Available online at www.sciencedirect.com

SCIENCE @ DIRECT®

ELSEVIER

Earth and Planetary Science Letters 223 (2004) 241–252

EPSL Frontiers

www.elsevier.com/locate/epsl

Frontiers

Planetary accretion in the inner Solar System

John E. Chambers*

Department of Terrestrial Magnetism, Carnegie Institution of Washington, 5241 Broad Branch Road N.W., Washington, DC 20015, USA

Received 15 December 2003; received in revised form 23 April 2004; accepted 27 April 2004
Available online 4 June 2004

Abstract

Unlike gas-giant planets, we lack examples of terrestrial planets orbiting other Sun-like stars to help us understand how they formed. We can draw hints from elsewhere though. Astronomical observations of young stars; the chemical and isotopic compositions of Earth, Mars and meteorites; and the structure of the Solar System all provide clues to how the inner rocky planets formed. These data have inspired and helped to refine a widely accepted model for terrestrial planet formation—the planetesimal hypothesis. In this model, the young Sun is surrounded by a disk of gas and fine dust grains. Grains stick together to form mountain-size bodies called planetesimals. Collisions and gravitational interactions between planetesimals combine to produce a few tens of Moon-to-Mars-size planetary embryos in roughly 0.1–1 million years. Finally, the embryos collide to form the planets in 10–100 million years. One of these late collisions probably led to the formation of Earth's Moon. This basic sequence of events is clear, but a number of issues are unresolved. In particular, we do not really understand the physics of planetesimal formation, or how the planets came to have their present chemical compositions. We do not know why Mars is so much smaller than Earth, or exactly what prevented a planet from forming in the asteroid belt. Progress is being made in all of these areas, although definitive answers may have to wait for observations of Earth-like planets orbiting other stars.

Keywords: earth; terrestrial planets; asteroids; accretion; solar nebula

1. Introduction

We are witnessing a revolution in planetary science. The discovery of about a hundred other planetary systems has provided a wealth of new information to a field that was previously focussed on only one. However, the new planets are probably all gas giants, akin to Jupiter and Saturn, so they tell us relatively little about the nature and origin of small, rocky planets like Earth. We know of only one other system of terrestrial planets. This is in a most unfamiliar place: orbiting a *pulsar*, the extinct remnant of a supernova explosion [1]. Remarkably, the pulsar planets bear a striking resemblance to the inner planets of the Solar System in terms of their orbits and masses, although they may have originated under very different conditions. Currently, we lack a way to detect terrestrial planets in orbit around ordinary stars [2], so we have almost no notion of how common or otherwise Earth-like planets may be. As tantalizing as the new planetary discoveries are, we must look elsewhere for clues to the origin of the Sun's terrestrial planets.

Astronomical observations of newborn stars show that many are surrounded by a disk-shaped region of

---

* SETI Institute, 2035 Landings Drive, Mountain View, CA 94043, USA. Tel.: +1-650-604-5514; fax: +1-650-604-6779.
E-mail address: john@mycenae.arc.nasa.gov (J.E. Chambers).

0012-821X/$ - see front matter © 2004 Elsevier B.V. All rights reserved.
doi:10.1016/j.epsl.2004.04.031

gas clouded with fine dust [3]. The disks generally have radii at least as large as the Sun's planetary system, and contain at least as much mass as our planets. These observations, together with the planar geometry of the Solar System, and the fact that the Sun's planets all orbit in the same direction, suggest the planets formed in a similar disk environment. This is often called the *protoplanetary disk* or *protoplanetary nebula*. If this disk had the same composition as the Sun, roughly 0.5% of the mass in Earth's locale would have existed in solid grains of rock and metal. The remaining 99.5% would have been gas: hydrogen, helium and volatile materials such as water and carbon monoxide. Since Earth is made almost entirely of rock and metal, it is clear that planet formation in the inner Solar System was a sideshow compared to the evolution of the more massive protoplanetary nebula itself, although the conventional viewpoint is just the opposite.

Interestingly, most known giant planets orbit stars (the Sun included) that contain above average amounts of dust-forming elements such as iron [4]. One way to interpret this correlation is that giant planets form most readily where solid materials are abundant. The same may be true of rocky planets. It is also apparent that planet formation is an inefficient process, because even stars with dust-poor disks contain enough material to form a respectable system of planets. Stars older than a few million years (Myr) apparently lack massive gas-rich disks [5]. If massive disks are essential in order to generate planets then planet formation must begin within a few million years. Some older stars possess tenuous disks containing some dust but apparently little gas [6,7]. Dust in these systems should be ground down to small sizes and pushed out of the system on time scales of 10^4–10^6 years by the gentle but insistent pressure of light from the star itself [8]. This may be second-generation dust formed by high speed collisions between solid bodies orbiting these stars or dust evaporating from the surface of comets [8]. If so, this suggests that dust is able to accumulate into large solid bodies in a variety of protoplanetary disks.

2. Physical and cosmochemical constraints

We can make a crude estimate of the minimum mass of the Sun's protoplanetary nebula by totalling up the rock, metal and ice that resides in the planets, and adding enough hydrogen and helium to give a composition similar to the Sun. This *minimum mass nebula* contains a few percent of a solar mass—a number that fits snugly within the range of values estimated for circumstellar disks. However, this material did not quietly metamorphose into the planets we see today. Most of the gas has gone, somewhere. Some of the dust has disappeared too, or at least it has moved around a great deal. The giant planets probably contain tens of Earth masses of rocky and icy material, while the vast expanse of the asteroid belt, between 2 and 4 astronomical units (AU) from the Sun, has only enough stuff to make a planet 1% the mass of Mercury. Just as curiously, 90% of the total mass of the inner Solar System now resides in a narrow strip between 0.7 and 1 AU.

