the giant planets.

In the inner parts of the system, remnant planetesimals and perhaps several dozen protoplanets continued to whiz about. Over the vast span of time we are discussing, collisions among these objects were inevitable. Giant impacts at this stage probably stripped Mercury of part of its mantle and crust, reversed the rotation of Venus, and broke off part of Earth to create the Moon (all events we discussed in other chapters).

Smaller-scale impacts also added mass to the inner protoplanets. Because the gravity of the giant planets could "stir up" the orbits of the planetesimals, the material impacting on the inner protoplanets could have come from almost anywhere within the solar system. In contrast to the previous stage of accretion, therefore, this new material did not represent just a narrow range of compositions.

As a result, much of the debris striking the inner planets was ice-rich material that had condensed in the outer part of the solar nebula. As this comet-like bombardment progressed, Earth accumulated the water and various organic compounds that would later be critical to the formation of life. Mars and Venus probably also acquired abundant water and organic materials from the same source, as Mercury and the Moon are still doing to form their icy polar caps.

Gradually, as the planets swept up or ejected the remaining debris, most of the planetesimals disappeared. In two regions, however, stable orbits are possible where leftover planetesimals could avoid impacting the planets or being ejected from the system. These regions are the asteroid belt between Mars and Jupiter and the Kuiper belt beyond Neptune. The planetesimals (and their fragments) that survive in these special locations are what we now call asteroids, comets, and trans-neptunian objects.

Astronomers used to think that the solar system that emerged from this early evolution was similar to what we see today. Detailed recent studies of the orbits of the planets and asteroids, however, suggest that there were more violent events soon afterward, perhaps involving substantial changes in the orbits of Jupiter and Saturn. These two giant planets control, through their gravity, the distribution of asteroids. Working backward from our present solar system, it appears that orbital changes took place during the first few hundred million years. One consequence may have been scattering of asteroids into the inner solar system, causing the period of "heavy bombardment" recorded in the oldest lunar craters.

14.4 COMPARISON WITH OTHER PLANETARY SYSTEMS

Learning Objectives

By the end of this section, you will be able to:

- Describe how the observations of protoplanetary disks provides evidence for the existence of other planetary systems
- Explain the two primary methods for detection of exoplanets
- Compare the main characteristics of other planetary systems with the features of the solar system

Until the middle 1990s, the practical study of the origin of planets focused on our single known example—the solar system. Although there had been a great deal of speculation about planets circling other stars, none had actually been detected. Logically enough, in the absence of data, most scientists assumed that our own system was likely to be typical. They were in for a big surprise.

Discovery of Other Planetary Systems

In **The Birth of Stars and the Discovery of Planets outside the Solar System**, we discuss the formation of stars and planets in some detail. Stars like our Sun are formed when dense regions in a molecular cloud (made of gas and dust) feel an extra gravitational force and begin to collapse. This is a runaway process: as the cloud collapses, the gravitational force gets stronger, concentrating material into a protostar. Roughly half of the time, the protostar will fragment or be gravitationally bound to other protostars, forming a binary or multiple star system—stars that are gravitationally bound and orbit each other. The rest of the time, the protostar collapses in isolation, as was the case for our Sun. In all cases, as we saw, conservation of angular momentum results in a spin-up of the collapsing protostar, with surrounding material flattened into a disk. Today, this kind of structure can actually be observed. The Hubble Space Telescope, as well as powerful new ground-based telescopes, enable astronomers to study directly the nearest of these *circumstellar disks* in regions of space where stars are being born today, such as the Orion Nebula (**Figure 14.14**) or the Taurus star-forming region.

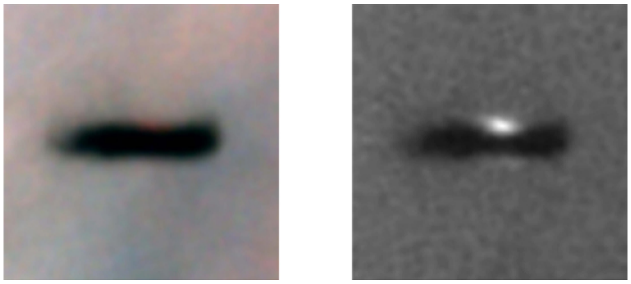

Figure 14.14 Protoplanetary Disk in the Orion Nebula. The Hubble Space Telescope imaged this protoplanetary disk in the Orion Nebula, a region of active star formation, using two different filters. The disk, about 17 times the size of our solar system, is in an edge-on orientation to us, and the newly formed star is shining at the center of the flattened dust cloud. The dark areas indicate absorption, not an absence of material. In the left image we see the light of the nebula and the dark cloud; in the right image, a special filter was used to block the light of the background nebula. You can see gas above and below the disk set to glow by the light of the newborn star hidden by the disk. (credit: modification of work by Mark McCaughrean (Max-Planck-Institute for Astronomy), C. Robert O'Dell (Rice University), and NASA)

