Although they are, in cosmic terms, mere scraps—insignificant to the grand narrative of heavenly expansion—planets are the most diverse and intricate class of object in the universe. No other celestial bodies support such a complex interplay of astronomical, geologic, and chemical and biological processes. No other places in the cosmos could support life as we know it. The worlds of our solar system come in a tremendous variety, and even they hardly prepared us for the discoveries of the past decade, during which astronomers have found more than 200 planets.

The sheer diversity of these bodies’ masses, sizes, compositions and orbits challenges those of us trying to fathom their origins. When I was in graduate school in the 1970s, we tended to think of planet formation as a well-ordered, deterministic process—an assembly line that turns amorphous disks of gas and dust into copies of our solar system. Now we are realizing that the process is chaotic, with distinct outcomes for each system. The worlds that emerge are the survivors of a hurly-burly of competing mechanisms of creation and destruction. Many are blasted apart, fed into the fires of their system’s newborn star or ejected into interstellar space. Our own Earth may have long-lost siblings that wander through the lightless void.
The study of planet formation lies at the intersection of astrophysics, planetary science, statistical mechanics and nonlinear dynamics. Broadly speaking, planetary scientists have developed two leading theories. The sequential-accretion scenario holds that tiny grains of dust clump together to create solid nuggets of rock, which either draw in huge amounts of gas, becoming gas giants such as Jupiter, or do not, becoming rocky planets such as Earth. The main drawback of this scenario is that it is a slow process and that gas may disperse before it can run to completion.

The alternative, gravitational-instability scenario holds that gas giants take shape in an abrupt whoosh as the prenatal disk of gas and dust breaks up—a process that replicates, in miniature, the formation of stars. This hypothesis remains contentious because it assumes the existence of highly unstable conditions, which may not be attainable. Moreover, astronomers have found that the heaviest planets and the lightest stars are separated by a “desert”—a scarcity of intermediate bodies. The disjunction implies that planets are not simply little stars but have an entirely different origin.

Although researchers have not settled this controversy, most consider the sequential-accretion scenario the most plausible of the two. I will focus on it here.
An Interstellar Cloud Collapses

Time: 0 (starting point of planet formation sequence)

Our solar system belongs to a galaxy of some 100 billion stars threaded with clouds of gas and dust, much of it the debris of previous generations of stars. “Dust” in this context simply means microscopic bits of water ice, iron and other solid substances that condensed in the cool outer layers of stars and were blown out into interstellar space. When clouds are sufficiently cold and dense, they can collapse under the force of gravity to form clusters of stars, a process that takes 100,000 to a few million years [see “Fountains of Youth: Early Days in the Life of a Star,” by Thomas P. Ray; SCIENTIFIC AMERICAN, August 2000].

Surrounding each star is a rotating disk of leftover material, the wherewithal for making planets. Newly formed disks contain mostly hydrogen and helium gas. In their hot and dense inner regions, dust grains are vaporized; in the cool and tenuous outer parts, the dust particles survive and grow as vapor condenses onto them.

Astronomers have discovered many young stars that are surrounded by such disks. Stars between one million and three million years old have gas-rich disks, whereas those older than 10 million years have meager, gas-poor disks, the gas having been blown away by the newborn star or by bright neighboring stars. This span of time delineates the epoch of planet formation. The mass of heavy elements in these disks is roughly comparable to the mass of heavy elements in the planets of the solar system, providing a strong clue that the planets indeed arose from such disks.

Ending point: Newborn star surrounded by gas and micron-size dust grains

COSMIC DUST BUNNIES

Even the mightiest planets have humble roots: as micron-size dust grains (the ashes of long-dead stars) embedded in a swirling disk of gas. The disk’s temperature falls with distance from the newborn star, defining a “snow line” beyond which water stays frozen. In our solar system, the snow line marks the boundary between the inner rocky planets and outer gas giants.