Much of what we know about the early history of the Solar System comes from studying *primitive meteorites*. These rocks come from asteroids that never became hot enough to melt. Thus, primitive meteorites and their parent bodies act as a kind of archaeological site, preserving the detritus formed in the first few million years of the Solar System. Most primitive meteorites are composed largely of *chondrules*—beads of rock typically about 1 mm in size. The composition and texture of most chondrules suggests they were once balls of dust floating in the solar nebula that were strongly heated and cooled over the space of a few hours [9]. The heating melted the chondrules but was not sufficiently protracted to allow all the more volatile elements such as sulfur to escape. Particular types of meteorite contain chondrules with distinctive sizes and compositions. This may mean that chondrules formed in small regions of the protoplanetary nebula in a series of separate events. Theories abound for the origin of chondrules [9]. Models in which dust is melted by shock waves in the nebula are currently in vogue [10,11], although the source of these shock waves is unclear. Chondrule formation may be intimately tied to other events in the Solar System. In particular, shock driven chondrule formation could require the early formation of Jupiter [11,12]. A small fraction of chondrules appear to have formed as a result of impacts on asteroids [13], which implies that large bodies had already accreted by the time these chondrules formed.

Primitive meteorites also contain refractory components, similar in size to chondrules. These *calcium aluminium rich inclusions* (CAIs) are a minor component of meteorites, and their origin is even more enigmatic than that of chondrules. They are important here because their age places timing constraints on planet formation. The relative ages of CAIs and chondrules can be estimated from the abundances of the decay products of short-lived isotopes. Particularly useful are, ^{41}Ca, ^{26}Al and ^{53}Mn, with half lives of 0.1, 0.7 and 3.7 Myr respectively [14]. The absolute ages of chondrules and CAIs can be calculated from the modern abundances of lead isotopes formed by the decay of long-lived isotopes of uranium [15].

According to the isotopes, CAIs are the oldest Solar System materials we possess. Most formed in an interval spanning only a few hundred thousand years [16] around 4.56 billion years ago [17]. Chondrules apparently formed 1–4 Myr later than this [17,18]. Thus, some CAIs survived in the nebula for millions of years before bedding down with much younger chondrules to form asteroids. It is conceivable that something similar happened in the region containing the terrestrial planets, in which case the early stages of planet formation spanned several million years at least.

Iron meteorites tell their own tale. These meteorites come from asteroids that became hot enough to melt and differentiate. The most plausible source of heat was the decay of short-lived isotopes, especially ^{26}Al. Melting must have occurred while was still abundant, which means these asteroids took something like 2 Myr to form [19,20]. Why did some asteroids melt when others did not? Presumably, different stages of planet and asteroid formation occurred concurrently in the same region of the nebula. Some objects formed earlier than others, and their subsequent thermal evolution was different as a result.

The terrestrial planets are also differentiated, with high density iron-rich cores and low density silicate-rich mantles. Earth's mantle is highly depleted in *siderophile* (iron loving) elements, when ratioed to silicon say, compared to the Sun. Presumably, these elements sank to the core along with the iron during core formation. The process of core formation is hard to disentangle from the process of accretion itself. It is likely that the two happened concurrently [21]. The time scale for core formation can be estimated using the decay of U isotopes to Pb, and also the short-lived isotope ^{182}Hf, which decays to ^{182}W with a half-life of 9 Myr. These isotope systems are useful because the parent nuclei are *lithophile* (silicate loving) while the daughter isotopes are more siderophile. Assuming core formation happened continuously and that accretion tailed off roughly exponentially over time, the lead isotopes indicate that Earth accreted/differentiated with a mean life of 15–40 Myr [22]. Somewhat confusingly, the Hf–W isotopes provide a shorter mean life of about 11 Myr [23]. The reason why these two systems give different results is unclear and the actual time scale probably lies somewhere in between [21].

The inner planets would have been mostly molten at the time they differentiated. Unfortunately, this melting erased much information about what happened to these bodies before their cores formed. We know rather more about subsequent events. In particular, Earth's mantle is blessed with more siderophile elements (gold etc.) than one would expect to find after its core formed. This is consistent with continued growth of the Earth after core formation ceased, although this *late veneer* constitutes less than 1% of Earth's total mass [24]. The mixture of osmium isotopes we see in Earth's mantle differs from carbonaceous chondrite meteorites (probably from the outer asteroid belt) but is similar to ordinary chondrites (probably from the inner belt) [24]. This is consistent with the late veneer coming either from material in the inner asteroid belt, or from the terrestrial-planet region itself.

Collisions shaped the inner Solar System in several ways. High speed impacts onto planetary surfaces supplied enough kinetic energy to cause melting. On small bodies, melted material tended to escape to space. Bodies the size of Ceres and larger were massive enough to hang on to some molten material, and impacts onto planet-sized bodies probably caused enough melting to trigger core formation [25]. The high density of Mercury may be the result of a violent collision with another large body, which removed much of Mercury's silicate mantle [26]. The Moon is highly depleted in both iron and volatile elements. This makes sense if the Moon formed from hot mantle material thrown into orbit around Earth after the planet was hit by another differentiated body [27]. The ancient surfaces of the Moon, Mercury and Mars bear the scars of numerous smaller impacts, although on the Moon at least, these collisions happened hundreds of millions of years after the planets formed [28].

Meteorites, together with rocks from Mars, the Moon and Earth, generally contain a similar mixture of isotopes, unlike dust grains that formed outside the Solar System. This suggests that material in the inner Solar System was thoroughly mixed on very fine scales at some point [29]. (Isotopes of oxygen do not obey this rule for reasons that are hotly debated.) Earth, Mars and the parent bodies of the various meteorite groups differ substantially in their chemistry however. Each object is made up of a different mixture of the major rock forming elements [24]. The spectral characteristics of modern asteroids vary in ways that are correlated with their distance from the Sun [30], and this is widely interpreted to reflect (literally) differences in their composition, as well as their thermal evolution. Hence, although different regions of the nebula probably exchanged a good deal of material, each of the inner planets and asteroids ultimately acquired a mixture of material unique to that body.

Compared to the Sun, many primitive meteorites are depleted in moderately volatile elements—those elements that condense and evaporate at temperatures between about 650 and 1250 K. The curious aspect is that these elements are depleted in a way that roughly correlates with their condensation temperature [31]. There is more than one way to interpret this correlation. The depletion pattern may represent a distant memory of an early hot phase in the history of the nebula [32], rather as the cosmic microwave background radiation provides a glimpse of the early history of the universe. If this interpretation is right, the inner few astronomical units of the nebula must have been hot enough to vapourize rock at some point. The CI group of primitive meteorites are not depleted in moderately volatile elements, which may mean they come from bodies that formed further from the Sun where temperatures were cooler. The depletion of moderately volatile elements can be interpreted in another way. Rather than indicating a globally hot nebula, the depletions could be caused by localized events such as those that generated chondrules [31]. In either case, planet-sized bodies probably suffered further depletions as a result of energetic collisions [33].

