LOW-MASS STAR AND PLANET FORMATION*

ALAN P. BOSS
Department of Terrestrial Magnetism, Carnegie Institution of Washington, 5241 Broad Branch Road
NW, Washington, DC 20015
Received 1989 June 22

ABSTRACT

Our current understanding of low-mass star and planet formation is reviewed through a brief comparison of the results of cosmogonical models with observations ranging from studies of star-forming regions to searches for planetary companions to low-mass stars. Five key phases are described, starting from the dense, interstellar cloud cores that form low-mass stars, through the protostellar collapse and fragmentation phase, to the formation of a protostellar object accreting gas from the surrounding protostellar disk and cloud envelope, on to the phase where planets are formed in the protostellar disk, and ending with the dissipation of the bulk of the protostellar disk and the appearance of an optically visible, pre-main-sequence star. While we appear to have developed a generally acceptable outline for the formation of low-mass stars and planetary systems, many important details remain uncertain at best, and suggestions are made for future theoretical and observational efforts.

Key words: star formation–planet formation

1. Introduction

The close association of young stars with clouds of gas and dust in the disk of our Galaxy implies that stars form by the compression and self-gravitational contraction of dense interstellar clouds. Laplace suggested that the planets of our solar system formed from a rotating, flattened cloud that also produced our Sun. Laplace’s nebula hypothesis explains the gross orbital characteristics of our planetary system, considerably narrows the possibilities for planet formation, and, of most importance for this review, the general acceptance of Laplace’s hypothesis means that the problems of star and planet formation are inextricably linked. Developing a self-consistent cosmogony for stellar and planetary systems will not only answer fundamental questions about the origin of our own solar system but also will give crucial insight into the possible prevalence of other planetary systems in our Galaxy.

Previous reviews in this series have described observations and theoretical models of winds and bipolar outflows (Pudritz 1986), the history of theoretical work on spherically symmetric, pre-main-sequence stars (Stahler 1988a), and observational aspects of low-mass star formation (Wilking 1989). The present review necessarily will be more of a survey than an in-depth, detailed examination, because of the great amount of research directed toward solving the coupled problems of star and planet formation; the last major conference on the entire subject produced a 1293-page book (Protostars and Planets II, Black and Matthews 1985). Additional details about star formation may be found in the conference proceedings edited by Pudritz and Fich (1988) and Dupree and Lago (1988) as well as in the following reviews: Bally (1986) and Shu, Adams, and Lizano (1987) on star formation in molecular clouds; Boss (1987a) on cloud collapse; Cameron (1988) and Boss, Morfill, and Tscharnuter (1989) on solar nebula models; Weidenschilling (1988) on early phases of planetary accumulation and Wetherill (1986) on late phases; and Lada (1985) on winds and outflows.

The plan for this review is to present a side-by-side comparison of theoretical models and observational studies of star and planet formation, with the obvious goal
being to assess how well our theoretical understanding agrees with astronomical reality. For organizational convenience, the review will be divided into five largely chronological phases that can be discussed individually, though some of these phases may overlap in time, and certainly all can only be fully understood in the context of the entire framework we are trying to assemble. Successive sections will thus deal with the dense interstellar clouds that produce low-mass stars, cloud collapse and fragmentation, formation of protostellar objects and disks, planet formation in protostellar disks, and finally, removal of the residual nebula and, thus, termination of the formation process.

2. Dense Interstellar Clouds

2.1 Observation

Because of the strong extinction of optical light in dusty star-forming clouds, investigation of the dense cloud phase of star formation usually involves millimeter and radio-wave observations of molecular species, such as CO and NH$_3$, that can be used to infer the properties of the molecular hydrogen forming the bulk of these clouds (and ~ 1% of the mass of the galaxy). Knowledge of dense cloud properties is essential for defining those physical processes likely to be important in subsequent phases of protostellar evolution and for determining the detailed initial conditions for the collapse phase.

Star formation occurs on widely varying scales, ranging from relatively isolated, low-mass (~ 1 $M_\odot$) dark clouds to giant molecular cloud complexes containing as much as ~ $10^6$ $M_\odot$ of matter. All of these objects are gravitationally bound, a condition obviously conducive to star formation, which requires self-gravity to overwhelm factors resisting collapse such as thermal pressure, magnetic fields, turbulent motions, and rotational effects. Though they are also sites of considerable low-mass star formation, giant molecular cloud complexes are dominated by high-luminosity, high-mass stars that have an appreciable effect on their surroundings. Our desire to focus on low-mass star formation directs our attention to the dark cloud complexes with masses $10^5$–$10^4$ $M_\odot$, mean densities of $10^2$–$10^3$ molecules cm$^{-3}$, and sizes on the order of 10 pc (e.g., Myers 1987). In dark cloud complexes (e.g., Taurus-Auriga, Ophiuchus, and Lupus), low-mass star formation occurs largely unaccompanied by high-mass stars (as implied by the term “dark cloud”). While high-mass star formation may require an external stimulus in the form of shocks from cloud-cloud collisions or expanding H II regions, low-mass star formation appears to be much more spontaneous, i.e., dependent on the internal evolution of dark cloud complexes (e.g., Lada 1978).

Dark cloud complexes contain clusters of localized mass concentrations, termed dark or dense cloud cores, with masses in the range 0.3–10 $M_\odot$, mean densities of $10^4$–$10^5$ cm$^{-3}$, sizes on the order of 0.1 pc, and temperatures close to 10 K. Unfortunately, dark cloud cores have sizes comparable to the angular resolution of single dish telescopes and, generally, are too faint to be detected by current interferometers (Myers 1987). Hence, we know little about their internal structure, such as the detailed density and velocity fields. Yet such details are crucial because dark cloud cores appear to be the most prevalent site of low-mass star formation. Infrared (IR) observations by the IR Astronomical Satellite (IRAS) have shown that roughly half of the dark cloud cores contain embedded IR (optically invisible) sources thought to be low-mass young stellar objects (Beichman et al. 1986). Undetected embedded IR sources could raise the fraction of cloud cores containing young stellar objects to much closer to unity. This observation is strong evidence that dark cloud cores or, at least their predecessors, define the initial conditions for the formation of many low-mass stars.

There is good evidence that dark-cloud envelopes are strongly centrally condensed (e.g., with power-law density profiles such as $\rho \propto r^{-2}$; Snell 1981; Arquilla and Goldsmith 1985), but considering that the majority of these clouds have already contracted to form dark cloud cores containing protostars, this strong envelope density gradient may be more a result of the star-formation process than an indication of initial conditions. Because of the limited spatial resolution of the observations, the dark cloud cores themselves are simply modeled as uniform density spheres for purposes of modeling the observations of dark cloud envelopes (Arquilla and Goldsmith 1985).

Observations of dark cloud complexes are also critical for determining which physical processes dominate star-forming cloud dynamics. There is evidence (though largely indirect, because of the great difficulty in directly measuring field strengths through the Zeeman effect) that dark cloud complexes are controlled primarily by magnetic fields (e.g., large-scale magnetic-field alignments inferred from polarization studies, Moneti et al. (1984), and the apparent need for magnetic support to avoid producing stars at an unacceptable rate). Individual dark clouds with mean densities of $\sim 1 - 5 \times 10^5$ cm$^{-3}$ tend to have their axes of rotation and minor axes (when appreciably flattened) aligned with the ambient magnetic-field direction (Heyer et al. 1987), as would be expected if magnetic fields are dynamically important. Furthermore, estimates of the magnetic-field contribution to molecular linewidths in large dark clouds also suggest the importance of magnetic fields for dark clouds (Myers and Goodman 1988), assuming that the nonthermal portion of the linewidth is due to turbulent motions associated with Alfvén (hydromagnetic) waves.

However, there is also evidence that the dynamical influence of the magnetic field decreases with decreasing cloud size (or increasing cloud density). Myers and Good-
man (1988) found that a constant magnetic-field strength of $\sim 30 \mu G$ was sufficient to account for the nonthermal portion of molecular linewidths for clouds ranging in size from 0.01 pc to 100 pc (Fig. 1). Because smaller clouds generally have higher densities, this implies that the magnetic field is not as important for the dynamics of small clouds, where thermal pressure and self-gravity begin to dominate. The decreased importance of magnetic fields with increased cloud density appears to be attributable to ambipolar diffusion (Mestel and Spitzer 1956), the continual slippage of field lines and ions (with fractional ionizations $\sim 10^{-7}$) past the neutral bulk of the cloud. This picture is strengthened by the recent Zeeman detection of a 27-$\mu G$ field in the dark cloud Barnard 1 (Goodman et al. 1989); previous Zeeman detections also support the idea of a $\sim 30-\mu G$ field being characteristic of cold dark clouds (Crutcher, Kazès, and Troland 1987). The picture is further substantiated by Heyer’s (1988) work, which found no correlation between the direction of the magnetic field and the rotational or minor axes of dense ($\sim 10^4 \text{ cm}^{-3}$) cloud cores, contrary to the results for dark clouds with densities $< 10^4 \text{ cm}^{-3}$. This lack of correlation implies that thermal pressure and self-gravity control the dynamics of dark cloud cores, rather than large-scale, ordered magnetic fields, though small-scale, disordered magnetic fields could still be important.