1. Grains collide, clump and grow.
2. Small grains are swept along by the gas, but those larger than a millimeter experience a drag force and spiral in.
3. At the snow line, local conditions are such that the drag force reverses direction. Grains tend to accumulate and readily coagulate into larger bodies called planetesimals.
The Disk Sorts Itself Out

Dust grains in the protoplanetary disk are stirred by nearby gas and collide with one another, sometimes sticking together, sometimes breaking apart. The grains intercept starlight and reemit lower-wavelength infrared light, ensuring that heat reaches even the darkest regions of the disk’s interior. The temperature, density and pressure of gas generally decrease with distance from the star. Because of the balance of pressure, rotation and gravity, gas orbits the star slightly slower than an independent body at the same distance would.

Consequently, dust grains larger than a few millimeters in size tend to outpace the gas, thereby running into a headwind that slows them down and causes them to spiral inward, toward the star. The bigger the grains grow, the faster they spiral. Chunks a meter in size can halve their distance from the star within 1,000 years.

As they approach the star, the grains warm up, and eventually water and other low-boiling-point substances, known as volatiles, boil off. The distance at which this happens, the “snow line,” lies between 2 and 4 AU (astronomical units) from the star, which in our solar system falls between the orbits of Mars and Jupiter. (The radius of Earth’s orbit is 1 AU.) The snow line divides the planetary system into an inner, volatile-poor region filled with rocky bodies and an outer, volatile-rich region filled with icy ones.

At the snow line itself, water molecules tend to accumulate as they boil off grains. This build-up of water triggers a cascade of effects. It produces a discontinuity in gas properties at the snow line, which leads to a pressure drop there. The balance of forces causes the gas to speed up its revolution around the central star. Consequently, grains in the vicinity feel not a headwind but a tailwind, which boosts their velocity and halts their inward migration. As grains continue to arrive from the outer parts of the disk, they pile up at the snow line. In effect, the snow line becomes a snowbank.

Crammed together, the grains collide and grow. Some break through the snow line and continue to migrate inward, but in the process they become coated with slush and complex molecules, which makes them stickier. Some regions are so thick with dust that the grains’ collective gravitational attraction also accelerates their growth.

In these ways, the dust grains pack themselves into kilometer-size bodies called planetesimals. By the end of the stage of planet formation, planetesimals have swept up almost all the original dust. Planetesimals are hard to see directly, but astronomers can infer their presence from the debris of their collisions [see “The Hidden Members of Planetary Systems,” by David Ardila; SCIENTIFIC AMERICAN, April 2004].

The embryos run out of raw material and stop growing.

The Rise of the Oligarchs

Billions of kilometer-size planetesimals, built up during stage 2, then agglomerate into moon- to Earth-size bodies known as embryos. Relatively few in number, embryos dominate their respective orbital zones; this “oligarchy” of embryos competes for the remaining material.

The embryos run out of raw material and stop growing.

Planetary Embryos Germinate

Time: 1 million to 10 million years

The cratered landscapes on Mercury, the moon and the asteroids leave little doubt that nascent planetary systems are shooting galleries. Collisions between planetesimals either build them up or break them apart. A balance between coagulation and fragmentation leads to a distri-
bution of sizes in which small bodies account for most of the surface area in the emerging system and large bodies account for most of its mass. The orbits may initially be elliptical, but over time, gas drag and collisions tend to make the paths around the star circular.

In the beginning the growth of a body is self-reinforcing. The larger a planetesimal becomes, the stronger the gravity it exerts, and the faster it sweeps up its less massive partners. When they attain masses comparable to our moon, however, bodies exert such strong gravity that they stir up surrounding solid material and divert most of it before they can collide with it. In this way, they limit their own growth. Thus, an “oligarch” arises—that is, a population of planetary embryos with similar masses that compete with one another for the residual planetesimals.

Each embryo’s feeding zone is a narrow band centered on its orbit. Its growth stalls once it acquires most of the residual planetesimals in the zone. By simple geometry, the size of the zone and the duration of feeding grow with distance from the star. At a distance of 1 AU, embryos plateau at about 0.1 Earth mass within 100,000 years. Out at 5 AU, they reach four Earth masses over a few million years. Embryos can grow even bigger near the snow line or on the edges of gaps in the disk, where planetesimals also tend to accumulate.

