portfolio of foundational solutions with the acquisition of Atotech, a critical provider of process chemistry solutions. The combination of MKS' laser drilling solutions, with Atotech's proprietary chemistry and plating equipment, positions MKS with a unique combination of foundational solutions to enable next generation interconnect formation—known as MKS' Optimize the Interconnect℠ offering.

The Semiconductor Devices and Process Technology Handbook presents the fundamental device physics, materials, and fabrication processes used to manufacture semiconductors, as well as the technologies, instruments, and equipment that are used to monitor, control, and automate the fabrication processes. This handbook was initially conceived as training material for our employees. As it has evolved, we want to share it with you, our customers, suppliers, and partners in the semiconductor industry, with the hope that you will find it informative and of value to your work.

This handbook came about thanks to what we learned from working alongside you for over half a century. In that spirit, we welcome your feedback and input to future editions. Rapidly evolving technology will continue to shape and transform this industry in the years to come. MKS Instruments looks forward to facing the opportunities and challenges this will present—together.

John T.C. Lee

President & Chief Executive Officer

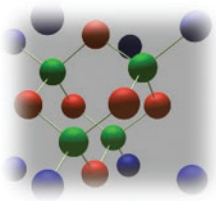
About MKS Instruments

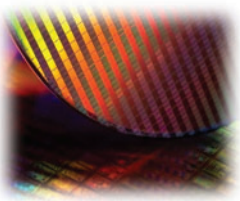
MKS Instruments enables technologies that transform our world. We deliver foundational technology solutions to leading edge semiconductor manufacturing, electronics, and packaging, and specialty industrial applications. We apply our broad science and engineering capabilities to create instruments, subsystems, systems, process control solutions, and specialty chemicals technology that improve process performance, optimize productivity, and enable unique innovations for many of the world's leading technology and industrial companies. Our solutions are critical to addressing the challenges of miniaturization and complexity in advanced device manufacturing by enabling increased power, speed, feature enhancement, and optimized connectivity. Our solutions are also critical to addressing ever-increasing performance requirements across a wide array of specialty industrial applications. Additional information can be found at www.mks.com.

Table of Contents

Section A

Semiconductor Device Physics, Materials, and Fabrication





I. Semiconductor Physics and Basic Device Structures

The Intel 4004 microprocessor was released in 1971 with approximately two thousand 10-micron (a millionth of a meter) transistors and a 500kHz clock frequency. The Intel Core i9-9990x released in 2018 has approximately six billion 14-nanometer (a billionth of a meter) transistors and a 4GHz clock frequency. This remarkable improvement in miniaturization and computing performance has required significant advancements in materials and device structures:

- Reduction of interconnect resistance and capacitance using new materials for metalization and dielectric materials.
- Control of leakage currents (maintaining high on/off current ratios) in transistors via high-K gate dielectrics and novel device structures (i.e., FinFET, Gate All Around, and nanosheet transistors).
- Vertical stacking of structures to enable higher density.

This section will provide a brief introduction to the underlying physics of the devices, materials used, and how they are structured/manufactured.

A. Semiconductor Physics

1. Electrical Characteristics of Solids

When it comes to electrical characteristics, solid matter falls into one of three categories (Table 1): it is either a conductor, an insulator, or a semiconductor, depending on whether or how well it conducts electricity. Conductors have many "mobile charge carriers," i.e., negative electrons. These are not strongly bound to individual atoms, and they can be easily made to flow through the conductor by the application of an electrical potential [1], [2]. Examples of solid conductors include all the familiar metals, along with semi-metals such as graphite, α-tin, bismuth, antimony, and arsenic. Insulators, on the other hand, have their electrons tightly bound to the atoms; it takes a great deal of energy (i.e., a very high electrical potential) to remove them from the atoms so that they can flow through the material. Examples of insulators include glass, ceramics, plastics, and wood.

Conductors	• Many free electrons, which can easily be made to flow through material • Examples: all metals, semi-metals such as carbon-graphite, antimony and arsenic
Insulators	• Very few free electrons • Examples: plastic, glass and wood
Semiconductors	• Between the extremes of good conductors and good insulators • Crystalline materials that are insulators when pure, but will conduct when impurity is added and/or in response to light, heat, voltage, etc. • Examples: elements such as silicon Si, germanium Ge, selenium Se; compounds such as gallium arsenide GaAs and indium antimonide InSb

Table 1. Conductors, insulators, and semiconductors.

Semiconductors lie between these extremes of electrical conductivity. They are high purity crystalline materials that conduct electricity under relatively low electrical potential (when compared to insulators), without the high electrical conductivity of metals. When certain impurities are added to the crystal matrix of a semiconducting material, they contribute mobile charge carriers (either electrons or holes) that can dramatically increase the electrical conductivity of the semiconductor under the proper conditions. Semiconducting materials include elements such as silicon and germanium, along with compound (i.e., multi-element) semiconductors such as gallium arsenide (GaAs), gallium nitride (GaN), silicon carbide (SiC), indium antimonide (InSb), and indium phosphide (InP). Figure 1 shows a Periodic Table of the Elements. The elemental semiconducting materials are all from Group IVA in this table, while compound semiconducting materials are formed through the combination of elements from Groups IIIA and VA (commonly referred to as III-V semiconductors) or Groups IIB and VIA (II-VI semiconductors).

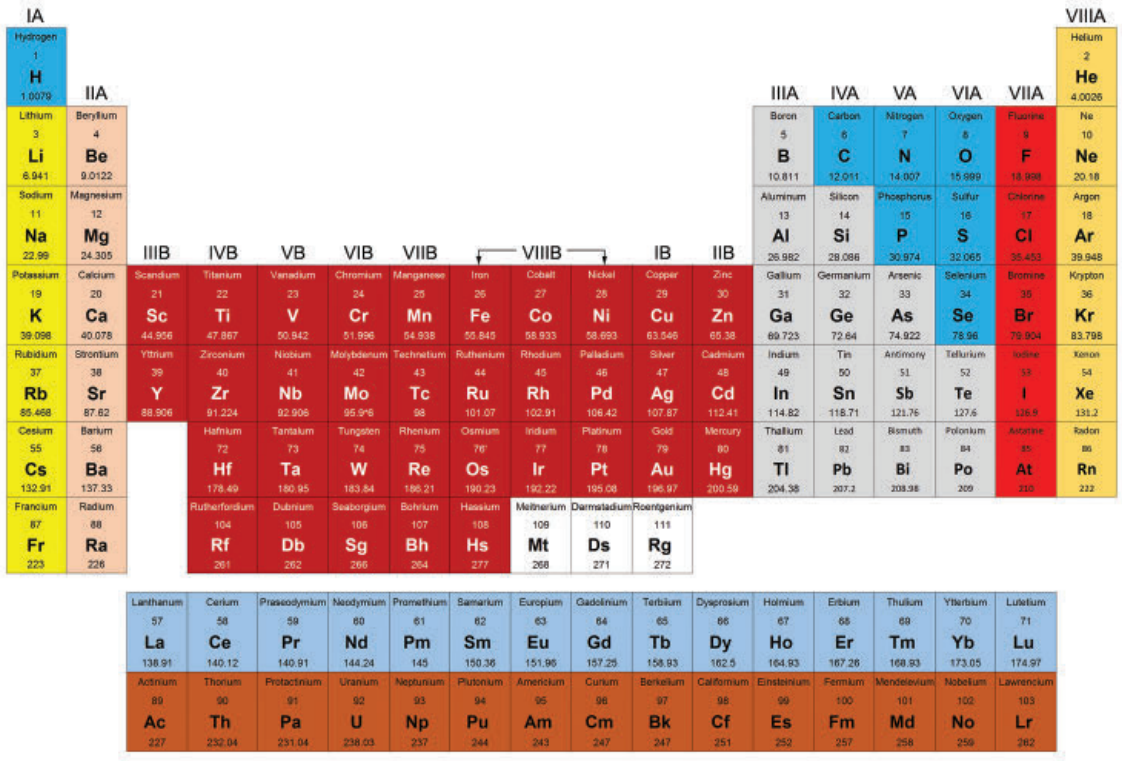


Figure 1. Periodic Table of the Elements.

