Unit 3　Integrated Circuits

3.1　Introduction

Integrated circuit (IC) (see Figure 3.1), also called microelectronic circuit, microchip, or chip, is an assembly of electronic components, fabricated as a single unit, in which miniaturized active devices (e.g., transistors and diodes) and passive devices (e.g., capacitors and resistors) and their interconnections are built up on a thin substrate of semiconductor material (typically silicon). The resulting circuit is thus a small monolithic "chip," which may be as small as a few square centimeters or only a few square millimeters. The individual circuit components are generally microscopic in size.

Integrated circuits have their origin in the invention of the transistor in 1947 by William B. Shockley and his team at the American Telephone and Telegraph Company's Bell Laboratories. Shockley's team (including John Bardeen and Walter H. Brattain) found that, under the right circumstances, electrons would form a barrier at the surface of a certain crystal, and they learned to control the flow of electricity through the crystal by manipulating this barrier. Controlling electron flow through a crystal allowed the team to create a device that could perform certain electrical operations, such as signal amplification, that were previously done by vacuum tubes. They named this device a transistor, from a combination of the words transfer and resistor. The study of methods of creating electronic devices using solid materials became known as solid-state electronics. **Solid-state devices proved to be much sturdier, easier to work with, more reliable, much smaller, and less expensive than vacuum tubes.** Using the same principles and materials, engineers soon learned to create other electronic components, such as resistors (see Figure 3.2) and capacitors. Now that electrical devices could be made so small, the largest part of a circuit was the awkward wiring between the devices.

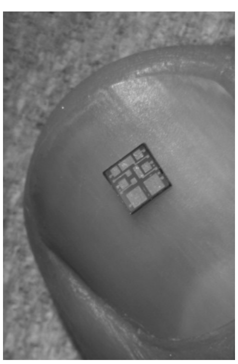


Figure 3.1　IC

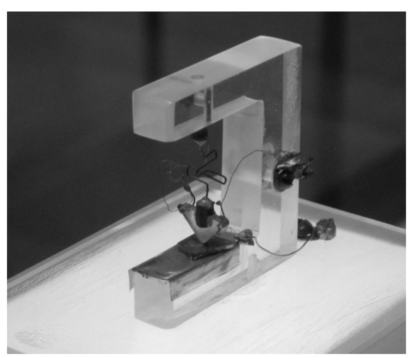


Figure 3.2　The first transistor

In 1958, Jack Kilby of Texas Instruments, Inc., and Robert Noyce of Fairchild Semiconductor Corporation independently thought of a way to reduce circuit size further. They laid very thin paths of metal (usually aluminum or copper) directly on the same piece of material as their devices. These small paths acted as wires. With this technique, an entire circuit could be "integrated" on a single piece of solid material and an integrated circuit (IC) was thus created. An IC can contain hundreds of thousands of individual transistors on a single piece of material the size of a pea. Working with that many vacuum tubes would have been unrealistically awkward and expensive. The invention of the integrated circuit made technologies of the information age feasible. ICs are now used extensively in all walks of life, from cars to toasters to amusement park rides.

There are several basic types of IC:

1. Analog Versus Digital Circuits

Analog, or linear, circuits typically use only a few components and are thus some of the simplest types of ICs. Generally, analog circuits are connected to devices that collect signals from the environment or send signals back to the environment. For example, a microphone converts fluctuating vocal sounds into an electrical signal of varying voltage. An analog circuit then modifies the signal in some useful way—such as amplifying it or filtering it of undesirable noise. Such a signal might then be fed back to a loudspeaker, which would reproduce the tones originally picked up by the microphone. Another typical use for an analog circuit is to control some device in response to continual changes in the environment. **For example, a temperature sensor sends a varying signal to a thermostat, which can be programmed to turn an air conditioner, heater, or oven on and off once the signal has reached a certain value.**

A digital circuit, on the other hand, is designed to accept only voltages of specific given values. A circuit that uses only two states is known as a binary circuit. Circuit design with binary quantities, "on" and "off" representing 1 and 0 (i.e., true and false), uses the logic of Boolean algebra. (Arithmetic is also performed in the binary number system employing Boolean algebra.) Figure 3.3 shows different logic circuits. These basic elements are combined in the design of ICs for digital computers and associated devices to perform the desired functions.

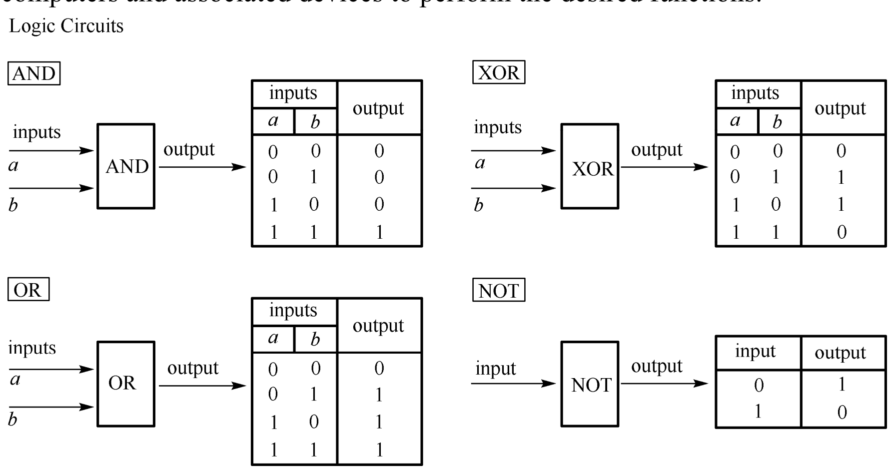


Figure 3.3 Different logic circuits

2. Microprocessor Circuits

Microprocessors are the most complicated ICs. They are composed of billions of transistors that have been configured as thousands of individual digital circuits, each of which performs some specific logic function. A microprocessor is built entirely of these logic circuits synchronized to each other. A microprocessor typically contains the central processing unit (CPU) of a computer.

Just like a marching band, the circuits perform their logic functions only on direction by the bandmaster. The bandmaster in a microprocessor, so to speak, is called the clock. The clock is a signal that quickly alternates between two logic states. Every time the clock changes state, every logic circuit in the microprocessor does something. Calculations can be made very quickly, depending on the speed (clock frequency) of the microprocessor.

Microprocessors contain some circuits, known as registers, which store information. Registers are predetermined memory locations. Each processor has many different types of registers. Permanent registers are used to store the preprogrammed instructions required for various operations (such as addition and multiplication). Temporary registers store numbers that are to be operated on and also the results. **Other examples of registers include the program counter (also called the instruction pointer), which contains the address in memory of the next instruction; the stack pointer (also called the stack register), which contains the address of the last instruction put into an area of memory called the stack; and the memory address register, which contains the address of where the data to be worked on is located or where the data that has been processed will be stored.**

Microprocessors can perform billions of operations per second on data. In addition to computers, microprocessors are common in video game systems, televisions, cameras, and automobiles.