Many of the circumstellar disks we have discovered show internal structure. The disks appear to be donut-shaped, with gaps close to the star. Such gaps indicate that the gas and dust in the disk have already collapsed to form large planets (**Figure 14.15**). The newly born protoplanets are too small and faint to be seen directly, but the depletion of raw materials in the gaps hints at the presence of something invisible in the inner part of the circumstellar disk—and that something is almost certainly one or more planets. Theoretical models of planet formation, like the one seen at right in **Figure 14.15**, have long supported the idea that planets would clear gaps as they form in disks.

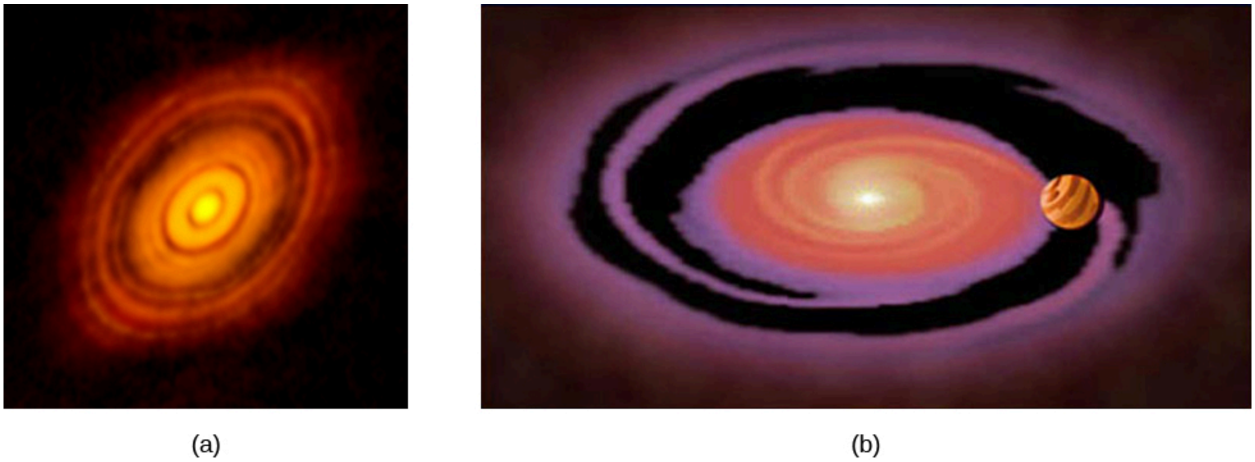

Figure 14.15 Protoplanetary Disk around HL Tau. (a) This image of a protoplanetary disk around HL Tau was taken with the Atacama Large Millimeter/submillimeter Array (ALMA), which allows astronomers to construct radio images that rival those taken with visible light. (b) Newly formed planets that orbit the central star clear out dust lanes in their paths, just as our theoretical models predict. This computer simulation shows the empty lane and spiral density waves that result as a giant planet is forming within the disk. The planet is not shown to scale. (credit a: modification of work by ALMA (ESO/NAOJ/NRAO); credit b: modification of work by NASA/ESA and A. Feild (STScI))

Our figure shows HL Tau, a one-million-year-old “newborn” star in the Taurus star-forming region. The star is embedded in a shroud of dust and gas that obscures our visible-light view of a circumstellar disk around the star. In 2014 astronomers obtained a dramatic view of the HL Tau circumstellar disk using millimeter waves, which pierce the cocoon of dust around the star, showing dust lanes being carved out by several newly formed protoplanets. As the mass of the protoplanets increases, they travel in their orbits at speeds that are faster than the dust and gas in the circumstellar disk. As the protoplanets plow through the disk, their gravitational reach begins to exceed their cross-sectional area, and they become very efficient at sweeping up material and growing until they clear a gap in the disk. The image of **Figure 14.15** shows us that a number of protoplanets are forming in the disk and that they were able to form faster than our earlier ideas had suggested—all in the first million years of star formation.