Electrical conduction in solids is best understood in terms of a model that physicists call "Band Theory." The next section provides an abridged description of how electrical conduction occurs in semiconductors; some knowledge of atoms, electrons, chemical bonds, and electron orbitals is required. Those unfamiliar or "rusty" on these concepts can find readily available texts as well as entertaining and instructive internet videos on the basics [3], [4], [5], [6], [7].

2. Electrical Conduction in Semiconductors

Pure semiconductor crystals are not particularly good electrical conductors, although their conductivities are much greater than those of insulators (Table 2). The electrical property that makes semiconducting materials, and especially silicon, so valuable in electronics and other device applications arises from the fact that their electrical conductivity can be continuously varied through the controlled incorporation of dopant atoms into the crystal lattice. This property allows doped and undoped silicon to be used for the control of electrical current in a multitude of electronic devices, including diodes, capacitors, and transistors. Transistors, especially, are extremely important in modern technology. They can be likened to a valve that controls the flow of electricity: a Metal-Oxide-Semiconductor Field Effect Transistor (MOSFET) device, Figure 2(a), behaves as an electrical On/Off switch, while a Bipolar Junction Transistor (BJT), Figure 2(b) behaves primarily as an amplifier (controlling signals) in its linear range. Such transistors can only be fabricated using semiconductor materials.

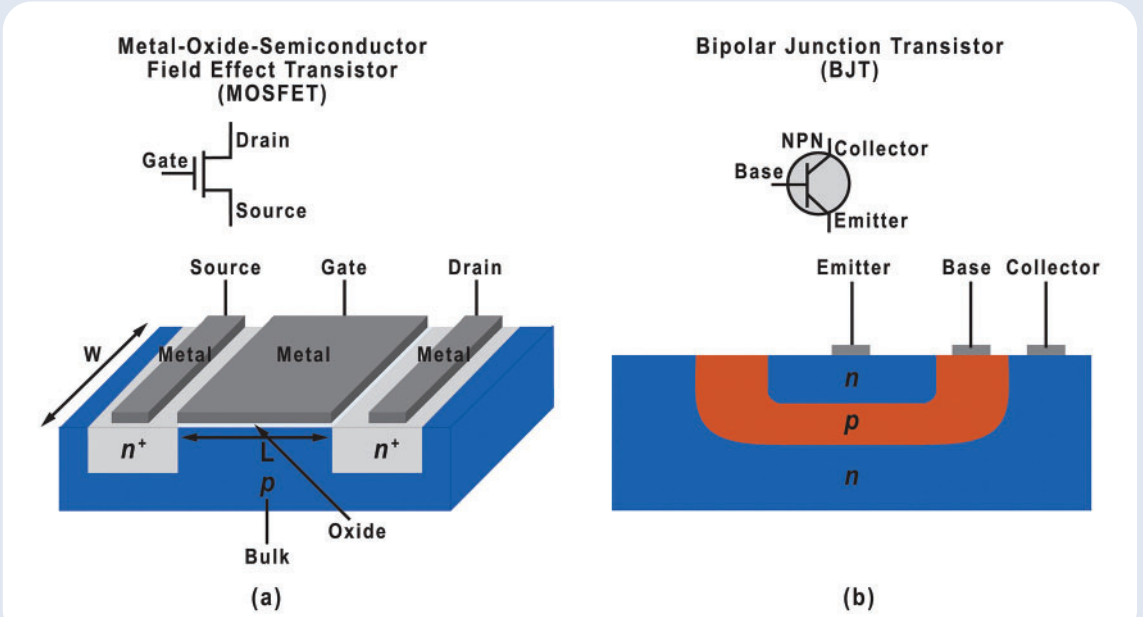


Figure 2. (a) Metal-Oxide-Semiconductor Field Effect Transistor (MOSFET) symbol and structure; (b) Bipolar Junction Transistor (BJT) symbol and structure.

Today silicon is the primary semiconducting material employed for large scale electronic device manufacturing and we will use it to illustrate the electrical properties of semiconductors. Keep in mind that the properties described for silicon can equally describe other semiconducting elements in Table 2, as well as the various compound semiconductors (GaAs, InP, etc.), that are used in more limited applications (albeit with some variation in the finer details).

Material	Resistivity (Ω-cm)	Conductivity (Ω^{-1}-cm^{-1})
Insulators		
Hard Rubber	1-100 x 10^{13}	1 × 10^{-15} to 1 × 10^{-13}
Glass	1-10000 x 10^{9}	1 × 10^{-13} to 1 × 10^{-9}
Quartz (fused)	7.5 x 10^{17}	1.33 × 10^{-18}
Semiconductors		
Carbon (Graphite)	3-60 x 10^{-5}	1.67 × 10^{3} to 3.33 × 10^{4}
Germanium	1-500 x 10^{-3}	2.0 to 1.00 × 10^{3}
Silicon	0.10-60	1.67 × 10^{-2} to 10
Metals		
Silver	1.63 x 10^{-8}	6.17 × 10^{7}
Copper (annealed)	1.72 x 10^{-8}	5.95 × 10^{7}
Aluminum	2.65 x 10^{-8}	3.77 × 10^{7}

Table 2. Relative electrical resistivities and conductivities of some metals.

Pure silicon and other elemental semiconductors are commonly referred to as intrinsic semiconductors [8], [9]. The term intrinsic means that electrical conductivity is an inherent property of the semiconducting material and independent of the presence of additives. In addition to intrinsic semiconductors, semiconductor device fabrication also makes use of extrinsic semiconductors that depend on the presence of dopant impurities to increase the electrical conductivity of the material. Table 2 shows the room-temperature resistivities and conductivities for several metals, intrinsic semiconductors, and insulators. Elemental silicon has a resistivity of 0.10-60 Ω-cm—significantly higher than metallic resistivities; however, it is many orders of magnitude smaller than the resistivity of a typical insulator.

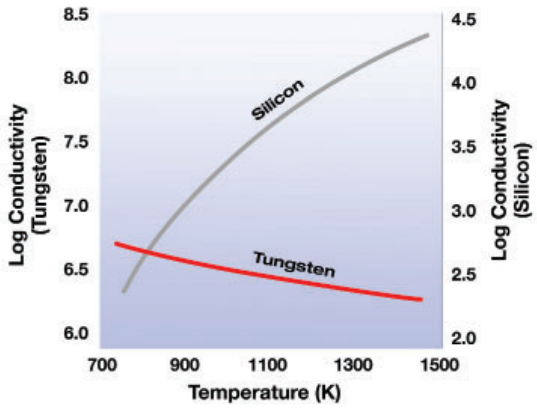


Figure 3. The electrical conductivity characteristics of a semiconductor (silicon) vs. those of a metal (tungsten) [10].

A further difference between the electrical characteristics of semiconductors and those of metals lies in the way their conductivities vary with temperature. At absolute zero (0 K), the electrical conductivity of a semiconductor has a value of zero (i.e., the conductivity is at its minimum) whereas a metal exhibits its maximum electrical conductivity at absolute zero; furthermore, conductivity increases with increasing temperature in a semiconductor, whereas it goes down with increasing temperature in a metal. Figure 3 shows a graph comparing the temperature variation of the electrical conductivity of a semiconductor (silicon) versus a metal (tungsten). The different behaviors of electrical conductivities seen in Figure 3 can be understood in terms of the different mechanisms that govern electrical conductivity in a metal versus a semiconductor.

Metals have many free electrons available to conduct electricity. In a metal, the valence band and the conduction band overlap, so electrons are free to move into many available and vacant energy levels (see Appendix A for a description of valence and conduction bands). When an electrical potential is applied across a piece of metal, these electrons can flow freely from higher to lower potential. The conductivity is limited only by the amount of electron scattering that occurs due to collisions between the flowing electrons and fixed atoms in the metal lattice. At 0 K, the atoms in the metal lattice are at rest and electron scattering is at its lowest possible value. As the temperature of the metal increases, thermal energy causes the atoms in the lattice to vibrate. These vibrations increase the effective diameter which an atom presents to the flow of electrons, increasing each atom's ability to scatter electrons.

