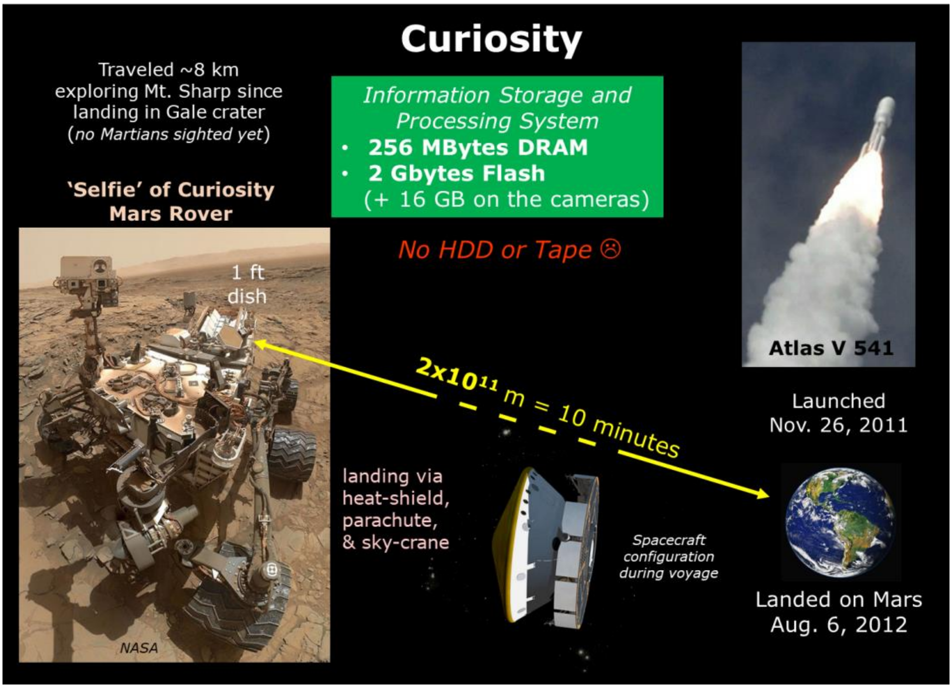
Chapter +11: Curiosity – A Road Trip on Mars



In January 2009, young twelve-year old Clara Ma of Kansas won the NASA essay competition to name the automobile-sized, one-ton, nuclear-powered Mars rover. The chosen name was "Curiosity". Certainly a very apt name - *curiosity* is by far the most important attribute for any scientist.

> "Curiosity is an everlasting flame that burns in everyone's mind. It makes me get out of bed in the morning and wonder what surprises life will throw at me that day. Curiosity is such a powerful force. Without it, we wouldn't be who we are today. When I was younger, I wondered, 'Why is the sky blue?', 'Why do the stars twinkle?', 'Why am I me?', and I still do. I had so many questions, and America is the place where I want to find my answers. Curiosity is the passion that drives us through our everyday lives. We have become explorers and scientists with our need to ask questions and to wonder. Sure, there are many risks and dangers, but despite that, we still continue to wonder and dream and create and hope. We have discovered so much about the world, but still so little. We will never know everything there is to know, but with our burning curiosity, we have learned so much." *Clara Ma, 2009*

The Curiosity rover stands in sharp contrast to Voyager and Galileo. It is a much more modern vehicle launched in 2011, more than twenty years after Galileo (1989) and more than thirty years after Voyager (1977) and also, notably, with only a mere two year launch delay. Unfortunately, from the perspective of the authors, Curiosity is devoid of any magnetic storage or even a magnetometer (Mars no longer has a liquid iron core and hence no magnetic field). So we take this chapter as the first opportunity to introduce solid-state memories - with which the rover is well endowed. However, before we embark on that task, it is well worthwhile to provide a brief description of Mars and of Curiosity's continuing highly-successful engineering and scientific mission.

Mars is desert planet but with large polar ice-caps. There are numerous canals that carry water from the icecaps to the temperate regions. The canals support broad swaths of cultivated land on either side that are visible from Earth. The canals are constructed and maintained by an ancient but advanced civilization known as the Barsoomians who come in several different colors.

Unfortunately, or perhaps fortunately, none of that last paragraph is true – except for the first sentence - Mars is indeed a desert planet with large polar icecaps. It is also similar to Earth in having a roughly the same length of day (24.66 hours) and similar axial tilt (25.19 degrees). But that is where the similarity ends. An unprotected tourist visiting Mars would die as certainly and as quickly as a visitor to Voyager's interstellar space or to the Galilean moons. Mars has almost no atmosphere (~1% of Earth's) and has nothing to prevent the radiation of the solar wind hitting the surface. Remarkably, around 1900, many astronomers really were reporting observations of canals on Mars. These turned out to be pure illusion brought on perhaps by some wishful thinking. Percival Lowell, a very distinguished businessman, author, and scientist, was perhaps the strongest proponent[1]. The idea of Martian canals fired the imagination of Edgar Rice Burrows (a prolific author also famous for 'Tarzan') who quickly populated the planet with fictional creatures and with exciting adventures and romances[2]. Sadly, the only real Martians likely to exist will be some extremophile bacteria sheltered well away from the surface.

