The observed properties of giant planets, models of their evolution, and observations of protoplanetary disks provide constraints on the formation of gas giant planets. The four largest planets in our solar system contain considerable quantities of hydrogen and helium; these gases could not have condensed into solid planetesimals within the protoplanetary disk. Jupiter and Saturn are mostly hydrogen and helium, but have larger abundances of heavier elements than does the Sun. Neptune and Uranus are primarily composed of heavier elements. The transiting extrasolar planet HD 149026 b, which is slightly more massive than Saturn, appears to have comparable amounts of light gases and heavy elements. The other observed transiting exoplanets are primarily hydrogen and helium, but may contain supersolar abundances of heavy elements. Spacecraft flybys and observations of satellite orbits provide estimates of the gravitational moments of the giant planets in our solar system, which in turn provide information on the internal distribution of matter within Jupiter, Saturn, Uranus, and Neptune. Atmospheric thermal structure and heat flow measurements constrain the interior temperatures of these planets. Internal processes may cause giant planets to become more compositionally differentiated or alternatively more homogeneous; high-pressure laboratory experiments provide data useful for modeling these processes. The preponderance of evidence supports the core nucleated gas accretion model. According to this model, giant planets begin their growth by the accumulation of small solid bodies, as do terrestrial planets. However, unlike terrestrial planets, the giant planet cores grow massive enough to accrete substantial amounts of gas before the protoplanetary disk dissipates. The primary question regarding the core nucleated growth model is under what conditions can planets develop cores sufficiently massive to accrete gas envelopes within the lifetimes of gaseous protoplanetary disks.

1. INTRODUCTION

The two largest planets in our solar system, Jupiter and Saturn, are composed predominantly of hydrogen and helium; these two lightest elements also comprise more than 10% of the masses of Uranus and Neptune. Moreover, most extrasolar planets thus far detected are believed (or known) to be gas giants. Helium and molecular hydrogen do not condense under conditions found in star-forming regions and protoplanetary disks, so giant planets must have accumulated them as gases. Therefore, giant planets must form prior to the dissipation of protoplanetary disks. Optically thick dust disks typically survive for only a few million years (chapters by Briceño et al. and Wadhwa et al.), and protoplanetary disks have lost essentially all their gases by the age of <10^7 yr (chapter by Meyer et al.), implying that giant planets formed on this timescale or less.

Jupiter and Saturn are generally referred to as gas giants, even though their constituents are not gases at the high pressures that most of the material in Jupiter and Saturn is subjected to. Analogously, Uranus and Neptune are frequently referred to as ice giants, even though the astrophysical ices such as H_2O, CH_4, H_2S, and NH_3 that models suggest make up the majority of their mass (Hubbard et al., 1995) are in fluid rather than solid form. Note that whereas H and He must make up the bulk of Jupiter and Saturn because no other elements can have such low densities at plausible temperatures, it is possible that Uranus and Neptune are primarily composed of a mixture of “rock” and H/He.

Giant planets dominate our planetary system in mass, and our entire solar system in angular momentum (contained in their orbits). Thus, understanding giant planet formation is essential for theories of the origins of terrestrial planets, and important within the understanding of the general process of star formation.

The giant planets within our solar system also supported in situ formation of satellite systems. The Galilean satellite system is particularly impressive and may contain important clues to the last stages of giant planet formation (Pollack and Reynolds, 1974; Canup and Ward, 2002; Mosqueira and Estrada, 2003a,b). Ganymede and Callisto are roughly half water ice, and Callisto has most of this ice mixed with rock. It follows that conditions must be appropriate for the condensation of water ice at the location where Ganymede formed, and conditions at Callisto have must have allowed formation of that body on a timescale exceeding about 10^5 yr, so that water ice would not melt and lead to a fully differentiated structure. The more distant irregular satellite systems
of the giant planets may provide constraints on gas in the outer reaches of the atmospheres of giant planets (Pollack et al., 1979; cf. Tsui, 2002).

The extrasolar planet discoveries of the past decade have vastly expanded our database by increasing the number of planets known by more than an order of magnitude. The distribution of known extrasolar planets is highly biased toward those planets that are most easily detectable using the Doppler radial velocity technique, which has been by far the most effective method of discovering exoplanets. These extrasolar planetary systems are quite different from our solar system; however, it is not yet known whether our planetary system is the norm, quite atypical, or somewhere in between.

Nonetheless, some unbiased statistical information can be distilled from available exoplanet data (Marcy et al., 2004, 2005; chapter by Udry et al.): Roughly 1% of Sun-like stars (late F, G, and early K spectral class main-sequence stars that are chromospherically quiet, i.e., have inactive photospheres) have planets more massive than Saturn within 0.1 AU. Approximately 7% of Sun-like stars have planets more massive than Jupiter within 3 AU. Planets orbiting interior to ~0.1 AU, a region where tidal circularization time-scales are less than stellar ages, have small orbital eccentricities. The median eccentricity observed for planets on more distant orbits is 0.25, and some of these planets travel on very eccentric orbits. Within 5 AU of Sun-like stars, Jupiter-mass planets are more common than planets of several Jupiter masses, and substellar companions that are more than 10x as massive as Jupiter are rare. Stars with higher metallicity are much more likely to host detectable planets than are metal-poor stars (Gonzalez, 2003; Santos et al., 2003), with the probability of hosting an observable planet varying as the square of stellar metallicity (Fischer and Valenti, 2005). Low-mass main-sequence stars (M dwarfs) are significantly less likely to host one or more giant planets with orbital period(s) of less than a decade than are Sun-like stars. Multiple-planet systems are more common than if detectable planets were randomly assigned to stars (i.e., than if the presence of a planet around a given star was not correlated with the presence of other planets around that same star). Most transiting extrasolar giant planets are predominantly hydrogen (Charbonneau et al., 2000; Burrows et al., 2003; Alonso et al., 2004), as are Jupiter and Saturn. However, HD 149026 b, which is slightly more massive than Saturn, appears to have comparable amounts of hydrogen + helium vs. heavy elements (Sato et al., 2005), making its bulk composition intermediate between Saturn and Uranus, but more richly endowed in terms of total amount of “metals” than any planet in our solar system.

Transit observations have also yielded an important negative result: Hubble Space Telescope photometry of a large number of stars in the globular cluster 47 Tucanae failed to detect any transiting inner giant planets, even though ~17 such transiting objects would be expected were the frequency of such planets in this low metallicity cluster the same as that for Sun-like stars in the solar neighborhood (Gilliland et al., 2000).

Various classes of models have been proposed to explain the formation of giant planets and brown dwarfs. Following Lissauer (2004) and consistent with current IAU nomenclature, these definitions are used in this chapter:

**Star:** self-sustaining fusion is sufficient for thermal pressure to balance gravity.

**Stellar remnant:** dead star — no more fusion, i.e., thermal pressure sustained against radiative losses by energy produced from fusion is no longer sufficient to balance gravitational contraction.

**Brown dwarf:** substellar object with substantial deuterium fusion — more than half the object’s original inventory of deuterium is ultimately destroyed by fusion.

**Planet:** negligible fusion (<13 Jupiter masses, or $M_{\text{Jup}}$), plus it orbits one or more stars and/or stellar remnants.

The mass function of young compact objects in star-forming regions extends down through the brown dwarf mass range to below the deuterium-burning limit (Zapatero Osorio et al., 2000; chapter by Luhman et al.; cf. Marley et al., 2006). This observation, together with the lack of any convincing theoretical reason to believe that the collapse process that leads to stars cannot also produce substellar objects (Wuchterl and Tschambrot, 2003; chapter by Whitworth et al.), strongly implies that most isolated (or distant companion) brown dwarfs and isolated high-planetary-mass objects form via the same collapse process as do stars.

By similar reasoning, the “brown dwarf desert,” a profound dip over the range ~10–50 $M_{\text{Jup}}$ in the mass function of companions orbiting within several AU of Sun-like stars (Marcy et al., 2004; chapter by Udry et al.), strongly suggests that the vast majority of extrasolar giant planets formed via a mechanism different from that of stars. Within our solar system, bodies up to the mass of Earth consist almost entirely of condensable material, and even bodies of mass ~15 $M_{\oplus}$ (Earth masses) consist mostly of condensable material. (The definition of “condensable” is best thought of as the value of the specific entropy of the constituent relative to that for which the material can form a liquid or solid. Hydrogen and helium within protoplanetary disks have entropies far in excess of that required for condensation, even if they are compressed isothermally to pressures on the order of 1 bar, even for a temperature of only a few tens of degrees. Thus, H and He remain in a gaseous state.) The fraction of highly volatile gases increases with planet mass through Uranus/Neptune, to Saturn, and finally to Jupiter, which is still enriched in condensables at least threefold compared to the Sun (Young, 2003). This gradual, nearly monotonic relationship between mass and composition argues for a unified formation scenario for all the planets and smaller bodies. Moreover, the continuum of observed extrasolar planetary properties, which stretches to systems not very dissimilar to our own, suggests that extrasolar planets formed in a similar way to the planets within our solar system.