Thus, as the temperature rises from 0 K, the metal atoms vibrate with continuously increasing amplitudes, increasing the degree of electron scattering and producing the observed reduction in conductivity seen in Figure 3.

Figure 4 can help in understanding how the mechanism of electrical current flow in a semiconductor differs from current flow in a metal. Figure 4(a) shows silicon atoms in a crystal, along with a crude schematic of how the valence electrons associated with each silicon atom are shared to produce four bonds containing eight electrons in every atom's valence shell. As you will recall from basic chemistry, covalently bound atoms require eight electrons in their valence shell to be stable. As Figure 4(a) shows, every valence electron in a silicon atom is involved in a covalent bond and so, unlike in a metal, none are free for electrical conduction; silicon, at absolute zero, is an insulator.

Figure 4(b) shows a crude schematic of the band structure (see Appendix A) of silicon at 0 K vs. the same structure at 298 K (25°C). At 0 K, the valence band is full and the conduction band empty and there is no thermal energy available to raise a valence electron into the conduction band. As the temperature increases, the silicon absorbs thermal energy. While this energy increases the thermal vibrations of the silicon atoms (producing increased atom-electron scattering), any loss in electrical conductivity due to this phenomenon is swamped by the increase in electrical conductivity due to the thermal promotion of electrons from the valence to the conduction band. This effect is not present in metals since the conduction band overlaps the valence band. Promotion of an electron in the valence band to the conduction band creates an electron-hole pair which constitutes mobile carriers in both the conduction band (electrons) and the valence band (holes). Electrical current will now flow in the material when a potential is applied. From Figure 4(b), we can see that in an intrinsic semiconductor, there is a balance between the number of

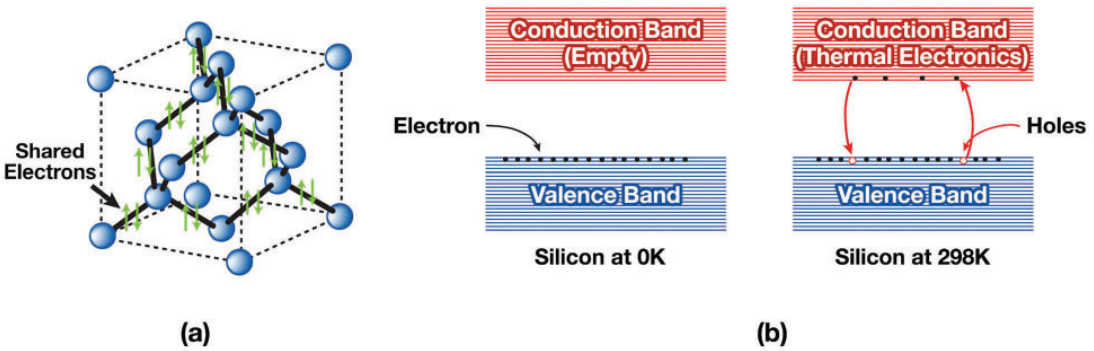


Figure 4. (a) Metal-Oxide-Semiconductor Field Effect Transistor (MOSFET) symbol and structure; (b) Bipolar Junction Transistor (BJT) symbol and structure.

electrons in the conduction band and the number of holes in the valence band. The concept of holes is important in discussions of electronic devices because electrical current is conventionally depicted as being due to the movement of holes. Since holes represent the absence of negative charges (electrons), it is useful to think of them as positive charges. By convention, electric fields [10], [11] (and therefore electric potential) are depicted as a vector extending outward from areas of positive charge. Therefore, when considering current flow, electrons move in a direction opposite to the applied electric field direction (i.e., toward the positive charge), while holes move in the direction of the electric field.

When compared with metals, semiconductor electrical conductivities are not very high, as can be seen from the values in Table 2. However, semiconductor conductivities can be significantly increased (and, in fact, fine-tuned to target values) by adding impurity atoms known as dopants. "Doped" semiconductor materials are commonly referred to as extrinsic semiconductors [8], [12]. Figure 5 shows the area of the Periodic Table that contains the elemental semiconductors and the dopant elements that can be used to create an extrinsic semiconductor. The dopants are

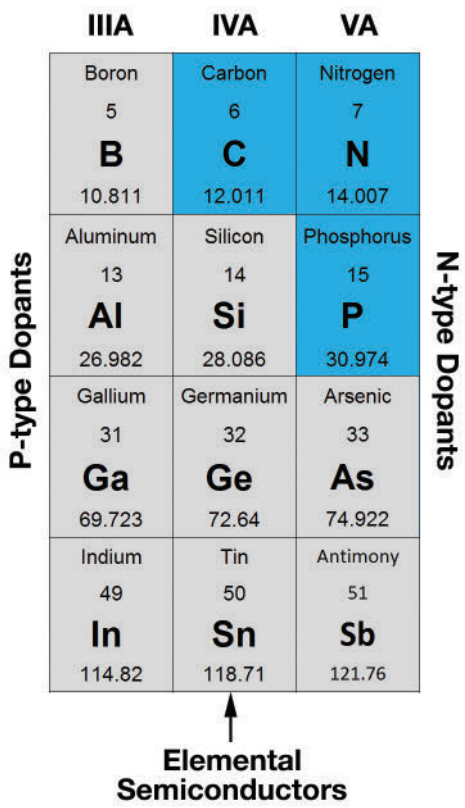


Figure 5. Elemental semiconductors and the most common n- and p-type dopants used.

those elements on either side of Column IVA in the Periodic Table. A dopant atom is similar in size to the intrinsic semiconductor atom in the same row but has either one fewer or one greater electron in its valence shell. The dopant atoms can easily replace the semiconductor atom in its crystal lattice. Figure 6 shows the insertion of either a Group VA (n-type dopant, phosphorus) or a Group IIIA (p-type dopant, boron) atom into the silicon crystal lattice as well as the effect that the substitution of the silicon atom by a dopant atom has on the band structure of the silicon.

When phosphorus, an n-type (donor) dopant is substituted for silicon, the extra electron in phosphorus's valence shell is at a higher energy level (the donor level in Figure 6(a)) than the electrons in the filled silicon valence band. This dopant electron can easily jump the small energy gap (0.045 eV for P, much less than the 1.1 eV band gap for pure silicon at room temperature; typical donor band gaps range from .039 to .054 eV) to the empty conduction band and become a free carrier. The concentration of free carriers (and therefore the electrical conductivity) in an n-type extrinsic semiconductor is roughly proportional to the dopant concentration in the material. Since the free electron generated from dopant atoms in an n-type semiconductor does not produce a corresponding hole in the filled valence band, the dominant charge carriers in an n-type semiconductor are electrons. In an analogous manner, boron can be easily substituted for silicon in the bulk crystal lattice, as shown in Figure 6(b). Since boron has one less electron than silicon, this creates a hole (an acceptor level) at an energy just greater than the top of the filled silicon