---

[1] https://en.wikipedia.org/wiki/Percival_Lowell, https://en.wikipedia.org/wiki/Martian_canal

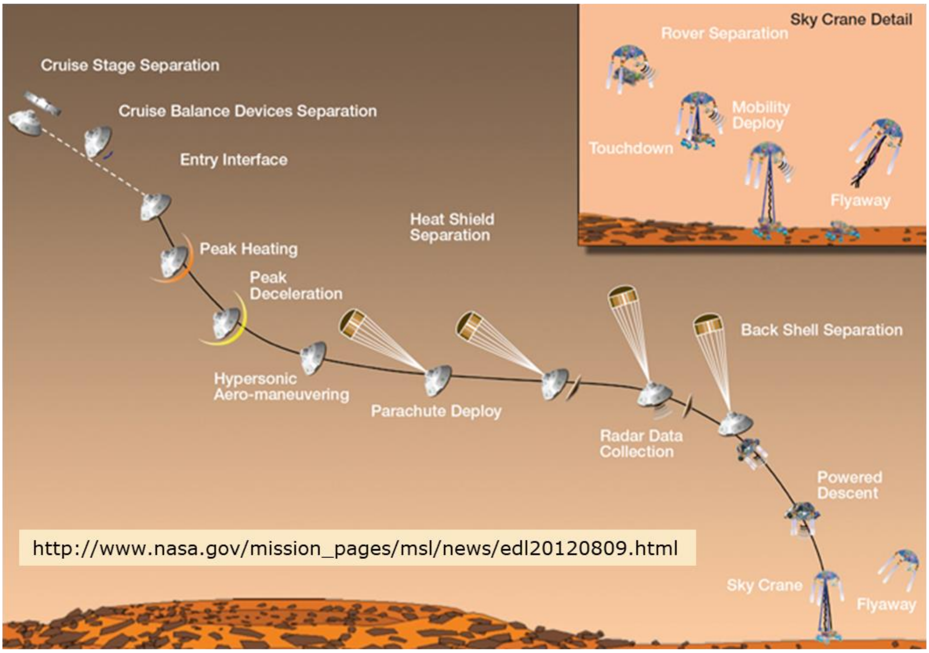
[2] https://en.wikipedia.org/wiki/Edgar_Rice_Burroughs https://en.wikipedia.org/wiki/Barsoom

Mars is actually a much smaller planet than Earth. Mars is 6779 km (4212 mi) in diameter. Earth is almost twice the size at 12,742 km (7918 mi) diameter. In fact, Mars is not very much larger than the largest moon, Ganymede, at 5,262 km (3274 mi) diameter. The small size has meant that Mars' core cooled and solidified early and there has been no significant tectonic activity for a billion years or more. The lack of a conducting liquid core has also meant that whatever protective magnetic field may have existed, it has now disappeared. Because of the low gravitational field (0.37g) Mars would have difficulty hanging on to a substantial atmosphere, but with the absence of a global magnetic field, whatever atmosphere was originally present has been largely stripped away by the solar wind. The lack of a substantial atmosphere not only makes Mars a decidedly unpleasant place to live, it also makes it a very difficult planet to land on safely.

How to land on Mars

The Curiosity spacecraft, broadly referred to as the Mars Science Laboratory (MSL) was launched on an Atlas V rocket from Cape Canaveral, Florida, on November 26th, 2011. A final stage liquid-fuel Centaur rocket boosted the MSL into a direct trajectory to Mars. After an eight month cruise phase, the MSL spacecraft, with Curiosity carefully cocooned inside, approached its high-speed rendezvous with Mars.

Missions to Mars have been notoriously unsuccessful. Altogether, 27 of 56 attempted missions have failed for a variety of reasons (as of Oct 2016). Some were due to human error, most notoriously, the Mars Climate Orbiter which got too close to Mars and burned up in the atmosphere because of confusion between metric and imperial units during trajectory control. But Mars is inherently difficult to land on. Earth has a relatively thick atmosphere at its surface, so parachutes can bring the descent speed down to just a few meters per second for a relatively gentle landing. This is not true on Mars where the surface pressure is only about 1% of Earth's. Parachutes on Mars can slow an object to perhaps 50 m/s (100 mph) but not nearly enough to land a large relatively fragile object like Curiosity. A variety of approaches have been used including giant airbags that cushioned the landings of the earlier rovers: Sojourner, Spirit, and Opportunity.



Landing sequence for the Curiosity rover (Mars Science Laboratory)

The landing sequence for NASA's MSL may seem like an improbable combination of events and techniques, but it did end in a complete success (in contrast to the European Space Agency's Schiaparelli[3] Mars Lander employing some similar techniques). Before entering the atmosphere the cruise-stage is separated and the aero-shell containing Curiosity is de-spun. Weights are ejected to shift the center of mass and then the attitude is carefully set, all in preparation for the aero-braking phase. On the leading, blunt end of the aero-shell is a huge 15 foot (4.5 m) diameter heat-shield. The aeroshell enters at a shallow angle to the atmosphere. It is shaped such that, with the correct orientation, a considerable amount of hydrodynamic lift is generated as it 'flies'. Small thrusters on the back of the aero-shell can change the orientation or attitude and thus the direction of the lift allowing the descending spacecraft to be actively steered. The spacecraft approaches Mars at 13,000 miles per hour (5.9 km/s). It 'flies' in at a descent angle of about 15 degrees executing a series of wide S-

[3] https://en.wikipedia.org/wiki/Schiaparelli_EDM_lander#Crash
Giovanni Schiaparelli (1835-1910) was a distinguished Italian scientist
https://en.wikipedia.org/wiki/Giovanni_Schiaparelli

shaped banking turns (like the Space Shuttle) to slow its descent and actively steer itself towards the desired landing site.