Models for the formation of gas giant planets were reviewed by Wuchterl et al. (2000). Star-like direct quasi-spherical collapse is not considered viable, both because of the observed brown dwarf desert mentioned above and theo-
retical arguments against the formation of Jupiter-mass objects via fragmentation (Bodenheimer et al., 2000a). The theory of giant planet formation that is favored by most researchers is the core nucleated accretion model, in which the planet’s initial phase of growth resembles that of a terrestrial planet, but the planet becomes sufficiently massive (several \(M_\oplus\)) that it is able to accumulate substantial amounts of gas from the surrounding protoplanetary disk.

According to the variant of the core nucleated accretion model analyzed by Pollack et al. (1996), Bodenheimer et al. (2000b), and Hubickyj et al. (2005), the formation and evolution of a giant planet is viewed to occur in the following sequence: (1) Dust particles in the solar nebula form planetesimals that accrete one another, resulting in a solid core surrounded by a low-mass gaseous envelope. Initially, runaway accretion of solids occurs, and the accretion rate of gas is very slow. As the solid material in the planet’s feeding zone is depleted, the rate of solids accretion tapers off. The gas accretion rate steadily increases and eventually exceeds the accretion rate of solids. (2) The protoplanet continues to grow as the gas accretes at a relatively constant rate. The mass of the solid core also increases, but at a slower rate. (3) Near this point, the rate of gas accretion increases in runaway fashion, and the protoplanet grows at a rapidly accelerating rate. The first three parts of the evolutionary sequence are referred to as the nebular stage, because the outer boundary of the protoplanetary envelope is in contact with the solar nebula, and the density and temperature at this interface are those of the nebula. (4) The gas accretion rate reaches a limiting value defined by the rate at which the nebula can transport gas to the vicinity of the planet. After this point, the equilibrium region of the protoplanet contracts, and gas accretes hydrodynamically into this equilibrium region. This part of the evolution is considered to be the transition stage. (5) Accretion is stopped by either the opening of a gap in the disk as a consequence of the tidal effect of the planet, accumulation of all nearby gas, or dissipation of the nebula. Once accretion stops, the planet enters the isolation stage. The planet then contracts and cools to the present state at constant mass.

Aside from core nucleated accretion, the only giant planet formation scenario receiving significant attention is the gas instability model, in which a giant planet forms directly from the contraction of a clump that was produced via a gravitational instability in the protoplanetary disk. Numerical calculations show that 1 \(M_{\text{Jup}}\) clumps can form in sufficiently gravitationally unstable disks (e.g., Boss, 2000; Mayer et al., 2002). However, weak gravitational instabilities excite spiral density waves; density waves transport angular momentum that leads to spreading of a disk, lowering its surface density and making it more gravitationally stable. Rapid cooling and/or mass accretion is required to make a disk highly unstable. Thus, long-lived clumps can only be produced in protoplanetary disks with highly atypical physical properties (Rafikov, 2005). Additionally, gas instabilities would yield massive stellar-composition planets, requiring a separate process to explain the smaller bodies in our solar system and the heavy element enhancements in Jupiter and Saturn. The existence of intermediate objects like Uranus and Neptune is particularly difficult to account for in such a scenario. Furthermore, metal-rich stars are more likely to host observable extrasolar planets than are metal-poor stars (Fischer and Valenti, 2005; chapter by Udry et al.); this trend is consistent with the requirement of having sufficient condensables to form a massive core, but runs contrary to the requirement of rapid disk cooling needed to form long-lived clumps via gravitational instabilities (Cai et al., 2006).

See the chapter by Durisen et al. for a more extensive discussion of the gas instability model.

We review the constraints on formation provided by the internal structure of giant planets in section 2. In section 3, we summarize recent models of giant planet growth via core nucleated accretion. These models have some important shortcomings, and the issues remaining to be resolved are highlighted in section 4. We conclude this chapter with a brief summary.

## 2. MODELS OF GIANT PLANETS

The central issues for giant planet models are these: Do they have cores (of heavy elements) and, if so, what do those cores tell us about how the planet formed? The existence of heavy element enrichments in the solar system’s four giant planets is not in doubt, because the mean densities of these planets are higher than the expected value for adiabatic bodies of solar composition. However, the existence of a core is less easily established, especially if the core is a small fraction of the total mass, as is likely in the case of Jupiter (Fig. 1).

Moreover, the presence or absence of a core does not automatically tell us whether or not a core existed at the time of planet formation. It is possible that the current core is an eroded remnant (less massive than the primordial core) or even enhanced because of rainout of heavy elements from the planet’s envelope.

Seismology is by far the best method for establishing the existence and nature of a core, but we lack this approach for the giant planets since (unlike the Sun) the normal mode excitation is expected to be too small to be detectable at present. Dynamical approaches exist (e.g., measurement of the precession constant, as used to determine Earth’s moment of inertia to high precision), but have not yet been implemented, since they require close-in, long-lived orbiters. We rely mostly on the old and very nonunique approach of
The region is thicker and the presence of a core is more certain. A similar structure applies for Saturn, except that the molecular ideal and relevant experiments are difficult. We still do not know whether hydrogen undergoes a first-order phase transition (the so-called molecular to metallic transition, although if it exists it cannot be described in such simple language). However, the shape of the pressure-density relationship remains uncertain even if one accepts (as most experts do) that there is no first-order phase transition at the temperatures relevant to the giant planets. One way to appreciate this difficulty is to ask what error in the equation of state for hydrogen corresponds to a 1 M$_\oplus$ error in heavy elements? Roughly speaking, this is in proportion to the corresponding fraction of the planet’s total mass, which is only 0.3% in the case of Jupiter and about 1% in the case of Saturn. Since the uncertainty in the equation of state is as much as several percent in the least well understood pressure range, the corresponding error in the estimated abundance of heavy elements may be as large as 5–10 M$_\oplus$.

Detailed reviews of giant planet structure include Hubbard et al. (2002) and Guillot (2005). The most complete modeling effort is the work of Guillot and collaborators. Simple coreless models of Jupiter are marginally capable of satisfying all the data. These models have a primordial solar hydrogen/helium ratio, but are enriched in heavy elements to the extent of about 10 M$_\oplus$. The most likely value for the mass of Jupiter’s core is in the range of 5–10 M$_\oplus$. To a first approximation, it does not matter (in terms of explaining the mean density) whether the heavy elements are in a core or distributed internally. This heavy element enrichment is possibly greater than the observed threefold enrichment in some heavy elements in the atmosphere. However, there is no reliable determination of oxygen (as water) and no direct way of detecting the rocky component remotely, since those elements condense and form a cloud deck far below the observable atmosphere. It is common practice in models to assume a uniform mixing of the heavy elements outside a core. In some models, a jump in composition is assumed at the hypothesized molecular-metallic hydrogen phase transition. The heat flow and convection-generated magnetic field suggest a well-mixed region that extends down to at least about the megabar region. However, there is no observational or theoretical requirement that the planet be homogeneous outside the hypothetical high-density (rock and ice) core. It is also important to realize that the common practice of placing a separate core of heavy elements at the centers of these planets is governed by simplicity, rather than by observation. To varying degrees, the “core” could have a fuzzy boundary with the overlying hydrogen-rich envelope.

The heavy element fraction of Saturn is larger than that of Jupiter and as a consequence we have a more confident conclusion despite somewhat less accurate data (see Table 1). The models indicate that there is indeed a core, several to 20 M$_\oplus$, with a preferred value of ~10 M$_\oplus$. The latest gravity data results from Cassini are consistent with this. There is an uncertainty for Saturn that does not arise for Jupiter: We do not know the rotation rate with high accuracy. Saturn kilometric radiation (SKR) emissions have changed their period by six minutes between Voyager (1980) and

![Diagram of the interior of Jupiter](image)

**Fig. 1.** Schematic cross-sectional view of the interior of Jupiter. A similar structure applies for Saturn, except that the molecular region is thicker and the presence of a core is more certain.
now, and there is currently no generally accepted understanding of the connection between this period and the rotation period of the deep interior (Stevenson, 2006).