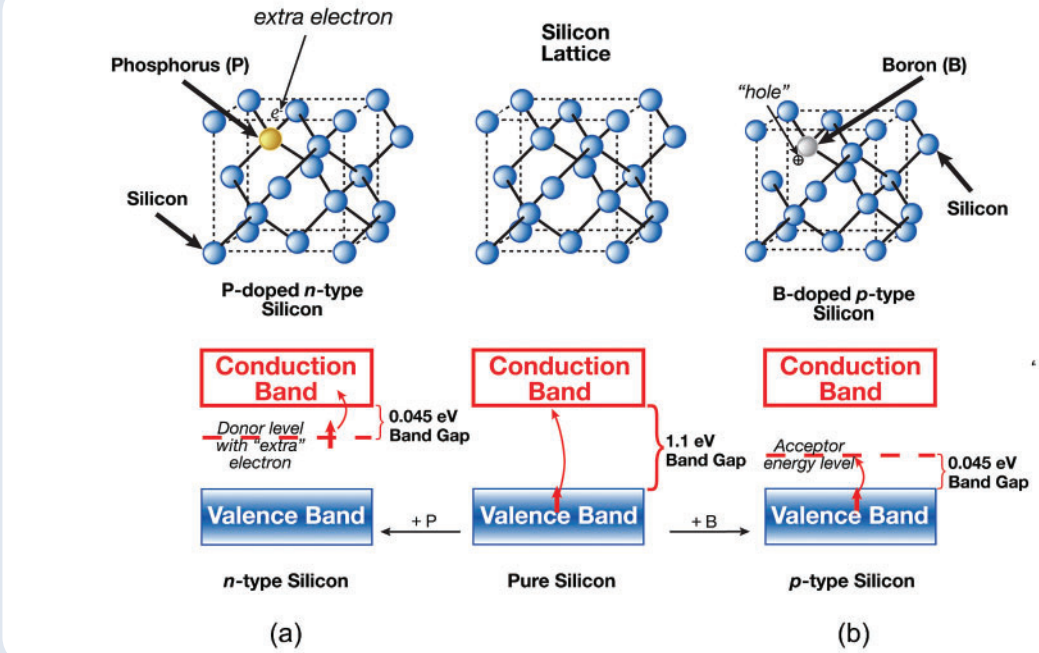


Figure 6. The effect of impurities on the band gap in extrinsic silicon [10].

valence band. Under the influence of an electric field, it takes little energy (0.045 eV for B; typical acceptor band gaps range from 0.045 to 0.160 eV) for an electron in the valence band to be promoted to this acceptor level, leaving behind a hole in the valence band. Other bound electrons in the valence band can then jump into this hole and the subsequent hole movement carries the electrical current. Thus, for p-type extrinsic semiconductors, holes are the majority charge carriers for current flow. Figure 7 shows the relationship between dopant concentration and resistivity (the inverse of conductivity) for n- and p-type dopants in crystalline silicon.

B. Basic Device Structures

1. P-N Junction

The ability to continuously vary a semiconductor's resistivity has only limited value when considered in isolation since a device composed of a single, doped semiconductor is simply a neutral, passive electrical component (i.e., a resistor). Rather, it is the combination of differently doped semiconductors with each other and with other materials that has leveraged the unique physical properties of semiconductors with the properties of more conventional insulators and conducting materials to create the solid-state electronic devices that enable much of modern technology. The most basic of these devices is the P-N junction [13], [14].

When p-type semiconductor and n-type semiconductor materials are placed in physical contact, the area around the contact (known as the junction) behaves differently than either of the two source materials. Figure 8 illustrates the unique electrical field and charge conditions that exist at a P-N junction along with the forces at play on the free carriers. In isolation, an n-type semiconductor material has a high concentration of free electrons while a p-type semiconductor has a high concentration of free holes. When the two materials are in contact, nature

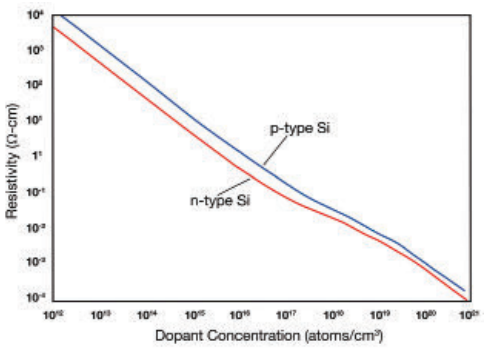


Figure 7. The dependency of the electrical resistivity at 298K on doping concentration.

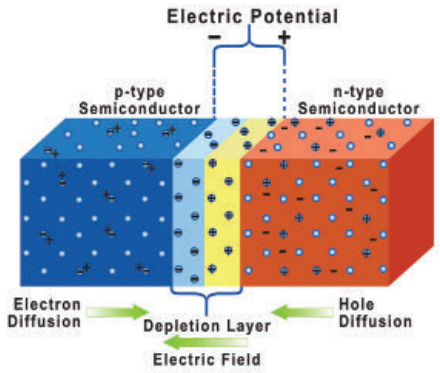


Figure 8. The P-N junction [16].

demands that the concentration of holes and electrons be the same throughout the two materials; this phenomenon is called diffusion [15]. If electrons and holes were uncharged, neutral species, diffusion of holes and electrons would eventually make the local concentration of holes and electrons constant throughout the entire body of the joined p- and n-type materials. However, when electrons migrate from the n-type semiconductor to the p-type semiconductor and holes migrate from the p-type semiconductor to the n-type semiconductor, each type of carrier leaves behind fixed charges; positive in the n-type material (positive Group VA cations in the silicon lattice) and negative in the p-type material (negative Group IIIA anions in the silicon lattice). This sets up an electric field across the junction that eventually prevents further migration. The region over which the electric field extends is known as the depletion region; it is devoid of free carriers and acts as an insulator that prevents further diffusion of electrons and holes. A built-in electrical potential, VD, is created by the electrical field at the P-N junction; in the case of silicon this potential has a value of about 0.7 V.

2. Diodes

Devices made of a simple P-N junction are solid state diodes (commonly referred to as a P-N Junction Diode) [13]. The characteristic response of a solid-state diode to an external electrical potential or to an energy source such as light forms the basis for devices such as bipolar junction transistors, solar cells, LEDs, lasers, and photodiodes. As discussed above and shown in Figure 8, an isolated P-N junction (i.e., not connected within a powered circuit) is in a state of equilibrium, with carrier generation, recombination carrier, diffusion, and drift all balanced in the presence of the electric field across the depletion region. If one were to connect the P and N terminals of the diode shown in Figure 8, zero voltage would be observed across the diode and there would be no current flow. However, when an external potential is applied to the diode terminals, the unique properties of the P-N junction result in one-directional current flow, the unique characteristic of a diode. Figure 9 shows a schematic of a silicon diode under conditions of zero, forward, and reverse bias. In the zero bias state Figure 9(a), there is no current flow, as described above.

In the forward bias state, Figure 9(b), the positive terminal of a power supply (i.e., a battery) is connected to the p-type material of the diode and the negative terminal of the battery to the n-type material. Under these conditions, the electric field due to the external voltage is oriented in the opposite direction of the electric field due to the depletion region (see Figure 8). When it is connected to a power supply in this manner, no current will flow through the diode until the external voltage exceeds the value of the built-in potential in the depletion region, i.e., 0.7 V in the case of a silicon P-N junction. Additionally, under forward bias conditions, the opposing E-field of the external voltage shrinks the depletion region of the P-N junction, making it much thinner; this has the effect of producing a very low resistance path through the P-N junction that will allow very large currents to flow through the diode with relatively small increases in voltage. This effect can be seen in the voltage/current curve for forward bias conditions that is

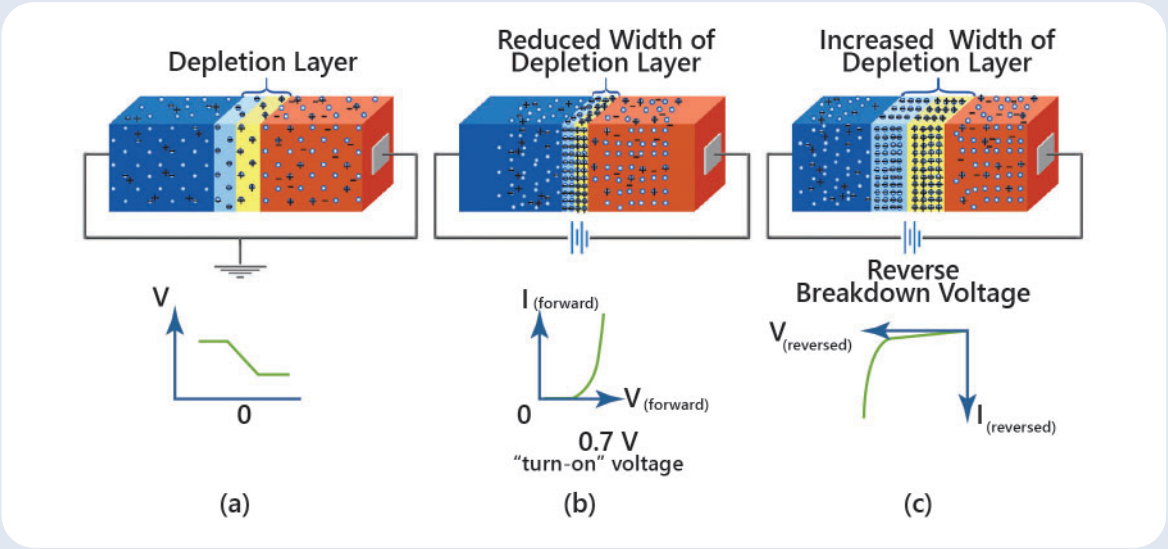


Figure 9. Current flow characteristics of a P-N junction diode [16].

shown in Figure 9. The effective limits to current flow through the P-N junction diode are so high that resistors are normally required in series with the diode to limit the current to values low enough to prevent thermal damage to the diode.