Oligarchic growth fills the system with a surplus of aspiring planets, only some of which will make it. The planets in our solar system may seem widely spaced, but they are as close together as they can be. Inserting another Earth-mass planet in the present-day space between the terrestrial planets would destabilize them all. The same is true of other known systems. If you come across a cup of coffee that is filled to the very rim, you can reasonably conclude that someone actually overfilled it and spilled some coffee; filling it exactly, without wasting a drop, seems unlikely. Similarly, planetary systems probably start with more material than they end up with. Bodies are ejected until the system reaches an equilibrium configuration. Astronomers have observed freely floating planets in young stellar clusters. Ending point: “Oligarchy” of moon- to Earth-mass planetary embryos.

4. A Gas Giant Is Born

Time: 1 million to 10 million years

Jupiter probably began as a seed comparable in size to Earth that then accumulated some 300 Earth masses of gas. Such spectacular growth hinges on various competing effects. An
embryo’s gravity pulls in gas from the disk, but the infalling gas releases energy and must cool off if it is to settle down. Consequently, the growth rate is limited by the cooling efficiency. If it is too slow, the star may blow away the gas in the disk before the embryo has a chance to develop a thick atmosphere. The main heat-transfer bottleneck is the flux of radiation through the outer layers of the emerging atmosphere, which is determined by the opacity of the gas (determined mainly by its composition) and the temperature gradient (determined largely by the embryo’s initial mass).

Early models indicated that embryos need to have a critical mass, about 10 times that of Earth, to allow for sufficiently fast heat transfer. Such large embryos can arise near the snow line, where material will have accumulated earlier. That may be why Jupiter is located just beyond the snow line. They can arise elsewhere if the disk contains more raw material than planetary scientists used to assume it would. In fact, astronomers have now observed many stars whose disks are a few times denser than the traditional estimate, in which case heat transfer poses no insurmountable problem.

Another factor working against gas giants is that the embryo tends to spiral inward toward the star. In a process known as type I migration, the embryo triggers a wave in the gaseous disk, which, in return, pulls on the embryo’s orbit gravitationally. The wave pattern follows the planet like the wake of a boat. The gas on the side that is farther from the star revolves more sluggishly than the embryo and acts to hold the embryo back, slowing it down. Meanwhile the gas interior to the orbit revolves more quickly and acts to pull the embryo forward, speeding it up. The exterior region, being larger, wins the tug-of-war and causes the embryo to lose energy and fall inward by several astronomical units over one million years. This migration tends to stall near the snow line, where the gas headwind turns into a tailwind and provides an extra boost to the embryo’s orbit. That may be yet another reason why Jupiter is where it is.

Embryo growth, embryo migration and gas depletion all occur at roughly the same rate. Which wins depends on the luck of the draw. In fact, several generations of embryos may start the process only to migrate away before they can complete it. In their wake, fresh batches of planetesimals from the outer regions of the disk move in and repeat the process, until eventually a gas giant forms successfully or the gas is lost and no gas giant is ever able to take root. Astronomers have detected Jupiter-mass planets around only about 10 percent of the sunlike stars they have examined. The cores of these planets may be the rare survivors of many generations of embryos—the last of the Mohicans.

The balance among the processes depends on the system’s original endowment of material. Nearly a third of stars that are rich in heavy elements are orbited by Jupiter-mass planets. Presumably these stars had denser disks that gave rise to larger embryos, which could evade the heat-transfer bottleneck. Conversely, fewer planets form around stars that are smaller or poorer in heavy elements.

Once growth takes off, it accelerates to a startlingly fast pace. Within 1,000 years, a Jupiter-mass planet can acquire half of its final mass. In the process, it dissipates so much heat that it can briefly outshine the sun. The planet stabilizes when it becomes massive enough to disturb the gaseous disk. Gas exterior to the planet’s orbit revolves faster than the planet, so the planet’s gravity tends to hold it back, causing it to fall toward the star—that is, away from the planet. Gas exterior to the planet’s orbit revolves slower, so the planet tends to speed it up, causing it to move outward—again, away from the planet. Thus, the planet opens up a gap in the disk and cuts off its supply of raw material. The gas tries to repopulate the gap, but computer simulations indicate that the planet wins the struggle if its mass exceeds about one Jupiter mass at 5 AU.