Uranus and Neptune are far less well understood than are Jupiter and Saturn. However, there is no doubt that they are mostly ice and rock, yet also possess ~2 $M_\oplus$ of gas each. Their atmospheres are estimated to have solar hydrogen to helium ratios, but the uncertainty is large because this determination is based on the pressure-induced absorption features of hydrogen, a method that has been unreliable for Jupiter and Saturn. The amount of hydrogen extractable from the ices is in principle about 20% of the total mass (assuming the hydrogen was delivered as water, methane, and ammonia), and this is marginally close to the hydrogen mass required by interior models. Moreover, there is the possibility that methane would decompose into carbon and hydrogen at extreme pressures. However, the atmospheres of Uranus and Neptune are highly enriched in methane (thus limiting the possibility of massive decomposition of this compound to very deep regions), and there is no experimental or theoretical evidence for extensive decomposition of water or ammonia under the conditions encountered inside these bodies. Consequently, it is not plausible to derive even 1 $M_\oplus$ of predominantly hydrogen gas from the breakdown of hydrogen-bearing ice or rock, even leaving aside the dubious proposition that such decomposed hydrogen would rise to the outer regions of the planet. This gas appears to have come from $H_2$ and He within the solar nebula. Uranus and Neptune (or precursor components massive enough to capture adequate amounts of gas) must have formed largely in the presence of the solar nebula, a very stringent constraint on the formation of solid bodies. While ice-rich embryos as small as ~0.1 $M_\oplus$ could conceivably have captured such gas mixed with steam within high mean molecular weight atmospheres (Stevenson, 1982, 1984; Lissauer et al., 1995), there remain many open questions about this process, and most models suggest that Uranus and Neptune reached a substantial fraction of their current masses prior to the dispersal of the solar nebula.

It is often supposed that the presence or absence of a core in Jupiter (for example) can be placed in one-to-one correspondence with the presence or absence of a nucleating body that caused the inflow of gas to form the much more massive envelope. However, there is no neat correspondence between mode of giant planet formation and current presence of a core. One could imagine core nucleated accretion even if there is no core remaining, because the core might become mixed into the overlying envelope by convective processes. One could also imagine making a core in the low-density protoplanet phase by rainout. We now discuss each of these processes in more detail.

Core mix-up (or erosion) can be thought of as analogous to the following simple fluid dynamical experiment. Suppose one took a pot that has a layer of salt at the bottom, and then gently (or not so gently) added water. One then heats the pot from below to stimulate convection. Under what circumstances will the salt end up fully mixed with the water (assuming saturation is not reached)? In analogy with giant planets, one must ignore diffusion across the depth of the pot, since the diffusion time within a planet is longer than the age of the universe. Under these circumstances, the relevant consideration is the work done by convection (or initial stirring) compared to the work that must be done to mix the material. The work done by convection is determined by the buoyancy flux integrated through time. In giant planets, this is dominated by cooling. (In terrestrial planets it is dominated by radiogenic heating.) In accordance with the virial theorem, the contraction of the planet changes the gravitational energy by about the same amount as the change in internal energy (work done against gravity); see Hubbard (1984). However, the buoyancy production is directly related to the amount by which the planet has cooled, since that cooling is expressed in luminosity and the luminosity comes mainly from convective transport of buoyant fluid elements.

If the relevant part of the planet cools by an amount $\Delta T$ and the coefficient of thermal expansion is $\alpha$, then the total available work is on the order of $Mg\alpha\Delta TH$, where $g$ is grav-

<table>
<thead>
<tr>
<th>TABLE 1. Mass fractions.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>Primordial Solar</td>
</tr>
<tr>
<td>Hydrogen</td>
</tr>
<tr>
<td>Helium</td>
</tr>
<tr>
<td>Volatile heavy elements (“ices”)</td>
</tr>
<tr>
<td>Involatile heavy elements (“rock”)</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Jupiter</td>
</tr>
<tr>
<td>Hydrogen</td>
</tr>
<tr>
<td>Helium</td>
</tr>
<tr>
<td>Volatile heavy elements (“ices”)</td>
</tr>
<tr>
<td>Involatile heavy elements (“rock”)</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Saturn</td>
</tr>
<tr>
<td>Hydrogen</td>
</tr>
<tr>
<td>Helium</td>
</tr>
<tr>
<td>Volatile heavy elements (“ices”)</td>
</tr>
<tr>
<td>Involatile heavy elements (“rock”)</td>
</tr>
</tbody>
</table>
ity and H is the height through which the buoyant elements rise. Since the salt (or ice and rock, in the case of the planets) has a very different density from the background fluid (hydrogen for the planets), the work done lifting a high-density mass $\Delta M$ is $\Delta M gH$. Accordingly, $\Delta M < M_x \Delta T$. The temperature drop over the age of the solar system is roughly comparable to the actual temperature now (e.g., the deep interior of Jupiter may have cooled from 40,000 K to 20,000 K). Deep within giant planets, $\alpha T \sim 0.05$. Consequently it is possible in principle to mix up on the order of 5% of the mass of the planet (15 $M_\oplus$ for Jupiter, 5 $M_\oplus$ for Saturn). In practice, the real amount is likely to be less than this by as much as an order of magnitude, for three reasons. First, the heat flow (or equivalently, the buoyancy production) at the top of the core is far less than that associated with the cooling of the overlying hydrogen (the main source of luminosity in these planets). For example, the heat content of a rock and ice mixture is about an order of magnitude less than the heat content of the same mass of hydrogen at the same temperature, because the latter has much lower molecular weight and hence much higher heat capacity. If one relied instead on radioactive heat, then the available heat flow would be only a few times the terrestrial value (per unit area), whereas the intrinsic heat flux of Jupiter (per unit area) is 30x greater. The second reason for lowering the expected erosion comes from consideration of the actual fluid dynamics of mixing. It is well established from laboratory experiments that this is typically an order of magnitude less efficient than the highest efficiency permitted from purely energetic arguments (Turner, 1973). Third, convective upward mixing of heavy molecules might be strongly inhibited by a compositional gradient. In those circumstances, the mixing will certainly be slower because of the lower diffusivity of the heavy atoms relative to the diffusivity of heat. This is the regime of “double diffusive convection” (Turner, 1973).

This discussion omits consideration of the difficult question of what happens when there are impacts of large embryos during the formation of the planet. The mixing that occurs during that early phase is not so readily analyzed by the arguments presented above and has been inadequately explored. The material is less degenerate, so that thermal differences have a potentially greater ability to cause compositional mixing. But the complexity of the fluid dynamics and shock processes make quantitative analysis difficult. In the corresponding problem of giant impacts during formation of terrestrial planets (and formation of the Moon), it has become apparent that one cannot rely on low-resolution smoothed particle hydrodynamics (SPH) simulations for quantitative assessment of the mixing. The reason is that the mixing involves Rayleigh-Taylor instabilities that grow most rapidly at length scales smaller than the resolution scale of these simulations. A similar problem will arise in the consideration of large ice and rock bodies hitting partially assembled giant planets.

It is possible that Uranus and Neptune provide the greatest insight into these issues of core formation and structure. This might seem surprising given our relatively poor understanding of their internal structure (Guillot, 2005). However, the ice and rock components of these planets are dominant relative to the hydrogen component, and plausible models that fit the gravity field clearly require some mixing of the constituents. End-member models consisting of a discrete rock core, ice mantle, and hydrogen/helium envelope are not permitted. To the extent that the formation of these planets is similar to that for Jupiter and Saturn (except of course for the lack of a late-stage large addition of gas), this would seem to suggest that there was considerable mixing even during the early stages.