At about 10 km (6 miles) above the surface and having slowed to about 500 meters per second (1,000 miles per hour) a mortar cannon shoots the supersonic parachute out of the rear of the aero-shell. This quickly slows the spacecraft to about 100 meters per second (200 miles per hour). At this point the heat-shield, having done its job, is allowed to fall away. At about a kilometer (0.6 miles) above the surface, the rover and final descent stage drop away from the aero-shell and parachute. The descent stage fires large rockets that slow the descent to about one meter per second. The Curiosity rover is now slowly released to dangle suspended by three ropes 8 meters (25 feet) below the descent stage. The rover's six wheels are deployed from their stowed positions to act as landing gear. As soon as the rover touches the surface, the ropes are severed and the descent module is directed to fly away and crash at some distance from the rover.

On August 6th, 2012, this entire sequence executed flawlessly. The actual landing was within 2.4 km (1.5 mi) of the center of the target area in Gale crater. Within moments of landing, Curiosity transmitted the signal that it was safely positioned on the surface of Mars. Fourteen minutes later at 10:32 pm, the signal was received at JPL and a pandemonium of joy erupted among the engineers in mission control.

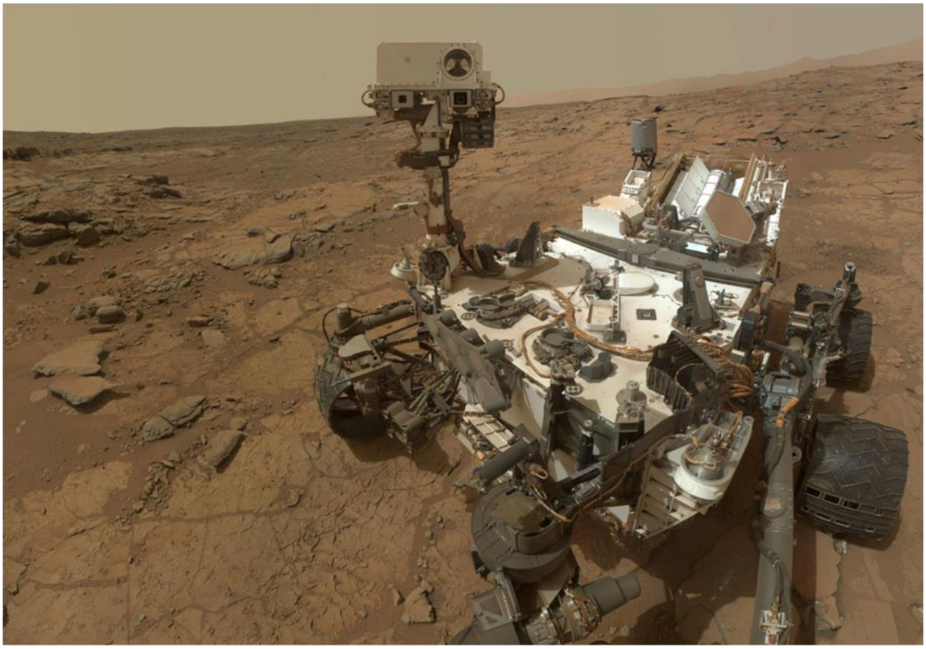
Curiosity Rover

Curiosity is roughly the size and weight (899 kg or 1,982 lbs) of a rather small, rather square automobile and it does move around using wheels. But that is where the similarity ends. There's obviously no place for passengers and no steering wheel. It sports six wheels, rather than the usual four. Its top speed is only 0.14 km/hr (0.09 mph). But, on the positive side, it will never need to stop at a petrol (gas) station. It is nuclear-thermo-electrically powered and the fuel should be sufficient to last for several decades.

Curiosity is actually a mobile, fully-automated, science laboratory[4] and a marvel of servo-mechanical engineering. The entire ensemble is powered by a Radioactive Thermonuclear Generator (RTG) similar to those used in the Voyager and Galileo missions. The generator provides about 100 watts

[4] https://en.wikipedia.org/wiki/Curiosity_(rover)
https://msl-scicorner.jpl.nasa.gov/Instruments/

of electric power (enough to light a standard incandescent light bulb) but also about 2 kilowatts of heat. The latter is put to very good use in warming the body of Curiosity and controlling the temperature of critical components via Freon-containing heat-pipes. Despite Gale crater lying almost on the equator, temperatures are decidedly chilly ranging from -90^o C (-130^o F) at night to 0^o C (32^o F) during the day (similar to our Antarctica). Unfortunately a mere 100 watts of electrical power is not really enough to run the rover and the experiments to full capability. So, at night, excess power from the RTG charges two large 42 amp-hour Lithium-ion batteries. The batteries themselves are an example of components that must be maintained within a narrow temperature-range for optimal functioning and reliability.



A **'selfie'** from Curiosity on the surface of Mars. This is a composite of several photographs taken between Feb. 3 and May 10, 2013. The camera is on the extended robotic arm held above the front left wheel (neither is seen). The RTG power supply and high-gain antenna (hexagonal) are towards the rear. MastCam and the inlets to the SAM and CheMin are towards the front.