This critical mass depends on the timing. The earlier a planet forms, the bigger it can grow, because plenty of gas remains. Saturn may have acquired a lower mass than Jupiter simply because it formed a few million years later. Astronomers have noticed a shortage of planets in the range of 20 Earth masses (Neptune’s mass) to 100 Earth masses (Saturn’s mass), which may be a clue to the precise timing.

**Ending point: Jupiter-size planet (or not)**

### 5. The Gas Giant Gets Restless

**Time: 1 million to 3 million years**

Oddly, many of the extrasolar planets discovered over the past decade orbit very close to their star, much closer than Mercury orbits the sun. These so-called hot Jupiters could not have formed in their current positions, if only because the orbital feeding zones are too small...
TIMELINE FOR WORLD-MAKING

Based on radiometric dating of meteorites and telescope observations of disks around other stars, planetary scientists have pieced together a rough timetable for planet formation.

0 to 100,000 years—star forms at center of disk and begins to undergo nuclear fusion

100,000 to 2 million years (Myr)—dust grains assemble into moon- to Earth-mass planetary embryos

2 Myr—first gas giant forms and clears out first-generation asteroids

10 Myr—gas giant triggers formation of other giant planets as well as terrestrial planets; most gas is lost by now

800 Myr—rearrangement of planets continues as late as a billion years after the process started

to provide enough material. Their presence appears to require a three-part sequence of events that for some reason did not occur in our own solar system.

First, a gas giant must form within the inner part of the planetary system, near the snow line, while the disk still has a considerable amount of gas. That requires a dense concentration of solid material in the disk.

Second, the giant planet must move to its present position. Type I migration cannot bring that about because it operates on embryos before they build up much gas. Instead type II migration must take place. The emerging giant planet opens a gap in the disk and suppresses the flow of gas across its orbit. In so doing, it must fight the tendency of turbulent gas in adjacent regions of the disk to spread. Gas never stops oozing into the gap, and its diffusion toward the central star forces the planet to lose orbital energy. This process is relatively slow, taking a few million years to shift a planet a few astronomical units, which is why the planet must start in the inner solar system if it is to end up hugging the star. As it and other planets migrate inward, they push along any residual planetesimals and embryos ahead of their paths, perhaps creating “hot Earths” in tight orbits.

Third, something must halt migration before the planet falls all the way into the star. The stellar magnetic field might clear gas from a cavity immediately around the star; without gas, migration ceases. Alternatively, perhaps the planet raises tides on the star, and the star, in turn, torques the planet’s orbit. These safeguards may not operate in all systems, and many planets may well fall all the way in.

Ending point: Tightly orbiting giant planet (“hot Jupiter”)

6. Other Giant Planets Join the Family

Time: 2 million to 10 million years

If one gas giant manages to arise, it facilitates the formation of subsequent gas giants. Many, perhaps most, known giant planets have siblings of comparable mass. In our solar system, Jupiter helped Saturn to emerge much faster than it would have by itself. It also lent a helping hand to Uranus and Neptune, without which they might never have grown to their present sizes; at their distance from the sun, the unaided formation process is so slow that the disk would have dissipated long before it could finish, leaving stunted worlds.

The pioneering gas giant has several helpful effects. At the outer edge of the gap that it opens up, material accumulates for much the same reasons it did at the snow line—namely, a pressure differential causes gas to speed up and act as a tailwind on grains and planetesimals, stopping their migration from more distant regions of the disk. Another effect of the first gas giant is that its gravity tends to fling nearby planetesimals to the outer reaches of the system, where they can form new planets.

The second-generation planets form out of the material that the first gas giant collects for them. The timing is critical, and fairly modest differences in timescales could lead to large differences in the outcome. In the case of Uranus and Neptune, the accumulation of planetesimals was too much of a good thing. The embryos became extra large, some 10 to 20 Earth masses, which delayed the onset of gas accretion—by which point little gas remained to be accreted. These bodies ended up with only about two Earth masses of gas. They are not gas giants but ice giants, which may in fact prove to be the more common type of giant.