Turning to the opposite problem of core rainout, making a core this way once the material is dense and degenerate is unlikely because the high temperatures and dilution make it thermodynamically implausible. Suppose we have a constituent of atomic (or molar) abundance $x$ relative to the overwhelmingly predominant hydrogen. Let the Gibbs energy cost of mixing this constituent to the atomic level in hydrogen be $\Delta G_x$. The physical origin of this energy is primarily quantum mechanical and arises from the mismatch of the electronic environments of the host and the inserted atom or molecule (including the work done in creating the cavity within the host). In the plausible situation of approximately ideal mixing, the solubility limit of this constituent is then $-e^{-\Delta G_x/kT}$, where k is Boltzmann’s constant (Steven-son, 1998). In the case of water (or oxygen), the expected average concentration relative to hydrogen is about 0.001. At 10,000 K, $kT \sim 1$ eV, and rainout could begin (starting from a higher temperature, undersaturated state) for $\Delta G_x$ (in eV) $\sim -\ln 0.001 \sim 7$. This is a large energy, especially when one considers that an electronically unfavorable choice (helium) has smaller $\Delta G_x$ of perhaps only a few eV and even neon (depleted by a factor of 10 in the atmosphere of Jupiter) apparently has a lower insertion energy than 7 eV. The basic physical point is that the least soluble constituents in the deep interior are expected to be atoms with very tightly bound electrons (the noble gases), and the insolubility of helium is aided by its higher abundance. In fact, the observed inferred rainout of helium is small for Jupiter, corresponding to $\sim 10\%$ of the helium mass, which is in turn about one-fourth of the total mass of Jupiter. The total helium rainout is accordingly about 5–10 $M_\oplus$. (This is unlikely to form a discrete core and is in any event not what is meant when one talks of a high-density core in these planets, since the density of helium is only roughly twice that of hydrogen.) Rainout of material suitable for a rock and ice core is less favorable both because it is much less abundant by atomic number and because it is electronically more compatible with the metallic state of hydrogen deep within the planet. Indeed, water is expected to be a metal at conditions not far removed from the metallization of hydrogen and is certainly ionic at lower pressures within Jupiter.

As with erosion, this discussion does not cover the potentially important case of very early nondegenerate conditions. In the early work on giant gaseous protoplanets, it was proposed that cores could rain out, much in the same way as rock or ice can condense and settle to the midplane of the solar nebula, i.e., under very low density conditions.
is able to advocate in some more recent work (e.g., Boss, 2002). The difficulty of this picture lies in the fact that as the material settles deeper and the protoplanet contracts, the combination of adiabatic heating and gravitational energy release is likely to cause the solid and liquid iron/silicates to undergo evaporation. This can be avoided only for rather low mass protoplanets. This nonetheless remains the most plausible way of forming a core in a planet that formed via gaseous instability. Note, however, that this will not produce a core that is more massive than that predicted by solar abundances, unless the appropriate amount of gas is lost at a later stage. Models of this kind have the appearance of special pleading if they are to explain the entire set of giant planets (including Uranus and Neptune).

It seems likely that whatever model one favors for giant planet formation, it should allow for the formation of a core, since Saturn probably has a core and one must in any event explain Uranus and Neptune. It would be contrived to attribute a different origin for Jupiter than for the other giant planets. It seems likely, therefore, that the formation of giant planets is closest to a “bottom-up” scenario that proceeded through formation of a solid embryo followed by the accumulation of gas.

3. GIANT PLANET FORMATION MODELS

The core nucleated accretion model relies on a combination of planetesimal accretion and gravitational accumulation of gas. According to this scenario, the initial stages of growth of a gas giant planet are identical to those of a terrestrial planet. Dust settles toward the midplane of the protoplanetary disk and agglomerates into (at least) kilometer-sized planetesimals, which continue to grow into larger solid bodies via pairwise inelastic collisions. As the planet grows, its gravitational potential well gets deeper, and when its escape speed exceeds the thermal velocity of gas in the surrounding disk, it begins to accumulate a gaseous envelope. The gaseous envelope is initially optically thin and isothermal with the surrounding protoplanetary disk, but as it gains mass it becomes optically thick and hotter with increasing depth. While the planet’s gravity pulls gas from the surrounding disk toward it, thermal pressure from the existing envelope limits accretion. For much of the planet’s growth epoch, the primary limit on its accumulation of gas is its ability to radiate away the gravitational energy provided by accretion of planetesimals and envelope contraction; this energy loss is necessary for the envelope to further contract and allow more gas to reach the region in which the planet’s gravity dominates. The size of the planet’s gravitational domain is typically a significant fraction of the planet’s Hill sphere, whose radius, \( R_H \), is given by

\[
R_H = \left( \frac{M}{3M_*} \right)^{1/3} r
\]

where \( M \) and \( M_* \) are the masses of the planet and star, respectively, and \( r \) is the distance between these two bodies. Eventually, increases in the planet’s mass and radiation of energy allow the envelope to shrink rapidly. At this point, the factor limiting the planet’s growth rate is the flow of gas from the surrounding protoplanetary disk.

The rate and manner in which a forming giant planet accretes solids substantially affect the planet’s ability to attract gas. Initially accreted solids form the planet’s core, around which gas is able to accumulate. Calculated gas accretion rates are very strongly increasing functions of the total mass of the planet, implying that rapid growth of the core is a key factor in enabling a planet to accumulate substantial quantities of gas prior to dissipation of the protoplanetary disk. Continued accretion of solids acts to reduce the planet’s growth time by increasing the depth of its gravitational potential well, but has counteracting affects by providing additional thermal energy to the envelope (from solids that sink to or near the core) and increased atmospheric opacity from grains that are released in the upper parts of the envelope. Major questions remain to be answered regarding solid-body accretion in the giant planet region of a protoplanetary disk, with state-of-the-art models providing a diverse set of predictions.

Because of the complexity of the physics and chemistry involved in giant planet formation, the large range of distance scales, the long time (compared to orbital and local thermal times) required for accumulation, and the uncertainties in initial conditions provided by the protoplanetary disks, detailed planet-growth models have focused on specific aspects of the problem, and ignored or provided greatly simplified treatments of other processes. The solid-accretion scenarios incorporated into envelope models to date have been quite simplified, and in some cases completely ad hoc. These issues are discussed in section 3.1.

A planet on the order of one to several \( M_{\oplus} \) is able to capture an atmosphere from the protoplanetary disk because the escape speed from its surface is large compared to the thermal velocity of gas in the disk. However, such an atmosphere is very tenuous and distorted, with thermal pressure pushing outward to the limits of the planet’s gravitational reach and thereby limiting further accretion of gas. The key factor governing the planet’s evolution at this stage is its ability to radiate energy so that its envelope can shrink and allow more gas to enter the planet’s gravitational domain. Evolution occurs slowly, and hydrostatic structure is generally a good approximation. However, the stability of the planet’s atmosphere against hydrodynamically induced ejection must be calculated. The basic physical mechanisms operating during this stage of growth appear to be qualitatively understood, but serious questions remain regarding the ability of planets to pass through this stage sufficiently rapidly to complete their growth while adequate gas remains in the protoplanetary disk. This timescale issue is being addressed by numerical simulations. Models of this phase of a giant planet’s growth are reviewed in section 3.2.

Once a planet has enough mass for its self-gravity to compress the envelope substantially, its ability to accrete additional gas is limited only by the amount of gas avail-
able. Hydrodynamic limits allow quite rapid gas flow to the planet in an unperturbed disk. But the planet alters the disk by accreting material from it and by exerting gravitational torques on it. Both of these processes can lead to gap formation and isolation of the planet from the surrounding gas. Hydrodynamic simulations lend insight into these processes, and are discussed briefly in section 3.3.

Radial motion of the planet and disk material can affect both the planet’s growth and its ultimate orbit. Much of a protoplanetary disk is ultimately accreted by the central star (chapter by Bouvier et al.). Small dust grains are carried along with the gas, but millimeter and larger particles can suffer a secular drag if they orbit within a gaseous disk that rotates slower than the Keplerian velocity because the gas is partially supported against stellar gravity by a radial pressure gradient (Adachi et al., 1976). Such gas drag can cause substantial orbital decay for bodies up to kilometer sizes (Weidenschilling, 1977). Once growing planets reach lunar to Mars size, their gravitational interactions with the surrounding disk can lead to substantial radial migration.Radial migration of a planet can have major consequences for its growth, ultimate orbit, and even survival. This process is reviewed in depth by the chapter by Papaloizou et al., but its relationship with planetary growth is briefly commented upon in section 3.4.

