Extended Abstracts of the 1992 International Conference on Solid State Devices and Materials, Tsukuba, 1992, pp. 1-5

Invited

One Hundred Years of Semiconductor-Device Developments (1874–1974)

S. M. Sze
Institute of Electronics
National Chiao Tung University
Hsinchu, Taiwan, R.O.C.

ABSTRACT

Semiconductor devices have formed the basis of the largest industry in the world -- the electronic industry, and have had unprecedented impact on our society. We consider the development of semiconductor devices over a period of one hundred years from the earliest investigation of metal-semiconductor contacts in 1874 to the first observation of resonant tunneling in 1974. We shall present major milestones in the historical development of bipolar, unipolar, microwave, and photonic devices and briefly discuss their important applications.

1. INTRODUCTION

The earliest systematic study of semiconductor devices is generally attributed to Ferdinand Braun[1)], who in 1874 noted the dependence of the total resistance on the polarity of the applied voltage and on the detailed surface conditions of the metal-semiconductor contacts. For the next three quarters of century, semiconductor devices gained some practical applications as photodiodes and rectifiers.

The major breakthrough came about in 1947 when a research team at AT&T Bell Laboratories, headed by William Shockley, invented the bipolar transistor. The bipolar transistor and its related semiconductor devices have formed the basis of the largest industry in the world -- the electronic industry with global factory sales of US$660 billions in 1991, and a projected sales of US$1500 billions in the year 2000.

Coincident with the growth of the electronic industry, semiconductor-device literature has also burgeoned and diversified. Figure 1 shows the annual publication of device papers.[2)] It is apparent that device literature was in its incubation phase from 1874 to 1947. After 1947, we moved to the growth phase. The initial rapid growth lasted over a quarter of a century (to 1974). During this period, the annual publication doubled every two years. By 1974, the annual publication reached 3000, and the total device papers published in the past one hundred years reached 13,000. After 1974, growth has continued but at a slower rate, doubling every twelve years. If the growth of device literature remains at the current rate, we expect to reach an annual publication volume of over 10,000 papers in 2000 and, by then, the grand total will approach 220,000 papers.

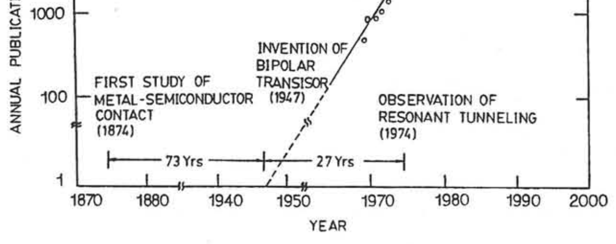

Fig. 1 Annual publication of semiconductor-device papers.

In this review, we consider key semiconductor devices from the incubation phase and the initial rapid-growth phase over a period of a hundred years (1874 to 1974). We have divided the devices into four groups: bipolar, unipolar, microwave, and photonic devices, and we shall consider them in subsequent sections.

It is interesting to point out that almost all major semiconductor devices were invented or developed in the aforementioned period. There are, however, two important exceptions -- the MODFET (modulation-doped field-effect transistor) and quantum-effect devices, which were developed after 1974. Nevertheless, the operational principle of the MODFET is virtually identical to that of the MOSFET (metal-oxide-semiconductor field-effect transistor), which was developed in 1960. The first paper on the observation of the resonant tunneling, on which the quantum-effect devices are based, was published in 1974.

2. BIPOLAR DEVICES

The bipolar transistor action was first observed in 1947. The experimental results of a point-contact transistor were presented by Bardeen and Brattain [3] in 1948. A schematic diagram of their device is shown in Fig.2. They also defined most of the key terminologies for the bipolar transistor, such as emitter, base, collector and current gain in their paper.

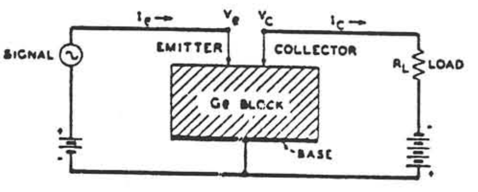

Fig. 2 Schematic of the point-contact bipolar transistor.