3. Memory Circuits

Microprocessors typically have to store more data than can be held in a few registers. This additional information is relocated to special memory circuits. Memory is composed of dense arrays of parallel circuits that use their voltage states to store information. Memory also stores the temporary sequence of instructions, or program, for the microprocessor.

Manufacturers continually strive to reduce the size of memory circuits—to increase capability without increasing space. In addition, smaller components typically use less power, operate more efficiently, and cost less to manufacture.

4. Digital Signal Processors

A signal is an analog waveform—anything in the environment that can be captured electronically. A digital signal is an analog waveform that has been converted into a series of binary numbers for quick manipulation. As the name implies, a digital signal processor (DSP) processes signals digitally, as patterns of 1 s and 0 s. For instance, using an analog-to-digital converter, commonly called an A-to-D or A/D converter, a recording of someone’s voice can be converted into digital 1 s and 0 s. The digital representation of the voice can then be modified by a DSP using complex mathematical formulas. For example, the DSP algorithm in the circuit may be configured to recognize gaps between spoken words as background noise and digitally remove ambient noise from the waveform. Finally, the processed signal can be converted back (by a D/A converter) into an analog signal for listening. **Digital processing can filter out background noise so fast that there is no discernible delay and the signal appears to be heard in "real time".** For instance, such processing enables "live" television broadcasts to focus on a quarterback's signals in an American gridiron football game.

DSPs are also used to produce digital effects on live television. For example, the yellow marker lines displayed during the football game are not really on the field; a DSP adds the lines after the cameras shoot the picture but before it is broadcast. Similarly, some of the advertisements seen on stadium fences and billboards during televised sporting events are not really there.

5. Application Specific ICs

An application specific IC (ASIC) can be either a digital or an analog circuit. As its name implies, an ASIC is not reconfigurable; it performs only one specific function. For example, a speed controller IC for a remote control car is hard-wired to do one job and could never become a microprocessor. An ASIC does not contain any ability to follow alternate instructions.

6. Radio-Frequency ICs

Radio-frequency ICs (RFICs) are widely used in mobile phones and wireless devices. RFICs are analog circuits that usually run in the frequency range of 3 kHz to 2.4 GHz (3,000 hertz to 2.4 billion hertz), and circuits operating at about 1 THz (1 trillion hertz) are under development. They are usually thought of as ASICs even though some may be configurable for several similar applications.

Most semiconductor circuits that operate above 500 MHz (500 million hertz) cause the electronic components and their connecting paths to interfere with each other in unusual ways. Engineers must use special design techniques to deal with the physics of high-frequency microelectronic interactions.

A special type of RFIC is known as a monolithic microwave IC (MMIC; also called microwave monolithic IC). These circuits usually run in the 2- to 100- GHz range, or microwave frequencies, and are used in radar systems, in satellite communications, and as power amplifiers for cellular telephones. **Just as sound travels faster through water than through air, electron velocity is different through each type of semiconductor material.** Silicon offers too much resistance for microwave-frequency circuits, and so the compound gallium arsenide (GaAs) is often used for MMICs. Unfortunately, GaAs is mechanically much less sound than silicon. It breaks easily, so GaAs wafers are usually much more expensive to build than silicon wafers.

Words and Expressions

microelectronic [ˌmaɪkroʊɪˌlek'trɑːnɪk] *adj*. 微电子（学）的；超小型电子的
miniaturize ['mɪnətʃəraɪz] *vt*. 使微型化；使成为缩影
substrate ['sʌbstreɪt] *n*. 基底；底物；底层；基层；衬底
semiconductor ['semikəndʌktər] *n*. 半导体；半导体装置

microscopic [ˌmaɪkrə'skɑːpɪk]	*adj*. 极小的；微小的；需用显微镜观察的
electron [ɪ'lektrɑːn]	*n*. 电子
barrier ['bæriər]	*n*. 垫垒；障碍；屏障；阻力
crystal ['krɪstl]	*n*. 结晶；晶体；晶振
vacuum tube	真空管
transistor [træn'zɪstər]	*n*. 晶体管
solid-state electronics	固态电子学；固体电子学
thermostat ['θɜːrməstæt]	*n*. 恒温器；温度自动调节器
binary ['baɪnəri]	*adj*. 二进制的（用 0 和 1 记数）；二元的 *n*. 二进制数
Boolean algebra	布尔代数
synchronize ['sɪŋkrənaɪz]	*v*. 使同步；（使）同步，在时间上一致，同速进行
program counter	程序计数器
stack pointer	堆栈指针；堆栈指示器；堆栈指针寄存器
digital signal processor(DSP)	数字信号处理器
capture ['kæptʃər]	*vt*. 俘获；捕获；把……输入计算机
ambient noise	环境噪声；氛围噪声；背景噪声
discernible [dɪ'sɜːrnəbl]	*adj*. 可辨的；看得清的；辨别得出的
digital effect	数字特效
reconfigurable [ˌrikən'fɪgjərəbl]	*adj*. 可重构的；可重配置的
remote control	遥控器；遥控
radio frequency IC(RFIC)	射频集成电路；射频芯片
microwave monolithic IC(MMIC)	微波单片集成电路
gallium arsenide	砷化镓（GaAs）
silicon wafer	硅晶圆；硅晶片；硅片

Notes

1. Integrated circuit (IC), also called microelectronic circuit, microchip, or chip, is an assembly of electronic components, fabricated as a single unit, in which miniaturized active devices (e.g., transistors and diodes) and passive devices (e.g., capacitors and resistors) and their interconnections are built up on a thin substrate of semiconductor material (typically silicon).

集成电路（IC），也被称为微电子电路、微芯片或芯片，集成了多个电子元件，制造为单个单元，其中微型有源器件（如晶体管和二极管）和无源器件（如电容器和电阻器）及其互连都建立在半导体材料（通常为硅）的薄衬底上。

2. Solid-state devices proved to be much sturdier, easier to work with, more reliable, much smaller, and less expensive than vacuum tubes.

事实证明，固态器件比真空管更坚固、更易用、更可靠、体积更小且成本更低。

3. For example, a temperature sensor sends a varying signal to a thermostat, which can be programmed to turn an air conditioner, heater, or oven on and off once the signal has reached a certain value.

例如，温度传感器向恒温器发送一个随温度变化的信号，一旦此信号达到某个特定值，预编程的恒温器就会自动打开或关闭空调、加热器或烤箱。

4. Other examples of registers include the program counter (also called the instruction pointer), which contains the address in memory of the next instruction; the stack pointer (also called the stack register), which contains the address of the last instruction put into an area of memory called the stack; and the memory address register, which contains the address of where the data to be worked on is located or where the data that has been processed will be stored.

寄存器的其他例子包括：程序计数器（也称指令指针），存储下一条指令在内存中的地址；堆栈指针（也称堆栈寄存器），记录放入堆栈的最后一条指令的地址；内存地址寄存器，其中包含要处理的数据所在的地址或已处理的数据将被存储的地址。

5. Manufacturers continually strive to reduce the size of memory circuits—to increase capability without increasing space. In addition, smaller components typically use less power, operate more efficiently, and cost less to manufacture.