The inner planets possess rather little in the way of highly volatile material such as water and the noble gases. This depletion can be attributed to high temperatures in the inner nebula-volatile materials simply didn’t condense while the planets were forming. However, the isotopic mixture of xenon on Earth and Mars implies that these planets have lost almost all their initial allotment of noble gases, and possibly a lot of other volatile material too [34]. The origin of Earth’s meagre volatile inventory is still unclear. Helium and neon leaking from the mantle hint that the planet might have captured a massive atmosphere directly from the nebula early in its history [35]. Volatile substances in the atmosphere could have entered the mantle while the planet was still molten. It is hard to explain the abundances of the other noble gases in the atmosphere today unless some volatiles came to Earth from another source, such as impacts by comets or asteroids [36,37]. The deuterium to hydrogen ratio in terrestrial seawater differs by a factor of two from the ratio measured in comets to date, so comets were probably less important in this respect than asteroids [38].

3. The planetesimal theory

The astronomical and cosmochemical data described above generally support a model for the formation of the planets known as the *planetesimal theory*. Crudely, the theory posits that dust grains in the nebula collided and stuck together to make aggregates; these collided to form bigger bodies, etc. until the largest objects were the size of planets. Objects tended to acquire most of their mass locally, so planets and asteroids forming in different regions of the nebula came to have somewhat different compositions. Once planet-sized bodies formed, it was mostly a matter of mopping up the remaining debris or removing it from the system. Something along these lines almost certainly happened in the inner Solar System, leading to the formation of Earth and the other terrestrial planets. As with all theories however, the devil is in the details. This is particularly apparent in the early stages of planetary accretion.

The planetesimal theory is often portrayed as a sequence of steps rather like acts in a play. The characters and scenery change with time, and the audience applauds at the end of each act. This division is partly an indication of how we think about complex problems, but it also reflects changes in the relative importance of physical processes at each stage of planetary accretion. Following convention, I will describe the stages in order, with the caveat that these stages probably overlapped in both time and space.

4. Formation of planetesimals

The story begins with the gas and dust of the Sun's protoplanetary disk. Gas pressure gives the disk a definite thickness in the vertical direction. Pressure decreases with distance from the Sun, which allows the gas to orbit the Sun slightly more slowly than a solid body moving on a circular orbit would. Dust grains feel little pressure support, so they tend to settle towards the midplane of the disk, sweeping up other grains en route to form loosely bound aggregates (see Fig. 1). In the absence of turbulence in the gas, a typical dust grain will reach the midplane in about 10^4 years [39].

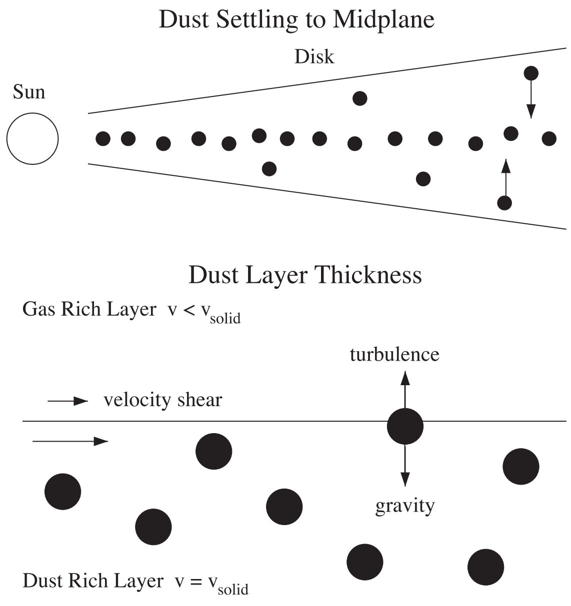

Fig. 1. Dust grains slowly settle to the midplane of the nebula due to the vertical component of the Sun's gravity, forming a solid-rich layer. This layer orbits the Sun slightly faster than the gas-rich layers above and below. The resulting wind shear generates turbulence, even if other sources of turbulence are absent. Thus, the solid-rich layer has a finite thickness.

As the dust becomes concentrated towards the midplane, the solid to gas ratio increases. The dust-rich layer begins to orbit the Sun with the speed of a solid body rather than the slightly slower speed of the gas. Gas in the dusty layer is herded along by the solid particles, moving faster than it would like to, while gas above and below the midplane moves more slowly as before. This differential velocity generates turbulence, which acts to puff up the dust-rich layer, even if no other sources of turbulence are present. A compromise is reached between gravity and turbulence, and this determines the thickness of the dust-rich layer [40].

At this point, the script for our play becomes hard to read. If the solid-rich layer becomes dense enough, the densest portions will be unable to resist the Narcissus-like attraction of their own gravity, becoming ever smaller and denser. Once this *gravitational instability* gets going, collapse can continue until solid bodies a few km in size are generated. Such bodies are dubbed *planetesimals*. Whether gravitational instability (GI) ever gets going is the subject of much debate [41–44]. It seems that GI can only occur if the ratio of solids to gas in a column of nebula material exceeds a critical value. Recent calculations suggest this critical value is several times the solid to gas ratio for material with solar composition, even when volatiles such as water and carbon monoxide have condensed [45]. In addition, bodies may have to grow to metre size or larger before conditions become right for GI [40].

Unabashed, theorists have thought of several ways to increase the solid to gas ratio and give GI a helping hand. This is done by either collecting solids in one place or removing some of the gas. Small solids have a tendency to move radially within the nebula and pile up at locations where there is a local maximum in the gas pressure [46], or where the concentration of solids is higher than average [47]. The local solid to gas ratio can also increase over time as small solids migrate inwards [42], or as gas escapes from the Solar System due to *photoevaporation* by ultraviolet light from the Sun. It remains to be seen whether these mechanisms operate with sufficient effect to permit GI.