### 3.1. Growth of the Core

Models of solid planet growth do a fairly good job of explaining the origin of terrestrial planets in our solar system (e.g., Agnor et al., 1999; Chambers, 2001), and can be applied with modification to the growth of planetary bodies at greater distances from the Sun and other stars (Quintana et al., 2002; Barbieri et al., 2002; Quintana and Lis, 2006). Most models of terrestrial planet growth start with a “minimum mass” disk, containing the observed heavy element components in the planets spread out smoothly into a disk, plus enough gas to make the disk’s composition the same as that of the proto-Sun. The disk is assumed to be relatively quiescent, with the Sun already largely formed and close to its current mass (Safonov, 1969). Micrometer-sized dust, composed of surviving interstellar grains and condensates formed within the protoplanetary disk, moves mostly with the dominant gaseous component of the disk. But it gradually agglomerates and settles toward the midplane of the disk. If the disk is laminar, then the solids can collapse into a layer that is thin enough for collective gravitational instabilities to occur (Edgeworth, 1949; Safonov, 1960; Goldreich and Ward, 1973); such instabilities would have produced planetesimals of ~1 km radius at 1 AU from the Sun. If the disk is turbulent, then gravitational instabilities are suppressed because the dusty layer remains too thick. Under such circumstances, continued growth via pairwise agglomeration depends upon (currently unknown) sticking and disruption probabilities for collisions among larger grains (Weidenschilling and Cuzzi, 1993). The mechanism for growth from centimeter to kilometer sizes remains one of the major controversies in terrestrial planet growth (Youdin and Shu, 2002; chapter by Dominik et al.). Nonetheless, theoretical models suggest that gravitational instabilities are more likely to occur farther from the star and that ices are stickier than rock. Moreover, many small to moderate-sized bodies are observed in the Kuiper belt beyond the orbit of Neptune (chapter by Cruikshank et al.) and probably smaller but still macroscopic bodies are inferred as parents to the observed dust seen in second-generation debris disks around Vega, Beta Pictoris, and many other stars (chapter by Meyer et al.). Thus, growth of solid bodies to multikilometer sizes in at least the inner portions of the ice-condensation region of most protoplanetary disks seems quite likely.

Once solid bodies reach kilometer size (using parameters that are appropriate for the terrestrial region of the protosolar disk), gravitational interactions between pairs of solid planetesimals provide the dominant perturbation of their basic Keplerian orbits. Electromagnetic forces, collective gravitational effects, and in most circumstances, gas drag, play minor roles. These planetesimals continue to agglomerate via pairwise mergers. The rate of solid-body accretion by a planetesimal or planetary embryo (basically a large planetesimal) is determined by the size and mass of the planetesimal/planetary embryo, the surface density of planetesimals, and the distribution of planetesimal velocities relative to the accreting body. Assuming perfect accretion, i.e., that all physical collisions are completely inelastic, this stage of growth is initially quite rapid, especially in the inner regions of a protoplanetary disk, and large bodies form quickly. The planetesimal accretion rate, $M_Z$, is given by

$$M_Z = \pi R^2 \sigma_Z \Omega F_g$$  \hspace{1cm} (2)

where $R$ is the radius of the accreting body, $\sigma_Z$ is the surface density of solid planetesimals in the solar nebula, $\Omega$ is the orbital frequency, and $F_g$ is the gravitational enhancement factor, which is the ratio of the total effective accretion cross section to the geometric cross section. If the velocity dispersion of the bodies is large compared to the Keplerian shear of the disk across the body’s accretion zone, the two-body approximation yields

$$F_g = 1 + \left( \frac{v_e}{v} \right)^2$$  \hspace{1cm} (3)

where $v$ is the velocity dispersion and $v_e$ is the escape velocity from the body’s surface. The evolution of the planetesimal size distribution is determined by the gravitationally enhanced collision cross section, which favors collisions between bodies having larger masses and smaller relative velocities.

Planetesimal growth regimes are sometimes characterized as either orderly or runaway. In orderly growth, particles containing most of the mass double their masses in about the same amount of time as the largest particle. When the relative velocity between planetesimals is comparable...
to or larger than the escape velocity, \( v \geq v_e \), the growth rate is approximately proportional to \( R^2 \), and there is an orderly growth of the entire size distribution. When the relative velocity is small, \( v \ll v_e \), the growth rate is proportional to \( R^4 \). In this situation, the planetary embryo rapidly grows much larger than any other planetesimal in its accretion zone. By virtue of its large, gravitationally enhanced cross section, this runaway particle doubles its mass faster than the smaller bodies do, and detaches itself from the mass distribution (Wetherill and Stewart, 1989; Ohtsuki et al., 2002).

Eventually a runaway body can grow so large that it transitions from dispersion-dominated growth to shear-dominated growth (Lissauer, 1987). Dynamical friction, which drives the distribution of planetesimal velocities toward a state of equipartition of kinetic energy of random motion (e.g., Stewart and Wetherill, 1988), reduces the random motions of the more massive bodies, so proximate embryos collide and merge. At this stage, larger embryos take longer to double in mass than do smaller ones, although embryos of all masses continue their runaway growth relative to surrounding planetesimals. This phase of rapid accretion of planetary embryos is known as oligarchic growth (Kokubo and Ida, 1998).

The self-limiting nature of runaway/oligarchic growth implies that massive planetary embryos form at regular intervals in semimajor axis. The agglomeration of these embryos into a small number of widely spaced terrestrial planets necessarily requires a stage characterized by large orbital eccentricities. The large velocities imply small collision cross sections (equation (3)) and hence long accretion times. Growth via binary collisions proceeds until the spacing of planetary orbits become dynamically isolated from one another, i.e., sufficient for the configuration to be stable to gravitational interactions among the planets for the lifetime of the system (Safronov, 1969; Wetherill, 1990; Lissauer, 1993, 1995; Agnor et al., 1999; Chambers, 2001; Laskar, 2000).

The early phases of growth from planetesimals are likely to be similar in the more distant regions of protoplanetary disks. However, the rate at which accretion of solids takes place depends upon the surface density of condensates and the orbital frequency (equation (2)), both of which decrease with heliocentric distance. Thus, the high-velocity final growth stage that takes \( O(10^8) \) yr in the terrestrial planet zone (Safronov, 1969; Wetherill, 1980; Agnor et al., 1999; Chambers, 2001) would require \( O(10^9) \) yr in the giant planet zone (Safronov, 1969). This is far longer than any modern estimates of the lifetimes of gas within protoplanetary disks, implying that giant planet cores must form via rapid runaway/oligarchic growth (chapter by Meyer et al.). Moreover, particles far from their stars are physically small compared to the size of their gravitational domains (Hill spheres), and giant planets eventually grow large enough that escape speeds from accreting planets exceed the escape velocity from stellar orbit at their locations.

For shear-dominated accretion, the mass at which an embryo becomes isolated from the surrounding disk is given by

\[
M_{\text{iso}} = \frac{8\pi\sqrt{3 r^2 \sigma Z}}{3 M_\odot^{1/2}}
\]

where \( r \) is the distance from the star (Lissauer, 1993). In the inner part of protoplanetary disks, Kepler shear is too great to allow the accretion of solid planets larger than a few \( M_\oplus \) on any timescale unless surface densities are considerably above that of the minimum mass solar nebula or a large amount of radial migration occurs. Larger solid planets are permitted farther from stars, but the duration of the final, high-velocity, stages of growth (Safronov, 1969) are far longer than the observed lifetimes of protoplanetary disks. The epoch of runaway/rapid oligarchic growth lasts only millions of years or less near 5 AU, and can produce \( \sim 10 \ M_\oplus \) cores in disks only a few times the minimum mass solar nebula (Lissauer, 1987). The masses at which planets become isolated from the disk, thereby terminating the runaway/rapid oligarchic growth epoch, are likely to be comparable only at greater distances from the star. However, at these large distances, random velocities of planetesimals must remain quite small for accretion rates to be sufficiently rapid for embryos to approach isolation mass within the lifetimes of gaseous disks. Indeed, if planetesimal velocities become too large, material is more likely to be ejected to interstellar space than accreted by the planetary embryos.

The fact that Uranus and Neptune contain much less H2 and He than Jupiter and Saturn suggests that Uranus and Neptune never quite reached runaway gas accretion conditions, possibly due to a slower accretion of planetesimals (Pollack et al., 1996). Theoretical difficulties with forming planets at Uranus/Neptune distances have been discussed in greater detail by Lissauer et al. (1995) and Thommes et al. (2003). New models are being proposed to address these problems by allowing rapid runaway accretion of a very small number of planetary embryos (cores) beyond 10 AU. In the model presented by Weidenschilling (2005), an embryo is scattered from the Jupiter-Saturn region into a massive disk of small planetesimals. The embryo is several orders of magnitude more massive than the individual planetesimals surrounding it, but still far less massive than the aggregate of the surrounding disk of planetesimals. Dynamical friction is thus able to circularize the orbit of the embryo without substantially exciting planetesimal eccentricities. Goldreich et al. (2004a,b) propose that (at least in the Uranus/Neptune region) planetesimals between growing embryos are ground down to very small sizes and are forced into low-inclination, nearly circular orbits by frequent mutual collisions. Planetary embryos can accrete rapidly because of their large, gravitationally enhanced collision cross sections in dynamically cold disks such as those in the models of Weidenschilling (2005) and of Goldreich et al. (2004a,b). Alternatively, Thommes et al. (2003) suggest that the cores and possibly also the gaseous envelopes of Uranus and Neptune accreted between or just exterior to the orbits of Jupiter and Saturn, and were subsequently scattered out to their current locations by gravitational per-
turbations of these two giant planets (see also Tsiganis et al., 2005). Alternatively/additionally, Uranus and Neptune may have avoided gas runaway as a result of the removal of gas from the outer regions of the disk via photoevaporation (Hollenbach et al., 2000).

