From Transistors to Gadgetry:
How the Transistor Has Shaped Modern Electronics

Dr. John Stewart

University Physics II

Section H001

Calley Martin

Spring 2012

Introduction

When most people turn on their laptop or smartphone, they likely think of circuitry, if they consider it at all, in terms of how the power gets from the battery to the phone, rather than of the simple fact that the circuitry *is* the phone. The capacitors and transistors of the integrated circuit are what run the processes initiated by the touch of a button. But what is a transistor? And how do they factor into technological development?

This paper will begin with the fundamental physics of transistor function and the basics of their design. In a brief history of the transistor I will discuss their discovery, early use, and the changes transistors have undergone over time (especially in size and fabrication). I will then discuss the importance of transistors to the circuitry of electronic devices and the correlation between past development in the two industries. Using this background, we may hopefully gain a scientific perspective on the future of transistor development and the, thereby, the future of mobile technology.

The Science of Transistors

As ability of a transistor to function stems directly from the properties of the semiconductor from which it is made, some understanding of the nature of these materials is a prerequisite for understanding the operation of a transistor. In a conductor, charge is free to move around due to the atomic structure of the material: because the atoms' valence energy levels are only partially full, electrons are able (and inclined) to move about within the material. In insulators, the opposite is true: the valence energy levels or 'bands' are full and so the electrons are not free to move around. This is because, as stated in the Pauli Exclusion Principle, no two electrons in the same system can occupy the same state at the same time (Feldman, 72); that is, there is nowhere for an electron to go when each state is full. In essence, a semiconductor is an insulator which can be made into a conductor under certain stimuli (ibid). Materials that can be said to act in this way are those who become less resistant, rather than more, as their temperature is increased. This effect can be greatly increased through ‘doping’, a process by which impurities that contain either electron deficiencies or superficiencies are introduced into the material in order to “upset the regularity of the atomic structure” (Thomas, 68). It is this irregularity that provides charge the incentive to move within a system.

A transistor is a device that utilizes the properties of semiconductors to amplify power in a circuit. Physically, transistors are made up of two semiconductor diodes separated by a base layer. The configuration of these can vary greatly by transistor, but all have three contact points; generally they are the collector, the emitter, and the base contact (see example in Fig. 1). In field effect transistors, the corresponding contacts are called source, drain, and gate contacts, and the semiconductor regions are separated from the base by a thin insulator layer (Thomas, 68). The purpose of the third contact of a transistor (unlike a simple resistor, which only has two) is to utilize the nature of their semiconductor components. In semiconductors, the resistance of a certain junction may be influenced by voltage in a nearby junction (Thomas, 68); thus, through the application of a small current or voltage to the base terminal of a transistor, we can control a much larger current through the emitter and collector terminals.

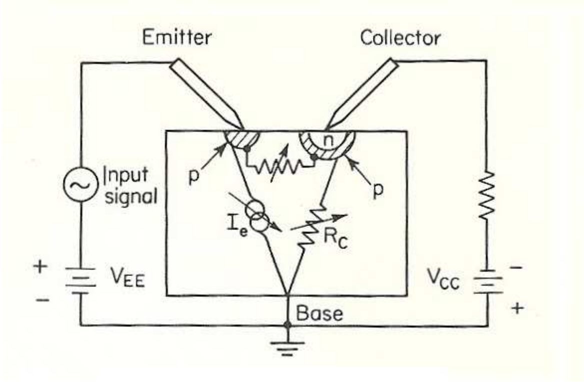


Figure 1. Circuit Diagram of a Point-contact Transistor

History of the Transistor

This “trans-resistance phenomenon” (Thomas, 68) was discovered by scientists attempting to utilize semiconductors to control the flow of electrons in solid matter. Already they could do so in a vacuum; triodes that controlled current within vacuum sealed tubes had been invented as early as 1918 (Haviland, 2002). However, the use of these triodes was limited by the ability to produce and maintain a vacuum within a tube, and not only was the size of any system made up of these tubes extremely large, but they took immense amounts of energy to run.