制造商不断努力缩小内存电路的尺寸，以在不增加物理空间的情况下增加存储容量。此外，较小的组件通常使用较少的功率，运行效率更高，制造成本更低。

6. Digital processing can filter out background noise so fast that there is no discernible delay and the signal appears to be heard in “real time”.

数字处理可以快速滤除背景噪声，因此没有明显的延迟，信号似乎是“实时”听到的。

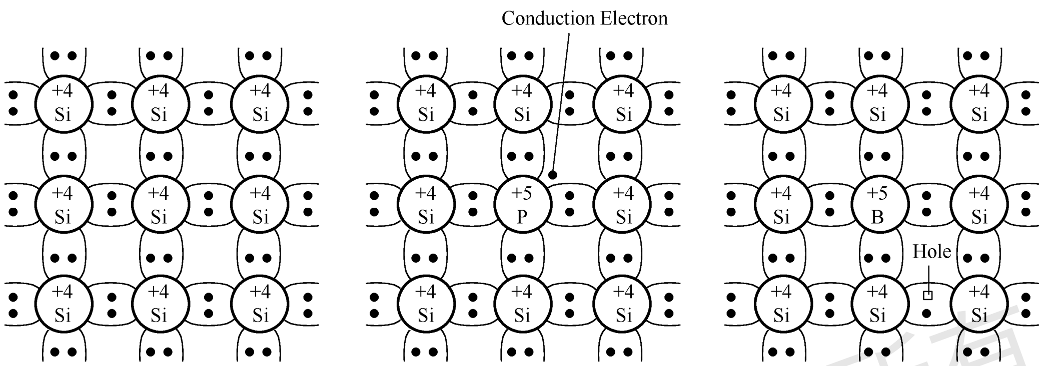
7. Just as sound travels faster through water than through air, electron velocity is different through each type of semiconductor material.

正如声音在水中的传播速度比在空气中的传播速度快一样，电子在不同半导体材料中的传播速度也不同。

3.2 PN Junction and Diode

Any material can be classified as one of three types: conductor, insulator, or semiconductor. A conductor (such as copper or salt water) can easily conduct electricity because it has an abundance of free electrons. An insulator (such as ceramic or dry air) conducts electricity very poorly because it has few or no free electrons. A semiconductor (such as silicon or gallium arsenide) is somewhere between a conductor and an insulator. It is capable of conducting some electricity, but not much.

Most ICs are made of silicon, which is abundant in ordinary beach sand. Pure crystalline silicon, as with other semiconducting materials, has a very high resistance to electric current at normal room temperature. However, with the addition of certain impurities, known as dopants, the silicon can be made to conduct usable currents. **In particular, the doped silicon can be used as a switch, turning current off and on as desired.** The process of introducing impurities is known as doping or implantation. Depending on a dopant’s atomic structure, the result of implantation will be either an N-type (negative) or a P-type (positive) semiconductor. An N-type semiconductor results from implanting donor atoms that have more electrons in their outer (bonding) shell than silicon. The resulting semiconductor crystal contains excess, or free, electrons that are available for conducting current. A P-type semiconductor results from implanting acceptor atoms that have fewer electrons in their outer shell than silicon. **The resulting crystal contains "holes" in its bonding structure where electrons would normally be located. In essence, such holes can move through the crystal conducting positive charges.** Figure 3.4 shows semiconductor bonds.



(a) Intrinsic with negligible impurities (b) N-type with donor(phosphorus) (c) P-type with acceptor(boron)

Figure 3.4 Semiconductor bonds

By creating a single zone of N material adjacent to a zone of P material, we wind up with the PN junction. The PN junction is arguably the fundamental building block of solid state semiconductor devices. The PN junctions can be found in a variety of devices including bipolar junction transistors (BJTs) and junction field effect transistors (JFETs). The most basic device built from the PN junction is the diode. Diodes are designed for a wide variety of uses including rectifying, lighting (LEDs) and photodetection (photodiodes).

Assuming the crystal is not at absolute zero, the thermal energy in the system will cause some of the free electrons in the N material to "fall" into the excess holes of the adjoining P material. This will create a region that is devoid of charge carriers (remember, electrons are the majority charge carriers in the N material while holes are the majority charge carriers in the P material). In other words, the area where the N and P materials abut is depleted of available electrons and holes, and thus we refer to it as a depletion region. This is depicted in Figure 3.5.

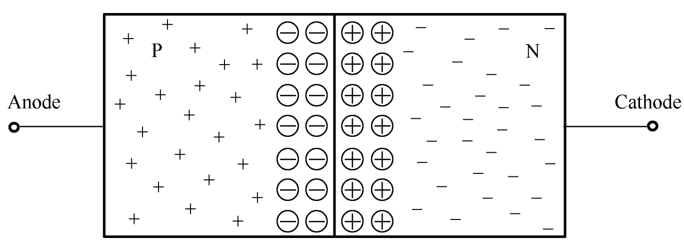


Figure 3.5 The PN junction

Now let's consider what happens if we were to connect this device to an external voltage source as shown in Figure 3.6. Obviously, there are two ways to orient the PN junction with respect to the voltage source. This version is termed forward-bias.

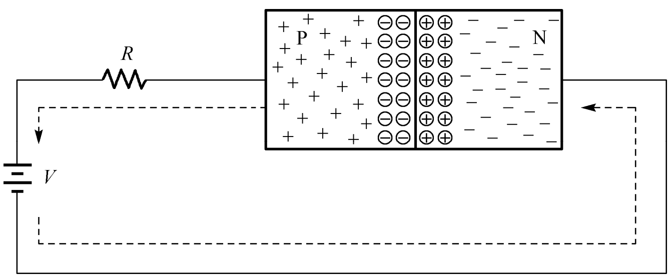


Figure 3.6 Forward-biased PN junction

The dotted line of Figure 3.6 shows the direction of electron flow (opposite the direction of conventional flow). First, electrons flow from the negative terminal of the battery toward the N material. In the N material, the majority carriers are electrons and it is easy for these electrons to move through the N material. Upon entering the depletion region, if the supplied potential is high enough, the electrons can diffuse into the P material where there are a large number of lower energy holes. From here, the electrons can migrate through to the positive terminal of the source, completing the circuit (the resistor has been added to limit maximum current flow). The "trick" here is to assure that the supplied potential is large enough to overcome the effect of the depletion region. That is, a certain voltage will be dropped across the depletion region in order to achieve current flow. This required potential is called the barrier potential or forward voltage drop. The precise value depends on the material used. **For silicon devices, the barrier potential is usually estimated at around 0.7 volts. For germanium devices, it is closer to 0.3 volts while LEDs may exhibit barrier potentials in the vicinity of 1.5 to 3 volts, partly depending on the color.**

If the voltage source polarity is reversed in Figure 3.6, the electrons in the N material will be drawn toward the positive terminal of the source while the P material holes will be drawn toward the negative terminal, creating a small, short-lived current. This has the effect of widening the depletion region and once it reaches the supplied potential, the flow of current ceases.

