## Chapter +10: Kepler



Johannes Kepler ${ }^{1}$ was the person who finally got it right! It is true that Nicolaus Copernicus set the scene by putting the Sun at the center and it is true that Isaac Newton finally introduced gravity to explain why it all worked. But it was actually Kepler who created the first simple accurate model of the solar system with all the correct shapes and timings for the planetary orbits.

The history of science is very much the history of our gradual enlightenment and understanding of the heavens. In some ways the planetary motions represent an ideal laboratory. Friction or damping is negligible and the interaction between planets is tiny. The paths of the planets in space are almost perfect and almost unchanging, hence the expression "the clockwork of the heavens". Progress in understanding often came from some really enormous leaps in perspective - one might almost say leaps of faith. Sometimes the progress was steady, but often key insights and concepts were lost for centuries before being rediscovered.

[^0]The ancients gazing up at the night sky could only wonder. They could see there were a bunch of tiny bright jewel-like white dots fixed on a black background that swept across the sky from east to west each night. Every night the entire pattern shifted slightly towards the west such that after one year the whole cycle repeated. Apart some transient events (meteors and comets), all the dots were fixed except, strangely, for just five that wandered slowly around the night sky in complex paths. The Greeks called them 'planetes' meaning wanderers. There were also the motions of the Sun and Moon to worry about, but their motions were much more regular and easier to predict.

The starting assumption was, of course, that the Earth is flat. The Sun and Moon also look like flat disks [the Moon is tidally locked and always presents the same face to us. The Sun appears featureless, so it too looks like a flat disk]. It is also very obvious to everyone that the Earth is fixed and stationary, otherwise there would be a lot of wind and it would be a bumpy ride and we'd probably fall off. (Had the Moon looked like a rotating sphere, things might have been quite different.) Anyway the Moon and Sun appear as just flat disks way high up in the sky. Similarly the five planets and the fixed stars are also a very long way away (no observable parallax). But since the Sun and Moon and the five planets are continually moving around, there must be some huge invisible supernatural beings up there to keep on pushing them, otherwise, obviously, they'd quickly come to a stop. To use the Roman names for these gods ${ }^{2}$ :

| Sol | - the Sun god (Greek: Helios) |
| :--- | :--- |
| Luna | - the Moon god (Greek: Selene) |
| Jupiter | - the king of the gods (also Jove) |
| Mars | - the god of war |
| Venus | - the goddess of love |
| Saturn | - the god of wealth and plenty |
| Mercury | - messenger of the gods |

Perhaps the first great leap of perspective was the realization by the ancient Greeks, Pythagoras and Aristotle among them, that the Earth was not flat but spherical. This realization came from several pieces of evidence. The further south one travels, the higher the Sun and Moon move in the sky but they don't get bigger in size. Also the pole star dips lower and lower towards the northern horizon but every 100 leagues (100

[^1]hours walking) changes the angle by exactly the same amount. So it couldn't be just due the change in position of an observer looking at a relatively nearby object. There were other clues too. Anxious merchants carefully scanning the horizon for their cargo ships to arrive, claimed to see the mast and sails long before the ship itself was visible. Also, during a lunar eclipse, the Earth's shadow on the moon appears as an arc of a much larger circle. With this leap of perspective to a spherical Earth, came the further realization that 'down' is not always the same direction but always points towards the center of the Earth. The sphere is a perfect object, so this is all very satisfactory and the center of the sphere must therefore be the center of the universe. The spherical Earth must then be centered inside a very large concentric sphere that rotates once a day. The inside of this outer 'celestial' sphere is painted black and sprinkled with tiny bright white spots or jewels for stars. All this is obvious.

The ancient Greeks were also able to estimate the distance to the Moon by making observations from various places on the Earth of the Moon's position against the background of stars. This effect is known as parallax. No parallax was observed for any other celestial objects, so the Moon must be closer than the other objects. So it must be pasted onto a separate intermediate transparent or 'crystal' sphere that has its own separate motion. Since the Sun and planets also have independent motions, they must also lie on similar crystal spheres. Next came the problem of explaining and predicting the motions of Sun and Moon and planets. The planets, in particular, are tricky. For example, Mars generally drifts eastward against the background stars but occasionally reverses direction and drifts west. It also has significant deviations north and south. This complex motion was modeled by assigning several smaller spheres to each uniform rate and with a certain position and orientation of axis. In this manner the complex 'epicycles' (cycles within cycles) of planetary motion could be explained. This was already the situation almost 2,000 years ago when Claudius Ptolemy ${ }^{3}$ in Alexandria, Egypt, wrote his famous treatise on planetary motion known as the Almagest. This document even contained convenient tables presenting the predicted positions of the planets. Copies of the original Almagest written in the original Greek survive to this day.

This so-called Ptolemaic model of nested crystal spheres held sway with various refinements until the $17^{\text {th }}$ century. Having a stationary spherical Earth at the center with everything else revolving around it seemed

[^2]perfectly natural. Everything in the system was mechanically supported and connected, albeit with transparent invisible spheres made of ethereal material. But best of all, it gave quite accurate predictions of planetary motions. In the $16^{\text {th }}$ century, a competing scheme with the Sun at the center (heliocentric) had been published by Nicolaus Copernicus ${ }^{4}$ a Prussian mathematician and astronomer. Copernicus was not the first to come up a heliocentric system, but he laid out the case very clearly in a book "De revolutionibus orbium coelestium" (On the Revolutions of the Heavenly Spheres) that deliberately paralleled Ptolomy's famous Almagest in structure and argument. The book was widely published and discussed, but still Ptolemy's original geocentric view prevailed. It was patently obvious that the Earth was the center of all gravitation and that it was not spinning wildly like a top or hurtling around the Sun at breakneck speed. Also, the dimensions of the planetary orbits in the Copernican system seemed quite unreasonably large. And, anyway, the Copernican model still required small messy epicycles to make small corrections in predicting the planetary motions.

