Theory of Star Formation

Christopher F. McKee
Departments of Physics and Astronomy, University of California, Berkeley, CA 94720; cmckee@astro.berkeley.edu

Eve C. Ostriker
Department of Astronomy, University of Maryland, College Park, MD 20742; ostriker@astro.umd.edu

Abstract
We review current understanding of star formation, outlining an overall theoretical framework and the observations that motivate it. A conception of star formation has emerged in which turbulence plays a dual role, both creating overdensities to initiate gravitational contraction or collapse, and countering the effects of gravity in these overdense regions. The key dynamical processes involved in star formation—turbulence, magnetic fields, and self-gravity—are highly nonlinear and multidimensional. Physical arguments are used to identify and explain the features and scalings involved in star formation, and results from numerical simulations are used to quantify these effects. We divide star formation into large-scale and small-scale regimes and review each in turn. Large scales range from galaxies to giant molecular clouds (GMCs) and their substructures. Important problems include how GMCs form and evolve, what determines the star formation rate (SFR), and what determines the initial mass function (IMF). Small scales range from dense cores to the protostellar systems they beget. We discuss formation of both low- and high-mass stars, including ongoing accretion. The development of winds and outflows is increasingly well understood, as are the mechanisms governing angular momentum transport in disks. Although outstanding questions remain, the framework is now in place to build a comprehensive theory of star formation that will be tested by the next generation of telescopes.

CONTENTS
INTRODUCTION ......................................................... 2
BASIC PHYSICAL PROCESSES ................................... 3
  Turbulence ......................................................... 3
  Self-Gravity ....................................................... 13
  Magnetic Fields ................................................ 19
MACROPHYSICS OF STAR FORMATION ......................... 24
  Physical State of GMCs, Clumps, and Cores .................. 24
  Formation, Evolution, and Destruction of GMCs .......... 34
  Core Mass Functions and the IMF ............................ 44
  The Large-Scale Rate of Star Formation ................. 50
1 INTRODUCTION

Stars are the “atoms” of the universe, and the problem of how stars form is at the nexus of much of contemporary astrophysics. By transforming gas into stars, star formation determines the structure and evolution of galaxies. By tapping the nuclear energy in the gas left over from the Big Bang, it determines the luminosity of galaxies and, quite possibly, leads to the reionization of the universe. Most of the elements—including those that make up the world around us—are formed in stars. Finally, the process of star formation is inextricably tied up with the formation and early evolution of planetary systems.

The problem of star formation can be divided into two broad categories: “microphysics” and “macrophysics”. The microphysics of star formation deals with how individual stars (or binaries) form. Do stars of all masses acquire most of their mass via gravitational collapse of a single dense “core”? How are the properties of a star or binary determined by the properties of the medium out of which it forms? How does the gas that goes into a protostar lose its magnetic flux and angular momentum? How do massive stars form in the face of intense radiation pressure? What are the properties of the protostellar disks, jets, and outflows associated with Young Stellar Objects (YSOs), and what governs their dynamical evolution?

The macrophysics of star formation deals with the formation of systems of stars, ranging from clusters to galaxies. How are giant molecular clouds (GMCs), the loci of most star formation, themselves formed out of diffuse interstellar gas? What processes determine the distribution of physical conditions within star-forming regions, and why does star formation occur in only a small fraction of the available gas? How is the rate at which stars form determined by the properties of the natal GMC or, on a larger scale, of the interstellar medium in a galaxy? What determines the mass distribution of forming stars, the Initial Mass Function (IMF)? Most stars form in clusters (Lada & Lada, 2003); how do stars form in such a dense environment and in the presence of enormous radiative and mechanical feedback from other YSOs?

Many of these questions, particularly those related to the microphysics of star formation, were discussed in the classic review of Shu, Adams, & Lizano (1987). Much has changed since then. Observers have made enormous strides in characterizing star formation on all scales and in determining the properties of the medium out of which stars form. Aided by powerful computers, theorists have been able to numerically model the complex physical and chemical processes associated with star formation in three dimensions. Perhaps most important, a new paradigm has emerged, in which large-scale, supersonic turbulence governs the macrophysics of star formation.