The transistor (a solid-state triode) grew out of not only the vacuum tube triode, but the solid-state rectifier, a device which was made from semiconductors like the transistor but only had two terminals. The rectifier, a device which converts alternating current to direct current, was invented by Ferdinand Braun (also the creator of the cathode ray tube) in 1874, Braun’s original solid-state design was unstable and therefore made obsolete by advances in vacuum tube technology. By the 1930’s, however, the limits of vacuum technology were apparent, and engineers were attempting to return to solid-state materials to develop alternate triode designs.

As a result, the first point-contact transistor (Fig. 1) was built and the transistor effect discovered in 1947 by John Bardeen and Walter Brattain of Bell Labs (Haviland, 2002).

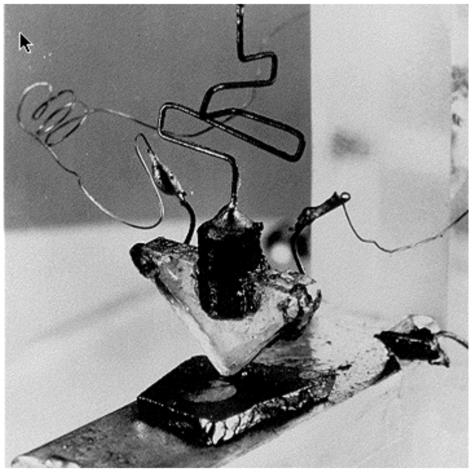


Figure 2. The First Transistor

Subsequently, William Shockley (also of Bell labs) developed the junction (or sandwich, as he called it) style transistor which became much more commercially successful due to the comparative simplicity of its production. He was the first to realize that, contrary to Bardeen and Brattain's belief, the electrons were being projected *through* the crystal and not moving along the surface (Brinkman, 97).

Initial semiconductor devices (such as Braun's 1874 rectifier) were made of either magnesium, copper oxide, or selenium, and were much less stable than what today's use of Germanium and Silicon (Ge and Si) can produce. The problem then was in the lack of sufficiently doped and refined crystals of these elements for production. Breakthroughs were made in the fifties, however, which allowed for successful refinement of even the high-temperature-requiring Silicon. The development of adequate doping processes took slightly longer, finally becoming reliable in the 1970's when sodium's detrimental effect on the silicon-silicon oxide interface was uncovered.

Today, progress in transistor development is marked by miniaturization. The inspiration for this trend is, in fact, consumer demand. Because the power of a computer can be directly related to the number of transistors available to preform computations, and consumers demand more portable and more powerful devices, transistors must be continually shrunk to meet the need (Bogatin, 97).

Transistors and the Integrated Circuit

The push for smaller, more powerful electronics spawned a new focus of research in the field: that of microelectronics. But miniaturizing any circuit is no easy task; the resistance of any wire is affected by its length and cross-sectional area, as is the capacitance of a capacitor. Because transistors connect both of these in complicated switching systems, keeping them in balance is both complex and essential (Coffey, 2003).

Three distinct approaches to the field evolved not long after the birth of the transistor in order to develop these miniature components: micro-assembly, thin film structures, and integrated circuits (Moore, 98). While all three produced smaller, lighter components, the integrated circuit has another draw: multiple transistors could be combined into a single unit by placing several triodes on a single piece of semiconducting crystal. In fact, borrowing techniques from the thin film and micro-assembly approaches, entire circuits of transistors, resistors, diodes, and capacitors could be combined into single units. The versatility of these 'chips' (yes, what we colloquially call computer chips today are actually intricate circuits of transistors laid out on a piece of semiconducting crystal) enabled the entire computer revolution. They were much more reliable than equivalent circuits put together piecemeal from individual components and could be produced rapidly and en masse. Today, all circuitry made up of transistors is in the form of integrated circuits with innumerable elements packed into tiny spaces. Take for example the heart or 'main board' (like a desktop motherboard) of an iPhone (Figure 3 below). Eighteen chips of hundreds of combined transistors, capacitors, etc. are all integrated onto a single board.