In the absence of gravitational instability, large bodies presumably form by the gradual aggregation of dust grains and small solids such as chondrules. Experiments show that irregular dust grains can stick together if they collide at speeds of up to a few tens of metres per seconds [48]. High collision speeds are more likely to cause grains to rebound or break apart, while low collision speeds are more likely to lead to sticking. Small solids are probably strongly coupled to the motion of the gas, so they typically undergo gentle collisions leading to accretion. Charge exchange between grains and the generation of electric dipoles also aids accretion, leading to the rapid formation of dust aggregates many centimetres in size [49].

Collisional accretion becomes more challenging when bodies reach 0.1 – 10 m in size. These objects are too large to be swept along at the same speed as the gas, but too small to be unaffected by it entirely. Because gas orbits the Sun more slowly than a solid body, boulder-sized objects feel a headwind. If solids in the dust-rich layer are effective at dragging the gas along with them, this headwind will be quite small [40]. Otherwise, the headwind will be around 50 m s^{-1}, similar to the wind speed in a hurricane [50].

The headwind affects boulder-size objects in two ways. Dust grains entrained in the gas strike large bodies with the same speed as the headwind. In principle, this increases the amount of material that can be swept up by large bodies. However, if the dust grains hit at high speeds they are more likely to cause erosion akin to sand blasting. Second, the headwind gradually robs large bodies of their orbital angular momentum, causing them to drift towards the Sun. This drifting due to *gas drag* can be extremely rapid for metre-sized objects—as fast as 1 AU in 500 years [40]. The ultimate fate of drifting bodies depends on the thermal structure of the nebula. If the inner nebula is hot, objects will evaporate when the temperature becomes high enough; otherwise, they fall into the Sun. Radial transport of solid material by gas drag may lead to substantial variations in both the solid to gas ratio and the chemical composition in different regions of the nebula [51].

The existence of gas drag would imply that solid bodies must grow rapidly until they are many metres in size if they are to survive. It seems reasonable that boulder-sized objects will only actually stick together during rare, low-velocity collisions. These objects may gain most of their mass by accreting smaller solids and dust grains. The nebula headwind might aid accretion in some cases by blowing small fragments from erosive collisions back onto metre-size bodies [52].

If the nebular gas is turbulent, small solids will not simply accumulate in a thin layer at the midplane. However, solids will become highly concentrated in stagnant regions. These solid-rich regions could evolve rapidly into planetesimals [53]. The efficiency of this *turbulent concentration* depends on the size and compactness of the solids. Chondrule-like particles seem particularly well suited in this respect [53], so it may be no accident that they form the major component of most primitive meteorites, while larger solid particles are not seen.

Despite substantial progress in understanding the earliest stage of planet formation, the origin of planetesimals must still be regarded as an unsolved problem. The audience watching our play could be forgiven for having serious misgivings at this stage in the performance. Fortunately, things proceed more smoothly in the next act.

5. Formation of planetary embryos

The second stage of planet formation begins when much of the solid material has formed into planetesimals a few kilometres in diameter. How these bodies formed is rather less important than how big they are. For the second stage to proceed, bodies must be large enough to gravitationally perturb their neighbours during close approaches. This stage of accretion has been examined extensively using theoretical models for two reasons: (i) the evolution depends on a small number of processes that are fairly well understood, and (ii) the number of planetesimals is huge. This means their evolution can be studied in a statistical sense, just as a gas composed of trillions of molecules can be described using kinetic theory.

The weakest link in the theory is understanding the outcome of collisions. Laboratory experiments have studied impacts involving planetary materials at a wide range of collision speeds, but these experiments are limited to bodies less than a metre in size. Most of what we know about planet-forming collisions comes from numerical simulations instead. To date, these simulations provide a rather sparse coverage of collisional phase space. That said, the results suggest most collisions lead to net accretion, unless the impact speed is subtantially higher than the target’s gravitational escape velocity or the impact is at a grazing angle [54,55].

A planetesimal accretes its smaller brethren at a rate that depends on the number of objects per unit volume and the planetesimal's velocity v_{rel} relative to other objects (see Fig. 2). If v_{rel} is large, a planetesimal collides only with objects that pass directly in front of it. If v_{rel} is small, a planetesimal's gravity will pull in material from further away. This *gravitational focussing* increases the frequency of collisions. More often than not, planetesimals approach each other without actually colliding, but their trajectories are altered by their gravitational interaction. The cumulative effect of many close encounters determines a planetesimal's velocity relative to other bodies in the same region of the nebula. Large bodies tend to acquire small relative velocities and vice versa, a state of affairs referred to as *dynamical friction*. All the while, gas drag is striving to make the orbits of the planetesimals circular and coplanar, effectively reducing v_{rel}.

Despite this apparent complexity, accretion is likely to proceed in one of only a few ways [56–58]. Initially, the largest planetesimals feed voraciously on smaller objects, while the collective gravitational effects of the small objects keeps v_{rel} low. This makes gravitational focussing highly effective. The largest bodies, termed *planetary embryos*, quickly outgrow all the others, a process known as *runaway growth*.

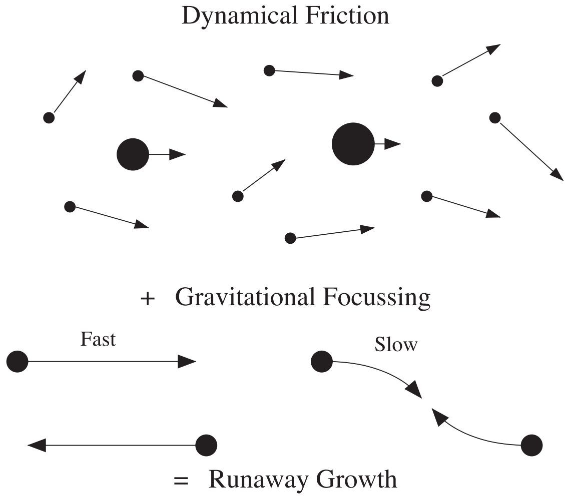

Fig. 2. The mechanics of runaway growth. Large bodies tend to have lower relative velocities than small objects as a result of numerous gravitational encounters. When large bodies pass close to each other, their trajectories are focussed by their gravitational attraction. Small bodies fly past each other too quickly to be significantly affected by their mutual attraction. Thus, large bodies grow faster than small ones.