However, the point-contact transistor was unreliable and non-reproducible. It quickly gave way to the junction-type bipolar transistor invented by Shockley [4], Fig.3. He not only established the theory of current flow in the p-n-p junction transistor, but also worked out the basic equation of the p-n junction. Shockley's paper [4] is the foundation upon which all modern devices are built.

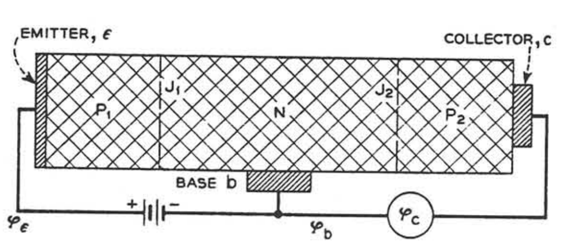

Fig. 3 Cross section of bipolar junction transistor.

To improve the emitter efficiency, Kroemer [5] in 1957, wrote an extremely important paper on the concept of using a wide-gap emitter. This is the first paper on the heterojunction bipolar transistor (HBT). The use of the heterojunction adds a new degree of freedom in transistor design and enables higher operating speeds. HBT is considered one of the most promising devices for ultrahigh-speed applications.

In 1951, Shockley et al. [6] proposed another important bipolar device: the multilayered p-n-p-n device, called a thyristor, Fig.4. The thyristor has two stable states (ON and OFF) and has low power dissipation in these states. It is a key device for switching applications. The detailed device principles and the first working thyristor were reported by Moll et al. [7], which served as the basis for all work in the study of thyristors.

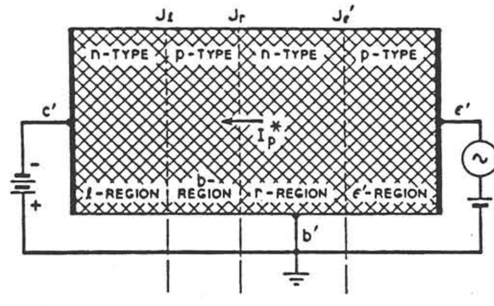

Fig. 4 A p-n-p-n transistor, also called the thyristor.

3. UNIPOLAR DEVICES

The first semiconductor device ever studied was a unipolar device -- the metal-semiconductor contact. As mentioned in the Introduction, in 1874 Braun discovered the rectification properties of metal-semiconductor contacts. It was later surmised that these properties might arise from the presence of a chemical layer on the surface of the semiconductor. In 1938, Schottky [8] suggested that the potential barrier in a metal-semiconductor contact could arise from stable space charge in the semiconductor alone without the presence of a chemical layer. The model arising from this idea is known as the Schottky barrier.

The fundamental current-transport process in metal-semiconductor contacts was derived by Bethe [9] in 1942, based on the thermionic-emission theory. This theory has been extended to include diffusion, electron-phonon scattering, quantum-mechanical reflection, tunneling, and minority-carrier effects. The metal-semiconductor contact is extremely important for direct current and microwave applications. When it is rectifying, the contact can serve as a microwave detector or as the gate electrode of an MESFET (metal-semiconductor field-effect transistor). The contact can also be ohmic, i.e., having negligible resistance regardless of the polarity of the applied voltage. All semiconductor devices need ohmic contact to make connections to other devices in an electronic system.

In 1952, Shockley [10] analyzed the junction field-effect transistor (the JFET), which is basically a voltage-controlled resistor. The first working JFET, shown in Fig.5, was reported by Dacey and Ross [11] in 1953. By replacing the p-n junction with a metal-semiconductor contact, Mead[12], in 1966, proposed the MESFET, shown in Fig.6. This device is a key device for high-speed applications, especially for monolithic microwave integrated circuits (MMIC).