Published simulations of the accumulation of giant planet atmospheres use simplified prescriptions for the planet’s accretion of solids. In some cases, the accretion rate of solids is assumed to be constant (Bodenheimer and Pollack, 1986; Ikoma et al., 2000). In others, an isolated planetary embryo grows by runaway accretion in a disk of much smaller planetesimals, as discussed in the following paragraph. The actual accretion of solids by a planet is more complex, variable in time and highly stochastic, and most likely including the occasional impact of a large body. But as discussed above, there are many open questions regarding the growth of solid cores at the locations of the giant planets within our solar system. Thus, more sophisticated models do not necessarily provide better approximations of actual core growth rates. Moreover, these simplified models illuminate several key aspects of how accretion of solids controls the rate of envelope (gas) accumulation.

The most sophisticated thermal models of the accumulations of massive gaseous envelopes by planets (Pollack et al., 1996; Bodenheimer et al., 2000b; Alibert et al., 2004, 2005; Hubickyj et al., 2005) assume runaway growth of an isolated (or nearly isolated) planet. An updated version (Greenzweig and Lissauer, 1992) of the classical theory of planetary growth (Safronov, 1969) is used, employing equations (2) and (3) with R replaced by R_{capt}, the effective (geometric) capture radius of the protoplanet for a planetesimal of a given size (including regions of the envelope sufficiently dense to capture planetesimals). These models begin with the growing protoplanet embedded in a disk of monodisperse planetesimal size and uniform surface density. The protoplanet’s feeding zone is assumed to be an annulus extending to a radial distance of about 4 R_{	ext{H}} on either side of its orbit (Kary and Lissauer, 1994). The feeding zone grows as the planet gains mass, and random scattering spreads the unaccreted planetesimals uniformly over the feeding zone. Radial migration of planetesimals into and out of the feeding zone is not considered in the models of Pollack et al. (1996), Bodenheimer et al. (2000b), and Hubickyj et al. (2005). However, some of the simulations by these authors terminate solids accretion at a predetermined core mass, thereby mimicking the effects of planetesimal accretion by competing embryos.

Alibert et al. (2004, 2005) incorporated planetary migration, thereby allowing the planet to move into regions of the disk with undepleted reservoirs of planetesimals. In some cases, they follow the simultaneous accumulation of multiple planets, and in these simulations one planet can migrate into a region already depleted of planetesimals as a consequence of accretion by another core. However, planetary orbits rapidly decay into the Sun in those simulations that include migration at rates predicted by theoretical models of interactions of planets with a minimum mass solar nebula. Thus, Alibert et al. arbitrarily reduce planetary migration rates by a factor of ~30; it isn’t clear that this is a better approximation than that of completely ignoring migration, as done by Hubickyj et al. (2005) and others.

In order for cores to reach the required masses prior to isolation from their planetesimal supplies (equation (4)), models that do not incorporate migration (e.g., Hubickyj et al., 2005) need to assume that the surface mass density of solids in Jupiter’s region was at least 2–3× as large as the value predicted by “classical” minimum mass models of the protoplanetary disk (Weidenschilling, 1977; Hayashi, 1981). This is fully consistent with disk observations, and with models suggesting both that the giant planets in our solar system formed closer to one another than they are at present (Fernandez and Ip, 1984; Hahn and Malhotra, 1999; Thommes et al., 1999; chapter by Levison et al.) and that a large number of icy planetesimals were ejected from the giant planet region to the Oort cloud as well as to interstellar space (e.g., Dones et al., 2004). Models in which cores migrate relative to the planetesimal disk (e.g., Alibert et al., 2005), or in which solids can be concentrated by diffusive redistribution of water vapor (Stevenson and Lunine, 1988), baroclinic instabilities (Klahr and Bodenheimer, 2006), or gravitational instabilities (Durisen et al., 2005), can form planets in lower-mass disks. But all models are subject to the stronger constraints of heavy element abundances in giant planets and disk lifetime.

Inaba et al. (2003) have performed simulations of giant planet growth that incorporate a more sophisticated treatment of solid-body accretion. In their model, multiple planetary embryos stir smaller planetesimals to high enough velocities that planetesimal collisions are highly disruptive. Inaba et al. include envelope thermal evolution (albeit using a more simplified treatment than that employed by the above-mentioned groups) and planetesimal accretion cross sections that are enhanced by the presence of the envelope (Inaba and Ikoma, 2003). As a result of the competition between nearby growing cores, they require an initial surface mass density at 5 AU of about twice that of Hubickyj et al. (2005) for core growth to occur on timescales consistent with observational constraints on disk lifetimes. Specifically, with a solid surface density 25 g cm^{-2} at 5 AU and assuming full interstellar grain opacity within the protoplanet’s atmosphere, they can form Jupiter possessing a ~20 M_{\oplus} core in <4 m.y. If they reduce the grain opacity by a factor of 100, they get a Jupiter with a 7 M_{\oplus} core in 5 × 10^{6} yr in a disk with surface density 12.5 g cm^{-2}. They are not able to form Saturn in either of these cases.

3.2. Gas Accretion: Tenuous Extended Envelope Phase

The escape velocity from a planetary embryo with M > 0.1 M_{\oplus} is larger than the sound speed in the surrounding gaseous protoplanetary disk at temperatures where ice can
condense, so such an embryo can begin to accumulate a quasistatic atmosphere. As the atmosphere/envelope grows, it becomes optically thick to outgoing thermal radiation, and its lower reaches can get much warmer and denser than the gas in the surrounding protoplanetary disk. It undergoes Kelvin-Helmholtz contraction as the energy released by the accretion of planetesimals and gas is radiated away at the photosphere. A thick atmosphere expands the accretion cross section of the planet, especially for small solid bodies. At this stage, the key processes are the accretion of solids and the radiation of thermal energy. Most detailed models of this phase are spherically symmetric (one-dimensional). The energy released by accretion of planetesimals and envelope contraction heats the envelope and regulates the rate of contraction. This in turn controls how rapidly additional gas can enter the domain of the planet’s gravitational reach and be accreted. Because the opacity is sufficiently high, much of the growing planet’s envelope transports energy via convection. However, the distended very-low-density outer region of the envelope has thermal gradients that are too small for convection, but is so large that it acts as an efficient thermal blanket if it is sufficiently dusty to be moderately opaque to outgoing radiation.

During the runaway planetesimal accretion epoch, the protoplanet’s mass increases rapidly (Fig. 2). The internal temperature and thermal pressure increase as well, preventing substantial amounts of nebular gas from falling onto the protoplanet. When the rate of planetesimal accretion decreases, gas falls onto the protoplanet more rapidly.

As a planet grows, its envelope mass is a sensitive function of the total mass, with the gaseous fraction increasing rapidly as the planet accretes (Pollack et al., 1996). Accretion initially proceeds slowly, governed by the growth of the mass of the solid core and release of thermal energy from the envelope. When the envelope reaches a mass comparable to that of the core, the self-gravity of the gas becomes substantial, and the envelope contracts when more gas is added, so further accretion is governed by the availability of gas rather than thermal considerations. The time required to reach this epoch of rapid gas accretion is governed primarily by three factors: the mass of the solid core (larger core mass implies more rapid accretion); the rate of energy input from continued accretion of solids (such energy keeps the envelope large and slows further accretion of gas); and the opacity of the envelope (low opacity allows the radiation of energy that enables the envelope to cool and shrink, making room for more gas to be accreted). These three factors appear to be key in determining whether giant planets are able to form within the lifetimes of protoplanetary disks. For example, in a disk with initial $\sigma_g = 10 \, \text{g/cm}^2$ at 5.2 AU from a 1 $\text{M}_\odot$ star, a planet whose atmosphere has 2% interstellar opacity forms with a 16 $\text{M}_\oplus$ core in $2.3 \times 10^6$ yr; in the same disk, a planet whose atmosphere has full interstellar opacity forms with a 17 $\text{M}_\oplus$ core in $6.3 \times 10^6$ yr; a planet whose atmosphere has 2% interstellar opacity but stops accreting solids at 10 $\text{M}_\oplus$ forms in 0.9 m.y., whereas if accretion of solids is halted at 3 $\text{M}_\oplus$, accretion of a massive envelope requires 12 m.y. (Hubickyj et al., 2005). Thus, if Jupiter’s core mass is significantly less than 10 $\text{M}_\oplus$, then it presents a problem for formation models.