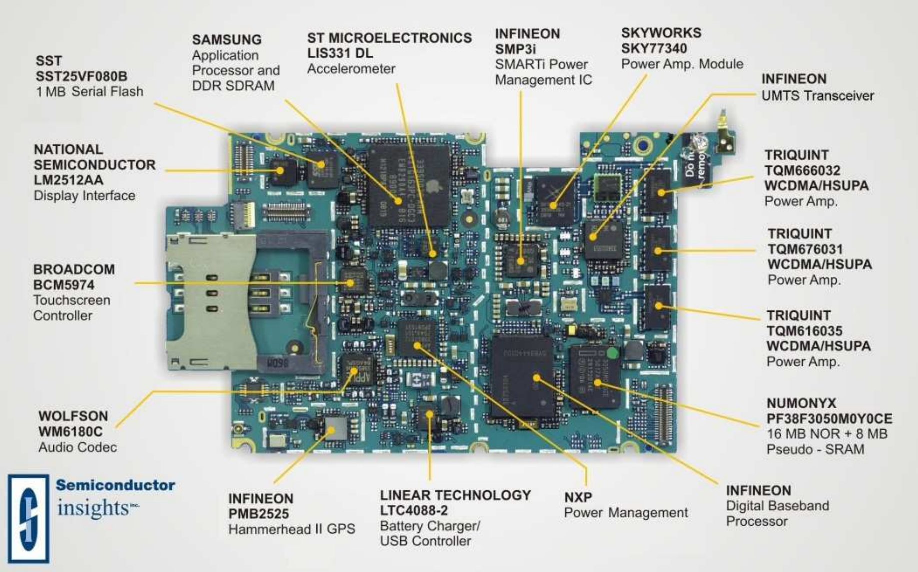


Figure 3. Main board of an iPhone

What's the Use?

The question we have so far failed to answer is that of why transistor development is the so key to technological development. The answer lies in the very function of a transistor: to perform logical operations through controlled current. It is common knowledge that computers run off a series of 1's and 0's, but more specifically, these digits translate into binary ("on" or "off") pulses through the switches that form a computer's circuitry. Transistors in pairs manipulate the current and voltage relationship in a circuit to preform 'calculations' based on Boolean logic. Each transistor controls one input to the logical AND or OR gate element of the circuit, and the output is based on the voltage felt by the base of the transistor. It should be noted that these operations are not based positive and negative voltages, but whether or not an element feels a voltage at all. The 'forward' and 'reverse' biases used in transistor function are not scaling or manipulating the amount or direction of current except to turn it on or off. Figure 4 shows a single example logic gate created with transistors.

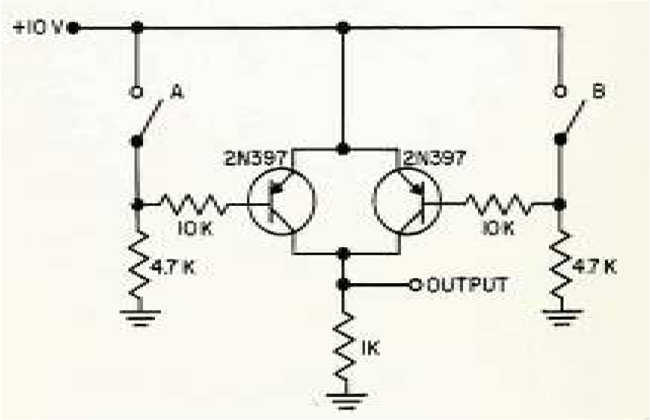


Figure 4. Example Two-transistor Logic Gate

Computers, and indeed all digital electronic devices, need vast quantities of these gates in order to perform the complex operations expected of them. So transistors became the backbone of logic in computing, and have in fact been called the "nerve cell of information technology"

("The Transistor", 2012). Even the first fully transistorized calculator (Fig. 5), the IBM 608 that was released in the late 1950's, required 3,000 of them. At the size of a washing machine and dyer combined, it was half that of its vacuum tube predecessor, and did not even have the capabilities of a modern TI-85.