A small amount of initial rotation can have a large effect on a cloud that has to increase in density by a factor of $\sim 10^{20}$ before forming a young star. However, even rapid rotational motions are hard to detect because the Doppler shift associated with rapid rotation is often comparable to thermal linewidths. Linewidths can be broadened by small-scale turbulence or large-scale motions not associated with rotation, though there is no evidence for global collapse of dark cloud complexes. Goldsmith and Arquilla (1985) have summarized the evidence for rotation (see also Arquilla and Goldsmith 1986) and found that the specific angular momentum $J/M$ decreases with decreasing cloud size $\ell$ ($J/M \propto \ell^{1.45}$) to values of $\sim 10^{11} \text{ cm}^2 \text{s}^{-1}$ for dark-cloud cores ($\ell \sim 0.1 \text{ pc}$). Because only the most rapidly rotating clouds show evidence for rotation, this estimate of $J/M$ must be considered as an upper bound for the specific angular momentum of typical dark cloud cores. The form of the rotation curve (e.g., uniform angular velocity or differential rotation) in dark cloud cores is even more uncertain than the magnitude of the angular velocity.

2.2 Theory

One long-standing mystery in star formation is identification of the process through which relatively massive interstellar clouds subdivide into stellar-sized objects. The discovery that dark cloud cores are the sites of substantial low-mass star formation means that processes operating within globally quasi-static, dark cloud complexes are capable of producing self-gravitating clumps with stellar masses. It has been suggested that supersonic turbulence prevents global collapse of dark cloud complexes, but that the turbulent spectrum is damped at smaller ($\sim 0.1 \text{ pc}$) scales (Bonazzola et al. 1987). A sufficiently large density perturbation could then drive a small-scale region into collapse, resulting in the formation of a dark cloud core and embedded protostar. However, supersonic turbulence quickly dies out unless it is being continually fed; Norman and Silk (1980) proposed that winds from T Tauri stars could replenish the supersonic winds in dark cloud complexes. Norman and Silk (1980) also attributed the existence of two distinct phases in dark cloud complexes (cold, dense clumps and a warmer inter-
clump medium) to collisions between clumps, leading to self-gravitational collapse of the clumps and, hence, T Tauri star formation, with the T Tauri stars then heating up the interclump medium.

Considering the observational evidence for magnetic support of dark cloud complexes, processes involving magnetic fields are also likely to be important for explaining fragmentation on the cloud complex level. Unfortunately, studying fragmentation in full detail requires a three-dimensional theoretical model, whereas the considerable amount of analytical and semi analytical work on magnetic fields in dark clouds is restricted to one or two dimensional approximations (reviewed by Mouschovias 1981, 1987). Some numerical work has been done on the three-dimensional collapse of isothermal clouds with frozen-in magnetic fields and has shown that dynamically important magnetic fields prevent fragmentation of the cloud in all cases, regardless of orientation of the magnetic field and the rotation axis, when rotation is present (Dorfi 1982; Benz 1984; Phillips 1986). When sufficiently strong, the magnetic field may even prevent the collapse of the cloud (Dorfi 1982). These numerical calculations are limited to frozen-in fields and relatively short time intervals (a few free-fall times at the initial cloud density), however, so it is not clear what would happen in a cloud with significant ambipolar diffusion over much longer time periods. Campbell and Mestel (1987) used the virial theorem to argue that no fragmentation occurs in a flux-frozen cloud, but that fragmentation could occur if ambipolar diffusion is significant. Clearly more theoretical work is needed to understand how dark cloud cores might develop in magnetically supported dark clouds and cloud complexes.

Frozen-in magnetic fields can be very efficient at magnetic braking—that is, enforcing corotation with the surrounding medium and thereby removing angular momentum from a contracting cloud via Alfven waves along the field lines, as outlined by Mestel (1965) and rigorously calculated by Mouschovias and Paleologou (1979, 1980; see also Pudritz and Silk 1987; Königl 1987). A numerical study of three-dimensional, frozen-in magnetic fields found that the time scale for magnetic braking is about ten times faster when the rotation axis is perpendicular to the magnetic-field direction than when the two are parallel (Dorfi 1982). This is because magnetic-field lines twist easier than they stretch. Components of rotation perpendicular to the magnetic-field direction thus are damped preferentially by magnetic braking, leaving behind the parallel component. Significant magnetic braking can occur at lower densities (and, hence, at earlier times in a contracting cloud) than the densities where ambipolar diffusion becomes important (Mouschovias, Paleologou, and Fiedler 1985). Hence, one expects to see rotation axes aligned with magnetic-field directions in magnetically dominated clouds. Because matter can always flow down field lines, one also expects to see clouds flattened into a disk perpendicular to a dominant magnetic field (e.g., numerically demonstrated by Black and Scott 1982). These theoretical conclusions form the basis for the interpretations of the observational work mentioned earlier (Moneti et al. 1984; Heyer et al. 1987) and, hence, are in agreement with observation by design; however, the validity of this picture of magnetic dominance of dark cloud complexes can be determined only by further measurements of magnetic-field strengths in dense molecular clouds.

The observational evidence for significant ambipolar diffusion and, consequently, a decreased role for magnetic fields on small size scales (Myers and Goodman 1988) and for densities $n > 10^4$ cm$^{-3}$ (Heyer 1988) is supported by the estimates of Mouschovias (1987; this paper also contains a more-detailed comparison of the predictions of magnetically dominated clouds with observations). In particular, the dark cloud field strengths found by Goodman et al. (1989) are in agreement with a constant-mass cloud evolution model of Mouschovias (1987). Furthermore, Mouschovias et al. (1985) have shown that ambipolar diffusion effects should become substantial once densities on the order of $10^4$–$10^6$ cm$^{-3}$ are reached, in good agreement with these observations, considering that there is likely to be some variation in this critical parameter from one star-forming region to another. Note that this result is significantly different from the predictions of Nakano and Umebayashi (1986), who suggest that significant loss of magnetic flux can only occur for very high-density clouds ($> 10^{12}$–$10^{19}$ cm$^{-3}$); much higher densities are implied here because the dominant process for flux dissipation is implicitly assumed to be Ohmic dissipation (Joule heating) rather than ambipolar diffusion (see also El-Nawawy, Aiaid, and El-Shalaby 1988).

Fleck (1987) has suggested that the source of molecular cloud angular momentum is turbulence in the interstellar medium and has shown that, in this case, the specific angular momenta of clouds should vary as $J/M \propto \ell^4$ to $J/M \propto \ell^{45}$, depending on the exact form of the turbulent spectrum, which compares well with the observational results ($J/M \propto \ell^{14}$) of Goldsmith and Arquilla (1985). Fleck (1987) also states that clouds larger than the turbulent correlation length $\ell_0$ should have strongly reduced $J/M$ because, on scales larger than $\ell_0$, the turbulent motions will average out to nearly zero. However, this prediction conflicts with the observations tabulated by Goldsmith and Arquilla (1985) which show evidence of increasing $J/M$ with size for large clouds where $\ell \gg \ell_0$ (e.g., Taurus cloud complex, Heiles Cloud 2); indeed, rapid rotation in these large clouds is necessary for establishing the exponent in the $J/M(\ell)$ relation. Multiple
correlation lengths in a single cloud might resolve this discrepancy (Fleck, private communication). Contraction of a cloud with conserved angular momentum and mass leads to conserved $J/M$, while contraction at conserved angular velocity yields $J/M \propto \ell^2$. The latter would be the case if an extremely strong magnetic field were able to enforce strict corotation with the surrounding medium during contraction. Evidently contraction in the presence of a magnetic field that is not quite able to enforce rigid rotation could result in a $J/M(\ell)$ relation that is also in agreement with that observed ($J/M \propto \ell^{1.5}$).