## Kepler's Three Laws

The controversy was not easily settled. The Copernican heliocentric model relying on circular orbits and small epicycles was still quite messy and was certainly counterintuitive. However Johannes Kepler ${ }^{5}$, a German mathematician and astronomer/ astrologer, particularly liked the Copernican perspective and believed he could put together a much more aesthetic and palatable version to explain the workings of the heavens. He was determined to prove that the structure of the universe was in fact heliocentric and that the sizes of the six known planetary orbits (including Earth) were simply linked by spheres inscribed and escribed around the five perfect Platonic (regular)
 solid shapes. The tetrahedron, octahedron, cube, dodecahedron, and icosahedron stacked one inside the other like Russian dolls. In 1596, he

[^3]published his ideas with the title "Mysterium Cosmographicum (The Cosmographic Mystery)".
Unfortunately, the real universe was not so obliging or perfect. In 1600, Kepler met in Prague with the renowned Danish astronomer, Tycho Brahe ${ }^{6}$. Kepler hoped Tycho Brahe's very accurate astronomical observations would confirm the structure described in Mysterium Cosmographicum. But Mars, more than any other planet, was a problem. It did not behave as if it were on a circular orbit. Kepler tried an ovoid (egg-shaped) orbit with the Sun near the pointy-end and that worked a little better. Finally, counterintuitively, he found that a perfectly elliptical orbit worked even better. But it only worked if the Sun was placed asymmetrically in the ellipse, at one of the two foci (Kepler's first law). This still left the question of what rate Mars should progress around its elliptical orbit. It seemed to move faster when it was closer to the Sun and slower when it was further away from the Sun (leading to Kepler's second law) ${ }^{7}$. In 1609, Kepler published "Astronomia Nova" (New Astronomy), explaining these first two laws.
The following year, in 1610, Galileo Galilei ${ }^{8}$ an Italian astronomer and professor of mathematics at University of Padua made two astounding observations that shifted the balance dramatically towards the heliocentric view. Galileo had improved upon the newly invented telescope and when he trained it on the planet, Jupiter, he noted 'stars' (actually moons) apparently orbiting around Jupiter. Later in the same year, he was able to see that Venus exhibited phases like the Moon. The first observation was evidence that not everything had to revolve around the Earth. The second observation was more damning and in direct contradiction of the Ptolemaic model which required that the crystal spheres of Venus and the Sun not intersect or overlap and that Venus therefore could never exhibit both crescent and gibbous ${ }^{9}$ phases (which it clearly did).

Buoyed by Galileo's observations and supported by Tycho Brahe's very precise observations on the remaining planets (whose orbits are closer to circular), Kepler took the final step linking the orbital periods of the planets to the size of their orbits (Kepler's third law). Again, counterintuitively, the period was determined only by the length of the major axis of the ellipse and was independent of the width of the minor

[^4]axis or the placement of the Sun within the ellipse. Kepler's treatise "Harmonices Mundi" (the Harmony of the World) was finally published in 1619 containing the three revolutionary laws governing the motion of planets (see diagram Kepler's three laws).


Kepler's three laws (shown for two planetary orbits):

1) The orbits are elliptical with the Sun at one of the two foci. (focus\#1 and \#2 for planet\#1 and focus \#1 and \#3 for planet\#2)
2) The planet sweeps out equal areas in equal times with respect to the sun (Area1 = Area2)
3) The square of the orbital period is proportional to the cube of the major axis, i.e. $(T 1 / T 2)^{2}=(L 1 / L 2)^{3}$

The magnitude of this achievement cannot be overstated. The task was to take observations of two-dimensional planetary positions referenced to the fixed background stars and referenced to the timing of celestial events and then to translate these into real motions in a fully three-dimensional space. This task is daunting enough even with modern computational tools. All this work predates any mechanical calculators (ignoring the abacus) or slide-rules or logarithms. It was all done with hand calculations and the tools of geometry (trigonometric tables did exist and actually first appeared in crude form in Ptolemy's Almagest).

So finally, after some two thousand years, Kepler had arrived at a simple mathematical and geometric model of the Solar system with very accurate predictive powers. Each planetary orbit was described by just six parameters describing the geometry (major and minor axes) and orientation (three angles) and timing (initial reference position) of the
orbit ${ }^{10}$. Unfortunately, it had lost, in the process, some of the beauty that Kepler and others had hoped to discover. For example, there were no simple numerical relationships between the orbits and no role could be found for the beloved Platonic solids. Furthermore, the mechanical basis had disappeared, so what was there to keep the planets on track and provided the motive power to drive them on their endless cycles?
It was left, of course, for Sir Isaac Newton ${ }^{11}$ at Cambridge University in England almost 70 years later in his "Principia" in 1687 to lay out the laws of motion and gravitation that finally explained the underlying mechanics and physics of the orbits. That final leap of recognition was embodied in Newton's famous equation $\mathrm{F}=\mathrm{G} \cdot \mathrm{M}_{1} \cdot \mathrm{M}_{2} / \mathrm{r}^{2}$. This equation contains both the concept that a spherically-symmetric gravitational field emanates from each celestial object in proportion to its mass as well as the idea that fieldstrength decayed with the inverse square of distance, $r$, from the object. In particular, the Earth was demoted from its position as the source of all gravitation. In one stroke, Newton had provided the most simple and ultimately most beautiful explanation of the truth in Johannes Kepler's three laws of planetary motion (ignoring some tiny effects from Einstein's relativity).