In the reverse bias state Figure 9(c), the positive terminal of a power supply is connected to the n-type material of the P-N junction diode while the negative terminal is connected to the p-type material. Under reverse bias, the positive voltage applied to the n-type material attracts the free electrons towards the positive terminal and the negative voltage applied to the p-type material attracts holes. This has the effect of widening the depletion region, producing a high resistance to current flow in the diode. The resistance produced in this manner is high enough that essentially no current can flow through the device (except for a small leakage current — see the curve in Figure 9(c)). This state exists under reverse bias conditions until the bias voltage becomes high enough to cause the diode's P-N junction to overheat and fail, producing a short in the electrical circuit.

3. Bipolar Junction Transistor

The P-N junction diode is the simplest electronic device that can be built using semiconducting materials (except for a resistor). If we take two P-N junction diodes and physically join them in a back-to-back manner as shown in Figure 10, we can produce the most fundamental of microelectronic devices — a transistor. Transistors such as the one shown in Figure 10 are called bipolar junction transistors [13], [17]. Bipolar junction transistors (BJTs) were invented in late 1947 by William Shockley, Walter Brattain, and John Bardeen at AT&T Bell Laboratories. They constituted what is arguably the first step in the microelectronics revolution and Shockley, Brattain and Bardeen were awarded the Nobel Prize in Physics in 1956 for this invention. The BJT is so named because its operation involves both electrons and holes. It can be configured as either a PNP transistor or an NPN transistor which can be constructed as either a three-layer sandwich as shown in Figure 10(a) or as a planar device on a semiconductor wafer as shown in Figure 10(b).

The BJT is fabricated with an emitter and collector of similar semiconductor type (i.e., either p-type or n-type) at either end with a very thin base of opposite polarity in between the emitter and collector. In BJTs, a small potential (for silicon devices, 0.7 V, enough to overcome the P-N junction potential) is applied to the base. This overcomes the

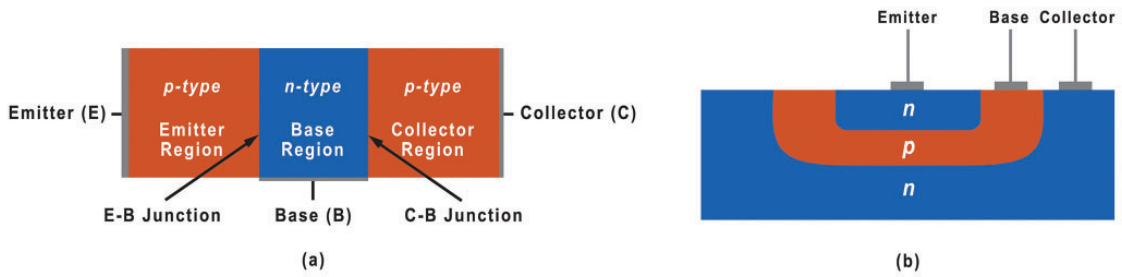


Figure 10. The bipolar junction transistor.

junction potential between the emitter and the base, producing a small current flow from the emitter to the base. This current flow, in turn, regulates the much larger current flow from the emitter to the collector. For those interested in a more in-depth understanding of how the BJT operates, reference [17] provides an excellent video description. BJTs can be described as "current regulating devices that control the amount of current flowing through them in proportion to the amount of biasing voltage applied to their base terminal, acting like a current-controlled switch" [18].

4. MOSFET

Another common form of transistor is the Metal Oxide Semiconductor Field Effect Transistor (MOSFET) [13], [19]. MOSFETs are planar surface devices that are the most used variant of Field Effect Transistors (FETs); the reader may also encounter Junction Gate Field Effect Transistors (JFETs) and Insulated Gate Field Effect Transistors (IGFETs). MOSFETs are a subset of IGFETs. FETs can replace BJTs in most electronic circuits and have advantages for use in microelectronics since they consume and dissipate less power and they can be made much smaller than equivalent BJTs. Indeed, MOSFET devices constitute the ubiquitous "bit" switch that is set to 0 ("Off" state) or 1 ("On" state) in microelectronic logic devices (i.e., computers). FETs differ from BJTs in the way that current flow through the device is controlled; the primary current flow through FETs (Field Effect Transistors) is controlled by a small voltage applied to one of the terminals rather than by a control current flow through any part of the device. Since MOSFETs are the most widely used FET, we will use these devices to describe the components and operation of this class of transistor. Figure 11 shows the elements of a typical MOSFET where the transistor is built into the surface of a silicon bulk substrate. MOSFETs can be built as either NMOS or PMOS transistors, depending on the polarities of the bulk, source and drain regions. Most carriers in NMOS devices are electrons while those in PMOS devices are holes. An NMOS device is built on a p-doped silicon substrate that has regions of n-type material which are created using ion implantation, Figure 11(a). These n-type regions are called the source and the drain. This situation is reversed in a PMOS device, Figure 11(b). The bulk material between the source and drain in a MOSFET is called the channel. A very thin insulating oxide layer covers the channel region; it is commonly referred to as a gate oxide. Finally, a conducting gate material, either a metal or highly doped polysilicon, is deposited on top of the gate oxide, creating the three-terminal device structure shown in Figure 11.

The operational characteristics of an nMOSFET transistor are shown in Figure 12; basic operating principles of a MOSFET device can be explained within the context of an NMOS device as shown in Figure 13. Figure 13(a) shows the device in the "Off" state with the gate, source, and drain voltages at zero and the bulk substrate connected to ground. Two P-N junctions exist between the n-type source/drain regions and the bulk p-type substrate. In operation, the potential between the drain and source, VDS, and that between the gate and source (VGS), are always positive. When a small voltage, VGS, is applied to the gate, the charge carrying holes in the p-type substrate are repelled away from

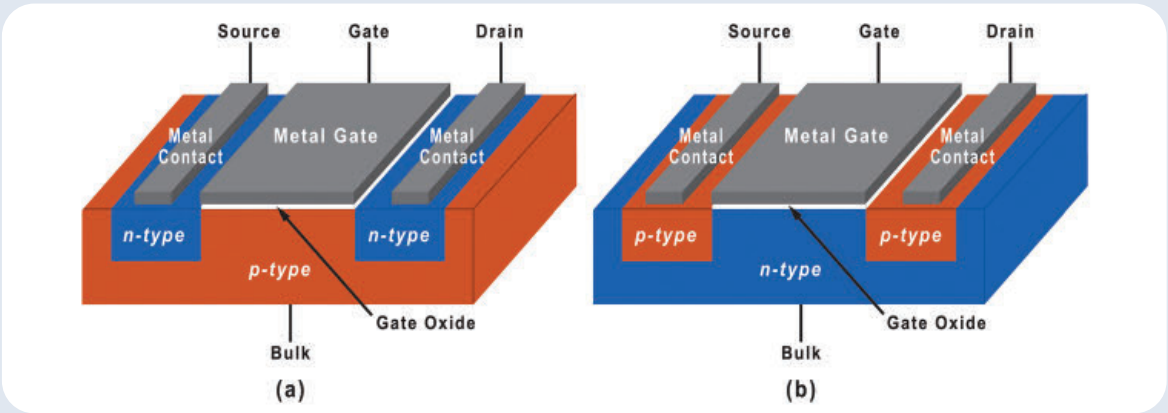


Figure 11. NMOS (a) and PMOS (b) MOSFETs.

the substrate surface. When VGS reaches a threshold value, VTH, the minimum gate to source voltage needed to turn the device on; this is less than the 0.7 V required in BJTs, typically 0.2-0.25 V in modern logic processors), the region under the gate becomes completely depleted of charge, producing a region in the substrate called the "depletion zone." Further increases in VGS attract electrons from the electron-rich source (VGS) and drain (VGD) regions into the region under the gate, producing an n+ region known as the "inversion layer," shown in Figure 13(b). This inversion layer is a conducting channel that connects the two n-type regions at the source and drain; it will allow electrons to flow from the source to the drain when there is a positive voltage, VDS, between the source and drain. To assure that the induced inversion channel extends all the way from source to drain, the MOSFET gate structure slightly overlaps the edges of source and drain (the latter is achieved by a method known as a self-aligned process [13], [20]).