The days of unfettered growth are numbered however. Runaway growth slows when planetary embryos become about 100 times more massive than a typical planetesimal. Now it is the gravitational perturbations of the embryos that determine (v_{rel}) rather than perturbations from the more numerous planetesimals [59]. Accretion enters a new self-regulated regime called *oligarchic growth* [55]. Planetary embryos continue to outgrow smaller planetesimals, but embryos in neighbouring regions of the disk are forced to grow at similar rates. Whenever one embryo gets too greedy, events conspire to allow nearby embryos to catch up. The more massive an embryo is, the more strongly it perturbs nearby planetesimals, thereby increasing v_{rel}. Thus, gravitational focussing is reduced, and the embryo grows more slowly than a smaller embryo would.

As in any good oligarchy, each embryo stakes out a region of influence, or *feeding zone* in the disk. A typical feeding zone in the inner Solar System is a roughly annular region of order 0.01 AU in width. A combination of dynamical friction and occasional gravitational interactions between neighbouring embryos acts to keep these bodies on widely spaced orbits. Each embryo accretes most of its mass from its own feeding zone, giving the embryos distinct chemical compositions.

The oligarchic growth stage lasts for 0.1–1 Myr from the time when planetesimals first appear in large numbers [58,60]. Oligarchic growth ends when the number of planetesimals dwindles so much that they can no longer restrain the actions of the planetary embryos. Our play has reached a moment of crisis. With the demise of the planetesimals, dynamical friction shuts down. The embryos stray beyond their feeding zones and the previous order collapses as the large bodies begin to interact strongly and collide with one another. Accretion of the planets now enters a prolonged terminal phase.

6. From embryos to planets

The final stage of planetary accretion involves a few dozen embryos with masses comparable to the Moon or Mars (0.01–0.1 Earth masses). Gravitational perturbations between embryos increase their relative velocities. Gravitational focussing becomes weak, and the accretion rate slows dramatically.

Over time, the embryos scatter one another inwards or outwards, and the radial ordering established during oligarchic growth becomes scrambled. Any primordial chemical and isotopic gradients are blurred as a result. The final planets are a mixture of material from a broad region of the inner Solar System, although each planet accretes more material from its own locale than elsewhere, so the mixture is different for each planet [61]. Earth and Venus are composites formed from a dozen or more embryos. Mars and Mercury contain material from only a few embryos, possibly as few as one in each case. Thus, these planets probably sampled rather less of the nebula than their larger siblings. The final stage of accretion is highly *chaotic*. That is, it depends sensitively on the outcome of individual events such as whether a close encounter between two embryos results in a collision or a near miss. To illustrate this, Fig. 3 shows the results of four numerical simulations of this stage of accretion, each beginning with the same total mass and number of embryos [62]. The planetary systems produced in each case are quite different.

The time scale for the final stage of accretion depends on whether nebula gas is still present. In the presence of a minimum-mass gas nebula, Earth may have formed in as little as 5 Myr [63], although this is hard to reconcile with the time scales derived from the U–Pb and Hf–W isotopes described above. In the absence of significant amounts of gas, numerical simulations suggest Earth took roughly 100 Myr to form, with the accretion rate declining approximately exponentially over time [62,64]. Small amounts of nebula gas can have a significant effect on late-stage accretion. In particular, the lingering presence of roughly 0.1% of a minimum-mass nebula may have helped to circularize the orbits of the inner planets as they neared completion [65,66].

The inner planets probably each accreted some planetesimals from the region that now contains the asteroid belt. These planetesimals have a different chemical composition, and are probably richer in volatile materials, than planetesimals that formed closer to the Sun. The asteroid belt may have been an important source of the water and other volatile substances that now exist on Earth [37]. Some primitive meteorites contain up to 10% water by mass, and

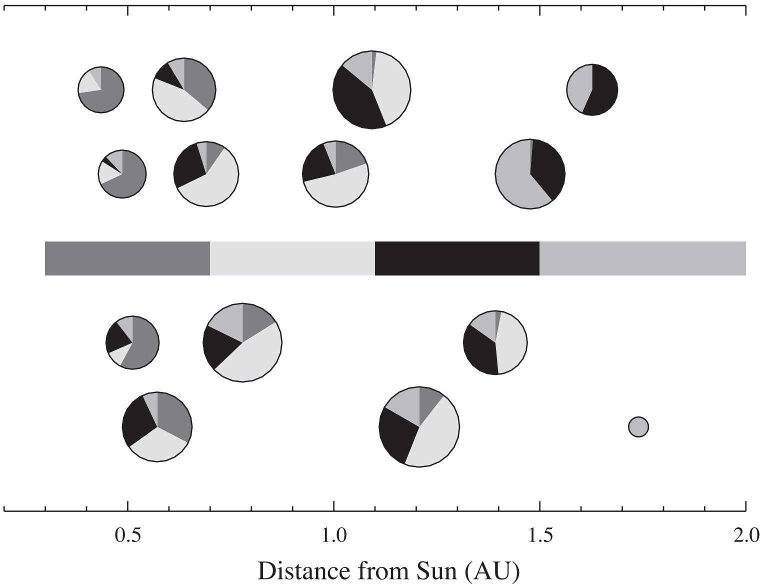

Fig. 3. The results of four numerical simulations of the final stage in the accretion of the inner planets. Each row of symbols shows one simulation, with symbol radius proportional to the radius of the planet. The segments in each pie chart show the fraction of material originating from each of the four zones of the nebula indicated by the shaded rectangles. In each simulation, the largest planet has a mass similar to Earth (results taken from Ref. [62]).

this water has a deuterium to hydrogen ratio similar to water on Earth. However, the bulk of this asteroidal matter must have arrived on Earth, or its precursors, before core formation was complete. These meteorites are rich in siderophile elements, so they would dominate the Earth's late veneer if they were added late, yet their Os isotopic and trace element compositions are distinct [24].

Late stage accretion is not a wholly efficient business. Some embryos fall into the Sun after straying into the asteroid belt, a region that contains unstable orbital resonances associated with Jupiter and Saturn. Up to 1/3 of the embryos that form within 2 AU of the Sun are likely to suffer this fiery fate [62]. High speed collisions between embryos can lead to fragmentation. Collision speeds are highest close to the Sun, which made Mercury especially vulnerable to disruptive impacts. This may explain why the innermost planet remained so small. The low mass of Mars compared to Earth and Venus is harder to fathom, and current theories have little to say on the subject. Chance events may have conspired to prevent Mars from accreting additional planetary embryos, but this explanation seems unsatisfactory. A significant fraction of collisions between embryos are likely to eject a substantial amount of material into orbit around the newly formed body [64]. It is likely that the Moon formed from such debris following the impact of a Mars-sized body onto Earth [27].