3. Cloud Collapse and Fragmentation

3.1 Theory

Because of the great observational difficulties associated with studying the cloud-collapse phase of low-mass protostellar evolution (e.g., relatively short time scales, as well as optical obscuration by dust, high optical depths in molecular lines, and large scales much less than 0.1 pc), and the considerable theoretical effort put into understanding collapse in the last two decades, theoretical models of this phase are considerably more advanced than observations and so will be discussed first. This section will concentrate on nonmagnetic collapse calculations, in large part because most of the theoretical effort has been directed toward the simpler problem of nonmagnetic collapse, and, consequently, models of nonmagnetic collapse have been much more fully developed. Magnetic-collapse calculations were mentioned in the previous section in the context of lower density phases of dark cloud evolution when magnetic fields are more likely to have been dominant.

Observations of dark cloud cores (or their predecessors) provide the initial conditions for theoretical models of cloud core collapse. With two important exceptions, it is remarkable how similar dark cloud core conditions are to the “classical” initial conditions used for protostellar collapse posed by Larson (1969): an isolated 1 $M_\odot$, spherical cloud of 10 K molecular hydrogen gas, at rest at a uniform density of $\sim 10^5$ cm$^{-3}$. The first important difference between these simplified initial conditions and dark cloud cores is in the degree of central condensation. Presumably the collapse phase begins once a quasi-equilibrium configuration is forced into collapse for some reason. A uniform density configuration cannot be in pressure equilibrium, so the theoretical initial conditions are obviously artificial in this respect. Uniform density does have the advantage of being the same assumption as is used in the analysis of the Jeans instability, where the critical mass is derived for which self-gravitation overcomes the stabilizing effects of thermal pressure. As previously discussed, dark cloud cores (or their predecessors) are probably centrally condensed, quasi-static objects. The second difference is the neglect of magnetic fields; presumably dark cloud cores are supported by a combination of thermal pressure and magnetic fields, and it is the progressive loss of the magnetic-field support through ambipolar diffusion that allows the collapse phase to begin. While the effect of starting from centrally condensed initial conditions instead of uniform initial conditions has been studied and will be discussed (e.g., Shu 1977; Boss 1987c, see Fig. 2), starting collapse from a partially magnetically supported configuration with significant ambipolar diffusion remains largely unstudied (but see Black and Scott 1982).

The history of theoretical progress on protostellar collapse has been one of increasing dimensionality. The pioneering calculations (e.g., Bodenheimer 1968; Larson 1969) assumed spherical symmetry, meaning that only one-dimensional equations needed to be solved. The resulting numerical simplification allowed Larson (1969) to follow the collapse of a 1 $M_\odot$ cloud all the way from the dense cloud phase to the formation of a pre-main-sequence star; in contrast, multidimensional collapse calculations have not yet been able to evolve a protostar through all these phases.

The classical initial conditions used by Larson (1969) are optically thin in the infrared (IR), allowing dust grains in the cloud to radiate away the compressional energy produced by collapse. As a result, the initial collapse phase occurs isothermally ($T \sim 10$ K). Larson (1969) found that the collapse of an initially uniform-density, nearly Jeans-stable cloud rapidly becomes nonhomologous, that is, the density profile does not remain invariant under a scale transformation. Instead, a pressure gradient develops in the envelope as the effects of the finite cloud size propagate inward, and this pressure gradient retards the collapse of the envelope. As a consequence, the density becomes higher at the center of the cloud than in the envelope, and because higher density regions collapse faster (the free-fall time $t_f \propto \rho^{-1/2}$), the central density begins to run away compared to the envelope. This runaway is halted when the center of the cloud becomes optically thick in the IR ($\rho \sim 10^{-15}$ g cm$^{-3}$); thereafter, compressional energy is trapped, the temperature becomes nonisothermal, and rising thermal pressure stops the collapse at the cloud center. The result is formation of a protostellar core surrounded by an infalling envelope with $\rho \propto r^{-2}$. This “outer core” accretes mass from the envelope and contracts to temperatures high enough ($\sim 3000$ K) for molecular hydrogen dissociation. Thermal energy is used to dissociate the hydrogen, and the consequent partial loss of thermal pressure support initiates another collapse phase. The center of the outer core collapses to nearly stellar densities ($\sim 0.01$ g cm$^{-3}$) before being halted again by rapidly rising thermal pressure once molecular hydrogen dissociation is complete, forming the “inner” protostellar core.

In the classical picture, the inner core is the final protostellar core on
Fig. 2—Final isodensity surfaces for three numerical models of the collapse of rotating, nonisothermal, protostellar clouds (from Boss 1987c). Each model started with 1 $\mathcal{M}_\odot$ of gas at 10 K and with equivalent ratios of thermal ($\alpha = 0.13$) and rotational ($\beta = 0.0016$) energies to gravitational energy; only the initial (spherically symmetric) density profile was varied. Model (a) started with uniform density, (b) with an exponential density profile 20 times higher at the center than at the boundary, and (c) with a power law ($p \propto r^{-1}$) profile. Region shown for each model is about 330 AU across. Densities of the plotted surfaces are $3 \times 10^{-14}$ g cm$^{-3}$ for (a), $5 \times 10^{-15}$ g cm$^{-3}$ for (b), and $1 \times 10^{-15}$ g cm$^{-3}$ for (c). Binary fragmentation is suppressed as the initial degree of central concentration is increased.
which the bulk of the cloud envelope eventually collapses. Larson's (1969) conclusions about spherically symmetric protostars have remained largely unchanged even after substantial improvements in the calculational techniques and physics (Winkler and Newman 1980; Stahler, Shu, and Taam 1981).

Shu (1977) has explored the alternative of starting collapse from a strongly centrally condensed configuration, in this case the singular isothermal sphere with \( \rho \propto r^{-2} \). Because the initial configuration is in equilibrium, albeit unstable, the initial collapse phase differs from the Larson (1969) picture. In the Shu (1977) model, collapse occurs from the "inside out". An external perturbation propagates inward and collapse begins at the center of the cloud, producing an expansion wave that moves outward. Successive layers of the envelope then begin to collapse once the expansion wave reaches their radius. Thereafter, the evolution is very similar to that in the classical model. In terms of the final stages of protostellar evolution, inside-out collapse is basically indistinguishable from the classical model (e.g., Stahler, Shu, and Taam 1980); the main difference is that the inner-core mass-accretion rate is roughly constant in the former case but rises and then falls in the latter case.

The next step in this theoretical progression was to consider the effects of rotation, which requires at least a two-dimensional calculation, typically assuming symmetry about the rotation axis ("axisymmetry"). Observationally, the magnitude of dark cloud rotation rates is still poorly known, so theoretical efforts have been directed toward understanding the effects of \( J/M \) varied between effectively zero and the maximum amount consistent with initial gravitational contraction. Larson (1972) and Black and Bodenheimer (1976) found that an initially rapidly rotating, uniform density cloud collapsed to form a ring around the rotation axis, in the cloud midplane. The axisymmetric ring forms because in rotating collapse, fluid elements move inward until all the gravitational attraction is needed to balance their centrifugal motion, at which point the fluid stops moving toward the rotation axis and may even rebound away from the axis and collide with other fluid elements. The result is a local density maximum that pulls in other matter through the self-gravity of the growing ring. Larson (1972) pointed out that an axisymmetric ring is likely to be unstable to fragmentation in a fully three-dimensional calculation and proposed ring breakup as a means of forming binary and multiple protostars.

The collapse of an axisymmetric rotating cloud with an initially singular density profile (\( \rho \propto r^{-2} \)), however, does not lead to ring formation (Terebey, Shu, and Cassen 1984); the strong initial central concentration results in formation of a flattened disk with maximum density at the center of the disk. If there is insufficient rotation to result in ring formation, even initially uniform density clouds form disks rather than rings (Boss and Haber 1982). Disks are considerably less likely to undergo fragmentation than rings.

The final step in this theoretical progression was to allow protostellar clouds to undergo arbitrary deformation, which requires a fully three-dimensional calculation ("nonaxisymmetric") but allows protostellar fragmentation to be studied directly. Nonaxisymmetric calculations showed that binary fragmentation could occur either through an intermediate ring phase, or directly in a cloud with a moderately large, initial bar-shaped density perturbation (e.g., Larson 1978; Boss 1980; Bodenheimer, Tohline, and Black 1980). Sustained isothermal collapse of an initially uniform density cloud always results in binary or multiple fragmentation, provided substantial rotation (\( \beta = F_{\text{rotational}}/E_{\text{gravitational}} \) > 0.02 or \( J/M > 2 \times 10^{30} \text{ cm}^2 \text{ s}^{-1} \)) is present initially, whereas slower rotating, nearly Jeans-stable clouds generally do not undergo fragmentation prior to outer core formation (Boss 1985). Fragmentation of an isothermal cloud with \( \alpha = E_{\text{thermal}}/E_{\text{gravitational}} \) ~ 0.2–0.7 typically occurs shortly after one free-fall time, once a rotationally flattened configuration has formed in the midplane. Though not yet directly demonstrated, hierarchical fragmentation also appears to be a likely outcome in the isothermal regime: Binary fragments have greatly reduced values of \( \alpha \) and may very well subdivide during their own collapse (Bodenheimer 1978; Boss 1986).