In its basic form, a diode is just a PN junction. It is a device that will allow current to pass easily in one direction but prevent current flow in the opposite direction. The schematic symbol for a basic switching or rectifying diode is shown in Figure 3.7. This is the ANSI standard which predominates in North America. The P material is the anode while the N material is the cathode. As a general rule for semiconductor schematic symbols, arrows point toward N material.

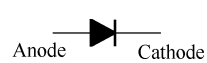


Figure 3.7 Diode schematic symbol (ANSI)

We can quantify the behavior of the PN junction through the use of an equation derived by William Shockley:

$$I_D = I_S(e^{\frac{V_D q}{nkT}} - 1)$$

where

I_D is the diode current.

I_S is the reverse saturation current.

V_D is the voltage across the diode.

q is the charge on an electron, 1.6E-19 coulombs.

n is the quality factor (typically between 1 and 2).

k is the Boltzmann constant, 1.38E-23 joules/kelvin.

T is the temperature in kelvin.

At 300 kelvin, q/kT is approximately 38.6. Consequently, for even very small forward (positive) voltages, the "−1" term can be ignored. **Also, I_S is not a constant. It increases with temperature, approximately doubling for each 10℃ rise.** For negative voltages (reverse-bias) the Shockley equation predicts negligible diode current. This is true up to a point. The equation does not model the effects of breakdown. When the reverse voltage is large enough, the diode will start to conduct. This is shown in Figure 3.8.

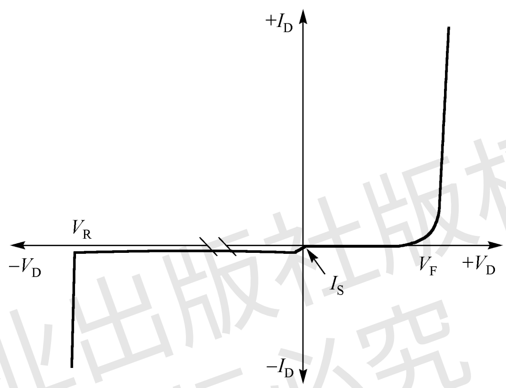


Figure 3.8 Simplified forward and reverse I_D-V_D curve for diode

V_F is the forward "knee" voltage (roughly 0.7 volts for silicon). I_S is the reverse saturation current (ideally zero but in reality a very small amount of current will flow). V_R is the reverse breakdown voltage. Note that the current increases rapidly once this reverse voltage is reached.

Words and Expressions

PN junction	PN 结
diode ['daɪoʊd]	*n*. 二极管
ceramic [sə'ræmɪk]	*n*. 陶瓷器；陶瓷制品
crystalline silicon	晶体硅
room temperature	室温
impurity [ɪm'pjʊrəti]	*n*. 杂质
dopant ['doʊpənt]	*n*. 掺杂物；掺杂剂
implantation [ˌɪmplæn'teɪʃn]	*n*. 离子注入；植入
donor atom	施主原子；供电子原子
outer shell	外电子层
acceptor atom	受主原子；接受体原子
hole [hoʊl]	*n*. 空穴
intrinsic semiconductor	本征半导体

bipolar junction transistor(BJT)	双极晶体管；双极结晶体管；双极结型晶体管
junction field effect transistor(JFET)	结型场效晶体管；结型场效应晶体管
rectify ['rektɪfaɪ]	*vt*. 整流；矫正；纠正
photodiode ['foʊtoʊˌdaɪoʊd]	*n*. 光电二极管
devoid [dɪ'vɔɪd]	*adj*. 缺乏；完全没有
majority charge carrier	多数载流子
depletion region	耗尽区
forward-bias ['fɔːrwərd'baɪəs]	*n*. 正向偏置
reverse-bias [rɪ'vɜːrs'baɪəs]	*n*. 反向偏置
diffuse [dɪ'fjuːs]	*v*. （使气体或液体）扩散，弥漫，渗透；（使光）模糊，漫射，漫散；传播；使分散；散布 *adj*. 弥漫的；扩散的；漫射的
migrate ['maɪgreɪt]	*v*. 迁移；转移
barrier potential	势垒电压；势垒电位
ANSI: American National Standards Institute	美国国家标准学会
anode ['ænoʊd]	*n*. 阳极；正极
cathode ['kæθoʊd]	*n*. 阴极；负极
reverse saturation current	反向饱和电流
quality factor	品质因数
Boltzmann constant	玻耳兹曼常数
reverse breakdown voltage	反向击穿电压

Notes

1. Any material can be classified as one of three types: conductor, insulator, or semiconductor.

任何材料都可以归为以下 3 种类型之一：导体、绝缘体或半导体。

2. In particular, the doped silicon can be used as a switch, turning current off and on as desired.

特别是，掺杂硅可以用作电流开关，根据需要关闭和打开。

3. The resulting crystal contains “holes” in its bonding structure where electrons would normally be located. In essence, such holes can move through the crystal conducting positive charges.

结果是，晶体的键合结构中原本是电子的位置被“空穴”取而代之。本质上，这样的空穴可以在晶体中移动，相当于传输了正电荷。

4. Assuming the crystal is not at absolute zero, the thermal energy in the system will cause some of the free electrons in the N material to “fall” into the excess holes of the adjoining P material.

假设晶体不是绝对零度，系统中的热能将导致 N 型材料中的一些自由电子“落入”相邻 P 型材料的空穴。

5. For silicon devices, the barrier potential is usually estimated at around 0.7 volts. For germanium devices, it is closer to 0.3 volts while LEDs may exhibit barrier potentials in the vicinity of 1.5 to 3 volts, partly depending on the color.

对于硅器件，势垒电压通常估计在 0.7 V 左右。对于锗器件，势垒电压接近 0.3 V，而 LED 的势垒电压可能在 1.5 V 至 3 V 附近，部分取决于 LED 的发光颜色。

6. Also, I_S is not a constant. It increases with temperature, approximately doubling for each 10℃ rise.

并且，I_S 不是常数。它随着温度的升高而增大，每升高 10℃，大约会增大一倍。

3.3 Transistor

A transistor is a semiconductor device used to amplify or switch electronic signals and electrical power. Transistors are one of the basic building blocks of modern electronics. It is composed of semiconductor material usually with at least three terminals for connection to an external circuit. A voltage or current applied to one pair of the transistor's terminals controls the current through another pair of terminals. Because the controlled (output) power can be higher than the controlling (input) power, a transistor can amplify a signal. Today, some transistors are packaged individually, but many more are found embedded in integrated circuits.

There are two broad classifications of transistors: bipolar junction transistors (BJT) and field effect transistors (FET).

3.3.1 BJT

A bipolar junction transistor consists of a three-layer "sandwich" of doped (extrinsic) semiconductor materials, either PNP in the Figure 3.9 (b) or NPN in the Figure 3.9 (d). Each layer forming the transistor has a specific name, and each layer is provided with a wire contact for connection to a circuit. The schematic symbols are shown in the Figure 3.9 (a) and (c).