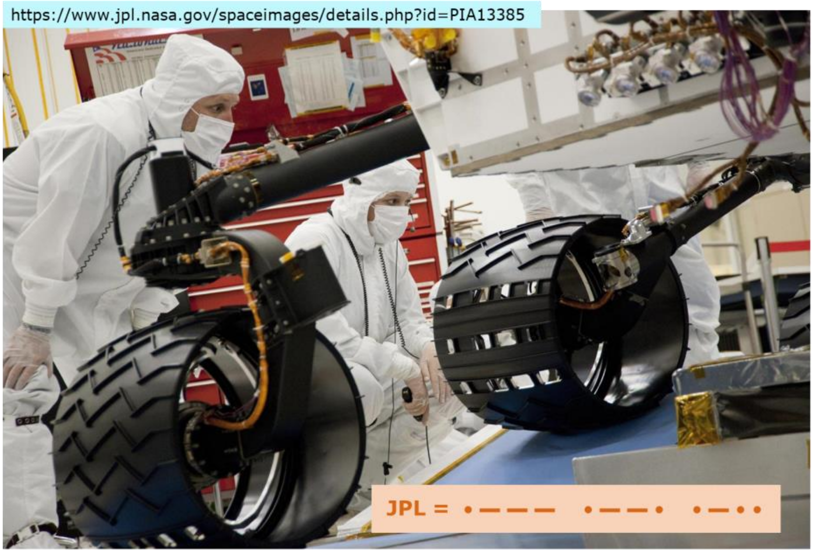
Let's examine the rover's activities on a typical Martian day (referred to as a 'sol' and, coincidentally, very close to 24 hours). Curiosity will wake up in the morning and first check for any new instructions from Earth. It will measure the local weather and will perhaps take a sniff of the morning "air" to test its composition (96% CO_2, 2% N_2, 2% Argon with only traces of O_2 and other gases). Then it will nap for a few hours till around noon when the Sun will have warmed its extremities to a more reasonable minus 30^{o}C. Then finally it is time for two or three hours of serious work.

Curiosity can drive autonomously. True, the roads on Mars are not in particularly good shape, but at least there is not a lot of traffic to deal with. The closest and, in fact, the only other active vehicle on Mars (as of 2018) is the Opportunity rover. The two vehicles are some 8500 km (5300 miles) apart and since both vehicles are limited to top speeds of just a few inches or centimeters per second, there is little chance of a serious collision. Curiosity is equipped with six extra-wide large-diameter 50 cm (20 in) wheels. These are similar in size to an automobile tire but obviously very different in design and not pneumatic. The rover has high clearance so it can drive happily over obstacles up to 60 cm (2 feet) tall. Each wheel is independently suspended and driven. The front and rear wheels can be steered independently to the extent that Curiosity can actually rotate in place without moving forward. The wheels have cleats for traction and the cleats also leave a distinct tread pattern on soft surfaces. The pattern allows the navigation cameras to gauge distance-travelled but the pattern just so happens to spell out JPL in Morse code[5]

During the next two or three hours of serious activity, Curiosity might traverse a few hundred feet (a hundred or so meters) following the general instructions received from Earth but navigating largely autonomously to avoid obstacles and difficult terrain. Curiosity has spectacularly good eyesight. There are a total of 17 cameras on board, 12 of which are dedicated to modeling the surrounding landscape in 3D for navigation and hazard avoidance. A tall mast over seven feet (2.2 m) high supports two high-resolution color cameras: a wide-angle 15^{o} camera that can resolve 22 mm per pixel at 100 meters and a narrow-angle 5.1° camera that can resolve 7.4 mm per pixel at 100 meters. As Curiosity progresses slowly on its journey, you may see it pause while the mast cameras slowly rotate through 360°. If Curiosity spots an interesting rock, there may be a sudden staccato burst of 50 or more 5 ns pulses from a high-power infrared laser aimed at the rock of interest. This is "CHEMCAM" in action. As the laser strikes the rock, a small quantity is vaporized and partially ionized. The

[5] https://mars.nasa.gov/msl/mission/rover/wheelslegs/

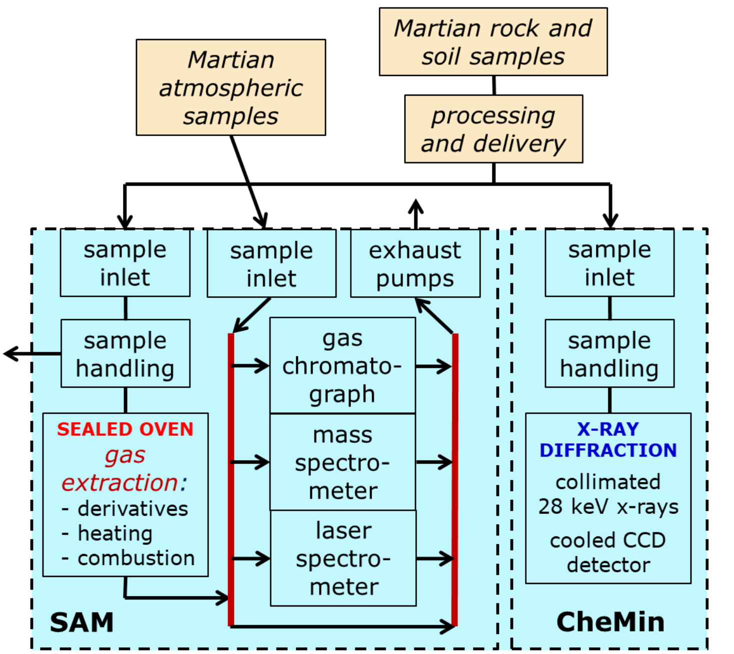
glowing plasma emits a rich spectrum of light from infrared to ultraviolet that is analyzed to reveal the composition of the rock.