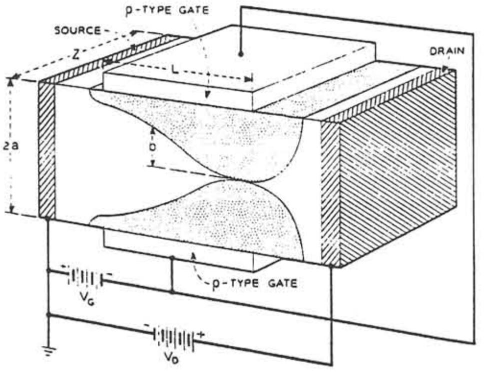

Fig. 5 Schmatic of junction field-effect transistor.

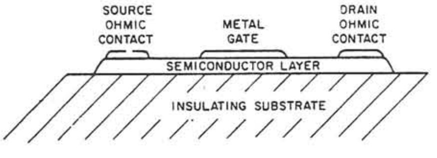

Fig. 6 Cross section of MESFET formed on an insulating substrate.

Currently, the most important semiconductor device is the MOSFET (metal-oxide-semiconductor field-effect transistor). In 1991, the worldwide sales of ICs (integrated circuits) was US$53 billions with over 70% MOSFET-based sales. The main reasons are the advantages of MOSFET in device miniaturization, low power dissipation, and high yield. In 1960, Kahng and Atalla [13] proposed and fabricated the first MOSFET using a thermally-oxidized silicon structure as shown in Fig.7. The device had a channel length of 25μm and a gate oxide thickness over 1000Å. Although present-day MOSFETs have been scaled down to submicron dimensions, the choice of silicon and thermally-grown SiO_2 used in the first MOSFET remains the most important combination.

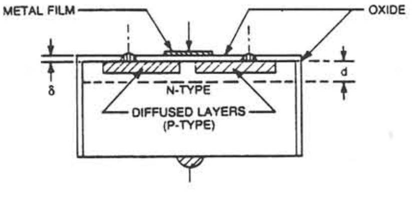

Fig. 7 Top view and cross section of the MOSFET.

As the complexity of ICs increases, we have moved from NMOS (n-channel MOSFET) to CMOS (complementary MOSFET) technology that employs both NMOS and PMOS (p-channel MOSFET) to form the logic elements. The advantage of CMOS technology is that the logic elements draw significant current only during the transition from one state to another and draw very little current between transitions, allowing power consumption to be minimized. The CMOS was proposed by Wanlass and Sah [14] in 1963. Today, CMOS technology is the most important technology for ultra-high-density integrated-circuit applications.

When the gate of a conventional MOSFET is modified so that semi-permanent charge storage inside the gate is possible, the new structure becomes a non-volatile memory device. The first non-volatile memory device was proposed by Kahng and Sze [15] in 1967 using a floating-gate concept, as shown in Fig.8. Because of its small cell area and its non-volatile property, this device and its related memory cells are key memory components in computer and communication systems.

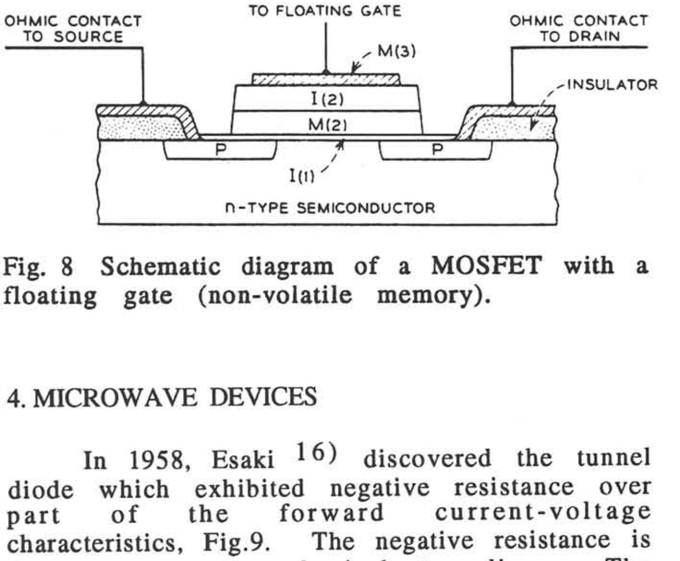

Fig. 8 Schematic diagram of a MOSFET with a floating gate (non-volatile memory).