## The Lagrange points

This chapter, the 10 million kilometer chapter, is actually quite a challenge to fill. In general, for each chapter, we're happy if we can find some size 'to the nearest order-of-magnitude'. So, strictly, for this chapter, it means something interesting between the $10 / \sqrt{ } 10$ and $10 \times \sqrt{ } 10$ million km scale. That is between 3.162 and 31.62 million km (approximately 2 to 20 million miles). For the previous chapter with a scale of 100 million km , Mars was an easy choice. Its distance from Earth varies between about 54 and 400 million km. Venus is quite a bit closer. Its distance from Earth varies from about 38 to 260 million km but even its closest approach is outside our liberal definition for this 10 million km chapter. The next nearest significant natural object is the Moon, but that is a mere 0.38 million km away, much smaller than our desired length-scale

[^5]However, there are also some invisible natural features associated with the Sun-Earth system. These are the so-called Lagrange points ${ }^{12}$. At these points in space, an object will orbit in exact synchrony with an identical orbital period to that of the Earth. There a five such points, the first three, $\mathrm{L}_{1}, \mathrm{~L}_{2}, \mathrm{~L}_{3}$, were recognized by the Swiss mathematician, Euler, and the last two, $\mathrm{L}_{4}, \mathrm{~L}_{5}$, were found by the Italian mathematician, Lagrange. The last two are of special interest in planetary science because they correspond to stable equilibria. Asteroids, rocks, dust, and other debris are attracted to these points and tend to collect in these locations. The $\mathrm{L}_{4}$ and $\mathrm{L}_{5}$, points for the massive Sun-Jupiter system have a collection of wellknown asteroids called the Trojans and similar examples are common throughout the solar system including within individual planetary systems.


Although the Lagrange points are precisely defined in space, an object orbiting in the vicinity of an $\mathrm{L}_{4}$ or $\mathrm{L}_{5}$ Lagrange point is only very weakly bound to that point and typically traces a large complex trajectory in three

[^6]dimensions around the Lagrange point (relative to the rotating frame of reference). "Large" here can mean more than half the separation of the principal bodies! Earth itself has a 300 -meter-diameter asteroid called 2010 TK7 ${ }^{13}$ that oscillates with a 400 year period around the leading $\mathrm{L}_{4}$ point in a rather-chaotic extremely-elongated tadpole-shaped trajectory. At its extremes, the trajectory takes it almost to the opposite side of the Sun and then, at its nearest, to within 20 million km of the Earth. So finally we seem to have identified an interesting object that approaches to within the 3 to 30 million km range we are looking for. However, its proximity to Earth is transitory and there are many other near-earth asteroids that make similar approaches to Earth. Nor are there any plans to send a spacecraft to explore 2010 TK7. It would also be quite difficult to find much to say about it. So 2010 TK7 is not, after all, a particularly good object to focus on for this chapter

The other three Lagrange points correspond to unstable equibria, so objects will tend to drift away from these points. Nevertheless the L1 and L2 points are of great interest for space probes and telescopes because they don't collect debris and because very little action (energy) is required for station-keeping. And, of course, it is very convenient that L1 and L2 are in fixed positions with respect to Earth and also positioned relatively close to Earth for ease of access and communication. Several spacecraft have used, or are using, or are planning to use these points. The L1 point is good for observing the Sun and also for observing the daylight side of the Earth. The L2 point is better for deep space telescopes since the Sun, Earth, and Moon are always positioned in a narrow region of the sky that can be avoided for most observations.
The giant James Webb space telescope ${ }^{14}$ to be launched in 2021 (greatly delayed) is destined for the L2 Lagrange point. James Webb was the scientist and visionary who led the fledgling NASA organization (and the Apollo lunar landing program) from 1961-68. The eponymous telescope is a monster. The primary mirror is 6.5 meters (about 21 feet) in diameter (Hubble's is a mere 2.4 meters or 7 feet). The James Webb design emphasizes the detection of infra-red light. For this reason, the mirror and critical detection components must be kept very cold. A very-large multilayer sunshade is to be deployed to continuously block the Sun's radiation and allow the telescope to settle to a temperature below 50 Kelvin ( -220 ${ }^{\circ} \mathrm{C}$ or $-370^{\circ} \mathrm{F}$ ). The telescope's mission is to probe to the very edge of the

[^7]visible universe. Light from objects so distant is weak and also very heavily red-shifted because of the recession-velocity associated with the expanding universe. This explains the giant mirror and the emphasis on detecting infrared radiation.
Unfortunately, having said all this, the L1 and L2 points are only 1.5 million kilometers ( 1 million miles) from Earth, a distance too short to qualify for this chapter. At the other extreme, the L4 and L5 points are roughly 150 million kilometers ( 93 million miles) from Earth, too far away to qualify for this chapter. The L3 point is even further away at about 300 million km ( 186 million miles). The L3 point is the ideal position for the "Counter Earth ${ }^{15 "}$ of science fiction novels since it is lies exactly opposite the Earth in its orbit and is always invisible being permanently blocked from us by the Sun. Perhaps disappointingly, numerous space-probes have shown that no such object exists.

## Detecting Exoplanets

The scientific community quickly accepted the heliocentric model. Then from various attempts to measure the vast distances to the fixed stars and compute absolute brightnesses and also from comparisons of the visible light spectra (color) from the Sun and from the stars, it gradually became accepted that the stars were in fact very distant suns in their own right. Furthermore, from detailed examination of brightness and color, our Sun turned out to be fairly typical. The question then immediately arose as to whether theses distant suns also had planetary systems orbiting around them? In other words - do exoplanets exist? Certainly, many binary systems could be observed with two stars orbiting each other. There were also examples where one or more small stars were in orbit around a much larger star. So it seemed quite likely that star systems with orbiting planets were in fact common - but how to detect them and could any of them be Earth-like?