A Wheel Check at NASA's Jet Propulsion Laboratory. Curiosity has six wheels each independently driven (*note JPL in morse-code formed into the wheel tread*)

If the rock proves to be of sufficient interest, Curiosity will roll to within a few feet of the rock, then carefully extend a 2.1 meter triple-jointed robotic arm. On the end of the arm is a collection of instruments, not least of which is a wire-brush used to clean off the surface of the rock. After that, the instrument set might rotate to bring to bear an optical microscope with a resolution of 14.5 micrometers per pixel to examine the fine rock structure in detail. The next instrument rotated into position will irradiate the rock with alpha particles (helium nuclei) and examine the X-rays returned. For the finale, the percussion drill can be brought to bear on the rock. This creates a 16 mm (5/8") diameter hole up to 50 mm (2") deep. The finely-ground material thus obtained from the interior of the rock can be captured for further processing. The rover may also choose to scoop a small sample of sand or dirt from the Martian surface for similar processing. In either case, the sample is subject to vibration and passed through a fine 1 mm sieve to the "SAM" analysis laboratory or to the "ChemMin" tool both housed in the body of the rover.

About half the body of Curiosity is devoted to the Sample Analysis at Mars or SAM instrument suite[6]. The three key instruments include a set of Gas Chromatographs (GC) to separate out different molecular species, a Quadrupole Mass Spectrometer (QMS) that can distinguish between different ions based on their charge to mass ratio, and a Tunable Laser Spectrometer (TLS) that is designed specifically to detect CO_2, CH_4, and H_2O and their isotope ratios. The diagram below shows the complex sample handling system. The samples can be from the atmosphere, from drilled rock samples, or from 'dirt' scooped from the Martian soil by the robotic arm. The solid samples can be heated in a helium atmosphere up to ~1000°C or treated with solvents then heated or reacted with oxygen - all with the objective of creating volatiles for analysis. The vacuum pump is critical in moving the gaseous samples around. The turbomolecular exhaust pump is 5 cm (2 inches) in diameter and runs at 100,000 rpm. It is capable of creating vacuums down to 10^{-6} Torr (10^{-4} Pascal)[7].



Curiosity Science Laboratory: a key aspect is the complex handling of both solids and gaseous materials both on the exterior or the rover and in the interior laboratory. ("SAM" stands for **S**ample **A**nalysis at **M**ars. "CheMin" stands for **Che**mistry and **Min**erology.)

---

[6] https://msl-scicorner.jpl.nasa.gov/Instruments/SAM/

[7] Definition: 1 torr = 1/760 of standard atmospheric pressure ~= 1 millimeter of mercury ~= 133.3 Pascals

The other key instrument to which powdered samples can be delivered is "ChemMin"[8]. In this instrument a collimated X-ray beam is directed at the sample. The regular arrays of atoms in any crystalline materials that may be present cause a small fraction of the X-rays to be deflected or 'diffracted' through certain angles depending on the interatomic spacings in the crystal. The resulting X-ray diffraction pattern is accumulated over several hours. Its characteristic pattern of circles allows identification of the crystalline components of the rock or dirt sample.

After this period of strenuous activity - creeping across the landscape with an occasional pause to collect and analyze a dirt or rock sample – the rover will stop and nap until evening-time when it's time to send its data to one of the orbiting Martian satellites to be relayed back to Earth. There are, in fact, a variety of redundant communications channels, including a low data-rate channel direct to Earth. However, the preferred, highest data-rate channels are via the orbiting relay satellites.

As we mentioned earlier, Curiosity differs from many of its predecessors in having no magnetic storage or memory or even a magnetometer. However, the rover employs a myriad of electric motors and actuators and valves and almost certainly these are all electromagnetic devices. Magnetism reigns supreme in this arena. It is true that similar devices can be built based on electrostatic principles (i.e. based on the attraction between charged surfaces). However, electrostatic devices have huge disadvantages in most applications. The lack of cheap high-permittivity materials (analogous to iron for electromagnetics), the requirement for high electrical voltages incompatible with battery and semiconductor technologies, and the limitations due to the risk of electrical breakdown and damage caused by current flow (there is nothing analogous to electrical current in a magnetic circuit) all count heavily against electrostatic devices in most applications.

Field-Effect Transistors

Curiosity does not include a tape-recorder or a hard disk drive and there is certainly no core-memory or plated-wire memory on board. What Curiosity does have instead is a large amount of several different types of radiation-hardened solid-state semiconductor memory. The dual-redundant computers each boast 256 kB of EEPROM, 256 MB of DRAM, and 2 GB

[8] https://msl-scicorner.jpl.nasa.gov/Instruments/CheMin/

of Flash memory[9]. In addition, each of the two MastCam cameras have a further 2 GB of flash memory capable of storing some 5500 raw color images. In this chapter, we will introduce the basics of solid-state memory including EEPROM and Flash memory (there will be more discussion of DRAM in the following chapter, E+10 meters, on the Kepler space-craft).