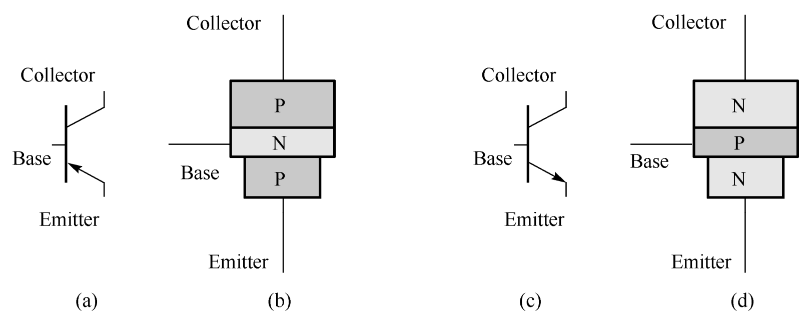


Figure 3.9 BJT: (a) PNP schematic symbol; (b) PNP physical layout; (c) NPN schematic symbol; (d) NPN physical layout

The functional difference between a PNP transistor and an NPN transistor is the proper biasing (polarity) of the junctions when operating. For any given state of operation, the current directions and voltage polarities for each kind of transistor are exactly opposite each other.

Bipolar junction transistors work as current-controlled current regulators. In other words, transistors restrict the amount of current passed according to a smaller, controlling current. The main current that is controlled goes from collector to emitter, or from emitter to collector, depending on the type of transistor it is (PNP or NPN, respectively). The small current that controls the main current goes from base to emitter, or from emitter to base, once again depending on the kind of transistor it is (PNP or NPN, respectively). This is shown in Figure 3.10.

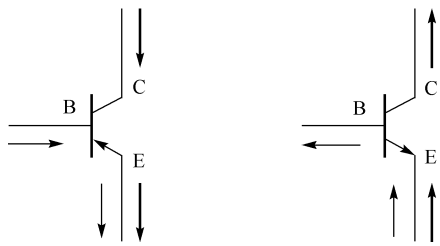


Figure 3.10 Small base-emitter current controls large collector-emitter current

Generally there are three different configurations of transistors and they are common base (CB) configuration, common collector (CC) configuration and common emitter (CE) configuration. The behaviors of these three different configurations of transistors with respect to gain are given below.

- Common base (CB) configuration: no current gain but voltage gain.
- Common collector (CC) configuration: current gain but no voltage gain.
- Common emitter (CE) configuration: current gain and voltage gain.

3.3.2 FET

A field effect transistor(FET) is a device utilizing a small voltage to control current, including the junction field effect transistor(JFET) and the insulated gate field effect transistor(IGFET). All field effect transistors are unipolar rather than bipolar devices. That is, the main current through them is comprised of either electrons through an N-type semiconductor or holes through a P-type semiconductor.

1. JFET

In a junction field effect transistor or JFET, the controlled current passes from source to drain, or from drain to source as the case may be. The controlling voltage is applied between gate and source. Note how the current does not have to cross through a PN junction on its way between source and drain: the path (called a channel) is an uninterrupted block of semiconductor material.

With no voltage applied between gate and source, the channel is a wide-open path for electrons to flow. **However, if a voltage V_{GS} is applied between gate and source of such polarity that it reverse-biases the PN junction, the flow between source and drain connection becomes limited or regulated.** Maximum gate-source voltage “pinches off” all current through source and drain, thus forcing the JFET into cutoff mode. **This behavior is due to the depletion region of the PN junction expanding under the influence of a reverse-bias voltage, eventually occupying**

the entire width of the channel if the voltage is great enough. Figure 3.11 shows the constructions and symbols for both configurations of JFETs.

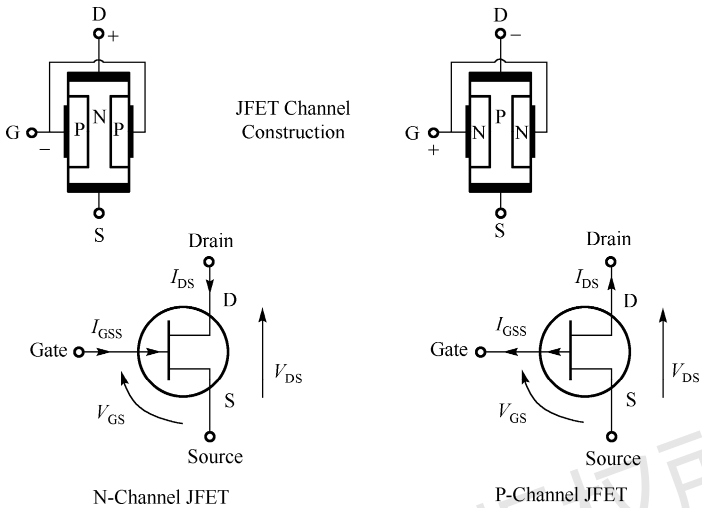


Figure 3.11 The constructions and symbols for both configurations of JFETs

Note how this operational behavior is exactly opposite of the bipolar junction transistor. Bipolar junction transistors are normally-off devices: no current through the base, no current through the collector or the emitter. JFETs, on the other hand, are normally-on devices: no voltage applied to the gate allows maximum current through the source and the drain. Also, take note that the amount of current allowed through a JFET is determined by a voltage signal rather than a current signal as with bipolar junction transistors. In fact, with the gate-source PN junction reverse-biased, there should be nearly zero current through the gate connection. For this reason, we classify the JFET as a voltage-controlled device and the bipolar junction transistor as a current-controlled device.

2. IGFET

Another type of field effect device—the insulated gate field effect transistor, or IGFET—exploits a similar principle of a depletion region controlling conductivity through a semiconductor channel, but it differs primarily from the JFET in that there is no direct connection between the gate lead and the semiconductor material itself. Rather, the gate lead is insulated from the transistor body by a thin barrier, hence the term insulated gate. **This insulating barrier acts like the dielectric layer of a capacitor and allows gate-source voltage to influence the depletion region electrostatically rather than by direct connection.**

The most common type of IGFET which is used in many different types of electronic circuits is called the metal-oxide-semiconductor field effect transistor, MOSFET for short, due to its metal (gate)-oxide (barrier)-semiconductor (channel) construction. In addition to a choice of N-channel versus P-channel design, MOSFETs come in two major types: depletion and enhancement.

- Depletion type: the transistor requires the gate-source voltage, V_{GS}, to switch the device “off ”. The depletion mode MOSFET is equivalent to a “normally closed” switch.

- Enhancement type: the transistor requires a gate-source voltage, V_{GS}, to switch the device "on". The enhancement mode MOSFET is equivalent to a "normally open" switch.

Figure 3.12 shows the constructions and symbols for both configurations of MOSFETs.