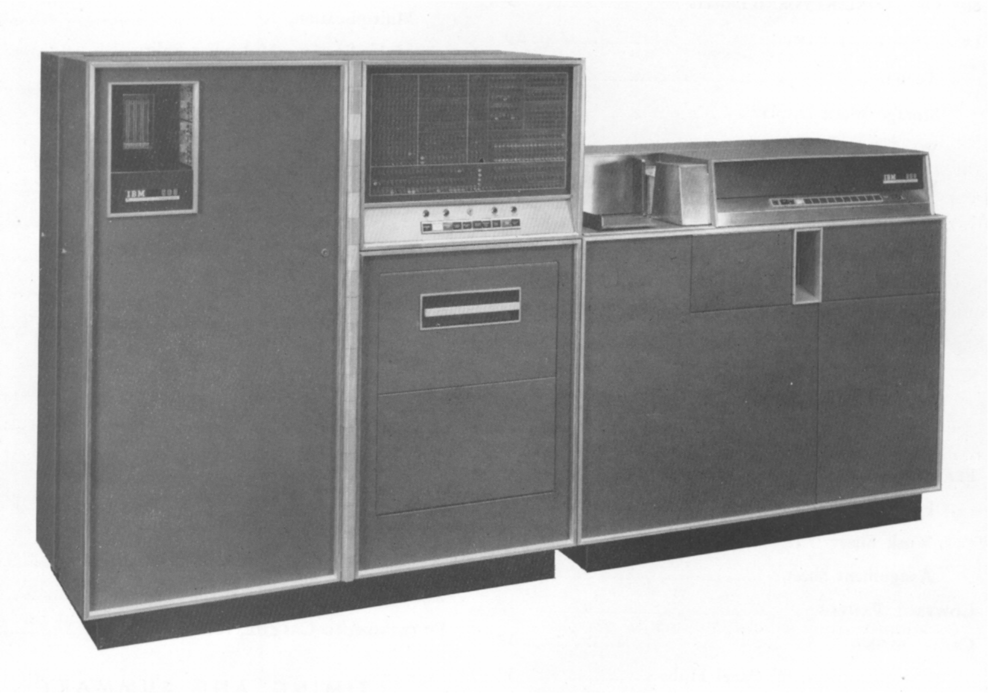
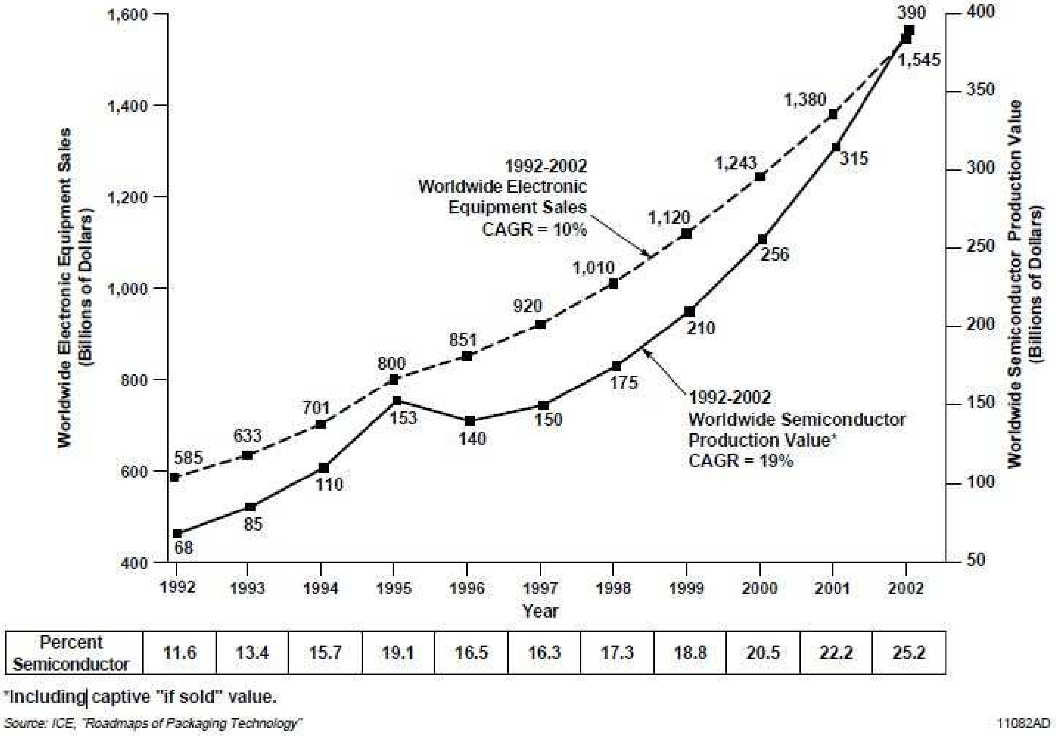