For our discussion, solid-state memory is synonymous with semiconductor memory is synonymous with silicon memory. There are exceptions, of course, but solid-state means no moving parts (no rotating disks) with access based on lithographically-defined lines and devices, and with semiconductor switching devices providing the access and storage mechanisms, and with Silicon as the preferred semiconductor material. We start the discussion with an explanation of the Field Effect Transistor (FET) or, more-fully, the Metal-Oxide-Silicon Field Effect Transistor (MOSFET) which is nowadays the basic building block of any solid state memory.

The FET is the solid-state analog of the vacuum tube or valve. In a vacuum tube, the flow of electrons from cathode to anode is controlled by the voltage on an intervening wire mesh or grid past which the electrons flow. The idea for a similar device but with the electrons flowing in a solid material rather than a vacuum had been around for some time. Such a device might avoid the waste of power and the reliability issues associated with the necessity of heating the cathode to 'boil off' the electrons. And with solid-state, maybe one could make much smaller devices. The first practical amplifying devices were created in 1947 by Bardeen, Brattain, and Shockley[10] at Bell Labs with the creation of germanium point-contact transistors and quickly followed by bipolar junction transistors. But it took another decade until the quality of semiconductor materials (Silicon quickly became the favored material) and the understanding of interfaces and surfaces had progressed sufficiently that working FET devices could be fabricated. This was first achieved by John Atalla and Dawon Kahng again at Bell Labs in 1959[11]. Since that time the FET has gradually come

---

[9] EEPROM = Electrically-Erasable Programmable Read-Only Memory
DRAM = Dynamic Random-Access Memory
Flash memory is a derivative of EEPROM with very high storage density (small cell size)

[10] "Inventing the transistor", Computer History Museum, Mountainview, California http://www.computerhistory.org/revolution/digital-logic/12/273

[11] http://www.computerhistory.org/siliconengine/metal-oxide-semiconductor-mos-transistor-demonstrated/

to replace bipolar transistors in almost all applications because of the simpler process and its ability to scale to very small dimensions.

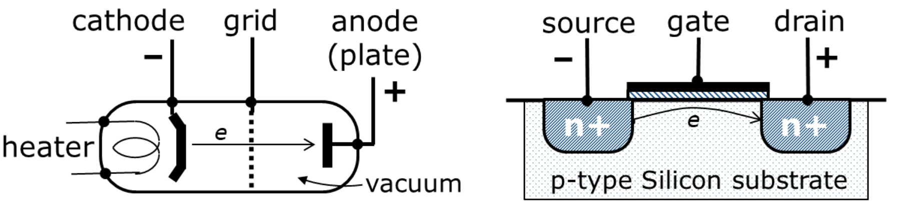
Silicon is a ‘semiconductor’. This means that it does conduct electricity, but rather poorly. In fact if the Silicon is very pure and in the form of a very high quality crystal, it will hardly conduct at all. Each silicon atom has four electrons in its outer shell that it shares with each of its four neighbors in the tetrahedral crystal structure. These shared electrons form strong localized covalent bonds and they cannot move around within the crystal. So there are no free electrons to conduct electricity. Pure crystalline Silicon at room temperature is effectively an insulator. However, if extra electrons can somehow be provided, these can move around relatively freely in response to electric fields. One way of introducing free electrons is simply to heat the silicon. The extra heat energy will disrupt a few of the tight covalent bonds and thus provide a source of free electrons that can wander around in the material. In this sense, pure crystalline silicon behaves much like the vacuum in a vacuum tube. The vacuum between two cold metal electrodes will completely prevent current flow, but if one electrode (cathode) is heated to ‘boil off’ some electrons, those electrons (being negative) will readily move through the vacuum towards the positive electrode (anode). Moving electrons constitute a flow of electric current, which, by an unfortunate convention, flows in a direction opposite to the electron movement.

Fortunately, with semiconductors, there are much easier ways of providing those extra electrons. Simply replacing a few of the Silicon atoms with Phosphorus atoms will do the job. Phosphorus has five electrons available for bonding. Four electrons will join in covalent bonds with the adjacent Silicon atoms, leaving the fifth electron relatively free to wander around. Replacing as few as one in a million of the Silicon atoms with Phosphorous atoms makes an enormous difference to the conductivity of Silicon. This process of deliberately introducing very dilute contaminant materials such as phosphorus is known as doping. Dopants like Phosphorus, Arsenic, and Antimony provide an extra electron while others like Boron or Aluminum have only three available electrons leaving one of the covalent bonds to the four adjacent Silicon atoms unsatisfied. The missing electron or ‘hole’ is similarly mobile and behaves much like a positively charged electron though not quite as mobile. Silicon doped to provide extra electrons is referred to as n-type while Silicon doped to yield missing electrons is referred to as p-type. Heavy doping (~1 in 10,000 atoms) is denoted by n+ or p+ and yields a highly conductive material.

The illustration below shows a simplified version of a Metal Oxide Silicon Field Effect Transistor (MOSFET) alongside a triode vacuum tube (valve). In both cases the current through the device is controlled by the voltage on an intermediate electrode placed very close to the path of the electrons. Either device can be used for amplifying analog signals, but generally in logic or memory circuits we are talking about binary operation, i.e. current is either flowing (ON) or not flowing (OFF). The FET shown is called an n-channel device but it is equally possible to build p-channel devices where the operation is based on holes rather than electrons. In fact CMOS logic involving both n-channel and p-channel devices in a complementary configuration is a favored form of logic in most applications (the ‘C’ in CMOS stands for complementary).