This review focuses on the advances made in star formation since 1987, with an emphasis on the role of turbulence. Recent relevant reviews include those on the physics of star formation (Larson, 2003), and on the role of supersonic turbulence
in star formation (Mac Low & Klessen, 2004; Ballesteros-Paredes et al., 2007). The chapter by Zinnecker & Yorke in this volume of ARAA provides a different perspective on high-mass star formation, while that by Bergin & Tafalla gives a more detailed description of dense cores just prior to star formation. Because the topic is vast, we must necessarily exclude a number of relevant topics from this review: primordial star formation (see Bromm & Larson, 2004), planet formation, astrochemistry, the detailed physics of disks and outflows, radiative transfer, and the properties of young stellar objects. We begin with an overview of the basic processes and scales in turbulent, self-gravitating, magnetized media, and we then review both the macrophysics and the microphysics of star formation.

2 BASIC PHYSICAL PROCESSES

2.1 Turbulence

As emphasized in §1, many of the advances in the theory of star formation since the review of Shu, Adams, & Lizano (1987) have been based on realistic evaluation and incorporation of the effects of turbulence. Turbulence is in fact important in essentially all branches of astrophysics that involve gas dynamics\(^1\), and many communities have contributed to the recent progress in understanding and characterizing turbulence in varying regimes. Here, we shall concentrate on the parameter regimes of turbulence applicable within the cold interstellar medium, and the physical properties of these flows that appear particularly influential for controlling star formation.

Our discussion provides an overview only; pointers will be given to excellent recent reviews that summarize the large and growing literature on this subject. General references include Frisch (1995), Biskamp (2003), and Falgarone & Passot (2003). A much more extensive literature survey and discussion of interstellar turbulence, including both diffuse-ISM and dense-ISM regimes, is presented by Elmegreen & Scalo (2004) and Scalo & Elmegreen (2004). A recent review focusing on the detailed physics of turbulent cascades in magnetized plasmas is Schekochihin & Cowley (2005).

2.1.1 SPATIAL CORRELATIONS OF VELOCITY AND MAGNETIC FIELDS

Turbulence is defined by the Oxford English Dictionary as a state of “violent commotion, agitation, or disturbance,” with a turbulent fluid further defined as one “in which the velocity at any point fluctuates irregularly.” Although turbulence is by definition an irregular state of motion, a central concept is that order nevertheless persists as scale-dependent spatial correlations among the flow variables. These correlations can be measured in many ways; common mathematical descriptions include autocorrelation functions, structure functions, and power spectra.

One of the most fundamental quantities, which is also one of the most intuitive to understand, is the RMS velocity difference between two points separated by a distance \(r\). With the velocity structure function of order \(p\) defined as \(S_p(r) \equiv \langle |v(x) - v(x + r)|^p \rangle\), this quantity is given as \(\Delta v(r) \equiv |S_2(r)|^{1/2}\). The autocorrelation function of the velocity is related to the structure function: \(A(r) \equiv \langle v(x) \cdot v(x + r) \rangle = \langle |v|^2 \rangle - S_2(r)/2\); note that the autocorrelation

\(^1\)Chandrasekhar (1949) presaged this development, in choosing the then-new theory of turbulence as the topic of his Henry Norris Russell prize lecture.
with zero lag is \( A(0) = \langle |v|^2 \rangle \) since \( S_2(0) = 0 \). The power spectrum of velocity, \( P(k) \equiv |\mathbf{v}(k)|^2 \), is the Fourier transform of the autocorrelation function. For zero mean velocity, the velocity dispersion averaged over a volume \( \ell^3 \), \( \sigma_v(\ell)^2 \), is equal to the power spectrum integrated with \( k_{\min} = 2\pi/\ell \). If turbulence is isotropic and the system in which it is observed is spatially symmetric with each dimension \( \approx \ell \), then the one-dimensional velocity dispersion along a given line-of-sight (a direct observable) will be related to the three-dimensional velocity dispersion by

\[
\sigma = \sigma_v(\ell)/\sqrt{3}.
\]

Analogous structure functions, correlation functions, and power spectra can also be defined for the magnetic field, as well as other fluid variables including the density (see §2.1.4). Delta-variance techniques provide similar information, and are particularly useful for reducing edge effects when making comparisons with observational data (Bensch, Stutzki, & Ossenskopf, 2001).