Once densities reach \( \sim 10^{-13} \text{ g cm}^{-3} \), compressional heating begins and thermal pressure starts to retard fragmentation. However, fragmentation during this nonisothermal collapse phase can still occur (Boss 1986). If the thermal energy of the cloud is low initially (\( \alpha < 0.2 \)), as is the case for fragments produced during the isothermal phase, then thermal pressure is unable to prevent the growth of the initial density perturbation, and fragmentation occurs within roughly a free-fall time. Nonisothermal fragments have increased values of \( \alpha \) and, hence, are unlikely to undergo hierarchical fragmentation; instead, they typically form rotationally flattened, quasi-equilibrium bodies that are the triaxial versions of the outer cores found in spherically symmetrical collapse. Low-mass (\( < 0.05 \mathcal{M}_\odot \)) clouds are unable to fragment in the nonisothermal regime because thermal pressure dominates their relatively smaller self-gravity. The cessation of fragmentation associated with outer core formation in the nonisothermal regime implies that the minimum protostellar mass formed through the hierarchical fragmentation of molecular clouds is \( \sim 0.01 \mathcal{M}_\odot \) (Boss 1986).

Transient binary fragmentation has been obtained following the collapse and formation of a rapidly rotating inner core (Boss 1989b), but in this case the nascent binary decays into a single object because of loss of orbital
angular momentum to trailing spiral arms through gravitational torques. The phenomenon of nascent binary system decay in quasi-static, rapidly rotating objects was first noticed in calculations of the dynamic fission instability (Durisen et al. 1986) and is a strong argument against producing binary stars through the fission mechanism (see Boss 1988).

The previous three-dimensional results are all based on starting from uniform density initial conditions, conditions that may be unduly restrictive. Starting collapse from a moderately centrally condensed configuration has only a slight effect on cloud fragmentation (Boss 1987c), as shown in Figure 2. However, starting collapse from a strongly centrally condensed, power-law density profile (e.g., Shu 1977) appears to completely stifle binary fragmentation (Fig. 2); the initial density “singularity” is too strong for fragmentation to overcome.

3.2 Observation

As mentioned earlier in this section, observations of the collapse phase of low-mass star formation are very difficult, so difficult in fact that there is as yet no indisputable evidence of low-mass cloud collapse, in spite of significant observational efforts (e.g., Myers 1980; see also Anglada et al. 1987). Direct evidence for collapse does exist for larger-scale clouds in the process of high-mass star formation: Welch et al. (1987) and Keto, Ho, and Reid (1987) have observed gravitational infall associated with ultracompact H II regions around massive stars. Mezger et al. (1988) claim to have detected several massive protostars still in the isothermal phase of collapse, though the detection does not include direct evidence for infall.

Also, as noted previously, much of the fragmentation responsible for producing dark cloud cores and molecular cloud clumps appears to be associated with quasi-static evolution rather than collapse, though certain relatively large-scale features have been identified as possible products of collapse (e.g., the ring from Schloerb and Snell 1984). Direct studies of the fragmentation theoretically anticipated to occur during the collapse phase of dark cloud cores (on scales much less than 0.1 pc) are hindered by the limited spatial resolution of radio telescopes; interferometric measurements will be necessary for progress here.

The degree of protostellar fragmentation can also be assessed indirectly through observations of binary and multiple stars. Main-sequence binary-star properties, such as frequency of companions and semimajor axis distributions, are reasonably well established and appear to be in basic agreement with theoretical predictions based on a limited amount of hierarchical fragmentation during dark cloud core collapse (see Boss 1988). However, knowledge about the properties of binary pre-main-sequence stars is just now becoming available (e.g., Reipurth 1988); clearly binary pre-main-sequence stars will provide the best means of testing the accuracy of our theoretical understanding of binary-star formation.

4. Protostellar Objects and Disks

4.1 Theory

In this section we consider the phase where a stable protostellar core and a rotationally supported, flattened disk have formed in the center of an infalling molecular cloud. The protostellar object and the disk are still gaining mass through ongoing accretion of gas from the cloud envelope, and the protostellar object also may be accreting substantial matter from the adjoining disk. The supersonic collapse of the cloud envelope is halted at an accretion shock enveloping both the protostar and the disk, and the subsequent heating due to the dissipation of kinetic energy produces high effective temperatures and luminosities, on the order of 3000 K and 10–30 L⊙ for a solar-mass star (e.g., Larson 1969). However, this accretion shock is buried deep within the optically thick, dusty cloud envelope and, as a consequence, the protostar appears as a strong source of IR emission with an effective temperature closer to ~200 K (Stahler et al. 1980). The spectrum resembles that of a blackbody during this phase (Bertout 1976). These models are spherically symmetric and do not consider the formation of a rotating protostar and disk; this complication will be discussed shortly.

The models of protostellar collapse and fragmentation discussed in the previous section imply that rotationally flattened disks are expected to be associated both with binary and with single protostar formation. As long as sufficient angular momentum is present in the initial dark cloud core, protostellar formation will involve the formation of a rotationally supported disk, even in the case of a cloud that fragments into a multiple system, though the subsequent evolution of the disk may be much more rapid in this case than in the case of a single protostar.

A heuristic estimate of disk size may be obtained by assuming that the collapsing cloud remains roughly spherical and uniform in density and does not lose mass or angular momentum. Balancing centrifugal and gravitational accelerations then implies a disk radius R_d on the order of 25(J/M)^1/2(4Gρ)^1/2. For a rapidly rotating, 1 M⊙ dark cloud core with J/M = 10^{21} cm^2 s^{-1}, R_d ~ 3000 AU. A more slowly rotating core with J/M = 10^{20} cm^2 s^{-1} could form a disk with R_d ~ 30 AU. Even in the case of a singular initial density profile, rapidly rotating protostellar collapse leads to the formation of a relatively large-scale, quasi-equilibrium disk in the protostellar envelope (Terebey et al. 1984; Boss 1987b).

The spectral appearance of a viscous accretion disk in Keplerian rotation was first calculated by Lynden-Bell and Pringle (1974), who found that the emission from the disk alone would have a spectrum of the form vF_v ~ v^{43} for small v. Beall (1987) has noted that at least some T Tauri
stars in the ρ Ophiuchus dark cloud complex have a spectrum consistent with this viscous accretion-disk prediction. In addition, there would be contributions from the protostar and the boundary layer between disk and star; Lynden-Bell and Pringle (1974) suggested that the disk contribution could account for IR excesses (Mendoza 1966), and the boundary layer for UV excesses (Walker 1972), in T Tauri stars.

Adams and Shu (1986) have developed a fast, approximate method for calculating the IR spectra of models where a spherical protostar is surrounded by an accretion disk; their procedure involves averaging in latitude over the flattened density structure produced in the Terebey et al. (1984) rotating isothermal collapse model, in order to simplify the envelope geometry to an “equivalent spherical envelope”. By varying parameters such as the protostellar mass-accretion rate, the cloud angular velocity and sound speed, and the efficiency of conversion of disk rotational energy into heat, Adams and Shu (1986) were able to reproduce much of the spectra of the IR source associated with Haro 6-10. The parameter values needed to explain typical IR sources tend to be quite plausible given our current understanding of low-mass star formation. However, the “equivalent spherical envelope” approximation has not yet been compared to traditional, one- and two-dimensional, radiative transfer methods (see Bertout and Yorke 1978), so some caution is appropriate.

Adams, Lada, and Shu (1987) used the parameterized models of Adams and Shu (1986) to construct an evolutionary sequence that reproduces many spectral features of young stellar objects, ranging from embedded protostars with negative spectral indices n (n = d log(Fv)/d log v) in the 1 to 10 μm region, to T Tauri stars with IR excesses and positive or nearly zero indices caused by accretion disks, to pre-main-sequence stars with n ~ 3 and without accretion disks. Adams et al. (1987) also divided the T Tauri-disk systems into two classes, those with “passive” disks that have no intrinsic luminosity but do reprocess 25% of the luminosity of the central star, and those with “active” disks where some intrinsic energy source produces emission in addition to that associated with reprocessing. The spectra of passive disk systems (n = 4/3 in the mid- to near-IR) require an accretion disk with a radial temperature profile of the form T \propto r^{-3/4}; the same spectral index is produced in a Keplerian accretion disk, which has very nearly the same temperature profile (Lynden-Bell and Pringle 1974).