Venus shines very brightly in our evening or morning sky at an apparent brightness or "magnitude" or around -4.5 . Venus is similar to Earth in size but has about twice the reflectivity (albedo) because of its full cloud cover. So, if one could peer through the dense Venusian clouds and see Earth from Venus, it would appear only half as bright as Venus from Earth. Since stellar magnitude is defined as minus $2.5 \log _{10}\left(\mathrm{~F} / \mathrm{F}_{\mathrm{Vega}}\right)$, where the brightness of the star Vega, $\mathrm{F}_{\text {Vega }}$, provides the reference point. (note the minus sign: large positive numbers mean very faint objects). Earth

[^8]viewed from Venus would be about magnitude $-4.5+-2.5 \log _{10}(1 / 2)=-$ 3.7, still a very bright object. But Venus and Earth are typically separated by about 1 A.U. ( 150 million km or 93 million miles) which is very close compared to stellar distances. The closest star, Alpha Proxima, is about $270,000 \mathrm{AU}$ ( 4.25 light years, 40,000 billion $\mathrm{km}, 25,000$ billion miles) away, so the light would be $(270,000)^{2}=72$ billion times dimmer. A factor of 72 billion means a change in magnitude of $2.5 \log _{10}(72,000,000,000)=27.1$. If Alpha Proxima were similar to the Sun, an earthlike planet would shine at Magnitude $-3.7+27.1=23.4$. Large modern telescopes can observe down to about magnitude 30. So, based on brightness, we should be able to directly observe Earth-like planets around stars in our vicinity.

The next question is whether a telescope can resolve two closely spaced objects such as a star and its planet. A separation of 1 A.U. at a distance of 270,000 A.U. is $1 / 270,000 \sim=4$ microradians (or 0.8 arc-seconds). For light at wavelength, $\sim 0.5$ micrometers, we would need a primary mirror of very roughly $0.5 \mu \mathrm{~m} * 270,000=14 \mathrm{~cm}$ ( 6 inches) in diameter. So we can apparently distinguish the two such objects with a very modest telescope. Unfortunately these crude measures of resolution assume we are trying to distinguish two object of similar brightness. In fact our Sun is some 2 billion times brighter than the Earth and, in practice, 'glare' caused by diffraction and scattering in the telescope optics make it impossible to do direct detection of an Earth-like exoplanet orbiting a Sun-like star.

However, there are a number of other indirect and perhaps somewhat surprising techniques that are being used very successfully to detect exoplanets. The first group of techniques depend on observations of the parent star alone. These are based on the knowledge that the star and planet revolve around their common center of mass. Because of its much greater mass, the motion of the star is very much smaller than the motion of the planet. Nevertheless, if such motions can be detected, much can be deduced about the planet's mass and orbital period including cases where there are multiple planets. There are two methods of detecting such motions. Both require measurements extending over years. The first method is simply to observe the position of the parent star with respect to distant background stars. This of course requires extreme precision and is limited to very massive planets orbiting far from their parent star. The second method is to look for Doppler shifts in the star's spectrum. These shifts are indicative of the star's motion along the observer's line of sight.


The Doppler method was responsible for most of the early successes in detecting exoplanets. It is perhaps surprising that motions along the line of sight can be detected more easily than motions across the line of sight. This a result of both the extreme accuracy with which time and frequency can nowadays be measured and the very narrow absorption line-widths observed in stars as light passes though the rarified atmospheres of the star. The "High Accuracy Radial velocity Planet Searcher" or HARPS spectrograph at the La Silla Observatory in Chile can detect stellar velocity changes as low as 0.3 meter/second. Detracting factors include the star's own rotation which blurs the spectral lines due to the spread in surface velocities between the advancing and retreating limbs of the rotating star.

As a reference point, the Earth alone causes the Sun to shift by only $\pm 450$ km in its annual journey and produces a maximum velocity of $0.09 \mathrm{~m} / \mathrm{s}$. The resulting Doppler-shift is taken with respect to the velocity of light and is $(0.09 \mathrm{~m} / \mathrm{s} / 300,000,000 \mathrm{~m} / \mathrm{s})=0.3$ parts per billion. In contrast, Jupiter is much more massive and much further away from the Sun and produces a huge shift of $\pm 742,000 \mathrm{~km}$. This means that Jupiter and the Sun orbit around a common center which is actually slightly outside the surface of the Sun. The maximum velocity induced in the Sun by Jupiter in its 12 -year cycle is $12.4 \mathrm{~m} / \mathrm{s}$ ( 41 parts per billion), well into the range
that can be detected and measured. Many exoplanets have been discovered using Doppler spectroscopy, but these tend to be massive, close-in planets orbiting smaller older stars. The velocity perturbations are larger if the parent star is small. Older stars tend to have slower rotation rates and less Doppler blurring. Current techniques (2018) can only detect Earth-like planets in the habitable-zone if they are associated with small stars.

In August 2016, the La Silla Observatory revealed that our closest neighboring star, a red-dwarf called Proxima Centauri, has an exoplanet orbiting within the habitable zone. The HARPS spectrograph had picked up a Doppler shift of about $2 \mathrm{~m} / \mathrm{s}$ with a period of 11 days. From the period plus an estimate of the star's mass (about $1 / 8^{\text {th }}$ the Sun's mass), we can deduce the distance of the planet from the star ( 0.05 AU or 7.5 million km ) and then from the velocity shift, we can deduce the mass of the planet (about 1.3 times Earth's mass). Unfortunately, 'habitable-zone' simply means that it's in the right temperature range for liquid water to exist. For a variety of other reasons the planet seems an unlikely home for life. The planet orbits very close-in to the parent red-dwarf star. It is almost certainly gravitationally locked (one side always faces its sun). Red-dwarf stars also have the unpleasant characteristic of producing copious X-rays and producing them in large bursts.