7. Formation of the asteroid belt

The story of planetary accretion 1 AU from the Sun has a happy ending but something clearly went wrong in the asteroid belt. Either planets never formed in this region or they survived only briefly. The imprint of short-lived isotopes seen in many meteorites implies the asteroids formed in a few million years. However, bodies the size of Ceres and Vesta can only have accreted this rapidly if the asteroid belt initially contained at least 100 times as much solid material as it does today [67], so the current low mass of the asteroid belt requires an explanation.

It is possible that the growth of large bodies in the asteroid belt was frustrated by the early formation of Jupiter. Gravitational perturbations from Jupiter would have increased the relative velocities of planetesimals in general, especially for bodies that are substantially different in size. This would delay the onset of runaway growth until bodies were larger than the largest asteroids that exist today [68]. Collisional erosion may have been an important process in the asteroid region, but this alone cannot explain the low mass of the asteroid belt today. Vesta currently sports a basaltic crust that formed in the first few million years of the Solar System. It is doubtful that this crust would have survived until now if >99% of the asteroid belt has been pummelled into dust [69]. Instead, the asteroid belt probably lost most of its bulk in another way.

Two other models are currently in the running. Each makes use of the fact that the asteroid belt is crisscrossed by a number of unstable orbital resonances associated with Jupiter and Saturn. Today, an asteroid entering a resonance is quickly forced onto a highly eccentric (elliptical) orbit, such that it typically falls into the Sun or is ejected from the Solar System in about 1 Myr [70]. While the nebula is still present, objects moving on eccentric orbits experience substantial gas drag. As the nebula disperses, some of the unstable resonances sweep across the asteroid belt, possibly more than once [71]. The combination of *resonance sweeping* and gas drag causes many bodies smaller than about 100 km to migrate inwards, leaching mass from the asteroid belt, and depositing it in the region where the terrestrial planets are forming [72].

Gas drag has its limits however. Bodies the size of the Moon or Mars are too massive to drift significantly. Left to their own devices, planetesimals in the asteroid region should eventually form planetary embryos, unless the giant planets form quickly. However, even planetary embryos as massive as Earth become vulnerable once Jupiter and Saturn form. Gravitational encounters between embryos cause their orbits to wander slowly through asteroid belt. Sooner or late, each embryo enters an unstable resonance where it is likely to be removed from the asteroid belt before another close encounter scatters it out of the resonance again. Numerical simulations show that all embryos in the asteroid region are likely to be lost in this way, along with the great majority of smaller asteroids, once the giant planets form [73]. All that remain are a few small objects too puny to continue the process any further.

There appears to be a fine line between forming a system of terrestrial planets and generating an asteroid belt. The outcome depends mainly on proximity to the giant planets and their unstable resonances [74,75]. When terrestrial planets do form, their characteristics are hard to predict ahead of time since the final stage of accretion is dominated by chance events involving a small number of planetary embryos. It is probably a matter of chance that the Solar System ended up with precisely four terrestrial planets, and that one of these now resides in a pleasantly habitable location [62]. A minor change at any stage in the formation of the planets could have produced a very different, perhaps equally fascinating, outcome.

8. Looking ahead

The coming years should see progress in a number of areas that will help our understanding of the origin of planets and asteroids. Astronomical observations of protoplanetary disks by NASA's Spitzer Space telescope and other programmes will provide better models for the structure and evolution of protoplanetary disks. Continuing searches for extrasolar giant planets will soon establish whether giants with orbits similar to Jupiter are rare or commonplace. There is also the exciting prospect of finding terrestrial planets orbiting Sun-like stars via NASA's upcoming Kepler mission. In cosmochemistry, we may soon see the resolution of several key questions including a clear understanding of the source(s) of short-lived isotopes in the early Solar System, and agreement on the time scale for Earth's accretion and differentiation, and the timing of the Moon-forming impact. The hugely successful ongoing programme to collect and analyse Antarctic meteorites is sure to throw up a few surprises in the years ahead. Finally, on the theoretical front, the time is ripe for breakthroughs on a number of vexing issues, including the origin of chondrules and planetesimals, an understanding of the physics of interactions between planets and protoplanetary disks, and the origin of water and other volatiles on the terrestrial planets. We have much to look forward to.

Acknowledgements

I am very grateful to Alan Boss, Lindsey Bruesch, Jeff Cuzzi, Alex Halliday, Helen Williams, Kevin Zahnle and an anonymous reviewer for providing comments that substantially improved this article and helped to avert a number of gaffs during its preparation. ***[AH]***

References

[1] M. Konacki, A. Wolszczan, Masses and orbital inclinations of planets in the PSR B1257+12 system, Astrophysical Journal 591 (2003) L147–L150.

[2] S. Seager, The search for extrasolar earth-like planets, Earth and Planetary Science Letters 208 (2003) 113–124.

[3] S.V.W. Beckwith, A.I. Sargent, R.S. Chini, R. Guesten, A survey for circumstellar disks around young stellar objects, Astronomical Journal 99 (1990) 924–945.

[4] D.A. Fischer, J.A. Valenti, Metallicities of stars with extrasolar planets, in: D. Deming, S. Seager (Eds.), Scientific Frontiers in Research on Extrasolar Planets, ASP Conference Series, vol. 294, ASP, San Francisco, 2003, pp. 117–128.

[5] K.E. Haisch, E.A. Lada, C.J. Lada, Disk frequencies and lifetimes in young clusters, Astrophysical Journal 553 (2001) L153–L156.

[6] C. Spangler, A.I. Sargent, M.D. Silverstone, E.E. Becklin, B. Zuckerman, Dusty debris around solar-type stars: temporal disk evolution, Astrophysical Journal 555 (2001) 932–944.

[7] J.S. Greaves, I.M. Coulson, W.S. Holland, No molecular gas around nearby solar-type stars, Monthly Notices of the Royal Astronomical Society 312 (2000) L1–L3.

[8] P. Artymowicz, M. Clampin, Dust around main sequence stars: nature or nurture by the interstellar medium, Astrophysical Journal 490 (1997) 863–878.