For isotropic turbulence, \( S_p \) and \( A \) are functions only of \( r = |\mathbf{r}| \), and \( P \) is a function only of \( k = |\mathbf{k}| \). The Fourier amplitude \( |\mathbf{v}(k)| \) is then (on average) only a function of \( \ell = 2\pi/k \), and can be denoted by \( v(\ell) \); to emphasize that these velocities are perturbations about a background state, the amplitude of a given Fourier component is often written as \( \delta v(k) \) or \( \delta v(\ell) \). When there is a large dynamic range between the scales associated with relevant physical parameters (see §2.1.3), correlations often take on power-law forms. If \( P(k) \propto k^{-n} \) for an isotropic flow, then

\[
v(\ell) \propto \sigma_v(\ell) \propto \Delta v(\ell) \propto \ell^q
\]

with \( q = (n - 3)/2 \). Sometimes indices \( n' \) of one-dimensional (angle-averaged), rather than three-dimensional, power spectra are reported; these are related by \( n' = n - 2 \).

The turbulent scaling relations reflect the basic physics governing the flow. The classical theory of Kolmogorov (1941) applies to incompressible flow, i.e. one in which the velocities are negligible compared to the thermal speed \( \sigma_{th} = (P_{th}/\rho)^{1/2} \) (where \( \rho \) is the density and \( P_{th} \) is the thermal pressure); \( \sigma_{th} \) is equal to the sound speed \( c_s = (\gamma P_{th}/\rho)^{1/2} \) in an isothermal \((\gamma = 1)\) gas. In incompressible flows, energy is dissipated and turbulent motions are damped only for scales smaller than the Reynolds scale \( \ell_v \), at which the viscous terms in the hydrodynamic equations, \( \sim \nu v(\ell)^2/\ell^2 \), exceed the nonlinear coupling terms between scales, \( \sim v(\ell)^2/\ell \); here \( \nu \) is the kinematic viscosity. At scales large compared to \( \ell_v \), and small compared to the system as a whole, the rate of specific energy transfer \( \dot{\mathcal{E}} \) between scales is assumed to be conserved, and equal to the dissipation rate at the Reynolds scale. From dimensional analysis, \( \dot{\mathcal{E}} \sim v(\ell)^3/\ell \), which implies \( n = 11/3 \) and \( q = 1/3 \) for the so-called “inertial range” in Kolmogorov turbulence. The Kolmogorov theory includes the exact result that \( S_3(\ell) = -(4/5) \dot{\mathcal{E}} \ell \).

Because velocities \( v(\ell) \sim \sigma_v(\ell) \sim \Delta v(\ell) \) in molecular clouds are in general not small compared to \( c_s \), at least for sufficiently large \( \ell \), one cannot expect the Kolmogorov theory to apply. In particular, some portion of the energy at a given scale must be directly dissipated via shocks, rather than cascading conservatively through intermediate scales until \( \ell_v \) is reached. In the limit of zero pressure, the system would consist of a network of (overlapping) shocks; this state is often referred to as Burgers turbulence (Frisch & Bec, 2001). Since the power spectrum corresponding to a velocity discontinuity in one dimension has \( P(k) \propto k^{-2} \), an isotropic system of shocks in three dimensions would also yield power-law scalings for the velocity correlations, with \( n = 4 \) and \( q = 1/2 \). Note that correlations can take on a power-law form even if there is not a conservative inertial cascade; a
large range of spatial scales with consistent physics is still required.