As estimates of the lifetimes of protoplanetary disks have decreased, a major concern has been whether or not giant planets can form faster than typical disks are dispersed, $\sim 2-5 \times 10^6$ yr. Planets can indeed form rapidly if they have sufficiently massive cores that accrete early and then stop growing and/or if the outer regions of their envelopes are transparent to outgoing radiation (have low opacities). But what are realistic values for these parameters? Observational

**Fig. 2.** Evolution of a giant protoplanet with $\sigma_{\text{init},Z} = 10 \, \text{g/cm}^2$ and grain opacity at 2% interstellar value. Details of the calculation are presented in Hubickyj et al. (2005). (a) The mass is plotted as a function of time, with the solid lines referring to the solids component of the planet, the dotted lines to the gaseous component, and the dot-dashed lines the total mass. Thick black curves: no solid accretion cutoff; thin black curves: solid accretion cutoff at 10 $\text{M}_\oplus$; gray curves: solid accretion cutoff at 5 $\text{M}_\oplus$. (b) The luminosity is plotted on a logarithmic scale as a function of time. Note that the cutoff runs are halted when the gas accretion rate reaches a limiting value defined by the rate at which the solar nebula can transport gas to the vicinity of the planet, whereas the planet in the run with no cutoff stops growing when $M_p = 1 \, \text{M}_\text{rup}$. The existence of a sharp peak in planetary luminosity during the phase of rapid gas accretion is physically plausible, but is likely to be somewhat lower and broader than shown in the plot because gas accretion almost certainly tapers off less abruptly than assumed for this calculation. Figure courtesy of O. Hubickyj.
constraints are quite weak. Limits upon the masses and locations of the heavy element components of the giant planets within our solar system were discussed in section 2. Atmospheric opacities and how the accretion rate of solids depends on time are quantities derived from planet formation models, and at present their values are quite ill constrained.

The ability of a planetary core to accrete gas does not depend strongly on the outer boundary conditions (temperature and pressure) of the surrounding disk, as long as there is adequate gas to be accreted (Mizuno, 1980; Stevenson, 1982; Pollack et al., 1996). The primary reason why giant planet formation is believed not to occur within a few AU of a star is the difficulty of forming a sufficiently massive core in the high Kepler shear environment of this region (Lissauer, 1987; Bodenheimer et al., 2000b).

The composition of the atmosphere of a giant planet is largely determined by how much heavy material was mixed with the lightweight material in the planet’s envelope. Accretion energy can lead to evaporation of planetary ices, and their mixing into the atmosphere can increase its mean molecular weight, allowing it to shrink and more gas to be trapped (Stevenson, 1982). As the envelope becomes more massive, late-accreting planetesimals sublimate before they can reach the core, thereby enhancing the heavy element content of the envelope considerably.

In the detailed thermal calculations of giant planet envelope accumulation performed to date, the accumulation of solids governs the accretion of gas. Yet apart from increasing the planet’s total mass, the effect of the extended gaseous envelope on the accretion rate of solids is minimal. But this would not be the case if most of the mass of condensed material in the disk near the planet’s orbit resided in very small solid bodies, nor if the planet migrated relative to solids in the disk (Kary et al., 1993).

### 3.3. Gas Accretion: Hydrodynamic Phase

As discussed in section 3.2, a protoplanet accumulates gas at a gradually increasing rate until its gas component is comparable to its heavy element mass. The rate of gas accretion then accelerates rapidly, and a gas runaway occurs (Pollack et al., 1996; Hubickyj et al., 2005). The gas runaway continues as long as there is gas in the vicinity of the protoplanet’s orbit.

The protoplanet may cut off its own supply of gas by gravitationally clearing a gap within the disk (Lin and Papaloizou, 1979). Such gaps have been observed around small moons within Saturn’s rings (Showalter, 1991; Porco et al., 2005). D’Angelo et al. (2003) used a three-dimensional adaptive mesh refinement code to follow the flow of gas onto accreting giant planets of various masses embedded within a gaseous protoplanetary disk. Bate et al. (2003) performed three-dimensional simulations of this problem using the ZEUS hydrodynamics code. In unperturbed disks, flows would increase with planet mass indefinitely. Using parameters appropriate for a moderately viscous minimum mass solar nebula protoplanetary disk at 5 AU, both groups found that <10 M\(_\oplus\) planets do not perturb the protoplanetary disk enough to significantly affect the amount of gas that flows toward them. Gravitational torques on the disk by larger planets drive away gas. Hydrodynamic limits on gas accretion reach to a few × 10\(^{-2}\) M\(_\oplus\) per year for planets in the ~50–100 M\(_\oplus\) range, and then decline as the planet continues to grow. An example of gas flow around/toward a 1 M\(_\text{Jup}\) planet is shown in Fig. 3. These calculations do not include the thermal pressure on the nebula from the hot planet, which is found to be the major accretion-limiting factor for planets up to a few tens of M\(_\oplus\) by the simulations discussed in section 3.2.

Calculations incorporating both hydrodynamic flows of gas in the disk and thermal physics of the planet are needed to fully understand the gas accretion rate by a growing planet. But it appears that the primary factor limiting growth of a planet smaller than a few dozen Earth masses is its ability to radiate energy, thereby allowing its envelope to shrink so that more gas can flow into the planet’s gravitational domain. For planets larger than ~100 M\(_\oplus\), thermal pressure from the envelope does not limit growth, but gravitational torques limit the flow of gas from the disk. Bate et al. (2003) find that gas accretion rates decline precipitously for planets more than a few times the mass of Jupiter, but that planets up to ~5 M\(_\text{Jup}\) can double in mass within a million years for nominal disk parameters. Thus, disks must be largely dispersed within ~10\(^{-4}\)–10\(^{6}\) yr after the onset of rapid accretion of gaseous envelopes by giant planets in order to explain the observed distribution (chapter by Udry et al.) of planetary masses.

### 3.4. Migration

A major uncertainty associated with the emergence of planets is their predicted orbital migration as a consequence of the gravitational torque between the disk and the planet (Goldreich and Tremaine, 1980; Ward, 1986; Bate et al., 2003). Planets that are more massive than Mars may be able to migrate substantial distances prior to the dispersal of the gaseous disk. Planets that are not massive enough to clear gaps around their orbits undergo Type 1 migration as a consequence of the difference between the repulsive torques that they exert on material interior and exterior to their orbits; in the linear regime these torques vary quadratically with planet mass, so migration rates are proportional to planetary mass. Planets that clear gaps around their orbits are subjected to Type 2 migration, by which they are dragged along with the evolving disk. Thus, it is quite possible that giant planets form several AU from their star and then migrate inward to the locations at which most extrasolar planets have been observed. Disk-induced migration is considered to be the most likely explanation for the “giant vulcan” planets with orbital periods of less than a week, because the Keplerian shear close to a star makes in situ formation of such objects quite unlikely (Bodenheimer et al., 2000b). Livio and Pringle (2003) find no basis to suggest that planetary migration is sensitive to disk metallicity, and conclude...
that the correlations between the presence of observable planets and stellar metallicity probably results from a higher likelihood of giant planet formation in metal-rich disks.

The difficulty with the migration models is that they predict that planets migrate too rapidly, especially Type 1 migration in the Earth- to Neptune-mass range that planetary cores grow through in the core nucleated accretion scenario. (Planets formed directly via gravitational instabilities would avoid the danger of Type 1 migration, but would be subject to a greater amount of Type 2 migration as a consequence of their early formation within a massive disk.) Moreover, because predicted migration rates increase as a planet moves inward, most migrating planets should be consumed by their star. However, a planet may end up in very close 51 Peg-like orbits if stellar tides can counteract the migration or if the disk has a large inner hole (Lin et al., 2000). Resolution of this rapid migration dilemma may require the complete and nonlinear analysis of the disk response to the protoplanet. Alternatively/additionally, planets may stop migrating if they approach a density enhancement interior to their orbits, which equalizes the positive and negative torques upon them in either a quasiequilibrium or stochastic manner that allows some “lucky” planets to survive (Laughlin et al., 2004), corotation torques might be able to slow down the migration of ~10 M⊕ objects (D’Angelo et al., 2003), and the small amounts of gas that leak into almost clear gaps may slow the migration of more massive planets. See Ward and Hahn (2000), Masset and Papaloizou (2003), Thommes and Lissauer (2005), and the chapter by Papaloizou et al. for more extensive discussions of planetary migration.