[9] R.H. Jones, T. Lee, H.C. Connolly, S.G. Love, H. Shang, Formation of chondrules and CAIs: theory vs. observation, in: V. Mannings, A.P. Boss, S.S. Russell (Eds.), Protostars and Planets vol. IV, University of Arizona Press, Tuscon, AZ, 2000, pp. 927–961.

[10] S.J. Desch, H.C. Connolly, A model of the thermal processing of particles in solar nebula shocks: application to the cooling rates of chondrules, Meteoritics 37 (2002) 183–207.

[11] S.J. Weidenschilling, F. Marzari, L.L. Hood, The origin of chondrules at jovian resonances, Science 279 (1998) 681–684.

[12] A.P. Boss, Shock-wave heating and clump formation in a minimum mass solar nebula, 31st Lunar Planetary Science Conference, Houston, Texas, 2000, abstract 1084.

[13] A.N. Krot, A.E. Rubin, Chromite-rich mafic silicate chondrules in ordinary chondrites: formation by impact melting, 24th Lunar Planetary Science Conference, Houston, Texas, 1993, pp. 827–828.

[14] J.N. Goswami, H.A.T. Vanhala, Extinct radionuclides and the origin of the Solar System, in: V. Mannings, A.P. Boss, S.S. Russell (Eds.), Protostars and Planets IV, University of Arizona Press, Tucson, 2000, p. 963.

[15] C.J. Allegre, G. Manhes, C. Gopel, The age of the Earth, Geochimica et Cosmochimica Acta 59 (1995) 1445–1456.

[16] M. Wadhwa, S.S. Russell, Timescales of accretion and differ-

entiation in the early solar system: the meteoritic evidence, in: V. Mannings, A.P. Boss, S.S. Russell (Eds.), Protostars and Planets IV, University of Arizona Press, Tucson, 2000, p. 995.

[17] Y. Amelin, A.N. Krot, I.D. Hutcheon, A.A. Ulyanov, Lead isotopic ages of chondrules and calcium–aluminum-rich inclusions, Science 297 (2002) 1678–1683.

[18] G.R. Huss, G.J. MacPherson, G.J. Wasserburg, S.S. Russell, G. Srinivasan, 26Al in CAIs and chondrules from unequilibrated ordinary chondrites, Meteoritics 36 (2001) 975–997.

[19] D.S. Woolum, P. Cassen, Astronomical constraints on nebular temperatures: implications for planetesimal formation, Meteoritics 34 (1999) 897–907.

[20] N. Sugiura, H. Hoshino, Mn–Cr chronology of five IIIAB iron meteorites, Meteoritics 38 (2003) 117–143.

[21] A.N. Halliday, Mixing, volatile loss and compositional change during impact-driven accretion of the Earth, Nature 427 (2004) 505–509.

[22] A.N. Halliday, Terrestrial accretion rates and the origin of the Moon, Earth and Planetary Science Letters 176 (2000) 17–30.

[23] Q. Yin, S.B. Jacobsen, K. Yamashita, J. Blichert-Toft, P. Telouk, A short timescale for terrestrial planet formation from Hf–W chronometry of meteorites, Nature 418 (2002) 949–952.

[24] M.J. Drake, K. Righter, Determining the composition of the Earth, Nature 416 (2002) 39–44.

[25] W.B. Tonks, H.J. Melosh, Core formation by giant impacts, Icarus 100 (1992) 326–346.

[26] W. Benz, W.L. Slattery, A.G.W. Cameron, Collisional stripping of Mercury's mantle, Icarus 74 (1988) 516–528.

[27] R.M. Canup, E. Asphaug, Origin of the Moon in a giant impact near the end of Earth's formation, Nature 412 (2001) 708–712.

[28] W.K. Hartmann, G. Ryder, L. Dones, D. Grinspoon, The time-dependent intense bombardment of the primordial Earth/Moon system, in: R.M. Canup, K. Righter (Eds.), Origin of the Earth and Moon, University of Arizona Press, Tucson, 2000, pp. 493–512.

[29] H. Becker, R.J. Walker, Efficient mixing of the solar nebula from uniform Mo isotopic composition of meteorites, Nature 425 (2003) 152–155.

[30] J. Gradie, E. Tedesco, Compositional structure of the asteroid belt, Science 216 (1982) 1405–1407.

[31] C.M.O'D. Alexander, A.P. Boss, R.W. Carlson, The early evolution of the inner solar system: a meteoritic perspective, Science 293 (2001) 64–68.

[32] P. Cassen, Nebular thermal evolution and the properties of primitive planetary materials, Meteoritics 36 (2001) 671–700.

[33] A.N. Halliday, D. Porcelli, In search of lost planets—the paleocosmochemistry of the inner Solar System, Earth and Planetary Science Letters 192 (2001) 545–559.

[34] K. Zahnle, Origins of atmospheres, in: C.E. Woodward, J.M. Shull, H.A. Thronson Jr. (Eds.), Proceedings of the International Conference at Estes Park, Colorado, 19–23 May, 1997.

[35] D. Porcelli, D. Woolum, P. Cassen, Deep Earth rare gases: initial inventories, capture from the solar nebula, and losses during Moon formation, Earth and Planetary Science Letters 237 (2001) 237–251.

[36] N. Dauphas, The dual origin of the terrestrial atmosphere, Icarus 165 (2003) 326–339.

[37] A. Morbidelli, J. Chambers, J.I. Lunine, J.M. Petit, F. Robert, G.B. Valsecchi, K.E. Cyr, Source regions and time scales for the delivery of water to Earth, Meteoritics 35 (2000) 1309–1320.

[38] F. Robert, The origin of water on Earth, Science 293 (2001) 1056–1058.

[39] S.J. Weidenschilling, Dust to planetesimals, Icarus 44 (1980) 172–189.

[40] J.N. Cuzzi, A.R. Dobrovolskis, J.M. Champney, Particle gas dynamics in the midplane of a protoplanetary nebula, Icarus 106 (1993) 102–134.

[41] S.J. Weidenschilling, J.N. Cuzzi, Formation of planetesimals in the solar nebula, in: E.H. Levy, J.I. Lunine (Eds.), Protostars and Planets III, University of Arizona, Tucson, 1993, pp. 1031–1060.