4. MICROWAVE DEVICES

In 1958, Esaki [16] discovered the tunnel diode which exhibited negative resistance over part of the forward current-voltage characteristics, Fig.9. The negative resistance is due to quantum-mechanical tunneling. The impact of the tunnel diode on the physics of semiconductors has been large, leading to important developments such as tunneling spectroscopy, and to increased understanding of tunneling phenomenon in solids.

A related device is the double-barrier structure whose energy diagram is shown in the insert of Fig.10. When an incident electron has an energy that exactly equals one of the discrete energy levels in the quantum well (e.g., E_1 or E_2), it will transit through the double barrier with a unity transmission coefficient. This is called resonant tunneling, which was observed by Chang et al. [17] in 1974 using a thin GaAs layer sandwiched between two AlGaAs barriers fabricated by the molecular-beam epitaxy technique. Figure 10 shows their results; the arrows indicate the observed voltages of singularities corresponding to these resonant states. This device is the basic building block of quantum-effect devices.

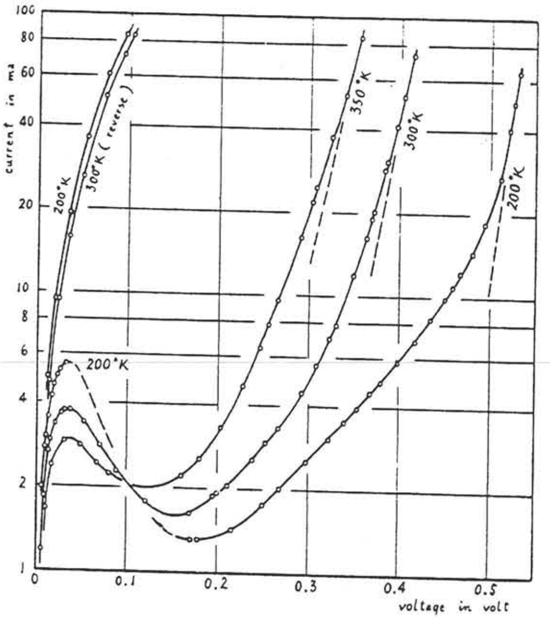

Fig. 9 Current-voltage characteristics of a tunnel diode measured at three temperatures.

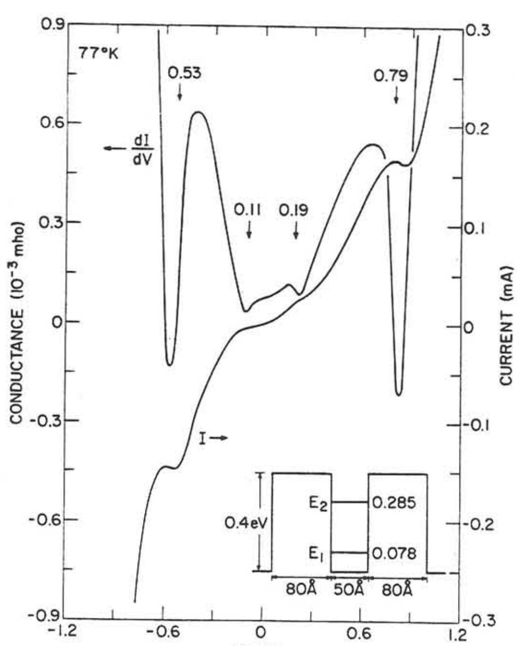

Fig.10 Current and conductance characteristics of a resonant-tunneling diode.Insert shows the energy diagram of the device.

IMPATT diode, proposed by Read [18] in 1958, can generate the highest cw (continuous wave) power at millimeter-wave frequencies, and is most extensively used in that frequency range (i.e., 30-300 GHz). The word IMPATT stands for "impact ionization and transit-time". IMPATT diodes employ impact ionization and transit-time properties of semiconductor structures to produce negative resistance at microwave frequencies. The experimental observation of the IMPATT oscillation was first reported by Johnston et al [19] in 1965.