Explaining the flattopped IR spectra of some T Tauri stars requires a slightly different physical model from that which explains many of the other young stars. Kenyon and Hartmann (1987) showed that a passive disk could produce a flat spectrum, provided that the disk had a modest flare, i.e., the disk scale height h increases with radius as h \propto r^{0.8}, allowing the disk to intercept more of the stellar photons at greater distances. This disk would also have a temperature distribution of the form T \propto r^{-3/4}. Kenyon and Hartmann (1987) also point out that limits on UV excesses imply that accretion from the disk onto the star is not an important contributor in most young stars. Adams, Lada, and Shu (1988) have proposed an alternative explanation for flattopped IR spectra, a flat disk with a temperature profile T \propto r^{-1/2}, a profile that is significantly flatter than the Keplerian accretion-disk profile (T \propto r^{-3/4}). Temperature profiles flatter than T \propto r^{-3/4} might be produced as a result of the thermostat effect associated with dust-grain evaporation and consequent large opacity decreases around 1500 K (Boss 1989a).

The alternative to these “parameterized” models of protostellar spectra is the “self-consistent” approach (Shu et al. 1987), where the spectra are based on the results of a hydrodynamical calculation of protostellar collapse and formation. This approach has the advantage of less free parameters, at least in principle; the only free parameters are the initial conditions for the collapsing cloud, and these are already constrained by dark cloud core properties and by the dynamical outcome of the collapse (e.g., single vs. binary star formation). In practice this approach has not been as successful as the parameterized method because of the great difficulties involved in a multidimensional hydrodynamical calculation; Tschammler (1987b) has been able to push a rotating cloud the farthest toward pre-main-sequence star formation, but even in this model the protostellar object is still deeply embedded in an opaque dust cloud, and so its spectral appearance is likely to be similar to that of a warm (~10 K–100 K) blackbody. The spectral and isophotal appearances of low-mass rotating protostellar disks have also been studied by Bodenheimer et al. (1988), who found that when viewed edge-on at 40 μm, emission from the disk is concentrated in two emission maxima above and below the disk midplane, because of the large optical depth in the midplane. The spectral indices of these disks in the 1 to 10 μm region are negative, as expected for protostellar objects (Bodenheimer et al. 1988; also see Bertout and Yorke 1978). Dent (1988) modeled the isophotal appearance of rotating protostellar disks (similar to those in Boss 1987b) at later times when a bipolar hole has opened through collapse; the emission at 1 μm forms a bipolar conical nebula above and below the disk midplane, similar in shape to the optical cometary nebulae.

We must note that the final stages of the classical model of protostellar core evolution may be in need of substantial revision, revisions that could greatly alter the observational appearance of protostellar objects. Recently, Tschammler (1987a) has studied the formation of the inner protostellar core in greater detail than previous workers. Rather than assuming that this core should be a...
hydrostatic object, Tscharnuter (1987a) allowed the core to rebound following its formation. Surprisingly, Tscharnuter found that the inner core undergoes a catastrophic expansion, driven primarily by the energy liberated by reassociation of atomic hydrogen and recombinations of molecular hydrogen (leading to an effective adiabatic exponent of the core, $\Gamma_1 < 4/3$, implying the possibility of dynamical instability). The inner-core accretion shock disappears in the explosion, and even the outer-core accretion shock can be swept up. The cloud is still doomed to collapse, of course, and eventually the envelope falls back in and reforms the inner core, but this second inner core can also catastrophically disrupt, producing a limited number of cycles (“hiccups”) of inner core formation and destruction. There are indications that in a fully three-dimensional protostar, the hiccups may be strongly asymmetric, and possibly jetlike (Boss 1989b), but rapid rotation may stabilize the hiccup instability. The full implications of protostellar hiccups are unknown as yet, but the absence of a stable protostellar accretion shock during the hiccup phase could account for the absence of an observational detection of an accreting low-mass protostar. However, whether or not hiccups can destroy the inner core is still questionable.

4.2 Observation

In spite of substantial efforts by observers to find evidence of collapse onto a low-mass protostellar core (e.g., Gee et al. 1985; Walker et al. 1986; Menten et al. 1987), there is still no generally accepted example of this key phase of protostellar evolution. Indeed, the search for an accreting protostar has been termed the Holy Grail of infrared astronomy (Wynn-Williams 1982). The detection of gravitational collapse is greatly complicated by the prevalence of strong outflows from many of the same young stellar objects where one hopes to find evidence for accretion, and linewidths broadened by infall generally cannot be distinguished from those broadened by outflow.

There is, however, evidence for embedded IR sources that could be identified as low-mass protostellar objects. Spherically symmetric models of protostellar evolution (Stahler et al. 1980) appear to be consistent with the luminosities and effective temperatures of embedded infrared sources in nearby molecular cloud cores (Beichman et al. 1986). The spectral appearance of these objects differs somewhat from that of spherically symmetric protostars (the objects have excess 12–25 μm emission compared to the models), and this difference appears to require the presence of a warm (> 200 K) disk (e.g., Adams and Shu 1996). Alternatively, the objects could be young T Tauri stars still embedded in a dense dust cloud (Beichman et al. 1986).

Several different lines of evidence indicate the prevalence of disks around young stellar objects. First, the large IR and ultraviolet (UV) excesses common in young stellar objects are difficult to explain without attributing the excesses to an active accretion disk and a boundary layer between disk and star, respectively (Strom et al. 1988; Basri 1988); the circumstellar dust responsible for IR excesses appears to reside in a flattened disk (Cohen and Witteborn 1985). Second, observations of mass loss from T Tauri stars imply the existence of opaque disks with masses of 0.01 to 0.1 M⊙, and sizes of order 100 AU, in order to explain the infrequent observation of red-shifted forbidden-line emission; usually only blueshifted lines are seen (Edwards et al. 1987). Third, the bipolar molecular gas flows often found to be emanating from young stellar objects (Snell, Loren, and Plancke 1980; reviewed by Lada 1985) are also often associated with large-scale (0.1–1 pc) rotating, flattened disks (Cantó et al. 1981; see references in Boss 1987b). The bipolar flows generally are aligned perpendicular to the major axis of the disk, suggesting that the disk has helped collimate the flow. While a variety of mechanisms have been advanced for creating bipolar flows (see Pudritz 1986), all of the mechanisms involve a flattened disk on either a very small (≈ R⊙ to AU) or large scale (≈ pc). Fourth, Hartmann and Kenyon (1987) have found evidence for accretion disks roughly in Keplerian rotation about FU Orionis stars; the rotation inferred from infrared CO lines is less than that inferred for optical, consistent with formation of the CO lines in the cooler, slower-rotating, outer regions of the disk. Finally, aperture synthesis maps (Fig. 3) of intermediate-scale (≈ 1000 AU) molecular gas disks have been produced with mm-wave interferometers (e.g., Sargent et al. 1988). The presence of a disk around HL Tauri (a 1-M⊙ young star) had been inferred from IR measurements (Cohen 1983; Grasdalen et al. 1984). The increased spatial resolution afforded by mm-wave arrays has allowed the 0.1 μm 2000 AU radius HL Tau disk not only to be mapped convincingly, but for Keplerian rotation of the disk to be demonstrated (Sargent and Beckwith 1987). A disk has even been found around a binary T Tauri star, T Tauri itself (Weintraub, Masson, and Zuckerman 1987).

5. Planetary Formation

5.1 Theory

In this section the terminology undergoes a subtle change: The protostellar disks we have been discussing will be referred to as the “solar nebula”, i.e., the gas and dust disk that produced our solar system some $4.5 \times 10^9$ yr ago. This reduction in scope is forced on us by our ignorance of planetary systems other than our own. Consequently, most of the work discussed in this section is specialized to our solar system, though we believe that the physical processes involved should be generally applicable to planet formation around other low-mass stars. In this sense the term “solar nebula” and the generic terms
"protostellar disk" and "protoplanetary disk" can be used interchangeably.

Because of the calculational difficulties of predicting the physical structure of a solar nebula formed through the collapse of an interstellar cloud, difficulties that are still not fully overcome, much of the progress in our understanding of planet formation has come from making plausible assumptions about the initial structure of the solar nebula (see Table 1 in Boss et al. 1989), and usually assuming as well the prior existence of the Sun (see the monumental treatise by Safronov (1969) and review by Wetherill (1980)). This approach assumes that whatever happened in earlier phases (e.g., prior to solar formation) had little effect on planet formation, thereby decoupling the two problems. The extent to which this decoupling is valid can only be ascertained by detailed study of the intermediate phases, when matter from the solar nebula was still accreting onto the protosun and the early phases of planet formation may have already started.