[42] A.N. Youdin, F.H. Shu, Planetesimal formation by gravitational instability, Astrophysical Journal 580 (2002) 494–505.

[43] W.R. Ward, On planetesimal formation: the role of collective particle behaviour, in: R.M. Canup, K. Righter (Eds.), Origin of the Earth and Moon, University of Arizona, Tucson, 2000, pp. 75–84.

[44] S.J. Weidenschilling, Radial drift of particles in the solar nebula: implications for planetesimal formation, Icarus 165 (2003) 438–442.

[45] M. Sekiya, Quasi-equilibrium density distributions of small dust aggregations in the solar nebula, Icarus 133 (1998) 298–309.

[46] N. Haghighipour, A.P. Boss, On pressure gradients and rapid migration of solids in a nonuniform solar nebula, Astrophysical Journal 583 (2003) 996–1003.

[47] J. Goodman, B. Pindor, Secular instability and planetesimal formation in the dust layer, Icarus 148 (2000) 537–549.

[48] T. Poppe, J. Blum, T. Henning, Analogous experiments on the stickiness of micron sized preplanetary dust, Astrophysical Journal 533 (2000) 454–471.

[49] J. Marshall, J. Cuzzi, Electrostatic enhancement of coagulation in protoplanetary nebulae, 32nd Lunar Planetary Science Conference, Houston, Texas, 2001, abstract 1262.

[50] S.J. Weidenschilling, Aerodynamics of solid bodies in the solar nebula, Monthly Notices of the Royal Astronomical Society 180 (1977) 57–70.

[51] T.F. Stepinski, P. Valageas, Global evolution of solid matter in turbulent protoplanetary disks, Astronomy and Astrophysics 319 (1997) 1007–1019.

[52] G. Wurm, J. Blum, J.E. Colwell, A new mechanism relevant to the formation of planetesimals in the solar nebula, Icarus 151 (2001) 318–321.

[53] J.N. Cuzzi, R.C. Hogan, J.M. Paque, A.R. Dobrovolskis, Size-selective concentration of chondrules and other small particles in protoplanetary nebula turbulence, Astrophysical Journal 546 (2001) 496–508.

[54] W. Benz, E. Asphaug, Catastrophic disruptions revisited, Icarus 142 (1999) 5–20.

[55] Z.M. Leinhardt, D.C. Richardson, T. Quinn, Direct N-body simulations of rubble pile collisions, Icarus 146 (2000) 133–151.

[56] E. Kokubo, S. Ida, Oligarchic growth of protoplanets, Icarus 131 (1998) 171–178.

[57] R.R. Rafikov, The growth of planetary embryos: orderly, runaway or oligarchic? Astronomical Journal 125 (2003) 942–961.

[58] E.W. Thommes, M.J. Duncan, H.F. Levison, Oligarchic growth of giant planets, Icarus 161 (2003) 431–455.

[59] S. Ida, J. Makino, Scattering of planetesimals by a protoplanet—slowing down of runaway growth, Icarus 106 (1993) 210–227.

[60] S.J. Weidenschilling, D. Spaute, D.R. Davis, F. Marzari, K. Ohtsuki, Accretional evolution of a planetesimal swarm, Icarus 128 (1997) 429–455.

[61] J.E. Chambers, P. Cassen, The effects of nebula surface density profile and giant-planet eccentricities on planetary accretion in the inner solar system, Meteoritics 37 (2002) 1523–1540.

[62] J.E. Chambers, Making more terrestrial planets, Icarus 152 (2001) 205–224.

[63] C. Hayashi, K. Nakazama, Y. Nakagawa, Formation of the Solar System, in: Protostars and Planets II, University of Arizona, Tucson, 1985, pp. 1100–1153.

[64] C.B. Agnor, R.M. Canup, H.F. Levison, On the character and consequences of large impacts in the late stage of terrestrial planet formation, Icarus 142 (1999) 219–237.

[65] J. Kominami, S. Ida, The effect of tidal interaction with a gas disk on formation of terrestrial planets, Icarus 157 (2002) 43–56.

[66] C.B. Agnor, W.R. Ward, Damping of terrestrial-planet eccentricities by density-wave interactions with a remnant gas disk, Astrophysical Journal 567 (2002) 579–586.

[67] G.W. Wetherill, An alternative model for the formation of the asteroids, Icarus 100 (1992) 307–325.

[68] S.J. Kortenkamp, G.W. Wetherill, Runaway growth of planetary embryos facilitated by massive bodies in a protoplanetary disk, Science 293 (2001) 1127–1129.

[69] D.R. Davis, E.V. Ryan, P. Farinella, Asteroid collisional evolution: results from current scaling algorithms, Planetary Science 42 (1994) 599–610.

[70] B.J. Gladman, F. Migliorini, A. Morbidelli, V. Zappala, P. Michel, A. Cellino, C. Froeschle, H.F. Levison, M. Bailey, M. Duncan, Dynamical lifetimes of objects injected into asteroid belt resonances, Science 277 (1997) 197–201.

[71] M. Nagasawa, H. Tanaka, S. Ida, Orbital evolution of asteroids during depletion of the solar nebula, Astronomical Journal 119 (2000) 1480–1497.

[72] F. Franklin, M. Lecar, On the transport of bodies within and from the asteroid belt, Meteoritics 35 (2000) 331–340.

[73] J.E. Chambers, G.W. Wetherill, Planets in the asteroid belt, Meteoritics 36 (2001) 381–399.

[74] G. Laughlin, J. Chambers, D. Fischer, A dynamical analysis of the 47 ursae majoris planetary system, Astrophysical Journal 579 (2002) 455–467.

[75] H.F. Levison, C.B. Agnor, The role of giant planets in terrestrial planet formation, Astronomical Journal 125 (2003) 2692–2713.

John E. Chambers is a research scientist at the SETI Institute, located in Mountain View, California. His work involves theoretical studies of the origin of planetary systems and habitable planets using computer modeling. Chambers received a Ph.D. in physics from the University of Manchester in the United Kingdom in 1995. He has since worked at the Carnegie Institution of Washington in Washington DC, Armagh Observatory in Northern Ireland, and NASA Ames Research Center in Moffett Field, California. He has been with the SETI Institute since 2002.