Another important microwave device is the transferred-electron diode (TED). In 1963, Gunn [20] discovered that coherent microwave output was generated when an electric field that exceeded a critical threshold value of several thousand volts per centimeter was applied across an n-type sample of GaAs or InP. The observed properties were due to field-induced transfer of conduction-band electrons from a low-energy, high-mobility valley to higher-energy, low-mobility valleys. Therefore, the name transferred-electron device. TED is a key device for millimeter-wave applications. The power output and efficiency of TEDs are generally lower than that of IMPATT diode. However, TEDs have lower noise, lower operating voltages, and relatively easier circuit designs.

5. PHOTONIC DEVICES

The electroluminescence phenomenon was discovered by Round [21] in 1907, who observed the generation of yellowish light from a crystal of carborundum when a potential of 10 volts was applied between two points on the crystal. Subsequently, two important electroluminescent devices were developed: the LED (light-emitting diode with a spectral line width typically 100Å) and the laser diode (light-amplification-by-stimulated-emission-of-radiation diode with line width of 0.1 to 1Å).

LEDs have a multitude of application in the essential information linkage between electronic instruments and their human users. A key application is as a light source in optical-fiber communication. The first LED specifically designed for this purpose is the surface-emitter developed by Burrus and Miller [22] in 1971, Fig.11.

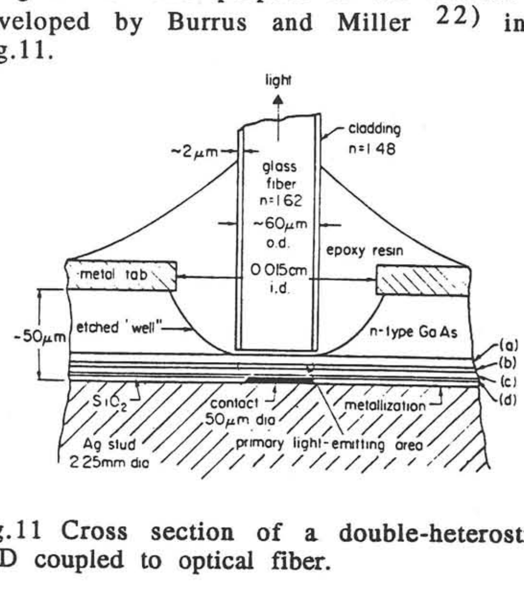

Fig.11 Cross section of a double-heterostructure LED coupled to optical fiber.

Semiconductor lasers are considered to be one of the most important light sources for optical-fiber communication. Semiconductor lasers are also important for applications in many areas of research and technology, such as high-resolution gas spectroscopy, pollution control, video recording, optical reading, and high-speed laser printing. The laser operation in semiconductor p-n junctions was achieved in 1962. Figure 12 shows the spectral distribution from a GaAs laser diode below and above the threshold obtained by Hall et al. [23)]

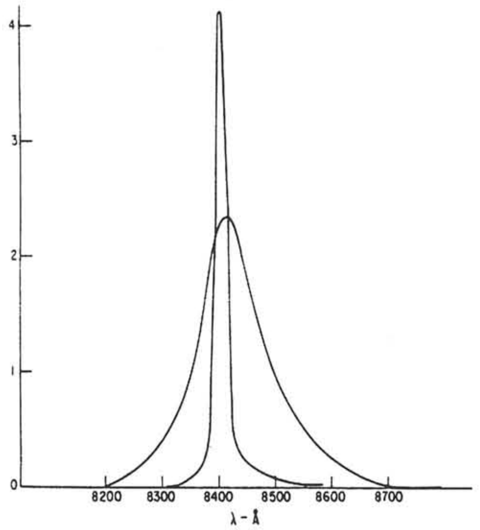

Fig. 12 Spectral distributions of a GaAs laser diode below (broad spectrum) and above (narrow spectrum) lasing threshold. Different vertical scales.