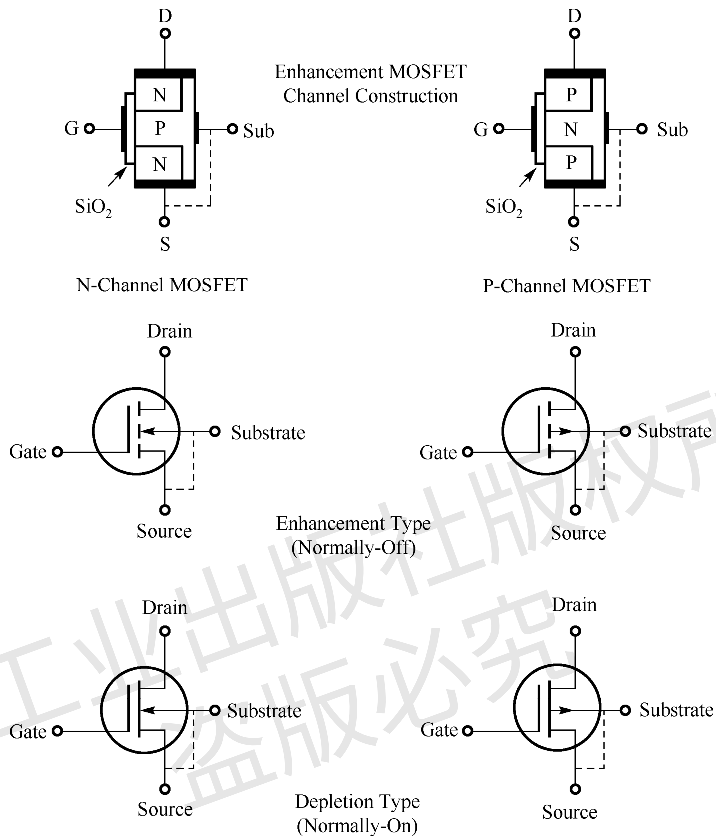


Figure 3.12 The constructions and symbols for both configurations of MOSFETs

The threshold voltage of a MOSFET is usually defined as the gate voltage where an inversion layer forms at the interface between the insulating layer (oxide) and the substrate (body) of the transistor. **As the gate terminal is electrically isolated from the main current carrying channel between drain and source, "No current flows into the gate" and the input resistance of the MOSFET is extremely high way up in the megohm (MΩ) region.**

MOSFETs are ideal for use as electronic switches or as common-source amplifiers as their power consumption is very small. Typical applications for metal-oxide-semiconductor field effect transistors are in microprocessors, memories, calculators, logic CMOS (complementary MOS) gates, etc.

Words and Expressions

amplify ['æmplɪfaɪ] *v*. 放大；增强（声音等）

package ['pækɪdʒ] *vt*. 封装；将……包装好

extrinsic [eks'trɪnzɪk] *adj*. 非本征的；非固有的；外来的

schematic symbol 电路符号；原理图符号

collector [kə'lektər] *n*. 集电极

base [beɪs] *n*. 基极

emitter [ɪ'mɪtər] *n*. 发射极

regulator ['regjuleɪtər] *n*. 调节器；调整器；校准器；监管者

insulated gate field effect transistor(IGFET) 绝缘栅场效[应]晶体管；绝缘栅场效[应]管

source [sɔːrs] *n*. 源极

drain [dreɪn] *n*. 漏极

gate [geɪt] *n*. 栅极

channel ['tʃænl] *n*. 沟道

cutoff ['kʌˌtɔf] *n*. 截止；切断；截止点；界限

dielectric [ˌdaɪɪ'lektrɪk] *n*. 电介质；绝缘体

electrostatically [ɪˌlektroʊ'stætɪkəli] *adv*. 静电地

metal-oxide-semiconductor field effect transistor(MOSFET) 金属-氧化物-半导体场效应晶体管

enhancement [ɪn'hænsmənt] *n*. 提高；增加；增强

threshold ['θreʃhoʊld] *n*. 阈值；阈；门限；门槛；界；起始点

inversion layer 反型层

complementary MOS(CMOS) 互补金属氧化物半导体

Notes

1. The functional difference between a PNP transistor and an NPN transistor is the proper biasing (polarity) of the junctions when operating.

PNP 晶体管和 NPN 晶体管之间的功能差异在于工作时的偏置电压极性不同。

2. Generally there are three different configurations of transistors and they are common base (CB) configuration, common collector (CC) configuration and common emitter (CE) configuration.

通常晶体管有 3 种不同的配置，它们是共基极（CB）配置、共集电极（CC）配置和共发射极（CE）配置。

3. However, if a voltage V_{GS} is applied between gate and source of such polarity that it reverse-biases the PN junction, the flow between source and drain connection becomes limited or regulated.

然而，如果栅极和源极之间施加 PN 结反向偏置电压 V_{GS}，则源极到漏极之间的电流将受到限制，或者说可以被调节。

4. This behavior is due to the depletion region of the PN junction expanding under the influence of a reverse-bias voltage, eventually occupying the entire width of the channel if the voltage is great enough.

这种行为是由于 PN 结的耗尽区在反向偏置电压的影响下扩展，如果电压足够大，最终耗尽区会占据沟道的整个宽度。

5. This insulating barrier acts like the dielectric layer of a capacitor and allows gate-source

voltage to influence the depletion region electrostatically rather than by direct connection.

该绝缘屏蔽层的作用类似于电容器的介电层，允许栅极-源极电压以静电方式影响耗尽区，而不是通过直接连接。

6. As the gate terminal is electrically isolated from the main current carrying channel between drain and source, "No current flows into the gate" and the input resistance of the MOSFET is extremely high way up in the megohm (MΩ) region.

由于栅极与漏源之间的主载流通道是电隔离的，因此，"没有电流流入栅极"，并且 MOSFET 的输入电阻高达兆欧（MΩ）量级。

3.4 FPGA

FPGA stands for field programmable gate array. At its core, an FPGA is an array of interconnected digital subcircuits that implement common functions while also offering very high levels of flexibility. But getting a full picture of what an FPGA is requires more nuance. **The following introduces the concepts behind FPGAs and briefly discusses what makes an FPGA different from a microcontroller in design, what logic gates are, and how to program an FPGA.**

3.4.1 FPGA Versus Microcontroller

Microcontrollers have become dominant components in modern electronic design. They're inexpensive and highly versatile, and nowadays they often serve as a person's first introduction to the world of electronics. As microcontrollers become increasingly powerful, there is less and less need to consider alternative solutions to our design challenges. Nonetheless, a microcontroller is built around a processor and processors come with fundamental limitations that need to be recognized and, in some cases, overcome.

So when would an engineer reach for an FPGA over a microcontroller? The answer comes down to software vs hardware. A processor accomplishes its tasks by executing instructions in a sequential fashion. **This means that the processor's operations are inherently constrained: the desired functionality must be adapted to the available instructions and, in most cases, it is not possible to accomplish multiple processing tasks simultaneously.**

The alternative is a hardware-based approach. **It would be extremely convenient if every new design could be built around a digital IC that implements the exact functionality required by the system: no need to write software, no instruction-set constraints, no processing delays, just a single IC that has input pins, output pins, and digital circuitry corresponding precisely to the necessary operations.** This methodology is impractical beyond description because it would involve designing an ASIC (application specific integrated circuit) for every board. However, we can approximate this methodology using FPGAs.

3.4.2 What Is a Programmable Gate Array?

An FPGA is an array of logic gates, and this array can be programmed (actually, "configured" is probably a better word) in the field, i.e., by the user of the device as opposed to the people who designed it.