LINK TO LEARNING

For an explanation of ALMA’s ground-breaking observations of HL Tau and what they reveal about plant formation, watch this **videocast (https://openstaxcollege.org/l/30eusoobhltavid)** from the European Southern Observatory.

Discovering Exoplanets

You might think that with the advanced telescopes and detectors astronomers have today, they could directly image planets around nearby stars (which we call **exoplanets**). This has proved extremely difficult, however, not only because the exoplanets are faint, but also because they are generally lost in the brilliant glare of the star they orbit. As we discuss in more detail in **The Birth of Stars and the Discovery of Planets outside the Solar System**, the detection techniques that work best are indirect: they observe the effects of the planet on the star it orbits, rather than seeing the planet itself.

The first technique that yielded many planet detections is very high-resolution stellar spectroscopy. The *Doppler effect* lets astronomers measure the star’s *radial velocity:* that is, the speed of the star, toward us or away from us, relative to the observer. If there is a massive planet in orbit around the star, the gravity of the planet causes the star to wobble, changing its radial velocity by a small but detectable amount. The distance of the star does not matter, as long as it is bright enough for us to take very high quality spectra.

Measurements of the variation in the star’s radial velocity as the planet goes around the star can tell us the mass and orbital period of the planet. If there are several planets present, their effects on the radial velocity can be disentangled, so the entire planetary system can be deciphered—as long as the planets are massive enough to produce a measureable Doppler effect. This detection technique is most sensitive to large planets orbiting close to the star, since these produce the greatest wobble in their stars. It has been used on large ground-based telescopes to detect hundreds of planets, including one around Proxima Centauri, the nearest star to the Sun.

The second indirect technique is based on the slight dimming of a star when one of its planets *transits*, or crosses over the face of the star, as seen from Earth. Astronomers do not see the planet, but only detect its presence from careful measurements of a change in the brightness of the star over long periods of time. If the slight dips in brightness repeat at regular intervals, we can determine the orbital period of the planet. From the amount of starlight obscured, we can measure the planet’s size.

While some transits have been measured from Earth, large-scale application of this transit technique requires a telescope in space, above the atmosphere and its distortions of the star images. It has been most successfully applied from the NASA Kepler space observatory, which was built for the sole purpose of “staring” for 5 years at a single part of the sky, continuously monitoring the light from more than 150,000 stars. The primary goal of Kepler was to determine the frequency of occurrence of exoplanets of different sizes around different classes of stars. Like the Doppler technique, the transit observations favor discovery of large planets and short-period orbits.

Recent detection of exoplanets using both the Doppler and transit techniques has been incredibly successful. Within two decades, we went from no knowledge of other planetary systems to a catalog of *thousands* of exoplanets. Most of the exoplanets found so far are more massive than or larger in size than Earth. It is not that Earth analogs do not exist. Rather, the shortage of small rocky planets is an observational bias: smaller planets are more difficult to detect.

Analyses of the data to correct for such biases or selection effects indicate that small planets (like the terrestrial planets in our system) are actually much more common than giant planets. Also relatively common are “super Earths,” planets with two to ten times the mass of our planet (Figure 14.16). We don’t have any of these in our solar system, but nature seems to have no trouble making them elsewhere. Overall, the Kepler data suggest that approximately one quarter of stars have exoplanet systems, implying the existence of at least 50 billion planets in our Galaxy alone.