A second group of techniques depend on observations of the parent star and the planet combined together. The most sensitive and successful of these techniques is the transit method. We are familiar with the transit of Venus across the Sun. This occasional phenomenon (the last one was in 2012 and the next one not until the year 2117) is visible to the naked eye (with a suitably heavy filter) as a tiny black dot passing across the Sun's disk. What is not so noticeable is the tiny drop in total sunlight reaching the Earth. From afar, the transit of Earth across the face of the Sun would be very similar to the transit of Venus. From a very long way away (lightyears), the black dot would not be visible but the tiny drop in Solar brightness might still be measurable. The Earth in transit produces a drop in solar illumination of (diameter of Earth / diameter of Sun) $)^{2}=(12,742$ $\mathrm{km} / 1,391,400 \mathrm{~km})^{2}=84$ parts per million with a characteristic maximum transit time of (diameter of Sun / velocity of Earth) $=(1,391,400 \mathrm{~km} / 30$ $\mathrm{km} / \mathrm{s}$ ) $=13$ hours every 12 months. Measurements obviously require extremely high-precision photometry down to a few parts per million and with very high-stability maintained over many years of observation. This kind of precision can in fact be attained and also with sufficient signal-tonoise to allow observations on stars thousands of light-years away. This means millions of stars - which is just as well because the chance of the
planet's orbital plane being oriented to allow a transit to be seen is roughly (diameter of the star / diameter of the orbit) $=(1,391,400 \mathrm{~km} \mathrm{/}$ $300,000,000 \mathrm{~km})=0.46 \%$, for a system like Earth-Sun. So on the positive side, the transit method is extremely powerful; on the negative side it completely misses the vast majority of planets!

## The Kepler Space Telescope

It is the Kepler space telescope that we have chosen as the title and topic for this chapter ${ }^{16}$. Kepler was launched specifically with the task of detecting exoplanets. It was launched on March $7{ }^{\text {th }} 2009$ in an 'Earthtrailing' orbit with a period of 372.5 days. This is just over seven days longer than the Earth's orbit of 365.25 days. As a result, every year Kepler lags behind Earth by an additional seven days or roughly seven degrees ( $\sim 0.125$ radians) which is approximately $0.125 \times 150,000,000=19$ million kilometers ( 12 million miles). This is the sort of scale we're looking for. For the first 18-months or so, Kepler was within our target 3 to 30 million km range from Earth. Since that time it has drifted far behind the Earth's position and passed through the L5 capture point. As of 2018, it is at a distance of 150 million km from Earth. Around 2060, the Earth will come sneaking up on Kepler from behind and shift Kepler into a lower faster orbit.

Kepler's task is to stare at a very large number of stars (about 150,000) in a particular region of the sky for a very long time to see if any of them 'blink ${ }^{\text {'17 }}$. That means do any of them show these tiny but regular diminutions in brightness that are indicative of an exoplanet. Kepler's primary mirror is 1.4 m ( 55 inch ) - quite modest compared to Hubble's 2.4 m mirror ( 94 inch ). However, Kepler's field-of-view is enormous (115 degrees-squared or 0.035 steradians ${ }^{18}$ ), corresponding to about a 12 -degree diameter view. This is similar to a soccer-ball held one meter away. Even so, this is still only $0.28 \%$ of the entire heavens. The telescope optics form a huge image on a giant $300 \mathrm{~mm} \times 300 \mathrm{~mm}$ ( 1 foot square) 95 MegaPixel CCD array ${ }^{19}$. (Compare with a cell-phone camera sensor that might be as small as 3 mm by 5 mm .) Since Kepler is interested in very accurate photometry rather than high resolution, the image is defocused very

[^9]slightly so that each star illuminates about seven adjacent pixels. To reduce noise, the CCD array is cooled by a large external radiator structure facing away from the Sun. The CCD array is designed to detect tiny changes in stellar flux (10-40 parts per million) over transit periods of 2 to 16 hours. Data from the CCD array is heavily compressed for storage in the 16 GigaByte local memory [more about both the CCD array and about Kepler's memory later]. Once a month, the spacecraft is re-oriented so that the fixed 0.8 m diameter high-gain antenna points towards Earth. This monthly memory download takes $8+$ hours at $4.3 \mathrm{Mb} / \mathrm{s}$ via a 32 GHz link. Once every three months, the spacecraft is rotated by 90 degrees to keep the solar panels pointing towards the Sun and keep the telescope, CCD array, and cooling radiator pointed away from the Sun.


The Kepler Space Telescope (courtesy NASA)
https://keplerscience.arc.nasa.gov/the-kepler-space-telescope.html https://www.nasa.gov/mission pages/kepler/news/keplerm-20132901.html

Kepler has not been without problems. The original plan was for a 3.5year program, but the measurements proved less accurate than desired and the natural variability of the stars greater than expected. For this reason, the mission was planned to be extended to six years. A critical aspect of the project is to keep the telescope pointing very accurately in the desired direction for very long periods. The main disturbances arise from the solar pressure (mainly photons) acting on the structure of the spacecraft. The disturbances in attitude (pointing) are corrected by reaction wheels. The
principle at play here is the conservation of angular momentum. If a reaction wheel is rotated clockwise, the spacecraft will rotate slightly counter-clockwise. Because the reaction wheels are so much smaller than the spacecraft itself, it takes many thousands or revolutions of a reaction wheel to make noticeable changes in the direction the spacecraft is pointing. However, the use of reaction wheels allows for especially fine and accurate control of the spacecraft attitude. However, because of the asymmetric structure of the spacecraft, the solar pressure can generate a net persistent torque. To counteract this, the reaction wheels have to generate a counteracting torque which they do by continuously increasing their rotation rate. So, every few days, small hydrazine thrusters must be fired to absorb the acquired angular momentum and bring the reaction wheels back down to a reasonable rpm.

