How would you have responded? As I remember it, my own response at the time was largely incoherent. But having had fifteen years to ponder the question, I'm going to try to frame an answer in these notes.

To begin with, take notice of the fact that the automobile as a technological artifact has converged in its major performance characteristics. Cars have been able to travel faster than any sane person would want to drive them for at least eighty years. With rare exceptions, they have for this length of time had four wheels, an internal combustion engine, a steering wheel, brakes, and so on. True, my 1955 Chevrolet had a carburetor and a choke that required endless fussing, and it didn't have antilock brakes or power steering, but it was basically the same machine I drive today, fifty years later. It is easy to think of many other machines that have similarly converged in all their essential performance characteristics, at least within factors of ten. For example: refrigerators, light bulbs, airplanes, the houses we live in, and so on. The usual kinds of machines seem always to have natural limits on performance and efficiency. Commercial aircraft can (and do) travel faster than sound, but it wouldn't make sense to have them travel, say, a hundred times the speed of sound. It might be possible to build a light bulb that lasted ten years, and was a little more efficient, but it would be much more expensive. I'm guessing that few consumers would care about a light bulb that lasted a thousand years.

## 2.2 The supercomputer in 1957

Computers are in a completely different category from these common everyday machines. There seems to be no limit on the utility of ever faster computers with more and more memory. I can't think of any other kind of artifact with kind of unlimited potential usefulness. I was driving my 1955 Chevrolet when I was working at my first summer job, in 1957, at IBM in Manhattan. The machine I was using, the IBM 704, was housed in a building on 57th Street, which IBM modestly called "World Headquarters". It was one of the last computers to use vacuum tubes (not transistors), and it filled the space of several large rooms (see Fig. 2.1). The tape drives (on the right) were the size of refrigerators, with glass doors the length of the cabinet and vacuum columns at the sides to take up the tape slack. They fitfully clicked and clacked as the tape jerked along. Only one person at a time could use the machine, and I punched a time clock to get on and off with my prepared deck of punched cards and large spools of tape. Running a program was an afternoon's project. When the line printer chugged off my results (I can still smell the ink and see the paper with alternating patterns of horizontal green lines to guide the eye), I collected my decks and output and took a cab headed downtown to my office to ponder my results. If there was a bug in the program, it meant another trip uptown and another afternoon. Today I'm impatient when the little hour glass spins for more than a second to run my program.

At the time, the 704 was one of the fastest machines in the world, by definition a "supercomputer". How fast was it? Fortunately, I saved my IBM 704 "Manual of Operation" [15].

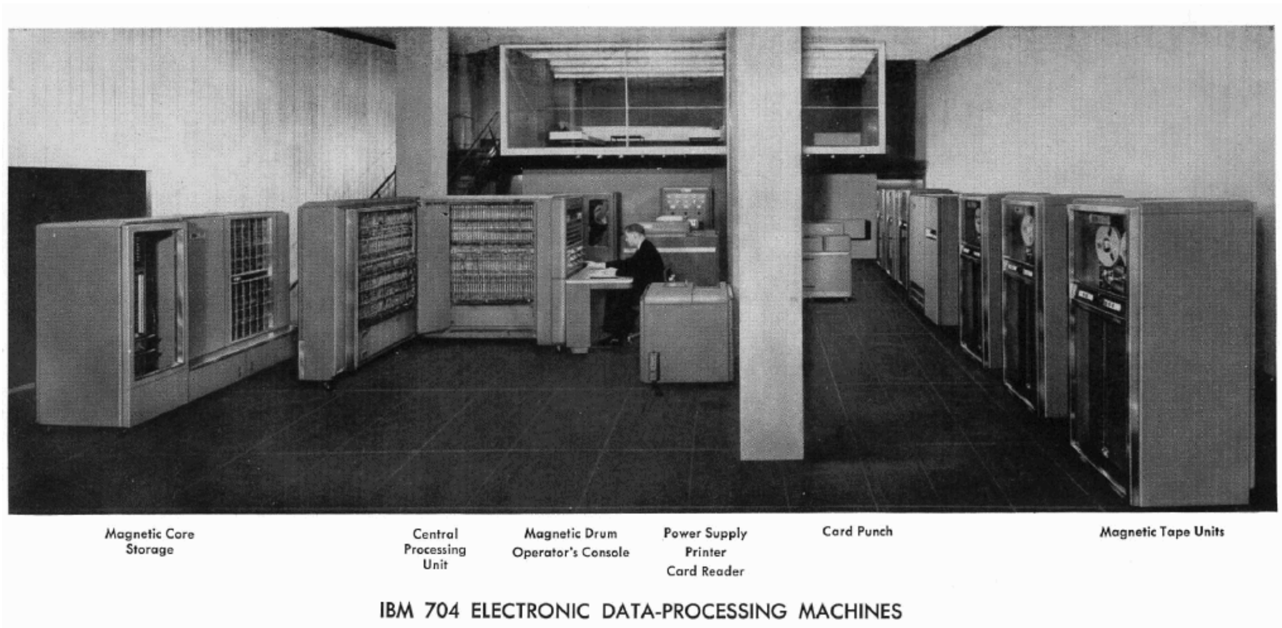


Figure 2.1: The IBM 704, ca. 1954. Source: [15], permission pending.