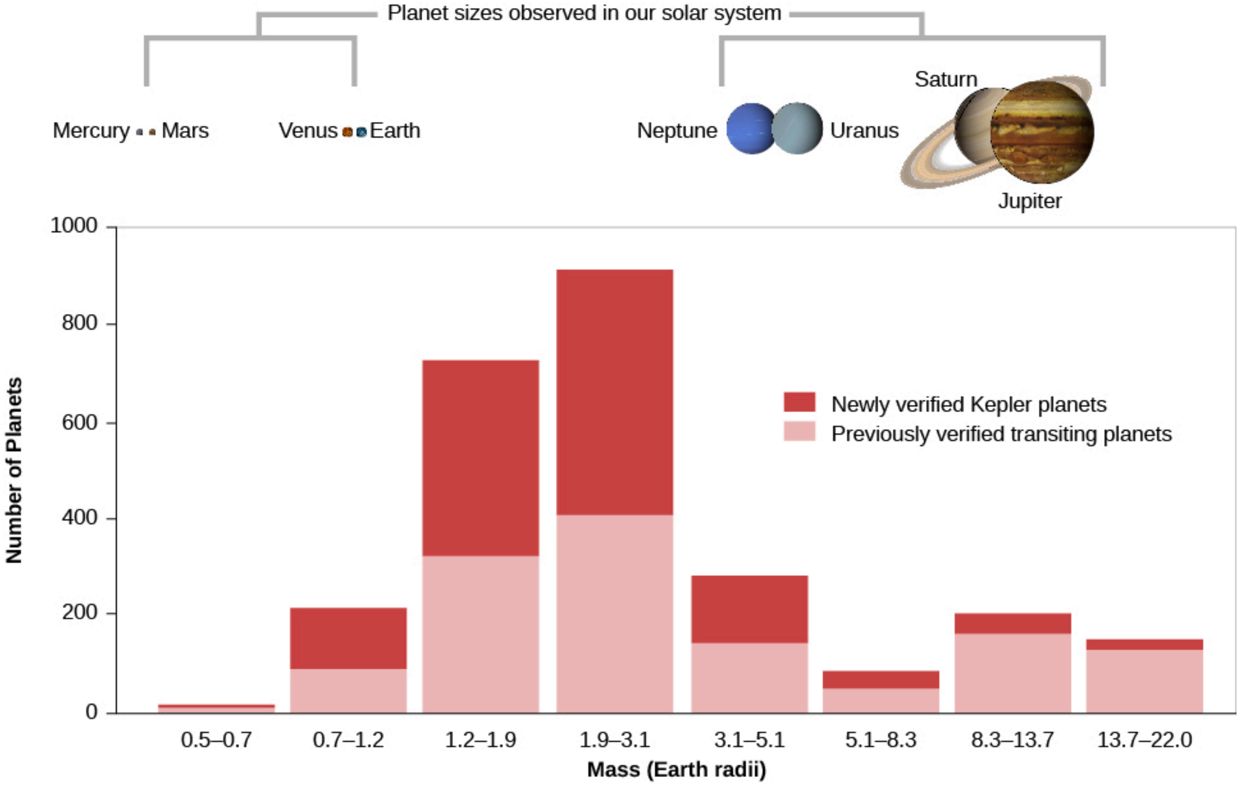

Figure 14.16 Transiting Planets by Size. This bar graph shows the planets found so far using the transit method (the vast majority found by the Kepler mission). The orange parts of each bar indicate the planets announced by the Kepler team in May 2016. Note that the largest number of planets found so far are in two categories that we don’t have in our own solar system—planets whose size is between Earth’s and Neptune’s. (credit: modification of work by NASA)

The Configurations of Other Planetary Systems

Let’s look more closely at the progress in the detection of exoplanets. Figure 14.17 shows the planets that were discovered each year by the two techniques we discussed. In the early years of exoplanet discovery, most of the planets were similar in mass to Jupiter. This is because, as mentioned above, the most massive planets were easiest to detect. In more recent years, planets smaller than Neptune and even close to the size of Earth have been detected.

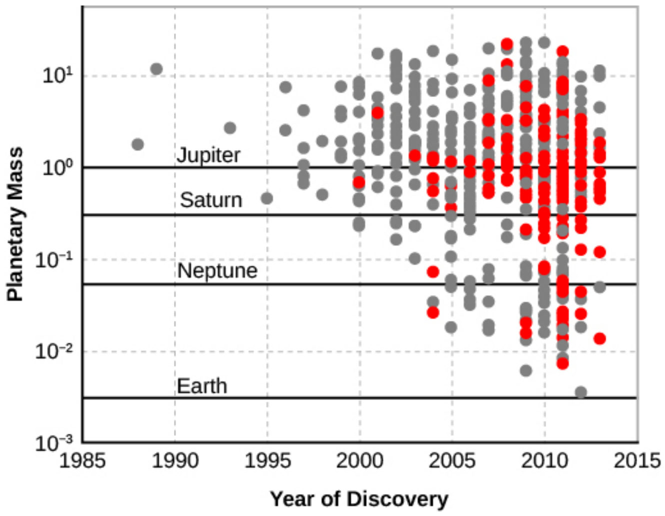

Figure 14.17 Masses of Exoplanets Discovered by Year. Horizontal lines are drawn to reference the masses of Jupiter, Saturn, Neptune, and Earth. The gray dots indicate planets discovered by measuring the radial velocity of the star, and the red dots are for planets that transit their stars. In the early years, the only planets that could be detected were similar in mass to Jupiter. Improvements in technology and observing strategies enabled the detection of lower mass planets as time went on, and now even smaller worlds are being found. (Note that this tally ends in 2014.)