For complete control of the spacecraft attitude (pointing), a minimum of three reaction wheels are required, corresponding roughly to the familiar pitch, yaw, and roll on an aircraft. Kepler has four reaction wheels so placed that it can continue to operate even if one fails. The first reaction wheel failure occurred three years into the mission. Ten months later, a second failure occurred. The failure of two reaction wheels was viewed as a "show-stopper". But once again some ingenuity from the ground crew millions of miles away allowed valuable work to continue. Kepler still had effective yaw and pitch control but the ability to use the reaction wheels to control roll (about the telescope axis) was gone. Noting that Kepler is almost symmetric about a plane passing through the telescope axis and the line of symmetry of the solar-cell array, the solution was to orient the telescope such that the Sun always lay in this plane of symmetry. In this way, the solar pressure is symmetric with respect to the roll axis greatly minimizing the torque disturbances on the roll axis. In this way, the two remaining reaction wheels together with occasional tiny bursts from the thrusters to correct roll were able to provide adequate pointing accuracy. The photometric accuracy in this new mode of operation is approaching that for the system with three working reaction wheels. The down-side is that the spacecraft is now lying with its telescope axis in the plane of the Earth's orbit (the ecliptic) and every 90 days or so it must be re-oriented to keep it pointing well away from the Sun's glare. In this mode, only single transit events can be detected and likely candidates must be confirmed with ground-based studies - especially with Doppler spectroscopy. As of July 2018, Kepler had discovered 2,650 confirmed exoplanets ${ }^{20}$. The

[^10]mission was extended to the point when the thrusters' hydrazine fuel will be exhausted in about 2018.

## Charge Coupled Device (CCD)

At the heart of the Kepler telescope is the giant 95-Megapixel CCD array upon which the star-field is imaged. There are 21 pairs of $2200 \times 1024$ pixel CCD wafers covering a total of 116 square degrees ( 0.035 steradians) field-of-view (As of July 2018, three pairs of the CCD wafers had failed). The CCDs respond to wavelengths from 400 nm to 900 nm , covering the visible range and extending slightly further into infra-red. The entire array is kept at a carefully controlled temperature of $-85^{\circ} \mathrm{C}$ with cooling by a radiator mounted on the side of Kepler that is always shaded from the Sun. The CCD array are designed to detect tiny changes in stellar flux (10-40 parts per million) over transit periods of 2 to 16 hours.


CCD Device illustrating charge movement with four-phase clocking
Surprisingly, Charge-Coupled-Devices (CCDs) were originally envisioned as data storage devices, but their primary application now is in electronic imaging. In a CCD, a long series of cells are chained together and operate
on the 'bucket-brigade' principle (yes, for putting out fires). In this case, the buckets (cells) are small capacitors that contain electrons rather than water. Upon command (a clock-cycle), every cell delivers its contents to its neighbor to one side ${ }^{21}$. At the same time, a new quantity of electrons can be fed into the first cell in the chain and correspondingly the electrons that get pushed out of the last cell can be amplified and utilized. The transfer efficiency at each step can be very close to perfect (e.g. $99.9999 \%$ ). A three or four-phase clock is required in order to keep the efficiency high and to define the direction of propagation ${ }^{22}$. A chain of such devices can be operated like a long shift-register. Binary data fed into one end will eventually appear at the other end after a long sequence of clock-pulses. As a data storage device, CCDs have the advantage of very simple architecture, but they are not random access and it may take many clock-cycles to reach the wanted data.

Semiconductors are sensitive to light. Visible photons have enough energy to create electron-hole pairs in the silicon. Normally these recombine quickly. However, if a lightly-doped junction between n-type (on top) and p-type silicon (below) is formed and a reverse-bias voltage applied, the electric field can be sufficient to separate the electrons and holes. The electrons are drawn towards the surface and the holes pushed away from the surface. The electrons near the surface can be collected under the CCD electrodes and used as a measure of the light-intensity that has struck that cell. The array is illuminated from the backside (opposite side to the electrodes) on a specially thinned wafer. Semiconductor light detectors have much higher efficiency than photographic film with as many as $70 \%$ of the incoming photons being counted (photographic film captures just a few percent). There are always a few spurious electrons generated by thermal excitation and other mechanisms and it is common in astronomical applications to cool the detector to reduce the "dark current" and improve the signal-to-noise ratio.

There are two phases of operation in using a CCD for imaging. In the first phase, the electrons generated by the incoming photons are gradually accumulated for each pixel. In the second phase, the CCD mechanism is

[^11]activated and the electrons from each cell are moved serially to the edge the array and thence to the corners where they are amplified and converted to digital form for further accumulation and manipulation. The second phase is much shorter to avoid smearing since the illumination is still present as the contents of the cells are being moved (this is referred to as a 'full-frame' CCD).


In operating mode, for Kepler, each of the large $25 \mu \mathrm{~m} \times 25 \mu \mathrm{~m}$ CCD pixels is read out after accumulating photons for six seconds. It is completely impractical to store or transmit information created at this rate, so Kepler does significant data-reduction locally. First, only pixels corresponding to target stars (about $6 \%$ of the total) are processed. Second, the six-second readouts are accumulated digitally for about 30 minutes. Finally these 30 -minute 'chunks' for each star are re-quantized and compressed for storage in the 16 GigaByte local DRAM memory.