The gravitational fields of the second-generation planets introduce an additional complication into the system. If the bodies form too close together, their interactions with one another and with the gaseous disk can catapult them into new, highly elliptical orbits. In our solar

[STAGE 5]

HOW TO HUG A STAR

In many systems, a giant planet forms and then spirals almost all the way into the star. The reason is that gas in the disk loses energy to internal friction and falls in, dragging the planet with it. Eventually the planet gets so close that the star exerts a torque on its orbit, stabilizing it.

Gas giant

© 2008 SCIENTIFIC AMERICAN, INC.
system, the planets all have nearly circular orbits and are spaced far enough apart to offer some immunity to one another's influence. In other planetary systems, however, elliptical orbits are the norm. In some, the orbits are resonant—that is, the orbital periods are related by a ratio of small whole numbers. Being born into this condition is highly improbable, but it can naturally arise when planets migrate and eventually lock onto one another gravitationally. The difference between these systems and our own may simply be the initial allotment of gas.

Most stars form in clusters, and more than half have binary companions. The planets may take shape in a plane that is not the same as the plane of the stellar orbit. In that case, the companion's gravity quickly realigns and distorts the planets' orbits, creating systems that are not planar, like our solar system, but spherical, like bees buzzing around a hive.

**Ending point: Coterie of giant planets**

### Earth-like Planets Assemble

**Time: 10 million to 100 million years**

Planetary scientists expect Earth-like planets to be more prevalent than gas giants. Whereas the gestation of a gas giant involves a fine balance of competing effects, formation of rocky planets should be fairly robust. Until we discover extrasolar Earths, however, we will have to rely on the solar system as our only case study.

The four terrestrial planets—Mercury, Venus, Earth and Mars—consist mostly of high-boiling-point material such as iron and silicate rocks, indicating that they formed inside the snow line and did not migrate significantly. At this range of distances, planetary embryos in a gaseous disk could grow to about 0.1 Earth mass, not much bigger than Mercury. Further growth required the embryos' orbits to cross so that they could collide and merge. That is easy enough to explain. After the gas evaporated, embryos gradually destabilized one another's orbits and, over a few million years, made them elliptical enough to intersect.

What is harder to explain is how the system stabilized itself again and what set the terrestrial planets on their present-day nearly circular orbits. A little bit of leftover gas could do the trick, but if gas were present, it would have prevented the orbits from becoming unstable to begin with. One idea is that after the planets nearly formed, a substantial swarm of planetesimals still remained. Over the next 100 million years, the planets swept up some of these planetesimals and deflected the rest into the sun. The

### BIGGEST AND BADDEST

Here are the record holders in extrasolar planetary systems as of March 2008. The planet masses are approximate because of measurement ambiguities.

- **Heaviest host star:** HD 13189 (4.5 solar masses)
- **Lightest host star:** GJ 317 (0.24 solar mass)
- **Tightest planet orbit:** OGLE-TR-56b (0.0225 AU)
- **Widest planet orbit:** PSR B 1620-26b (23 AU)
- **Heaviest planet:** NGC 4349 No 127b (19.8 Jupiter masses)
- **Lightest planet:** PSR 1257+12b (0.02 Earth mass)
planets transferred their random motion to the doomed planetesimals and entered into circular or almost circular orbits.

Another idea is that the long-range influence of Jupiter’s gravity caused the emerging terrestrial planets to migrate, bringing them into contact with fresh material. This influence would have been strongest at special resonant locations, which moved inward with time as Jupiter’s orbit settled into its final shape. Radiometric dating indicates that the asteroids formed early (four million years after the sun did), followed by the formation of Mars (10 million years after), then Earth (50 million years after)—as if a wave instigated by Jupiter was sweeping through the solar system. If unchecked, its influence would have pushed all the terrestrial planets to the orbit of Mercury. How did they avoid this unhappy outcome? Maybe they grew too massive for Jupiter to move them significantly, or maybe they were knocked out of Jupiter’s range of influence by giant impacts.

That said, most planetary scientists do not think Jupiter’s role was decisive in the formation of rocky planets. Most sunlike stars lack Jupiter-like planets, yet they still have dusty debris, indicating the presence of planetesimals and planetary embryos that could assemble into Earth-like worlds. A major question that observers need to answer over the coming decade is how many systems have Earths but not Jupiters.