Logic gates (AND, OR, XOR, etc.) are the basic building blocks of digital circuitry. However, an FPGA is not a vast collection of individual Boolean gates. It's an array of carefully designed and interconnected digital subcircuits that efficiently implement common functions while also offering very high levels of flexibility. The digital subcircuits are called configurable logic blocks (CLB), and they form the core of the FPGA's programmable-logic capabilities.

The CLBs include look-up tables, storage elements (flip-flops or registers), and multiplexers that allow the CLBs to perform Boolean, data-storage, and arithmetic operations. The CLBs need to interact with one another and with external circuitry. For these purposes, the FPGA uses a matrix of programmable interconnects and input/output (I/O) blocks. The FPGA's "program" is stored in SRAM cells that influence the functionality of the CLBs and control the switches that establish the connection pathways. An I/O block consists of various components that facilitate communication between the CLBs and other components on the board. These include pull-up/pull-down resistors, buffers, and inverters. Figure 3.13 shows field programmable gate array.

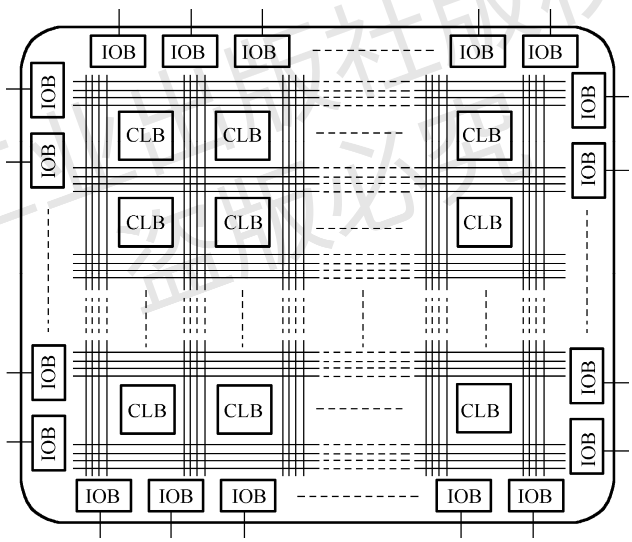


Figure 3.13 Field programmable gate array

3.4.3 How Do You Program an FPGA?

How do we go about turning an array of CLBs into a digital circuit that does precisely what we want it to? At first glance, it seems like a rather complicated task. Indeed, FPGA implementation is generally considered more difficult than programming a microcontroller. However, FPGA development does not require thorough knowledge of CLB functionality or painstaking arrangement of internal interconnects, just as microcontroller development does not require thorough knowledge of a processor's assembly-language instructions or internal control signals.

Actually, it is somewhat misleading to present an FPGA as a standalone component. **FPGAs are always supported by development software that carries out the complicated process of converting a hardware design into the programming bits that determine the behavior of interconnects and CLBs.**

People have created languages that allow us to “describe” hardware. They’re called (very appropriately) hardware description languages (HDL), and the two most common are VHDL and Verilog. Despite the apparent similarity between HDL code and code written in a high-level software programming language, the two are fundamentally different. **Software code specifies a sequence of operations, whereas HDL code is more like a schematic that uses text to introduce components and create interconnections.**

Figure 3.14 shows a digital circuit example and Figure 3.15 is the corresponding VHDL code.

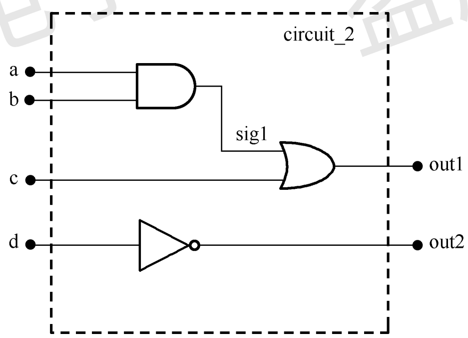


Figure 3.14 A digital circuit example

```
1    library ieee;
2    use ieee.std_logic_1164.all;
------------------------------------
3    entity circuit_2 is
4    port(a: in std_logic;
5         b: in std_logic;
6         c: in std_logic;
7         d: in std_logic;
8         out1: out std_logic;
9         out2: out std_logic);
10   end circuit_2;
------------------------------------
11   architecture Behavioral of circuit_2 is
12        signal sig1: std_logic;
13   begin
14        sig1 <= (a and b);
15        out1 <= (sig1 or c);
16        out2 <= (not d);

17   end Behavioral;
```

Figure 3.15 VHDL code of the example circuit

Lines 1 and 2: These lines add the required library and package to the code. Since the “std_logic” data type is used, we have to add the “std_logic_1164” package.

Lines 3-10: These lines specify the name of the module along with its input/output ports.

Lines 11-17: This part of the code describes the operation of the circuit. As you may have noticed, there is one internal node in Figure 3.14; it is labeled “sig1”. We use the “port” statement from “entity” to define the input/output ports, but how can we define the internal nodes of a circuit? For this, we use the “signal” keyword.

In line 12 of the above code, the “signal” keyword tells the synthesis software that there is a node in the circuit labeled “sig1”. Similar to the definition of the ports, we use the keyword “std_logic” after the colon to specify the required data type. Now we can assign a value to this node (line 14) or use its value (line 15).

Note: The “std_logic” data type is a commonly used data type in VHDL. It can be used to describe a one-bit digital signal which can actually take on nine values.

‘U’: Uninitialized

‘1’ : The usual indicator for a logic high, also known as ‘Forcing high’

‘0’: The usual indicator for a logic low, also known as ‘Forcing low’

‘Z’: High impedance

‘-’: Don’t care

‘W’: Weak unknown

‘X’: Forcing unknown

‘H’: Weak high

‘L’: Weak low

Among these values, we commonly use ‘0’, ‘1’, ‘Z’, and ‘-’.

In 2016, long-time industry rivals Xilinx (now part of AMD) and Altera (now an Intel subsidiary) were the FPGA market leaders. They controlled nearly 90 percent of the market. Both Xilinx (now AMD) and Altera (now Intel) provide proprietary electronic design automation software for Windows and Linux (ISE/Vivado and Quartus) which enables engineers to design, analyze, simulate, and synthesize (compile) their designs. **Modern FPGAs are sophisticated, high-performance devices that can be somewhat intimidating for those who are accustomed to using microcontrollers for gathering data, controlling ASICs, and performing mathematical operations.** You might find, however, that in some applications the improved performance and versatility are worth the additional design effort.