Numerical models of the collapse of a dense cloud to form the protosun and the solar nebula have been produced with both axisymmetry (Tscharnuter 1978, 1987) and nonaxisymmetry (Boss 1989a). Even with slow rotation initially (in order to avoid binary fragmentation), a large fraction of the cloud mass typically ends up in the solar nebula, prevented by its angular momentum from accreting onto the protosun. Hence, physical processes have been sought to transport mass inward, which effectively requires transporting angular momentum outward. The four most-likely physical processes are viscous stresses, gravitational torques, magnetic fields, and acoustic waves; comparatively little attention has been paid to the latter two processes (see, however, Stepinski and Levy (1988) and Larson (1989), respectively).

Viscous accretion-disk models were pioneered by Lynden-Bell and Pringle (1974) and have become the dominant type of solar nebula model (e.g., Cameron 1978; Lin and Papaloizou 1980; Cassen and Moosman 1981; Morfill and Völk 1984). Viscous accretion-disk models depend on the existence of turbulence to generate an effective vis-
Convective instability in the direction perpendicular to the nebula midplane is the best-understood mechanism for generating turbulence (Lin and Papaloizou 1980). Viscous stresses result in the desired outward transport of angular momentum through the frictional shear between adjacent fluid parcels on Keplerian orbits. Vertically driven convection may not be very efficient at producing radial mass transport, however (Cabot et al. 1987; see the discussion in Boss et al. 1989); time scales for angular-momentum transport may be longer than $2 \times 10^6$ yr. Also, convective turbulence may be terminated because of decreased opacity and enhanced radiative cooling caused by dust-grain coagulation, or simply because of insufficient heating at the nebula midplane.

Gravitational torques are caused by gravitational forces between elements of nonaxisymmetric solar nebulae (Larson 1984), for example, between the inner and outer regions of trailing spiral arms. Gravitational torques result in outward transport of angular momentum in this case. Time scales for angular-momentum transport by gravitational torques can be as short as $10^4$ yr or less even in low-mass nebulae (Boss 1989a); the required nonaxisymmetry can be generated by coupling to the collapse motions and need not require the presence of a gravitationally unstable, massive nebula (Cassen et al. 1981).

As far as we know, there are just two means for forming planets in a solar nebula, planetesimal accumulation and gravitational instability of the nebular gas disk. By far the most likely means for forming the terrestrial planets and the rock and ice cores of the giant planets is through planetesimal accumulation, which requires first dust-grain coagulation, then formation of intermediate-sized bodies termed "planetesimals", and, finally, collisional accumulation of planetesimals into planets.

Little dust-grain coagulation can occur while the nebula is turbulent (Weidenschilling 1988) because vigorous turbulence keeps the dust grains well mixed with the gas. Once turbulence ceases, however, dust grains sediment to the midplane of the nebula on quite short time scales ($\sim 10^5$ yr). Dust grains sediment down to form a thin subdisk in the nebula midplane because the gas, unlike the dust grains, is supported by thermal pressure and forms a much thicker disk. The enhanced spatial density of dust grains during this sedimentation phase also leads to rapid growth by coagulation to sizes on the order of 1 cm.

Planetesimal accumulation is traditionally thought to involve gravitational instability of the dust disk (Safronov 1969; Goldreich and Ward 1973). When the surface density of the dust disk becomes large enough through ongoing sedimentation, self-gravity of the dust disk breaks up the disk and leads to the formation of a large number of planetesimals roughly 1 km in size. The occurrence of this gravitational instability has been strongly questioned by Weidenschilling (1988), on the basis that even in the absence of convectively driven turbulence, turbulence caused by the shear between a Keplerian dust disk and a slightly slower rotating gas disk will prevent further sedimentation and forestall the process short of the critical density. Growth in this size range might then require continued collisional coagulation, with differential motion caused by gas drag, producing 1-km objects in perhaps $10^4$ yr (Weidenschilling 1988). Calculations of the subsequent phases of planet formation typically assume the existence of a swarm of about $10^{12}$ planetesimals in the terrestrial planet region alone.

Subsequent planetesimal growth required collisions between these appreciably self-gravitating, 1-km bodies. Two distinct phases arise (e.g., Wetherill 1980), the first involving accumulation of planetesimals on nearby orbits and ending in about $10^7$ yr at 1 AU with the depletion of all nearby planetesimals. This phase probably involved the runaway growth of a number of planetesimals (e.g., Wetherill and Stewart 1989); runaway growth occurs in part because gravitational attraction produces an effective increase in the collisional cross section, so that the largest body grows the fastest. The first phase produced planetesimals about 500 km in size on nearly circular orbits. The second phase involves collisions between these more widely separated planetesimals. The requisite large increase in orbital eccentricity is caused by the stochastic effects of gravitational forces between the planetesimals, especially during close encounters. In contrast to the first phase, the second phase probably did not involve runaway accretion, at least not in the terrestrial planet region; rather, accumulation proceeded through the mutual collision of more nearly equal mass bodies. As a consequence, spectacular impacts between planetary-sized bodies must have been commonplace (Fig. 4). A glancing impact between the Earth and a Mars-sized body is the favored means for explaining the formation of the Earth-Moon system (Benz, Slattery, and Cameron 1987). The second phase produces the final planetary system, with the terrestrial planets forming in about $10^7$ to $10^8$ yr (Wetherill 1986).

The key problem with the planetesimal accumulation theory is explaining the rapid formation of the giant planets. In order to accrete H,He-rich envelopes, the cores of Jupiter and Saturn must have formed prior to solar nebula removal, that is, within $10^6$–$10^7$ yr after solar formation (Walter et al. 1988; Strom et al. 1989). However, planetesimal accumulation times for the $\sim 10$ Earth mass cores of the giant planets can be $10^8$ years or longer. Rapid formation of giant planet cores only appears to be possible if the solar nebula was substantially more massive in the giant planet region than is usually assumed (which is a problem; see Boss 1989a) and if runaway accretion produced the cores within about $10^6$ years (e.g., Lissauer 1987).
The only serious alternative to forming planets by planetesimal accumulation is a gravitational instability of the gaseous portion of the solar nebula (Cameron 1978). A cold, massive solar nebula can be gravitationally unstable, and break up into a number of "giant gaseous protoplanets", on a time scale of about ten years (Cassen et al. 1981). Evidently this process could produce giant planets well before solar nebula removal.

However, the existence of a number of severe problems has greatly lowered the likelihood that any of our planets formed by gaseous gravitational instability (other planetary systems might well be different, of course). First, the instability requires a cold nebula (cf. Boss 1989a) much more massive (∼1 \(M_\odot\)) than obtained (0.01–0.1 \(M_\odot\)) by reconstituting the planets to solar composition (Weidenschilling 1977). Even the strongest stellar wind from a 1 \(M_\odot\) star could not remove 1 \(M_\odot\) of nebula to infinity. Second, there is a problem associated with the miscibility inferred for silicates in the high-pressure H,He envelopes of the giant planets; sedimentation of silicates to form the rocky cores could not have taken place (Stevenson 1982). Finally, even if rocky cores formed by some means, the gaseous envelope must be stripped off to form terrestrial planets from giant gaseous protoplanets. However, this stripping must have occurred only in the inner and outer solar nebula, not at Jupiter and Saturn, in order to explain the rock and ice compositions of the terrestrial planets and Uranus and Neptune.
Neptune, respectively. The two possible stripping processes, solar tidal forces and thermal evaporation by the Sun, would not produce this peculiar effect.

5.2 Observation

Theoretical work on planet formation has not uncovered any reasons to believe that our solar system is a unique object, and hence we must assume that planetary systems are the rule rather than the exception. Searches for other planetary systems are extremely difficult, however, and it must be admitted that undisputed observational evidence for other planetary systems is minimal at present.

The circumstellar dust detected around the "2 M_☉", main-sequence star β Pictoris (Smith and Terrile 1984) provided the first direct evidence of a strongly flattened dust disk around another star. The small size inferred for the dust grains implies that the dust is subject to orbital decay by the Poynting-Robertson effect and hence must be continually refreshed, and, coupled with the spectral reflectivity of the dust, the observations point to the need for a large reservoir of asteroids and/or comets around β Pic to provide a source of dust (Gradie and Hayashi 1987). Larger planets could also be present but remain unseen.