When a drain-source bias, VDS, is applied to an NMOS device in the above threshold conducting state, electrons move in the channel inversion layer from source to drain. At relatively small values of VDS, the I/V characteristics of the device are linear with ID (drain current) increasing with increasing VD (drain voltage), as shown in Figure 13(b). Any change in the gate-source voltage, VGS, will alter the electron density in the inversion layer and, in this way, changes to VGS can also control the device current. For this reason, characteristic I-V curves for NMOS devices typically depict a family of curves at different VGS, as shown in Figure 12. When the drain voltage is increased to a value known as the saturation voltage, VSAT, the charge and current flow characteristics in an NMOS device evolve, as depicted in Figure 13(c). The inversion layer under the gate becomes wedge shaped, wider (or deeper) near the source and essentially disappears (zero thickness) at the drain. This

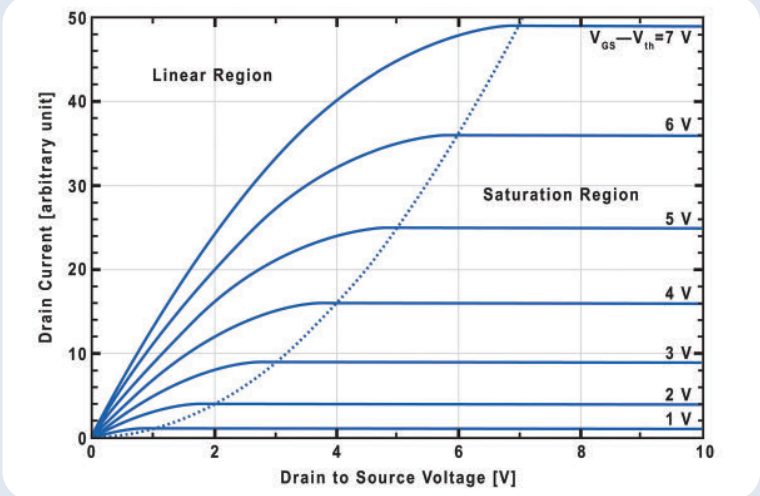


Figure 12. nMOSFET I-V characteristics [21].

phenomenon is known as "pinch-off" [22] and the point where the inversion layer thickness is reduced to zero is called the "pinch-off point." Pinch-off occurs because, at VSAT, the effective potential between the gate and substrate at the source end of the channel (Veff = VGS) is greater than the potential between the gate and the substrate at the drain end of the channel, which is just the potential needed to form the inversion layer called the threshold voltage (Veff = VGS-VSAT = VTH). Any higher voltage on the drain will cause the gate-to-substrate voltage to be reduced below the threshold voltage and the inversion layer will not be formed, creating the pinch-off point where there are no longer any mobile electron carriers in the channel. When the voltage applied to the drain is increased beyond VSAT, the pinch-off point moves further towards the source, reducing the effective channel length, Leff, as shown in Figure 13(d). Under these conditions, the area between the pinch-off point and the drain is fully depleted with no inversion layer. Since this region has no positive free carriers, there is no possibility for electron-hole recombination if an electron enters the region from the electron-rich source and, if there is an electric field across the depletion zone, the electron can freely transit to the drain. (Interested readers can find a more detailed discussion of this phenomenon in reference [22].) As can be seen from Figure 12, the current through the device becomes controlled solely by the gate voltage under drain saturation conditions.

5. FinFETs

As MOSFET devices have continued to shrink in size, certain limitations have begun to impact their performance. Specifically, at nanometer dimensions, subthreshold current leakage becomes a significant factor (Ioff becomes unacceptably high) and the potential on the drain begins to dominate the electrostatics in the channel region of

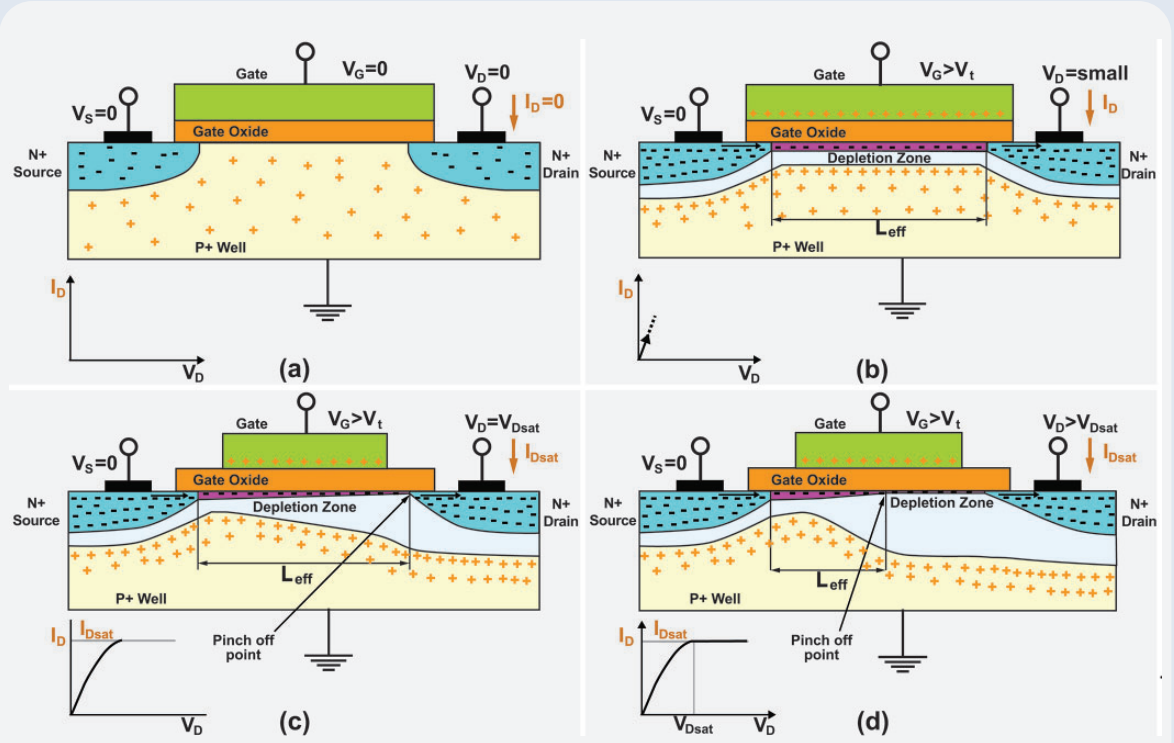


Figure 13. Operating characteristics of an NMOS field effect transistor.

the device. This causes a loss of gate control over the device current and an increase in Ioff; these characteristics are known as short channel effects. Increases in Ioff have obvious consequences for device performance. The problem especially impacts the devices used in mobile applications where it can lead to an unacceptable drain on battery power. Over the years, this problem has been overcome using ever thinner gate oxides and the deployment of high-k dielectric materials that increase the gate-channel capacitance. These approaches have been effective up to about a decade or so ago when the continued thinning of gate oxides produced irremediable problems with gate leakage and gate-induced drain leakage in nanometer-scale devices. These problems have been resolved using an approach known as multi-gate FETs (MGFETs), a three-dimensional departure from traditional planar MOSFET designs.