Fig. 3. The surface mass density of a gaseous disk containing a Jupiter-mass planet on a circular orbit located 5.2 AU from a 1 M∗ star. The ratio of the scale height of the disk to the distance from the star is 1/20, and the dimensionless viscosity at the location of the planet is \( \alpha = 4 \times 10^{-3} \). The distance scale is in units of the planet’s orbital distance, and surface density of 10^{-4} corresponds to 33 g/cm^2. The inset shows a closeup of the disk region around the planet, plotted in cylindrical coordinates. The two series of white dots indicate actual trajectories (real particle paths, not streamlines) of material that is captured in the gravitational well of the planet and eventually accreted by the planet. Figure courtesy of G. D’Angelo. See D’Angelo et al. (2005) for a description of the code used.
Many of the known extrasolar giant planets move on quite eccentric (0.2 < e < 0.7) orbits. These orbital eccentricities may be the result of stochastic gravitational scatterings among massive planets, some of which have subsequently merged or been ejected to interstellar space (Weidenschilling and Marzari, 1996; Levison et al., 1998; Ford et al., 2001), by perturbations of a binary companion (Holman et al., 1997), or by past stellar companions if the now-single stars were once members of unstable multiple-star systems (Laughlin and Adams, 1998). However, as neither scattering nor migration offer a simple explanation for those planets with nearly circular orbits and periods from a few weeks to a few years, the possibility of giant planet formation quite close to stars should not be dismissed (Bodenheimer et al., 2000b).

Most of the observed extrasolar giant planets orbit between a few tenths of an AU and a few AU from their star, i.e., they are located much closer to their stars than Jupiter is from our Sun. These planets may have formed farther from their star and migrated inward, but without a stopping mechanism, which isn’t known at these distances, they would have fallen into the star. Lissauer (2001) suggested that the orbits could be explained if disks cleared from the inside outward, leaving the planets stranded once they were too far interior to the disk for strong gravitational coupling to persist. Observations of the 2:1 resonant planets orbiting GJ 876 by Marcy et al. (2001; see also Rivera et al., 2005) support such a model, as do data that imply that the star CoKu Tau/4 has a disk with an inner hole (Forrest et al., 2004).

4. OUTSTANDING QUESTIONS

What are the data on giant planet composition and structure telling us? Are the enhancements in heavy elements in the atmospheres of the giant planets within our solar system the result of mixing of material throughout the planet, and thereby reflective of the planets’ bulk compositions, or were they produced by a late veneer of planetesimal accretion or accretion of gas from a nebula depleted in H2 and He? Guillo and Hueso, in preparation)? And why do they appear to be dominated by very low condensation temperature planetesimals, which seem required to produce the comparable enrichments of gasses of different volatilities (Owen et al., 1999)? What are the masses of the planetary cores,  and are these reflective of core masses during the accretionary epoch, or have they been increased by settling or reduced by convective mixing? Progress on answering this question depends mainly on improvements in our understanding of the high pressure behavior of hydrogen.

The core nucleated accretion model provides a sound general framework for understanding the formation of giant planets. According to this scenario, giant planets begin their growth as do terrestrial planets and smaller bodies, but they become massive enough to gravitationally accrete substantial amounts of the abundant light gases prior to the dispersal of the protoplanetary disk. However, many first-order questions remain:

How rapid do solid cores accrete in the giant planet formation region? The solid core provides a gravitational potential well for the gas to fall into. Counteracting this tendency, ongoing accretion of solids provides additional heating that expands the planet’s envelope, limiting accretion of gas, especially if the solids sink deeply into the gravitational potential well, down to or near the core (Pollack et al., 1996). Additionally, if continued accretion of solids provides a substantial amount of small grains that persist in the planet’s radiative atmosphere, the resulting thermal blanket reduces planetary luminosity.

Are the atmospheres of growing giant planets good thermal blankets (high optical depth to outgoing radiation caused by the presence of abundant small grains) or nearly transparent? Models suggest that the ability to radiate energy is a key factor in determining how rapidly an atmosphere contracts, thereby allowing the planet to continue to grow. Low-opacity atmospheres allow giant planets to form much more rapidly and/or with significantly smaller cores than do high-opacity atmospheres. Small grains are provided to the planet both from the disruption and ablation of accreted planetesimals and entrained in the accreted gas, but the amounts and residence times are quite uncertain. Such grains are not present in large quantities in the atmospheres of giant planets in our solar system, nor are they detected in cool brown dwarfs (chapter by Marley et al., but do they settle downward fast enough to allow the atmosphere to be transparent during the formation epoch (Podolak, 2003)?

How does a (growing or fully formed) giant planet interact with the surrounding protoplanetary disk? Models of a planet gravitationally clearing a gap around itself and accretion of material through a partially formed gap give “reasonable” results. But predicted migration rates are simply too rapid for the survival of as many giant planets as are observed within our solar system and around nearby Sun-like stars. Either giant planets form much more readily than predicted by models (perhaps because disks are significantly more massive) and the survivors that we see are a tiny fraction of the bodies formed, or migration rates have been substantially overestimated.

5. SUMMARY

The smoothness of the distribution of masses of young M stars, free-floating brown dwarfs, and even free-floating objects somewhat below the deuterium-burning limit argues strongly that these bodies formed in the same manner, i.e., via collapse, in some cases augmented by fragmentation. In contrast, the mass gap in nearby companions to Sun-like stars (the brown dwarf desert) is convincing evidence that most if not all of the known giant planets formed in a different manner.

Various models for giant planet formation have been proposed. According to the prevailing core nucleated accretion model, giant planets begin their growth by the accumulation of small solid bodies, as do terrestrial planets. However, unlike terrestrial planets, the growing giant planet
cores become massive enough that they are able to accumulate substantial amounts of gas before the protoplanetary disk dissipates. The primary question regarding the core accretion model is whether planets can accrete very massive gaseous envelopes within the lifetimes of typical gaseous protoplanetary disks. Another important question is whether or not proto-Jupiter’s core was sufficiently massive to capture large quantities of hydrogen and helium.

The main alternative giant planet formation scenario is the disk-instability model, in which gaseous planets form directly via gravitational instabilities within protoplanetary disks. The formation of giant planets via gas instability has never been demonstrated for realistic disk conditions. Moreover, this model has difficulty explaining the supersolar abundances of heavy elements in Jupiter and Saturn, and it does not explain the origin of planets like Uranus and Neptune. Nonetheless, it is possible that some giant planets form via disk instability, most likely in the regions of protoplanetary disks distant from the central star, where Keplerian shear is small and orbital timescales are long. Additionally, a few planets probably form via fragmentation of molecular cloud cores during collapse or are captured via exchange reactions involving (usually young) binary/multiple stars.

Most models for extrasolar giant planets suggest that they formed as Jupiter and Saturn are believed to have (in nearly circular orbits, far enough from the star that ice could condense), and subsequently migrated to their current positions, although some models allow for in situ formation. Gas giant planet formation may or may not be common, because the gas within most of protoplanetary disks could be depleted before solid planetary cores grow large enough to gravitationally trap substantial quantities of gas. Additionally, an unknown fraction of giant planets migrate into their star and are consumed, or are ejected into interstellar space via perturbations of neighboring giant planets, so even if giant planet formation is common, these planets may be scarce.

While considerable progress toward understanding the internal structure and formation of giant planets has been made recently, major questions remain. As we continue to place new data and simulation results into the jigsaw puzzle, some present pieces will surely need to be repositioned or discarded. With the wealth of new information being provided, we expect the picture to become clearer in the near future.

Acknowledgments. We thank O. Hubickyj and an anonymous referee for providing us with constructive comments on the manuscript. This work was partially supported by the NASA Outer Planets Research Program under grant 344-30-99-02 (J.J.L.), and by NASA’s Planetary Geology and Geophysics Program (D.J.S.).

REFERENCES

Kary D. M. and Lissauer J. J. (1994) In Numerical Simulations in Astro-