Illustrating the analogy between a triode vacuum tube and a field-effect transistor (FET). Both devices rely on electric fields influencing electron flow

Vacuum tube (triode) **Field-Effect Transistor (FET)**



In a vacuum tube, electrons will flow from cathode to anode. The flow can be controlled by the voltage on the grid.

In an n-channel FET, electrons will flow from source to drain. The flow can be controlled by the voltage on the gate

The diagram tries to illustrate the analogy between the vacuum tube and the field-effect transistor. The tube is turned off by applying a negative voltage to the grid which repels the electrons and prevents them reaching the anode. The FET operates a little differently. It is normally off and this is achieved by designing in a light p-type doping of the substrate assuring there are no free electrons around. The gate is made of metal separated from the Silicon by a thin layer of insulating silicon dioxide, SiO_2. Raising the voltage on the metal gate creates an electric potential that attracts electrons in a thin layer (channel) close to the surface of the Silicon providing a path for current flow. Generally the source and substrate in an FET are connected and kept at the same voltage. A positive voltage is applied to the drain and the doping levels are designed to make the gate threshold voltage (from non-conducting to conducting) roughly midway between the source and drain voltages (much more convenient than with a vacuum tube).

There is a big advantage to separating the gate from the substrate with the thin insulating layer of naturally-formed native oxide. This means that DC current cannot flow and the input gate resistance or impedance is extremely high. If a positive voltage is applied to the gate and then disconnected, the remaining positive charge on the gate can take a very long time to dissipate - a feature very useful for memory devices! In fact if the gate is isolated and not connected to any external circuitry and completely surrounded by high-quality insulating oxide, the charge can take many years or decades to dissipate. This is the virtue of the *floating gate*.

Flash Memory

Flash memory, taking advantage of the floating gate, was invented by Fujio Masuoka[12] while working for Toshiba around 1980. The name 'flash' apparently being suggested by Masuoka's colleague, Shōji Ariizumi, because the erasure process for resetting the memory contents reminded him of the flash of a camera. Flash memory differs from EEPROM mainly in memory addressability. The latter can be addressed down to individual bytes whereas flash memory can only be read and updated in large 'pages'. Flash memory employing a NAND[13] configuration of cells offers the highest capacity per die (chip). Starting around 2000, NAND-flash memory cards[14] and USB "thumb" drives[15] started becoming available. The flash memory cards became ubiquitous in digital cameras and cell-phones. Thumb drives quickly replaced removable storage devices such as floppy disks and CD-R and DVD-R for transferring and storing data between computers. In larger capacities Flash memory can be fitted with an electrical interface and mechanical form factor that mimics a hard disk drive (HDD). These "Solid State Drives" (SSD) have largely replaced disk drives in applications that favor low-power and mechanical robustness (e.g. laptop and notebook computers), or fast data access (e.g. transaction processing for Enterprise applications).

---

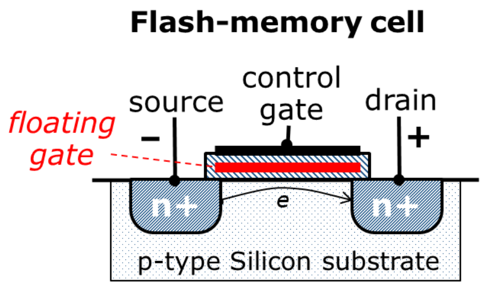
[12] https://www.forbes.com/global/2002/0624/030.html#4ca044db3da3

[13] NAND refers to the memory configuration. Multiple transistors are connected in series. The chain conducts (and output goes low) only if all gates are high

[14] Flash memory cards are also referred to as (Micro-) Secure Digital (SD) cards.

[15] Thumb Drive also referred to as USB flash drive, flash stick, jump drive, memory stick

The storage mechanism is provided by electrons trapped on a floating gate that is completely surrounded by a high-quality insulator. The stored charge can persist for many years or even decades. The charge can be modified by applying high voltages to force electrons to cross the thin insulating layers.



Flash-memory cell is an FET with an additional floating gate that can store charge (electrons) for very extended periods (years)

The key question, of course, was how to modify the charge on a floating gate to which, by definition, one cannot connect to any circuitry (because the charge will leak away). The diagram above shows a flash memory cell. Interposed between the control gate and the Silicon channel is a floating gate completely surrounded by SiO_2. The current through the channel responds to the voltage potential provided by the combined effect of the voltage on the control gate and the electron charge on the floating gate. The charge on the floating gate can thus be determined by checking what control gate voltage is required to start the channel conducting. So we have way of reading back how much charge is on the floating gate, but still no answer to the question of how to change that charge so we can actually store information.

The answer basically is *brute force*. If the voltage is high enough, the thin insulating oxide layer breaks down and electrons can flow to or from the floating gate. The expression 'breaks down' is an exaggeration and simplification. A huge effort goes into selecting materials, optimizing the cell geometry and fabrication process and the values and timing of the voltage sequence required to program the cell. The actual physics is described by the Fowler–Nordheim equations which combine field emission and quantum tunneling[16]. Field emission means the electric field is strong enough to free electrons from a surface and quantum tunneling means the barrier is thin enough that an electron has a finite probability of being found on the other side of the barrier. The voltages used to set or program a cell are an order of magnitude higher (~10 volts) than those used in reading (~1 volt).

[16] https://en.wikipedia.org/wiki/Field_electron_emission

For NAND flash memory (the most common form), raising the voltage on the control gate to 10 or 20 volts and keeping the substrate at zero serves to add electrons to the floating gate, while raising the substrate to 10 to 20 volts and keeping the control gate at zero similarly serves to remove electrons from the floating gate. In this manner, the electron charge on the floating gate can be set (programmed) thus storing information.