Figure 5. IBM 608, The First Transistorized Calculator

Developmental Trends

Considering the dependency of computational devices on transistor technology, it is no surprise that the development of one has closely followed that of the other. As consumer desire for convenience has pushed electronic devices to become smaller and more powerful, transistor and semiconductor industries have been blazing the way for the miniaturization of hardware and electronic components. Figure 5 below employs sales data to illustrate the closeness of the relationship between the electronic device and component industries.



Percent Semiconductor	11.6	13.4	15.7	19.1	16.5	16.3	17.3	18.8	20.5	22.2	25.2

*Including captive "if sold" value.

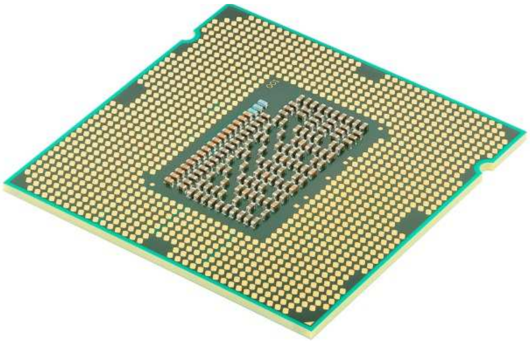
Source: ICE, "Roadmaps of Packaging Technology"

Figure 6. Industry Sales Trends (1992-2001)

Now, one side of this relationship can be plotted fairly reliably, thanks in part to Gordon Moore. In 1965 he made a prediction (that he returned to and revised ten years later) and developed what would come to be known as Moore's Law: that, based on the going rate of "device complexity" over time, transistor density could be expected to double every eighteen (originally twelve) months (Borsuk, 2003). Moore's Law was less a general law of nature and more like a self-fulfilling prophesy. Recognizing the need for sustaining such development or else risk falling behind, the rate given by Moore became an industry-wide standard, a goal of sorts. In fact, in the 1990's, representatives of the semiconductor industry, as well as others began to collaborate on major planning documents known as Roadmaps. These roadmaps identified goals and the best strategies to meet them in order that developmental expectations would continue to be met (Grier, 2006). They are what have enabled downscaling to the point where Intel’s 2011 IvyBridge microprocessor (pictured below) can contain 1.4 billion transistors in only 160 square millimeters of space (Shimpi, 2011).

Moore made no claim as to the progression of electronics as a whole, only to certain component technologies, the semiconductor and transistor. But since growth in these fields has so entirely formed the basis for advances in computer capability and information technologies (Borsuk, 2003), it is often extended to a much more general claim. What happens, then, if Moore’s law fails and we can no longer continue to shrink the size of transistors? It is not yet known what, if any, minimum size there might be for a transistor, but it is feasible that, if we continue to thin the layers of their semiconductor composition, eventually there may be too few atoms for the material to maintain its crystalline state. In fact, extrapolating with Moore’s Law indicates that the scale of transistors should reach these atomic dimensions by 2050 (Borshuk, 2003).