Currently, the dominant non-planar transistor device geometry is the Fin Field Effect Transistor or FinFET [23]. Figure 14 shows a comparison between the structure of a conventional planar MOSFET and that of a FinFET. It can be seen that all of the familiar components of a MOS device are present in the FinFET (drain, source, gate, gate dielectric) and that the device more effectively separates the source and drain from the substrate silicon (FinFETs can be constructed on both bulk silicon and on buried oxide layers) and has a multiple gate configuration. The source and drain regions are now part of a vertical fin structure, with the gate dielectric and gate electrode wrapped around it to produce multiple gates, one on either side of the fin as well as one on top of the device. FinFETs can also be extended to create a gate all around the device.

FinFET designs produce much better control of the gate current by using more than one gate to control a single channel. They have much lower Ioff leakage characteristics and can conduct significantly more current in the On state than comparable MOSFET designs.

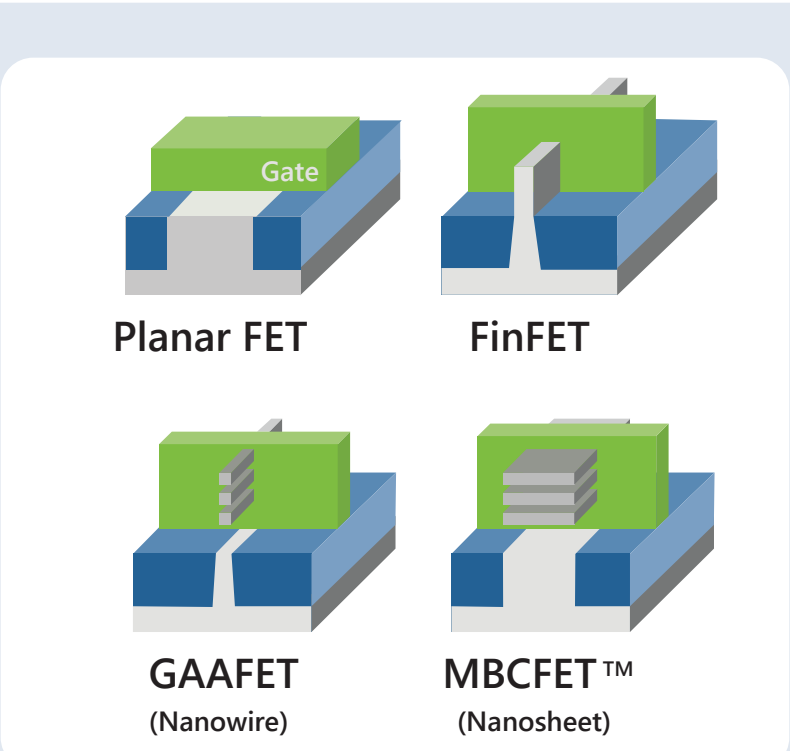


Figure 14. A comparison between the structures of a planar MOSFET transistor and newer structures that reduce leakage current from drain to source.

The next generation of MOSFET structures involve wrapping the gate completely around the channel to further reduce the leakage current and enable further reduction in size – these structures are known as Gate All Around (GAAFET) and Multi-Bridge Channel FET (MBCFET).

6. Other Common Semiconductor Devices

In addition to the semiconductor devices that have been described in detail above, there are many other device types that the reader may encounter in interactions with semiconductor manufacturers. Below are some representative devices with graphic illustrations of their structures.

Flash transistors [24] are used in memory devices. Their design is similar to that of MOSFETs with the main design difference being that they employ two gates rather than one in a planar MOSFET design (Figure 15). When a potential is applied to the control gate of the flash transistor, current flows between the source and drain. It differs from a conventional MOS device in that some electrons tunnel through the gate oxide to produce a permanent charge on the floating gate. The presence of this charge constitutes a logic "1" in the memory device. The charge on the floating gate can be removed by placing a negative potential on the control gate. When the floating gate has no charge, it constitutes a logic "0" in the memory device. Flash memory transistors are thus a fundamental component of most non-volatile memory devices, which are devices that maintain their charge when power is turned off. In addition to the Floating Gate (FG) method, the charge can be stored in an insulating layer – this is known as Charge Trap Flash (CTF).

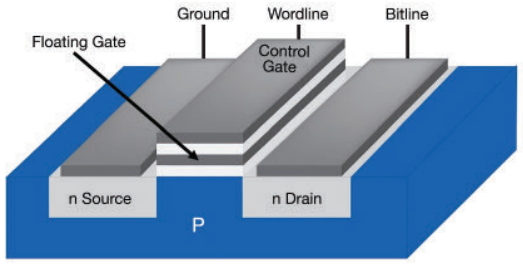


Figure 15. Flash transistor structure.

Modern Flash memory employs innovative schemes to pack billions of bits into a small package. By carefully controlling the amount of charge stored in a cell, one can represent multiple bit patterns (see [25], https://en.wikipedia.org/wiki/Multi-level_cell). As an example, the Multi-Level Cell (MLC) approach stores two bits of information by quantizing the stored charge into four levels:

- 00 = no charge
- 01 = x units of charge
- 10 = 2x units of charge
- 11 = 3x units of charge

This scheme can be further extended to store 3, 4, and 5 bits per cell, but the quantization scheme becomes more difficult to differentiate the charge levels in the presence of noise and electron loss.

Another method for increasing storage capacity is to arrange the memory cell array into a 3-dimensional structure known as 3D NAND Flash or VNAND (Vertical NAND). An example 3D NAND Flash cell is shown in Figure 16. The electrons are stored in the silicon nitride layer and programming is performed using a metal gate to inject the charge across the Tunnel Oxide. These cells are created in a vertical array by etching high aspect ratio structures called Channel Holes into an alternating stack of silicon dioxide and silicon nitride (see Figure 17).

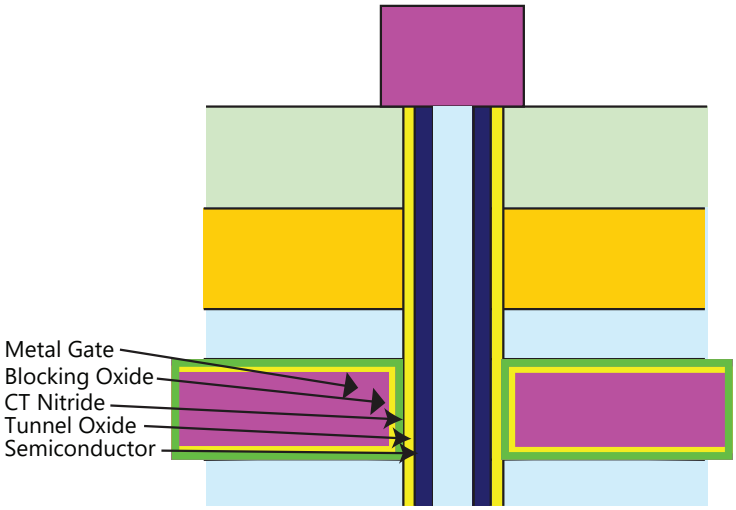


Figure 16. Charge Trap 3D NAND Flash cell.

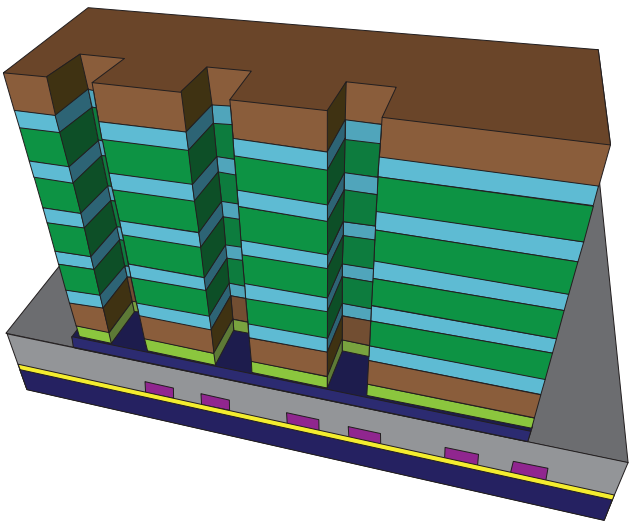


Figure 17. Cross-section of channel hole structures in a 3D-NAND memory.