Campbell, Walker, and Yang (1988) have searched for substellar companions to 16 solar-type stars, using an extremely precise means of measuring radial-velocity changes of absorption lines in the primary star spectrum caused by the orbit of the primary around the common center of mass. Seven of the 16 stars studied showed statistically significant variations consistent with the presence of a "1 to 9 Jupiter-mass companion; this study appears to be our best evidence for extrasolar planetary systems (the well-known claim for a companion to Barnard’s star has not been confirmed).

Substellar companions with masses less than about 10 Jupiter masses are properly called “planets”, but objects in the range between "0.01 M_☉" ≈ 10 Jupiter masses and "0.08 M_☉" are called “brown dwarfs”, reflecting a theoretical difference in the origin ascribed to these objects. Planets form by secondary processes in disks around stars; considering that such disks should be smaller in mass than the central stars, and that the entire disk is unlikely to be used to form a single planet, planetary masses must be considerably smaller than stellar masses. Brown dwarfs are stars that are too low in mass to burn hydrogen (reviewed by Liebert and Frobst 1987) but are produced by the same process that forms main-sequence stars: collapse and fragmentation of an interstellar cloud, which is thought to lead to a minimum mass of about "0.01 M_☉" (Boss 1986 and references therein). Evidence for the existence of specific brown dwarfs has been extremely controversial, though the claims of brown dwarfs around white dwarf stars by Zuckerman and Becklin (1987) and Becklin and Zuckerman (1988) and evidence for a rising initial mass function toward very low masses (Hawkins and Bessell 1988) both suggest that the existence of brown dwarfs might yet become generally accepted. Recently, a new IR camera has been used to detect nine brown-dwarf candidates in Taurus (Forrest et al. 1989); if the objects are brown dwarfs, their masses are inferred to be in the range 0.005 to 0.02 M_☉.

6. Termination of Formation

6.1 Observation

In this section we discuss the phase when the remaining gas and dust are removed from the vicinity of the pre-main-sequence star and planetary system—removal of the raw material for star and planet formation clearly defines the end of the formation process. During this phase the pre-main-sequence star makes its first optical appearance as a young star.

The omnipresent evidence for strong stellar winds associated with young stellar objects suggests that winds may play an important role in removing nebula gas and thus terminating formation. Much of our information about early stellar winds comes from observations of young T Tauri stars, which are low-mass, pre-main-sequence, optically visible, variable stars with model ages around 10^6 yr (e.g., Cohen and Kuhi 1979). T Tauri stars have long been known to possess winds with mass-loss rates on the order of 10^{-5} M_☉ yr^{-1} (Herbig 1958; Kuhi 1964). Strings of Herbig-Haro (HH) objects (consisting of knots of shocked gas) have been found aligned on both sides of T Tauri stars, moving with proper motions directed away from the central star (Herbig and Jones 1981; reviewed by Schwartz 1983). A steady, spherically symmetric wind from a T Tauri star is unable to explain the acceleration of these objects (e.g., Lada 1985), implying that young T Tauri stars are able to produce even more energetic (though possibly transient) winds than those normally associated with pre-main-sequence low-mass stars. The infrequent (though possibly periodic) occurrence of an FU Orionis outburst, where a T Tauri star suddenly increases in luminosity by a factor of 10 to 100 (Herbig 1977), can also lead to a mass-loss rate a factor of 100 to 1000 times higher than that of a normal T Tauri (Crosswell, Hartmann, and Avrett 1987). An FU Orionis star has been found associated with a HH object (Graham and Frogel 1985), demonstrating the occurrence of both of these energetic events in a single young star. Edwards et al. (1987) showed that stellar winds around a number of T Tauri stars are best modeled by a wind that is latitude-dependent, being higher in velocity and lower in density at the pole than toward the equator, where the presence of an optically thick disk is inferred.

There is strong evidence for energetic mass loss during even earlier phases, when low-mass stars are still embed-
ded in molecular cloud cores. The frequent occurrence of large-scale, bipolar flows of molecular gas around embedded young stellar objects (Snell et al. 1980) implies that perhaps all low-mass stars experience a similar phase during their first $10^5$ yr (Lada 1985). Bipolar flows involve large amounts of gas ($\sim 0.1$ to $100 \, M_\odot$) moving at high speed ($\sim 10$ to $100 \, \text{km s}^{-1}$); explaining the delivery of the momentum needed to drive such a wind has been difficult (e.g., Lada 1985), though there are indications that a neutral (H i) component may be sufficient to account for the momentum (Rodríguez and Cantó 1983; Lizano et al. 1988). Collimation factors (ratios of length/diameter) for bipolar flows usually range between 1 and 3, with some factors being as high as 6 (Lada 1985), and in many cases collimation factors may be even higher in reality because of beam smearing. Smaller-scale, optically visible jets are even more highly collimated (Mundt 1988), implying that the degree of collimation decreases with increasing radius from the star. Perhaps of most importance for our purposes in this section, Myers et al. (1988) have shown that molecular outflows have sufficiently high energy, momentum, frequency, longevity, and hydrodynamic coupling to dark cloud cores in order “to be the main agent of core dispersal”.

As previously mentioned, classical T Tauri stars have IR and UV excesses that are attributed to the presence of circumstellar matter in the form of a disk. Classical T Tauri stars are thus thought to harbor significant gas and dust and, hence, could still be accreting mass from the disk or forming planets in the disk. Another class of T Tauri stars has been identified on the basis of their X-ray emission, the “naked” T Tauri stars (Walter 1986). These stars show greatly diminished IR and UV excesses, yet have ages ranging from $10^5$ to $4 \times 10^7$ yr (Walter et al. 1988), implying that the formation process must effectively terminate around some low-mass stars within $\sim 10^5$ yr. The statistics for the gradual disappearance with time of IR excesses in both classical and naked T Tauri stars implies that circumstellar disks are generally removed within $3 \times 10^6$ to $10^7$ yr (Strom et al. 1989).

6.2 DeCampli (1981) examined several different possibilities for generating mass-loss rates of $\sim 10^{-8} \, M_\odot \, \text{yr}^{-1}$ in T Tauri stars. After ruling out a stellar wind driven simply by a very hot, solar-type corona (which would also produce an unrealistically high optical luminosity), DeCampli suggested that Alfvén waves propagating through the outer layers of the convection zone could dump sufficient energy in the stellar atmosphere to explain the mass loss. This hydromagnetic mechanism also implies surface magnetic fields on the order of $\sim 300 \, \text{G}$; attempts at measuring the Zeeman effect in T Tauri stars so far have produced only upper limits of $\sim 500 \, \text{G}$ on the possible surface magnetic-field strength (Johnstone and Penston 1987). In spite of the difficulty of observational confirmation, Alfvén waves remain the favored means of explaining T Tauri mass-loss rates. Shu and Terebey (1984) have hypothesized that the convection caused by the initiation of deuterium burning leads to solid-body rotation in T Tauri stars, and the rotational energy dissipated in turning a differentially rotating star into a rigidly rotating star could be used to power the wind; how the energy is transmitted to the wind is not specified. In order to explain the episodic, enhanced energy losses associated with FU Ori outbursts, Larson (1980) proposed a rotational instability similar to the fission instability (e.g., Durisen et al. 1986), while Hartmann and Kenyon (1985) advocated the episodic accretion of abnormally large amounts of matter from an accretion disk.

Strong stellar winds are also thought to be energetically capable of driving the bipolar molecular gas flows, and if neutral wind components are common, strong stellar winds (with mass-loss rates of $\sim 3 \times 10^{-6} \, M_\odot \, \text{yr}^{-1}$) may also be able to supply the required momentum (Lizano et al. 1988). The primary alternative to stellar winds for driving bipolar flows is the hydromagnetic disk wind, variants of which have been proposed by Pudritz and Norman (1983; see also Pudritz 1986) and Uchida and Shibata (1984). In the former model, a large-scale ($> 10^{15} \, \text{cm}$) disk is heated at its base (by either friction from ambipolar diffusion or by an embedded protostar), producing a wind that is then centrifugally accelerated by magnetic fields to the velocities observed in bipolar flows, assuming the field can maintain constant angular velocity in the wind over large distances. Hydromagnetic disk models have been criticized by Shu et al. (1987), partially on the grounds that powering a strong wind by a disk is energetically inferior to powering by a protostar, because the gravitational energy available in the protostar is several orders of magnitude larger than in the disk.