We also know that many exoplanets are in multiplanet systems. This is one characteristic that our solar system shares with exosystems. Looking back at Figure 14.15 and seeing how such large disks can give rise to more than one center of condensation, it is not too surprising that multiplanet systems are a typical outcome of planet formation. Astronomers have tried to measure whether multiple planet systems all lie in the same plane using astrometry. This is a difficult measurement to make with current technology, but it is an important measurement that could help us understand the origin and evolution of planetary systems.

Comparison between Theory and Data

Many of the planetary systems discovered so far do not resemble our own solar system. Consequently, we have had to reassess some aspects of the “standard models” for the formation of planetary systems. Science sometimes works in this way, with new data contradicting our expectations. The press often talks about a scientist making experiments to “confirm” a theory. Indeed, it is comforting when new data support a hypothesis or theory and increase our confidence in an earlier result. But the most exciting and productive moments in science often come when new data *don’t* support existing theories, forcing scientists to rethink their position and develop new and deeper insights into the way nature works.

Nothing about the new planetary systems contradicts the basic idea that planets form from the aggregation (clumping) of material within circumstellar disks. However, the existence of “ hot Jupiters”—planets of jovian mass that are closer to their stars than the orbit of Mercury—poses the biggest problem. As far as we know, a giant planet cannot be formed without the condensation of water ice, and water ice is not stable so close to the heat of a star. It seems likely that all the giant planets, “hot” or “normal,” formed at a distance of several astronomical units from the star, but we now see that they did not necessarily stay there. This discovery has led to a revision in our understanding of planet formation that now includes “planet migrations” within the protoplanetary disk, or later gravitational encounters between sibling planets that scatter one of the planets inward.

Many exoplanets have large orbital eccentricity (recall this means the orbits are not circular). High eccentricities were not expected for planets that form in a disk. This discovery provides further support for the scattering of planets when they interact gravitationally. When planets change each other’s motions, their orbits could become much more eccentric than the ones with which they began.

There are several suggestions for ways migration might have occurred. Most involve interactions between the giant planets and the remnant material in the circumstellar disk from which they formed. These interactions would have taken place when the system was very young, while material still remained in the disk. In such cases, the planet travels at a faster velocity than the gas and dust and feels a kind of “headwind” (or friction) that causes it to lose energy and spiral inward. It is still unclear how the spiraling planet stops before it plunges into the star. Our best guess is that this plunge into the star is the fate for many protoplanets; however, clearly some migrating planets can stop their inward motions and escape this destruction, since we find hot Jupiters in many mature planetary systems.

14.5 PLANETARY EVOLUTION

Learning Objectives

By the end of this section, you will be able to:

- Describe the geological activity during the evolution of the planets, particularly on the terrestrial planets
- Describe the factors that affect differences in elevation on the terrestrial planets
- Explain how the differences in atmosphere on Venus, Earth, and Mars evolved from similar starting points in the early history of the solar system

While we await more discoveries and better understanding of other planetary systems, let us look again at the early history of our own solar system, after the dissipation of our dust disk. The era of giant impacts was probably confined to the first 100 million years of solar system history, ending by about 4.4 billion years ago. Shortly thereafter, the planets cooled and began to assume their present aspects. Up until about 4 billion years ago, they continued to acquire volatile materials, and their surfaces were heavily cratered from the remaining debris that hit them. However, as external influences declined, all the terrestrial planets as well as the moons of the outer planets began to follow their own evolutionary courses. The nature of this evolution depended on each object’s composition, mass, and distance from the Sun.

Geological Activity

We have seen a wide range in the level of geological activity on the terrestrial planets and icy moons. Internal sources of such activity (as opposed to pummeling from above) require energy, either in the form of primordial heat left over from the formation of a planet or from the decay of radioactive elements in the interior. The larger the planet or moon, the more likely it is to retain its internal heat and the more slowly it cools—this is the “baked potato effect” mentioned in **Other Worlds: An Introduction to the Solar System**. Therefore, we are more likely to see evidence of continuing geological activity on the surface of larger (solid) worlds (**Figure 14.18**). Jupiter’s moon Io is an interesting exception to this rule; we saw that it has an unusual source of heat from the gravitational flexing of its interior by the tidal pull of Jupiter. Europa is probably also heated by jovian tides. Saturn may be having a similar effect on its moon Enceladus.