For our planet, a defining moment occurred 30 million to 100 million years after the formation of the sun, when a Mars-size embryo knocked into the proto-Earth and threw out huge amounts of debris that coagulated into the moon. Such a giant impact is unsurprising given the amount of material careening around the early solar system, and Earth-like planets in other systems may have moons, too. Giant impacts also had the effect of ejecting the tenuous primitive atmosphere. The present-day atmosphere of Earth mostly came from gas that was trapped in the planetesimals that formed it and was later vented by volcanoes.

**Ending point:** Terrestrial planets

---

**Noncircular Reasoning**

In the inner solar system, planetary embryos cannot grow by swooping up gas but must collide with one another. To do so, their orbits must intersect, and to intersect, something must disturb them from their original circular orbits.

When embryos form, they have circular or nearly circular orbits, which do not intersect.

Gravitational interactions among the embryos or with a giant planet disturb the orbits.

The embryos agglomerate into an Earth-size planet. The planet then returns to a circular orbit by stirring up the remaining gas and scattering leftover planetesimals.

---

**Stage 7**

**Mop-Up Operations Commence**

**Time:** 50 million to 1 billion years

By this point, the planetary system is almost done. A few effects continue to fine-tune it: the disintegration of the wider star cluster, which may destabilize the planets’ orbits gravitationally; internal instabilities that develop after the star clears out the last of its gaseous disk; and
Meteors are not just space rocks but space fossils—planetary scientists’ only tangible record of the origin of the solar system. Planetary scientists think that they come from asteroids, which are fragments of planetesimals that never went on to form planets and have remained in deep freeze ever since. The composition of meteorites reflects what must have happened on their parent bodies. Intriguingly, they bear the scars of Jupiter’s early gravitational effects.

Iron and stony meteorites evidently originated in planetesimals that had melted, thereby allowing their iron and rocky silicate material to separate from each other, the heavy iron sinking to the core and the lighter silicates becoming concentrated in the outer layers. Researchers believe that this heating was brought about by the radioactive isotope aluminum 26, which has a half-life of 700,000 years. A supernova explosion or nearby star probably seeded the protosolar cloud with this isotope, in which case the first generation of planetesimals in our solar system contained plenty of it.

Yet iron and stony meteorites are very rare. Most meteorites consist instead of chondrules, which are millimeter-size pebbles that predate the formation of planetesimals and cannot survive melting. It therefore seems that most asteroids are not left over from the first generation of planetesimals. That generation must have been cleared out, presumably by Jupiter. Planetary scientists estimate that the region now occupied by the main asteroid belt used to have 1,000 times as much material as it does now. The few grains that eluded Jupiter’s clutches, or later drifted into the region of the belt, collected into new planetesimals, but little radioactive aluminum 26 was left by then, so these bodies never fully melted. The isotopic composition of chondrules in meteorites dates them to about two million years after the solar system started forming.

The glassy texture of the chondrules suggests that before being incorporated into planetesimals, they were abruptly heated, turned to molten rock and allowed to cool. The waves that drove Jupiter’s early orbital migration should have evolved into shock fronts, which could account for this flash heating.

—D.N.C.L.

No Grand Design

Before the age of discovery of extrasolar planets, our solar system was the only case study we had. Although it provided a wealth of information on the microphysics of important processes, it also narrowed our vision of how other systems could develop. The surprising planetary diversity discovered in the past decade has enormously expanded our theoretical horizons. We have come to realize that extrasolar planets are the last-generation survivors of a sequence of protoplanetary formation, migration, disruption and ongoing dynamic evolution. The relative orderliness of our solar system does not reflect any grand design.

Theorists have shifted their focus from providing scenarios to account for the relics of solar system formation to the construction of theories with some predictive power to be tested by forthcoming observations. Up to now, observers have seen only Jupiter-mass planets around sunlike stars. With a new generation of detectors, they will search for Earth-size planets, which the sequential-accretion scenario suggests are common. Planetary scientists may have only begun to see the full diversity of worlds in this universe.

more to explore


For the most up-to-date list of planet discoveries, go to http://exoplanet.eu