## Dynamic Random Access Memory (DRAM)

Scientific and engineering data acquired by Kepler is stored in a 16 GigaByte solid-state recorder capable of holding up to 60 days of processed and compressed data. The memory is synchronous dynamic random access memory (DRAM) with simultaneous read and write capability. DRAM is a dynamic volatile memory. "Dynamic" implies that
the data needs to be continually refreshed otherwise it disappears over a time-scale of seconds or minutes (the industry standard is one refresh at least every 64 milliseconds). "Volatile" means that it has to be supplied with a continuous source of power. The continuous power is provided in the Kepler spacecraft by ensuring that the Solar panels are always illuminated and, in case attitude control fails, that the on-board Lithiumion battery is kept fully charged.

DRAM forms the main active memory in all modern computers including gaming computers and cell-phones and many other devices. The advantages of DRAM over Flash memory include much faster writing and reading processes, bit- or word-level addressability, and the absence of major wear mechanisms that cause long-term degradation. Flash memory can achieve much higher storage density and lower cost per MegaByte but tends to take a role more similar to that of a Hard Disk Drive (HDD). A processor will typically talk routinely to the DRAM memory but only occasionally transfer large blocks of data into or out of the HDD or Flash memory (The acronym SSD, for Solid State Drive, is used to refer to a large unit of flash memory configured with an HDD interface). DRAM is fast compared with flash memory, but still much slower than the highspeed storage registers (Static-RAM) that are built as an integral part of the processor chip.


The essence of DRAM technology is a cell with single storage capacitor accessed by a single field-effect transistor (FET) switch. The simplicity of the design allows for a very high density of cells on the Silicon chip or die.

Data is stored in a particular cell by switching on the appropriate FET with the address line and allowing the voltage on the bit-line to reach the capacitor and charge or discharge it so it has the same voltage as the bitline. High voltage may signify a binary 'one' and low voltage a binary ' 0 '. The capacitor is physically very small with a very thin oxide or nitride insulating layer to maximize capacitance. Nevertheless the charge on the capacitor does gradually leak away through the insulation and through the transistor switch. The time-scale for this may range from seconds to hours depending on the process and process variability and especially on the ambient temperature
Reading is accomplished by pre-charging a balanced pair of bit-lines (typically one line connected to the string of odd cells and the other line connected to the even cells) and then addressing/enabling one cell on one of the two lines. The high or low voltage on the addressed capacitor shifts the voltage balance on the pair of bit lines. The imbalance is sensed by a latching comparator. Here 'latching' refers to the idea that whatever imbalance is sensed, there is internal positive feedback within the comparator that reinforces the imbalance and quickly latches itself into a well-defined negative (zero) or positive (one) state. This latching includes the input sense lines which are also forced strongly into the state initially sensed. Because the transistor on the cell being read is still switched on (closed), the read operation automatically refreshes the charge on the capacitor.
The word synchronous is also used in the description of the Kepler DRAM. This implies that the device is driven by an external clock. This mode allows for significant clocked logic and a state-machine to be included into the device and for operations to be pipelined. This greatly improves the device capabilities and throughput. Modern DRAM invariably takes advantage of synchronous operations within the device. DRAM chips are extreme examples of very large scale integration (VLSI) incorporating up to 16 Gbits (i.e. 16 billion transistors and capacitors) in a single die (as of 2016). Reading and Writing access times are typically in the low tens of nanoseconds.

Data stored in DRAM is subject to occasional errors. This is especially true in space environments where high energy particles can cause charge to leak or can cause permanent damage to a cell. Radiation-hardened devices tend to use special sapphire substrates and larger cell size (older technology). Kepler includes a radiation-hardened processor and DRAM and a large cell-size in the CCD imager. The processor is a thirdgeneration radiation-hardened version of the PowerPC chip that is used on
some models of Macintosh computers. In addition, the DRAM or the DRAM controller frequently includes some error-correction capability.

The simple example below shows an extended Hamming code that has the capability of correcting any single-error in a codeword and also detecting (but not correcting) whenever two errors occur in a codeword.

| Toy example of a single errorcorrecting double error-detecting Hamming code $(8,4,4)$ covering just four bits of data. | Data |  | Codeword |
| :---: | :---: | :---: | :---: |
|  | 0000 | $\rightarrow$ | 00000000 |
|  | 0001 | $\rightarrow$ | 00011011 |
|  | 0010 | $\rightarrow$ | 00101110 |
| Four parity bits (shown in blue) are added to the four data bits. In this | 0011 | $\rightarrow$ | 00110101 |
| simple example the storage required | 0100 | $\rightarrow$ | 01000111 |
| would be doubled, but with longer | 0101 | 7 | 01011100 |
| data- and code-words, the fraction of | 0110 | $\rightarrow$ | 01101001 |
| overhead becomes much less. | 0111 | 7 | 01110010 |
| It can be seen by inspection that a | 1000 | $\rightarrow$ | 10001101 |
| given codeword differs from all other | 1001 | $\rightarrow$ | 10010110 |
| codewords in at least four bits. This | 1010 | $\rightarrow$ | 10100011 |
| means that a single-bit error leaves a codeword within one bit of the | 1011 | $\rightarrow$ | 10111000 |
| correct codeword (and at least three | 1100 | $\rightarrow$ | 11001010 |
| bits from any incorrect codword) and | 1101 | $\rightarrow$ | 11010001 |
| can therefore be corrected. An error | 1110 | $\rightarrow$ | 11100100 |
| on two bits can be detected but not necessarily corrected. | 1111 | $\rightarrow$ | 11111111 |

Much more complex coding and detection/decoding schemes are employed in communications channels and in hard disk drives. A typical format in an HDD will expend about $15 \%$ of the capacity on redundant bits for error-correction. The fraction on solid state drives (SSD) and especially on deep-space channels can be very much larger.