Figure 7. Intel's IvyBridge Microprocessor



So, if we can no longer go smaller, what do we do? Something similar has already happened in the case of “mixed-signal devices”, those that combine analog and digital components for processes such as converting a sound file to sound waves. Due to “various physical reasons,.. you can’t simply shrink these circuits the way you shrink digital circuits” (Grier, 2006), without losing the properties that allow them to function. This does not mean the end of growth in the mixed-signal industry, however. Instead of shrinking their product, companies have had to come up with clever designs and more efficient manufacturing processes. Upon reaching a lower limit to the size of their product, the transistor industry could follow this example as a model of where to go next.

Conclusion

The question then becomes, what would happen to the rest of the electronics industry? After relying on the ability to provide more power simply by increasing circuit density for so long, will it reach a similar stall when transistors cease to shrink? The DOD (Department of Defense) is afraid of such an occurrence, having come to depend on such rapid advances in order to “maintain technological superiority” (Borshuk, 2003). If transistor and integrated circuit engineering does reach such a point, perhaps research efforts could be better spent endeavoring to discover something entirely knew, rather than reefing the old. As Underhill said in 1933 about then-obsolete technologies, “when a machine has reached its highest stage of development it is ready for [either] the scrap heap or the museum. Each is a step toward something entirely different which will perform the same general function in a much better and more efficient manner.” Such was the transistor to the vacuum tube triode, and surely not even the sky is the limit to where we can go from here.

Bibliography

Brinkman, William F., Douglas E. Haggan, and William W. Troutman. Dec 1997. "A History of the Invention of the Transistor and Where It Will Lead Us*." IEEE Journal of Solid-State Circuits, 3*2 no. 12.

Bogatin, Eric. 1997. "End User Roadmaps." In *Roadmaps of Packaging Technology,* edited by Dick Potter and Laura Peters. Scottsdale, AZ: Integrated Circuit Engineering. http://smithsonianchips.si.edu/ice/cd/pkg_bk/chapt_02.pdf.

Coffey, Timothy and Gerald M. Borsuk. July 2003. "Moore's Law: A Department of Defense Perspective." *Defense Horizons*, no. 30: 1-10.

Davies, P. C. W. 1986. *The Forces of Nature.* Avon: The Bath Press.

Encyclopædia Britannica Online. "Electricity." Accessed April 6, 2012. http://www.britannica.com/EBchecked/topic/182915/electricity.

Feldman, James M. 1972. *The Physics and Circuit Properties of Transistors*. New York: John Wiley & Sons, Inc.

Grier, David Alan. Feb 2006. "The Innovation Curve." *Computer*, 39 no. 2: 8-10.

Haviland, David B. Nobelprize. 2002. "The Transistor in a Century of Electronics." Accessed Feb 18, 2012. http://www.nobelprize.org/educational/physics/transistor/history/.

Kilby, Jack S. July 1976. "Invention of the Integrated Circuit." IEEE Transactions on Electron Devices. 23 no. 7: 648-654.

Lowry, H. R. 1960. *General Electric Transistor Manual*. New York: General Electric Company.

Moore, Gordon E. Jan 1998. "Cramming More Components onto Integrated Circuits." Proceedings of the IEEE, 86 no. 1: 82-85.

Nobelprize.org. 2002. “The Transistor.” Accessed Feb 18, 2012. http://www.nobelprize.org/educational/physics/transistor/function/intro.html.

Pizzo, Patrick P. 1997. “Semiconductor Materials.” Last revision Dec. 12, 2011. http://www.engr.sjsu.edu/WofMatE/Semiconductors.htm

Shimpi, Anand Lal. 2011. “Ivy Bridge: 1.4B Transistors.” Accessed April 22, 2012. http://www.anandtech.com/show/4798/ivy-bridge-148b-transistors.

Thomas, Harry E. 1968. *Handbook of Transistors, Semiconductors, Instruments, and Microelectronics.* Englewood Cliffs, NJ: Prentice-Hall, Inc.

Underhill, Charles R. 1933. *Electronics at Work.* New York: McGraw Hill