Unfortunately, the process is not perfect and the high voltages and energies involved in programming the cell can damage the insulating SiO_2 barrier layer. Atoms may shift position causing defects in the crystal structure through which electrons may leak away and electrons may get trapped in the oxide layer which shifts the switching threshold. This gradual degradation in characteristics is referred to as ‘wear’. The number of programming cycles that can be endured ranges from just a few hundred cycles to many hundreds of thousands of cycles depending on the type of flash memory. We will talk more a about multi-level flash and 3D flash in the coming chapters.

Flash memory is also susceptible to radiation damage in much the same way as from the effects of wear. High energy ions or electrons can change the contents (charge) of a memory cell or, worse still, can cause permanent physical damage to the oxide layers causing charge leakage. Curiosity has a total of over 6 GB of flash memory which will gradually degrade with time due to both wear and radiation damage. There have been some issues with the flash memory on the main computer requiring switching temporarily to the backup computer and memory. However as of July 2018, the flash memory is still operational.

This chapter is meant to represent the scale of 100,000,000 km, yet we are talking about flash memory cells with dimensions of a few tens of nanometers. In the final chapters of the book, we will come back to this topic of how such tiny features are created both in the context of hard disk drives and in the context of solid state memory.

The next chapter takes us down another factor of ten to the 10,000,000 km (~7 million miles) scale. This is much more than the distance to the moon yet much less than the distance to Mars. It is not so easy to find an object at this distance from Earth.

Further Reading

"Curiosities Revolutionary Experiments" Dec 14, 2012 https://www.youtube.com/watch?v=3kHBPc943sg

Rob Manning and William Simon, "Mars Rover Curiosity", Smithsonian Books, Washington, DC, 2014

Rod Pyle, "Curiosity", Promethius Books, Amherst, New York, 2014

http://www.computerhistory.org/siliconengine/metal-oxide-semiconductor-mos-transistor-demonstrated/

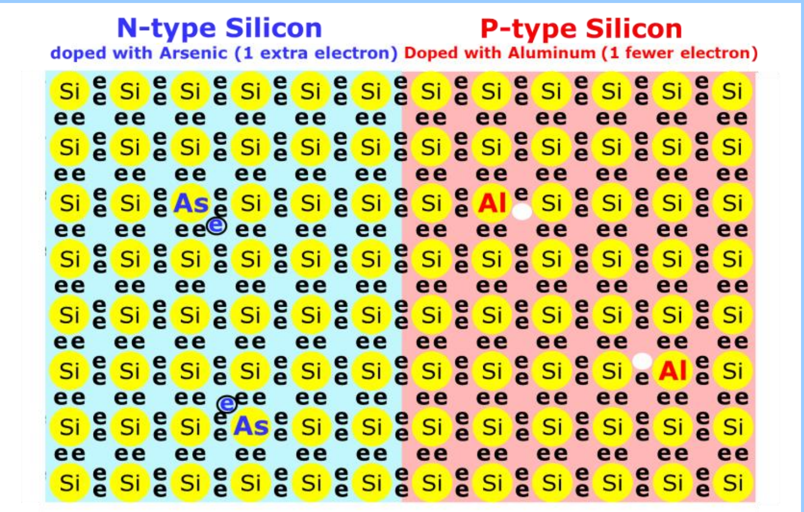
Chris Woodford, "Flash memory", June 18, 2018. https://www.explainthatstuff.com/flashmemory.html#null

Chih-Tang Sag, "Evolution of the MOS Transistor — From Conception to VLSI", Proc. IEEE, Vol. 76, No. 10, October 1988, pp. 1280-1326.

K. Choi, "NAND Flash memory", (Samsung), May 7th, 2010, https://www.ece.umd.edu/~blj/CS-590.26/nand-presentation-2010.pdf

Semiconductors

Silicon is an example of a semiconductor. Very pure crystalline silicon is not a good conductor and has high resistance to current flow. However, if impurities are present or at elevated temperatures, silicon behaves more like a metal and conducts electricity well. Silicon has four electrons in its outer shell of electrons. In a silicon crystal, each atom has exactly four neighbors with which it happily shares electrons in covalent bonds so that each atom has effectively eight electrons (eight is a very stable configuration, as in the inert gasses). With all the electrons in this stable configuration, none are free to move and the electrical resistance of the silicon crystal is very high. However, at elevated temperatures thermal energy can be sufficient to break some of the bonds and the free electrons can conduct electricity causing the resistance to drop. Another way of achieving good conduction is to introduce impurities (doping) - see below.



Semiconductor P-N junction forming a diode that will conduct electricity only from P to N (electrons flow in opposite direction).
Note: silicon actually forms a tetrahedral lattice in 3D, not the 2D square lattice shown

Silicon doped with just a few arsenic (As) atoms (5 available electrons) will provide extra electrons that are relatively free to move through the crystal (n-type). Silicon doped with just a few aluminum (Al) atoms (3 available electrons) will provide ‘missing electrons’ or ‘holes’ that are also relatively free to move through the crystal (p-type). If a p-n junction is voltage biased so that electrons and holes flow towards the junction and are able cancel each other, then electrical current will flow. Current will not flow in the reverse direction. A semi-conductor p-n junction forms a diode. This is a basic electrical circuit element.