The basic cycle time of the machine was 12 microseconds, but even the fastest instructions required two cycles. Floating-point operations took about 20 cycles, so the basic speed of the computer was about 4 Kflops, where a flop is a floating-point operation per second, and a Kflop is 1000 flops. And just how much faster is the fastest machine today? As I write, the fastest supercomputer is the IBM Blue Gene/L, currently benchmarked at 135 Tflops, where a Tflop is a trillion ($10^1$2) flops.[2] Thus, the Blue Gene/L is about 10 billion times faster than the 704. One second of computation on the Blue Gene/L translates to 317 years on the 704.

While we went through this speed comparison for the fastest computers of their times, even current laptops are impressively faster than those of 1957. A modern laptop that has a clock speed of 2 GHz (corresponding to a cycle time of 0.5 nsec), might be able to do about a billion flops—the laptop can pipeline the floating-point operations, and may even have multiple processors, so it doesn't need 20 machine cycles for floating-point instructions. This makes it about 250,000 times faster than the 704, which means that 1 second on my laptop would need about 3 days on the IBM 704 I used in 1957. Not bad for a four-pound device that I can throw in my backpack (refer to Fig. 2.1).

The progress in memory has also been startling. The random-access memory (RAM) for the 704 used magnetic core memory,[3] which was very expensive and very bulky, and even the largest

[2] "IBM's Blue Gene/L Sets New Speed Record", Aaron Ricadela, Courtesy of *InformationWeek*, March 28, 2005, http://informationweek.networkingpipeline.com/news/159907121.

[3] The "core" in "core dump" comes to us from this era. I guess future generations will also be using the term

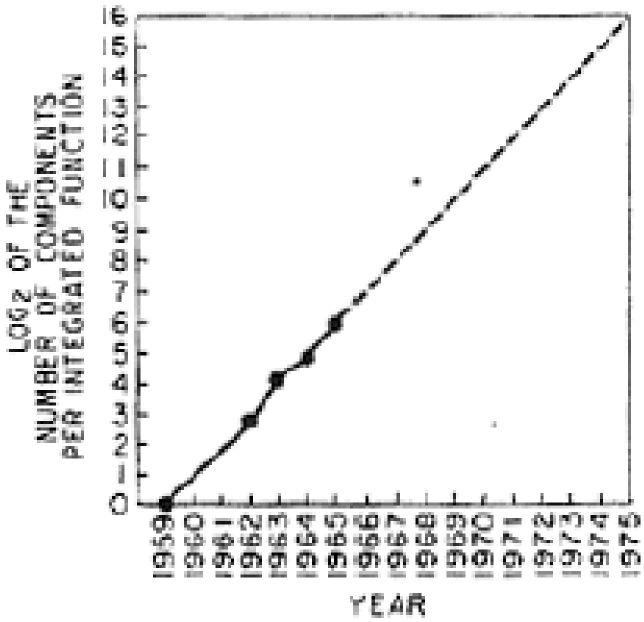


Figure 2.2: Moore's meager data, from his 1965 article[26]. Permission pending.

machines were always strapped for RAM. The largest size available was 32 Kwords, where each word had 36 bits, which means that the largest of these gigantic machines had a little less than 150 Kbytes of RAM, about one-hundredth the size of my digital camera's built-in memory—the memory that is much too small for a vacation without the few hundred Mbytes on a tiny card.

I won't belabor the point further. The technological advances have been astronomical. But the really interesting phenomenon is the *rate* of the progress, its roots, and whether it can be sustained. Does the history of the computer fit neatly into Provost Goldfeld's conception of the convergence to a refined product and a saturated market? Should we expect the improvements in computers to grind to a halt when some technical or economic limits have been reached? Or is there something fundamentally different going on here? To examine this question more closely, we go back to the remarkable article that Gordon E. Moore wrote for an electronics magazine in 1965 [26].

## 2.3 Moore's Law

Predicting future technology is a tricky job, and pictures of the future are often very far from the mark. First of all, it's difficult enough to see what might be directions in which it will

"dial tone".