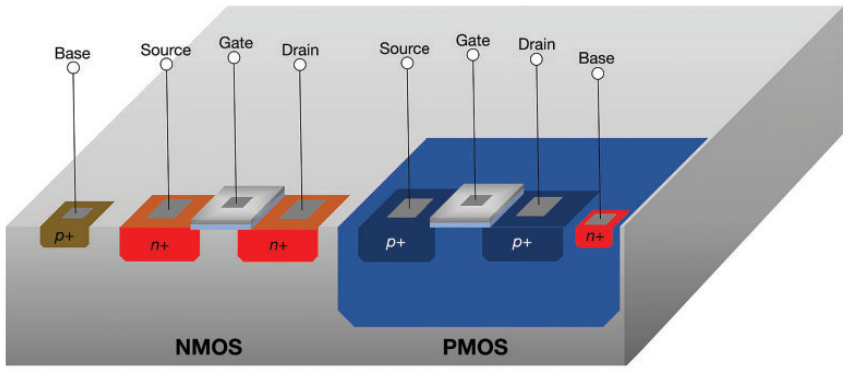


Figure 18. CMOS device structure.

CMOS

The acronym CMOS describes a Complementary Metal Oxide Semiconductor device [13], [25]. These composite devices combine NMOS and PMOS switches to achieve unique characteristics that allow a full suite of logic devices to be built [13], [26]. The device structure for a generic CMOS device is shown in Figure 18. CMOS devices have high immunity to noise and low power consumption, thus generating less waste heat. It is the most used technology for very large scale integrated (VLSI) circuits.

Rectifier

Semiconductor rectifiers are p-n junction diodes specifically designed to rectify an alternating current i.e., to give a low resistance to current flow in one direction and a very high resistance in the other direction.

Zener Diode

A Zener diode is a voltage regulator that is a p-n junction diode having a precisely tailored impurity distribution that produces a well-defined breakdown voltage. It can be operated in the reverse direction to serve as a constant voltage source, as a reference voltage for a regulated power supply, and as a protective device against voltage and current transients.

Varactor Diode

A varactor (variable reactor) is a device whose reactance, a measure of opposition to a change in current or voltage, can be varied in a controlled manner with a bias voltage. It is a p-n junction with a special impurity profile, and its capacitance variation is very sensitive to reverse-biased voltage. Varactors are widely used in parametric amplification, harmonic generation, mixing, detection and voltage-variable tuning applications.

Tunnel Diode

A tunnel diode is a single p-n junction in which both the p and n sides are heavily doped. The depletion layer is extremely narrow and under forward bias electrons, can tunnel or pass directly through the junction, producing a negative resistance effect (i.e., the current decreases with increasing voltage). The tunnel diode is used in special low-power microwave applications, such as local oscillators and frequency-locking circuits.

Schottky Diode

A Schottky diode has a metal-semiconductor contact (i.e., an aluminum layer in intimate contact with an n-type silicon substrate). It is electrically similar to a p-n junction and is used extensively for high-frequency, low-noise mixer and switching circuits. This metal-semiconductor junction has a lower forward bias voltage than a semiconductor/semiconductor junction and leads to higher switching speeds.

References: Semiconductor Physics and Basic Device Structures

[1] E. M. Purcell and D. J. Morin, Electricity and Magnetism 3rd Edition, Cambridge University Press, 2013.

[2] D. Owens, "Physics 12.4.1a - Electric Potential and Potential Difference," [Online]. Available: https://www.youtube.com/watch?v=wT9AsY79f1k.

[3] R. Chang, Essential Chemistry: a core text for general chemistry - 2nd Edition, McGraw-Hill Higher Education, 2000.

[4] "General Chemistry," [Online]. Available: https://www.youtube.com/watch?v=FSyAehMdpyI&list=PL8dPuuaLjXtPHzzYuWy6fYEaX9mQQ8oGr.

[5] "Energy and Chemistry - Change, Chapter 6," [Online]. Available: https://www.youtube.com/watch?v=GqtUWyDR1fg.

[6] "Electronic structure of atoms," [Online]. Available: https://www.youtube.com/watch?v=rcKilE9CdaA.

[7] "Chemical Bonding," [Online]. Available: https://www.youtube.com/watch?v=a8LF7JEb0IA.

[8] H. T. Grahn, Introduction to Semiconductor Physics, Singapore: World Scientific Co. Pte. Ltd., 1999.

[9] "Intrinsic and Extrinsic Semiconductors," [Online]. Available: https://www.youtube.com/watch?v=o-PPbmMm0eA.

[10] J. Walker, "Part 3," in Haliday & Resnick Fundamentals of Physics Extended 9th Edition , John Wiley & Sons Inc.

[11] "The Electric Field Around a Charge," [Online]. Available: https://www.youtube.com/watch?v=vaDT4GwAZ2I&list=PL6io3nDcMmqAwIL5TeBx_cvBGpiH1XRj.

[12] "Intrinsic Silicon and Extrinsic Silicon," [Online]. Available: https://www.youtube.com/watch?v=uFaZiVy3bZo&index=59&list=PL8TDSRZRF0t9wYCFIEToRXkBqbIXqzC3r.

[13] B. Anderson and R. L. Anderson, Fundamentals of Semiconductor Devices, McGraw-Hill Education, 2005.

[14] "PN Junctionh," [Online]. Available: https://www.youtube.com/watch?v=ar7xDMR4P_U.

[15] "Passing Gases: Effusion, Diffusion and the Velocity of a Gas," [Online]. Available: https://www.youtube.com/watch?v=TLRZAFU_9Kg.

[16] "PN Junction Theory," [Online]. Available: http://www.electronics-tutorials.ws/diode/diode_2.html.

[17] "How a Transistor Works," [Online]. Available: https://www.youtube.com/watch?v=DXvAlwMAxiA.

[18] "Bipolar Transistors," [Online]. Available: http://www.electronics-tutorials.ws/transistor/tran_1.html.

[19] "How MOSFETs and Field-Effect Transistors Work," [Online]. Available: https://www.youtube.com/watch?v=tz62t-q_KEc&index=2&list=PLUmzZYg1_YlAGG2tw28fW7wrPFYDlnn6n.

[20] "Self-Aligned Gate," [Online]. Available: https://en.wikipedia.org/wiki/Self-alig

[21] "Current-Voltage Characteristics," Wikipedia, [Online]. Available: https://en.wikiped
Current%E2%80%93voltage_characteristic.

[22] "Why MOSFET Pinchoff Occurs," [Online]. Available: http://electronics.stackexchange.com/questions/77198/why-mosfet-pinchoff-occurs.

[23] D. Bhattacharya and N. K. Jha, "FinFETs: From Devices to Architecture," in Advances in Electronics, 2014.

[24] P. Pavani, R. Bez, P. Olivio and E. Zanoni, "Flash memory cells-an overview," Proceedings of the IEEE, vol. 85, no. 8, pp. 1248-1271.

[25] S. S. Kim, S. K. Yong, W. Kim, S. Kang, H. W. Park, K. J. Yoon, D. S. Sheen, S. Lee and C. S. Hwang, "Review of Semiconductor Flash Memory Devices for Material and Process Issues," Advanced Materials, p. 2200659, 2022.

[26] "CMOS," Wikipedia, [Online]. Available: https://en.wikipedia.org/wiki/CMOS.

[27] "Building logic gates from MOSFET transistors," [Online]. Available: https://www.youtube.com/watch?v=1rZyGL1K5QI&t=31s.