One advantage of the hydromagnetic disk model for bipolar flows is that it produces a collimated flow intrinsically, because the wind starts from the surface of the disk, in an effectively plane parallel atmosphere. In comparison, an initially isotropic wind from a protostar must be collimated extrinsically (vs. Torbett (1984), who proposed the generation of an intrinsically collimated wind at the boundary layer between the disk and protostar). The most likely means for collimating an isotropic wind is through the effects of the surrounding molecular gas disk, because of the preference for the wind to escape in the direction parallel to the maximum gradient in the disk density (i.e., perpendicular to the disk midplane; Barrai and Cantó 1981). Such a wind might be further accelerated by the formation of a structure similar to a de Laval nozzle (Königl 1982). Preferential flow in the direction perpendicular to the disk midplane might also be aided by the fact that during the collapse leading to formation of the
protostar and disk, the matter along the cloud rotation axis is strongly depleted by accretion onto the protostar, whereas the matter away from the rotation axis is partially supported by rotation and, hence, is more likely to remain suspended in a quasi-equilibrium, large-scale disk (Boss 1987b). Protostellar collapse should thus result in evacuation of a bipolar-like cavity awaiting the arrival of the protostellar wind; this result could have been anticipated ten years ago, and surely the theoreticians missed a rare opportunity by failing to suggest the possible existence of bipolar winds prior to their discovery by Snell et al. (1980).

Shu et al. (1987) pointed out that bipolar flows appear to characterize a phase of evolution intermediate between a protostar undergoing only accretion and an optically visible, pre-main-sequence star undergoing only outflow. That is, a protostar may begin to drive a strong wind even while accretion onto the protostar from the disk and onto the disk from the cloud core proceeds. This wind will preferentially escape along the rotation axis, forming a collimated bipolar flow. As time proceeds, the outflow will become less collimated, either through accretion of higher \( J/M \) gas leading to a wider conical opening in the constraining disk or through erosion of the opening by the wind (Boss 1987b).

The widening of the bipolar flow into a more nearly isotropic wind leads to a phase where the wind will begin to remove the last vestiges of the protostellar disk; that is, the wind will remove the solar nebula within which planetary materials have presumably already been accumulated into large-enough bodies to withstand subsequent wind erosion. The hydrodynamical efficiency of solar nebula removal by a strong wind has been estimated by Horedt (1978), who found that a \( 10^{-8} \, \text{M}_\odot \, \text{yr}^{-1} \) wind blowing over a concave nebula would entrain sufficient gas and dust to remove 0.1 \( \text{M}_\odot \) of nebula matter to infinity in \( 10^7 \) yr. Elmegreen (1978) calculated that a stellar wind blowing over a nebula would raise waves on the nebula surface, creating turbulence that could act as a source of effective viscosity in the nebula. The nebula would then evolve like a viscous accretion disk, with the matter being transported onto the protostar. These two calculations evidently represent the extremes for nebula removal—nebula transport either to \( r \sim 0 \) or to \( r \sim \infty \). UV radiation has also been suggested as an aid to nebula removal (Horedt 1982), though it appears to be considerably less efficient than a gaseous wind. Considering the evidence from naked T Tauri stars for nebula removal on time scales possibly as short as \( 10^5 \) yr (Walter et al. 1988), the hydrodynamics of nebula removal deserves further research. It is worth noting that because wind erosion is likely to be inefficient (the wind can always follow the path of least resistance and avoid the nebula), low-mass stellar winds are unlikely to be able to remove more than \( \sim 0.01 \, \text{M}_\odot \) from a solar nebula (Boss et al. 1989); the bulk of a more massive nebula presumably must be dumped onto the protostar by processes intrinsic to the nebula (i.e., evaporation through viscous, gravitational, or magnetic torques).

Once the wind has completed nebula removal, no further matter can accrete onto the protostar, and the formation process is terminated. Shu and Terebey (1984) have suggested that the cutoff of accretion by wind initiation is what determines stellar masses. While stellar winds are likely to be an important part of the termination process, as outlined above, other factors must come into play as well, such as the angular-momentum distribution of the infalling matter, which will be critical in determining which matter reaches the protostar almost directly, which is processed through the disk, and which is returned to the interstellar medium through the actions of energetic stellar winds.

At this point the young star is an optically visible, pre-main-sequence star, quasi-statically contracting toward the main-sequence. Stahler (1983, 1988b) has determined the appearance of protostars at the moment when accretion is terminated and they first become optically visible (Fig. 5), and has defined the location in an H-R diagram of the stellar “birthline”, above which protostars are still accreting, below which they are visible. As seen in Figure 5 this birthline is in strikingly good agreement with observations of both naked (Walter et al. 1988) and classical T Tauri stars (Cohen and Kuhi 1979). Considering that Stahler’s models do not include the effects of rotation, magnetic fields, or Tschernuthe’s (1987a) protostellar core hiccups, the implication is that all three of these severe complications are of little importance for determining the final protostellar (or initial stellar) appearance of low-mass stars. While this is reassuring, the success of the birthline also means that deciphering what happened during earlier phases when these complications presumably were of greater importance will be a much harder task than would have been the case if T Tauri stars provided direct evidence for part of the solution of, e.g., the angular momentum and magnetic flux problems of star formation.

7. Challenges

Our understanding of star and planet formation has become considerably more robust in the last two decades, and the framework of a reasonably self-consistent body of theoretical and observational work is now becoming evident. However, many critical challenges remain, a number of which are listed in this concluding section.

7.1 Theoretical Challenges

1) More models of magnetic dark cloud collapse need to be calculated, including the effects of detailed ionization calculations and ambipolar diffusion of the field lines.
Extending this type of calculation to three dimensions may be especially important for understanding dark cloud core formation.

(2) In lieu of or in addition to (1) above, calculations of nonmagnetic collapse should use initial conditions based on models of isothermal, magnetic equilibria (rotating: Tomisaka, Ikeuchi, and Nakamura 1988), nonrotating: Lizano and Shu 1989) in order to best simulate the likely conditions for the initiation of the collapse phase.

(3) The physical reality of and observational implications of Tscharnuter's (1987a) protostellar core instability need to be more generally tested and explored, respectively. Do protostellar hiccups really mean that accreting protostars are unlikely to be observed?

(4) Multidimensional collapse calculations still have not been able to rigorously evolve a protostar all the way from the dark cloud phase to the formation of a pre-main-sequence star. Fulfilling this challenge will require a
resolution of the question of angular-momentum transport in protostellar disks.

(5) A direct demonstration of hierarchical fragmentation during the isothermal phase of protostellar collapse would improve our understanding of the formation of hierarchical multiple stellar systems and small clusters of stars.

(6) At present the most uncertain phase of planetary evolution appears to be the phase where planetesimals grow in size from perhaps cm-sized objects to ~ 1-km objects—did a global gravitational instability of the dust subdisk really occur, or did growth through this phase follow some other path?

(7) The outstanding problem in planetary cosmogony is forming Jupiter and the other giant planets prior to nebula removal (~ 10^6–10^7 yr).

(8) The mechanism that generates winds in young stellar objects powerful enough to drive bipolar flows is quite uncertain yet of obvious importance for understanding both pre-main-sequence evolution and clearing of the residual gas and dust from protoplanetary disks. How and when were the last vestiges of the solar nebula removed?

7.2 Observational Challenges

(1) There is a continual need for higher spatial-resolution observations in order to better resolve the density and velocity structure of dark cloud cores; clearly, mm-wave interferometers will be increasingly useful here. Evidence for or against fragmentation during the collapse of dark cloud cores would be particularly useful.

(2) Improved observations will require the presence of optically thin amounts of molecules that can be excited into emission at very high densities in order to probe the innermost regions of dense (~ 10^10–10^12 g cm^-3) protostellar disks.

(3) Considering the abundance of young stellar objects found in dark cloud cores, the discovery and characterization of the predecessors (precollapse) to dark cloud cores should be undertaken.

(4) Unambiguous evidence for gravitational collapse of dark cloud cores and for an accreting low-mass protostar will remain as major observational goals.

(5) The dynamical properties of binary pre-main-sequence stars should be determined and compared to those of main-sequence stars—interesting evolutionary effects (e.g., eccentricity changes) may be found.

(6) Extrasolar planetary-system detection and the coupled problem of brown-dwarf detection will continue to challenge observers; if a few examples can be agreed on, the next challenge will be to enlarge the sample of each type of object and thus provide a better understanding of the range of their characteristics.

(7) Observations of the effects of stellar winds (possibly highly variable) on their immediate surroundings (i.e., within 10 AU) will be important for timing the removal of nebula gas and dust relative to planet formation.

I thank my cosmogonical colleagues at DTM, John A. Graham, Mark H. Heyer, George W. Wetherill, and Harold A. Williams, for many lively discussions about star and planet formation and John A. Graham for his improvements to the manuscript. Writing of this review was partially supported by National Science Foundation grant AST 88-17334 and by National Aeronautics and Space Administration grant NAGW-1410.

REFERENCES

Cameron, A. G. W. 1978, Moon Planets, 18, 5.