In 1963, Kroemer [24)] proposed the use of a double heterojunction for carrier confinement to reduce the threshold current density for lasing. The objective of continuous operation at room temperature was realized in 1970 by Hayashi et al. [25)] using a GaAs-AlGaAs double heterojunction laser structure.

Another important photonic device is the solar cell. Solar cells furnish the most important long-duration power supply for satellites and space vehicles. The solar cell is also a major candidate to obtain energy from the sun for terrestrial applications. The first solar cell was developed by Chapin et al. [26)] in 1954 using a diffused silicon p-n junction and obtained a 6% conversion efficiency.

6. DISCUSSION

Semiconductor devices will continue to be the key components for advanced electronic systems in the foreseeable future. As pointed out previously, almost all major semiconductor devices were developed before 1974. Since that time, numerous improvements and substantial advances have been made in the semiconductor-device field. New device concepts (e.g., bandgap engineering and functional devices) [27)], new materials (e.g., GeSi alloy and modulation-doped superlattice), and new technologies (e.g., rapid thermal annealing and micromachining technique) have been applied to design and to fabricate devices with enhanced performances. In the year 2000, we will have smaller, faster, more functional, and more reliable devices for the electronic industry which, in turn, will have a profound affect on the world economy.

REFERENCES

1) F. Braun, Ann. Phy. Chem. 153 (1874) 556.
2) S.M. Sze: Semiconductor Devices: Pioneering Papers (World Scientific - Singapore, 1991) p.xvii.
3) J. Bardeen and W.H. Brattain, Phys. Rev. 74 (1948) 230.
4) W. Shockley, Bell Syst. Tech. J. 28 (1949) 435.
5) H. Kroemer, Proc. IRE, 45 (1957) 1535.
6) W. Shockley, M. Sparks, and G.K. Teal, Phys. Rev. 83 (1951) 151.
7) J.L. Moll, M. Tanenbaum, J.M. Goldey, and N. Holonyak, Proc. IRE 44 (1956) 1174.
8) W. Schottky, Naturwissenschaften, 26 (1938) 843.
9) H.A. Bethe, MIT Radiat Lab. Rep. (1942).
10) W. Shockley, Proc. IRE 40 (1952) 1365.
11) G.C. Dacey and I.M. Ross, Proc. IRE 41 (1953) 970.
12) C.A. Mead, Proc. IEEE 54 (1966) 307.
13) D. Kahng and M.M. Atalla, IRE Device Research Conf., Pittsburgh, PA in June 1960.
14) F.M. Wanlass and C.T. Sah, ISSCC Dig. Tech. Pap. (1963) 32.
15) D. Kahng and S.M. Sze, Bell Syst, Tech. J. 46 (1967) 1283.
16) L. Esake, Phys Rev. 109 (1958) 603.
17) L.L. Chang, L. Esaki, and R. Tsu, Appl. Phys. Lett. 24 (1974) 593.
18) W.T. Read, Bell Syst. Tech. J. 37 (1958) 401.
19) R.L. Johnston, B.C. Deloach, and B.G. Cohen, Bell Syst. Tech. J. 44 (1965) 369.
20) J.B. Gunn, Solid State Commun. 1 (1963) 88.
21) H.J. Round, Electrical World, Feb. 9, 19 (1907) 309.
22) C.A. Burrus and B.I. Miller, Opt. Commun. 4 (1971) 307.
23) R.N. Hall et al, Phys. Rev. Lett. 9 (1962) 306.
24) H. Kroemer, Proc. IEEE, 51 (1963) 1782.
25) I. Hayashi, M.B. Panish, P.W. Foy, and S. Sumski, Appl. Phys. Lett. 17 (1970) 109.
26) D.M. Chopin, C.S. Fuller, and G.L. Pearson, J. Appl. Phys. 25 (1954) 676.
27) F. Capasso, S. Sen, and F. Beltram: High Speed Semiconductor Devices, ed. S.M. Sze (Wiley, New York, 1990) p.465.