Words and Expressions

nuance ['nuːɑːns]	*n*. 细微的差别
versatile ['vɜːrsətl]	*adj*. 多功能的；多用途的；多才多艺的
sequential [sɪ'kwenʃl]	*adj*. 按次序的；顺序的；序列的
approach [ə'proʊtʃ]	*n*.（处理问题、完成任务的）方法
methodology [ˌmeθə'dɑːlədʒi]	*n*. 方法论；（从事某一活动的）方法，原则
XOR: exclusive OR	异或
look-up table	查找表
flip-flop ['flɪp flɑːp]	*n*. 触发器
multiplexer ['mʌltiˌpleksər]	*n*. 多路复用器
SRAM: static random access memory	静态随机存储器
buffer ['bʌfər]	*n*. 缓冲器；缓存区；缓冲存储器 *vt*. 缓存；缓冲；存储
inverter [ɪn'vɜːrtər]	*n*. 反相器；逆变器；变换电路；转换开关；

非门

hardware description language(HDL)	硬件描述语言
synthesis ['sɪnθəsɪs]	*n*. 综合；合成；综合体
indicator ['ɪndɪkeɪtər]	*n*. 标志；指示器；指针
impedance [ɪm'piːdns]	*n*. 阻抗
proprietary [prə'prɑɪəteri]	*adj*. 专有的；专利的

Notes

1. The following introduces the concepts behind FPGAs and briefly discusses what makes an FPGA different from a microcontroller in design, what logic gates are, and how to program an FPGA.

下面介绍 FPGA 的相关概念，并简要讨论 FPGA 与微控制器在设计上的区别、什么是逻辑门，以及如何对 FPGA 进行编程。

2. This means that the processor's operations are inherently constrained: the desired functionality must be adapted to the available instructions and, in most cases, it is not possible to accomplish multiple processing tasks simultaneously.

这意味着处理器的操作本身受到限制：所需的功能必须适应可用的指令，并且在大多数情况下，不可能同时完成多个处理任务。

3. It would be extremely convenient if every new design could be built around a digital IC that implements the exact functionality required by the system: no need to write software, no instruction-set constraints, no processing delays, just a single IC that has input pins, output pins, and digital circuitry corresponding precisely to the necessary operations.

如果每种新的系统设计都能基于一个数字集成电路来精确实现所需功能，设计者不需要编写软件，没有指令集约束，也没有处理延迟，只需要一个具有输入引脚、输出引脚并能精确完成所需操作的数字集成电路，那么将会是非常方便的。

4. FPGAs are always supported by development software that carries out the complicated process of converting a hardware design into the programming bits that determine the behavior of interconnects and CLBs.

FPGA 总是由开发软件支持，该开发软件可以将硬件设计转换为编程数据，以定义互连和 CLB 的行为。

5. Software code specifies a sequence of operations, whereas HDL code is more like a schematic that uses text to introduce components and create interconnections.

软件代码指定一系列操作，而 HDL 代码更像是一个原理图，它使用文本来引入电路组件并创建互连。

6. Modern FPGAs are sophisticated, high-performance devices that can be somewhat intimidating for those who are accustomed to using microcontrollers for gathering data, controlling ASICs, and performing mathematical operations.

现代 FPGA 是一种复杂的高性能设备，对那些习惯于使用微控制器收集数据、控制 ASIC 和执行数学运算的人来说，这可能有点惊人。

Exercises

1. Match the terms (1)−(6) with the definitions A−F.

(1) depletion region	A. a two-terminal electronic component that conducts current primarily in one direction
(2) VHDL	B. a three-terminal electronic device used to control the flow of current by the voltage applied to its gate terminal
(3) FET	C. a language that describes the behavior of electronic circuits, most commonly digital circuits
(4) diode	D. it is formed from a conducting region by removal of all free charge carriers, leaving none to carry a current
(5) breakdown voltage	E. an empty spot with a positive charge in the crystal lattice
(6) hole	F. the threshold voltage, beyond which an insulator starts behaving as a conductor and conducts electricity

2. Translate into Chinese.

(1) In April 2019, two of the world's largest semiconductor foundries—Taiwan Semiconductor Manufacturing Company Limited (TSMC) and Samsung Foundry—announced their success in reaching the 5 nm technology node, propelling the miniaturization of transistors one step further to a new age.

(2) Since digital circuits involve millions of times as many components as analog circuits, much of the design work is done by copying and reusing the same circuit functions, especially by using digital design software that contains libraries of prestructured circuit components.

(3) In general, MOSFET amplifiers tend to have good high frequency performance, offer low noise and exhibit low distortion with modestly sized input signals. Compared to BJTs, their voltage gain magnitude is lower.

3. Translate into English.

（1）发光二极管通电时可以发光，而光电二极管在光照下产生电流。

（2）MOSFET 的尺寸是用栅极长度来衡量的，也就是俗称的特征尺寸或特征长度，用符号 *L* 来表示。

（3）随着晶体管尺寸的缩小，它们的工作速度也随之提高。

4. Read the following article and write a summary.

The debate of whether Moore's law is "dying" (or already "dead") has been going on for years. It has been discussed by pretty much everyone. But before we can give an aswer to that, let's first clarify the meaning of Moore's law.

Moore's law stems from the observation of Gordon Moore, co-founder and chairman emeritus of Intel, made in 1965. At the time, he said that the number of transistors in a dense integrated circuit had doubled roughly every year and would continue to do so for the next 10 years. In 1975, he revised his observation to say that this would occur every two years indefinitely. Moore's observation became the driving force behind the semiconductor technology revolution that led to the proliferation of computers and other electronic devices.

Over time, the details of Moore's law were amended to reflect the true growth of transistor density. First, the doubling interval was increased to two years and then decreased to around 18 months. The exponential nature of Moore's law continued and created decades of opportunity for the semiconductor industry and the electronics that use them.

The issue for Moore's law is the inherent complexity of semiconductor process technology, and these complexities have been growing. Transistors are now three-dimensional, and the small feature size of today's advanced process technologies has required multiple exposures to reproduce these features on silicon wafers. This has added extreme complexity to the design process and has "slowed down" Moore's law.

This slowing down has led many to ask, "Is Moore's law dead?" The simple answer to this is no, Moore's law is not dead. While it's true that chip densities are no longer doubling every two years (thus, Moore's law isn't happening anymore by its strictest definition), Moore's law is still delivering exponential improvements, albeit at a slower pace. The trend is very much still here.

Intel's CEO Pat Gelsinger believes that Moore's law is far from obsolete. As a goal for the next 10 years, he announced in 2021 not only to uphold Moore's law, but to outpace it. There are many industry veterans who agree with this. Mario Morales, a program vice president at IDC, said he believes Moore's law is still relevant in theory in an interview with TechRepublic.

"If you look at what Moore's law has enabled, we're seeing an explosion of more computing across the entire landscape," he said, "It used to be computing was centered around mainframes and then it became clients and now edge and endpoints, but they're getting more intelligent, and now they're doing AI inferencing, and you need computing to do that. So, Moore's law has been able to continue to really push computing to the outer edge."

While the consensus is that Moore's law is slowing down and that it might soon be augmented, it is still driving improvements in processing technology and the amount of progress that follows these improvements. If it were dead, it simply couldn't do this.

5. Language study: Describing components

Two questions we may need to answer when we describe components are:

(1) What is it called?

(2) What does it do?

In other words, we need to be able to:

(1) label components.

(2) describe their functions.

We can use these ways of labeling components:

(1) It is called a photodiode.

(2) It is known as a Schottky transistor.

We can describe the functions of components like these:

(1) An FPGA **provides** a way to implement digital logic circuits in a short amount of time.

(2) Batteries **convert** chemical energy into electrical energy.