The discussion in this chapter started with how mankind understanding of the solar system gradually developed and then finished with an explanation of how modern DRAM devices work. The central excuse for all this was the Kepler space telescope which, at one time (in about 2010) was about 10 million kilometers from Earth. So our next logarithmic leap should bring us to one million kilometers. We argue that the Moon at about 0.4 million kilometers falls within our loose definition (nearest order of magnitude). For the next chapter we also take a large leap back in time - about 50 years to the year 1969 and to the Apollo 11 moon landing. This
is fascinating time - from all the politics and the personalities, to the mission and the engineering, and then all the details of the computer and memory systems - all, of course, designed in the 1950s and 60s.

## Further Reading:

Claudius Ptolemy, [W. Donahue (Ed.), B. Perry (Tr.), "The Almagest: Introduction to the Mathematics of the Heavens", Green Lion Press, December 7, 2014
"Johannes Kepler", World Heritage Encyclopedia, July 2018
http://central.gutenberg.org/articles/johannes_kepler
NASA Press Kit "Kepler: NASA's First Mission Capable of Finding Earth-Size Planets", February 2009, https://www.nasa.gov/pdf/314125main_Kepler_presskit_2-19_smfile.pdf
P. Amico and J. Beletic (Editors) "Scientific Detectors for Astronomy", Kluwer Academic Press, 2004
(See also the tutorial box below in blue)

## Photo-electric Effect in Semiconductors

In a silicon crystal all the electrons are held in the covalent bonds between adjacent atoms. If sufficient energy is available, the bond can be broken and both the displaced electron and the corresponding missing-electron or 'hole' can move relatively freely in the crystal. The energy required to break the bond is welldefined at exactly 1.11 electron-Volts ( eV ) or $1.78 \times 10^{-19} \mathrm{Joules}$.


All electromagnetic radiation (including visible-light and infrared) is quantized. The quanta are called photons. The energy of a photon is given strictly by its frequency. The Planck-Einstein relationship states that the energy of a photon is given by the Planck constant ( $\mathrm{h}=4.14 \times 10^{-15} \mathrm{eV} / \mathrm{Hz}$ ) times its frequency, f .

When a photon enters a silicon crystal, the electrons respond strongly to the photon and in doing so, reduce its velocity by about a factor of four (i.e. the refractive index is about 4). However, as long as the photon energy is less than that required to break any of the covalent electron bonds, the photon does not lose any energy. But, above the critical energy level, the photon can break the electron bonds and create electron-hole pairs. The electron and hole from each photon can be separated by an electric field and counted or used by suitable external electronics. This is the basis for the CCD or CMOS chip that forms images in your camera or the photo-voltaic Solar cells that may cover your roof.

The critical threshold energy for silicon is 1.11 eV corresponding to an optical frequency of $2.68 \times 10^{+14} \mathrm{~Hz}$. This gives a wavelength, $\lambda=\mathrm{c} / \mathrm{f}$, in free space of 1120 nm (infra-red). At wavelengths longer than this, the silicon is lossless and transparent. At wavelengths shorter than 1120 nm , the photons start creating electron-hole pairs and the silicon appears opaque (lossy) with a metallic luster.


[^0]:    ${ }^{1}$ http://www.keplersdiscovery.com

[^1]:    ${ }^{2}$ The Greek is often used to form a prefix: e.g. heliocentric, selenology

[^2]:    ${ }^{3}$ http://www.self.gutenberg.org/articles/eng/ptolemy

[^3]:    ${ }^{4}$ http://central.gutenberg.org/articles/nicolas copernicus
    5 https://en.wikipedia.org/wiki/Johannes Kepler

[^4]:    ${ }^{6}$ http://central.gutenberg.org/articles/tycho brahe
    7 Today we would recognize this as conservation of angular momentum.
    ${ }^{8}$ http://central.gutenberg.org/articles/galileo_galilei
    ${ }^{9}$ Crescent moon = less than a half moon; Gibbous moon = more than a half moon

[^5]:    ${ }^{10}$ In general a seventh parameter is required describing the period or, equivalently, the mass of the central object (e.g. Sun, Earth, Jupiter etc.)
    ${ }^{11}$ http://central.gutenberg.org/articles/issac newton

[^6]:    ${ }^{12}$ E. Howell, "Lagrange Points: Parking Places in Space", August 21, 2017 https://www.space.com/30302-lagrange-points.html

[^7]:    ${ }^{13}$ https://en.wikipedia.org/wiki/2010 TK7
    ${ }^{14}$ https://www.jwst.nasa.gov/

[^8]:    ${ }^{15}$ https://en.wikipedia.org/wiki/Counter-Earth

[^9]:    ${ }^{16}$ https://www.nasa.gov/mission pages/kepler/main/index.html https://en.wikipedia.org/wiki/Kepler (spacecraft)
    ${ }^{17}$ https://keplerscience.arc.nasa.gov/the-kepler-space-telescope.html
    ${ }^{18}$ A steradian is a measure of solid angle. There are $4 \pi$ steradians in a complete sphere
    ${ }^{19}$ http://www.specinst.com/What Is A CCD.html https://en.wikipedia.org/wiki/Charge-coupled device

[^10]:    ${ }^{20}$ https://en.wikipedia.org/wiki/List of exoplanets discovered using the Kepler space craft

[^11]:    ${ }^{21}$ Typically the clocking is done with three or four partially overlapping phases of voltage that guide the electrons along: https://en.wikipedia.org/wiki/Chargecoupled device\#/media/File:CCD charge transfer animation.gif
    ${ }^{22}$ http://hamamatsu.magnet.fsu.edu/articles/fourphase.html