Turbulence in a magnetized system must differ from the unmagnetized case because of the additional wave families and nonlinear couplings involved, as well as the additional diffusive processes – including resistive and ion-neutral drift terms (see §2.1.3). When the magnetic field $B$ is strong, in the sense that the Alfvén speed $v_A \equiv B/\sqrt{4\pi\rho}$ satisfies $v_A \gg v(\ell)$, a directionality is introduced such that the correlations of the flow variables may depend differently on $r_\parallel$, $r_\perp$, $k_\parallel$, and $k_\perp$, the displacement and wavevector components parallel and perpendicular to $\vec{B}$. For *incompressible* MHD turbulence, Goldreich & Sridhar (1995) introduced the idea of a critically-balanced anisotropic cascade, in which the nonlinear mixing time perpendicular to the magnetic field and the propagation time along the magnetic field remain comparable for wavepackets at all scales, so that $v_A k_\parallel \sim v(k_\perp, k_\parallel)k_\perp$. Interactions between oppositely-directed Alfvén wavepackets travelling along magnetic fields cannot change their parallel wavenumbers $k_\parallel = \vec{k} \cdot \vec{B}$, so that the energy transfers produced by these collisions involve primarily $k_\perp$; i.e. the cascade is through spatial scales $\ell_\perp = 2\pi/k_\perp$ with $v(\ell_\perp)^3/\ell_\perp \sim \text{constant}$. Combining critical balance with a perpendicular cascade yields anisotropic power spectra (larger in the $k_\perp$ direction); at a given level of power, the theory predicts $k_\parallel \propto k_\perp^{2/3}$. Magnetic fields and velocities are predicted to have the same power spectra.

Unfortunately, for the case of strong compressibility ($c_s \ll v$) and moderate or strong magnetic fields ($c_s \ll v_A \lesssim v$), which generally applies within molecular clouds, there is as yet no simple conceptual theory to characterize the energy transfer between scales and to describe the spatial correlations in the velocity and magnetic fields. On global scales, the flow may be dominated by large-scale (magnetized) shocks which directly transfer energy from macroscopic to microscopic degrees of freedom. Even if velocity differences are not sufficient to induce (magnetized) shocks, for trans-sonic motions compressibility implies strong coupling among all the MHD wave families. On the other hand, within a sufficiently small sub-volume of a cloud (and away from shock interfaces), velocity differences may be sufficiently subsonic that the incompressible MHD limit and the Alfvénic cascade approximately hold locally.

Even without direct energy transfer from large to small scales in shocks, a key property of turbulence not captured in classical models is intermittency effects – the strong (space-time) localization of dissipation in vortex sheets or filaments, which can occur even with a conservative energy cascade. (Shocks in compressible flows represent a different class of intermittent structures.) Signatures of intermittency are particularly evident in departures of high-order structure function exponents from the value $p/3$, and in non-Gaussian tails of velocity-increment probability distribution functions (PDFs) (e.g. Sreenivasan & Antonia, 1997; Lis et al., 1996). Proposed methods to account for intermittency in predicting correlation functions for incompressible, unmagnetized turbulence have been discussed by She & Leveque (1994) and Dubrulle (1994). Boldyrev (2002) proposed an adaptation of this framework for the compressible MHD case, but omitted direct dissipation of large-scale modes in shocks. Research on formal turbulence theory is quite active (see Elmegreen & Scalo, 2004 for a review of the recent theoretical literature relevant to the ISM), although a comprehensive framework remains elusive.

Large-scale numerical simulations afford a complementary theoretical approach
to model turbulence and to explore the spatial correlations within flows. Numerical experiments can be used to test formal theoretical proposals, and to provide controlled, quantitative means to interpret observations – within the context of known physics – when formal theories are either nonexistent or limited in detail. In drawing on the results of numerical experiments, it is important to ensure that the computational techniques employed adequately capture the relevant dynamical processes. For systems in which there are steep gradients of velocities and densities, grid-based methods are more accurate in following details of the evolving flow (such as development of instabilities) than SPH methods, which have been shown to have difficulty capturing shocks and other discontinuities (Agertz et al., 2006).

Spatial correlations within turbulent flows have been evaluated using numerical simulations in a variety of regimes. Overall, results are consistent with theoretical predictions in that power-law scalings in the velocity and magnetic field power spectra (or structure functions) are clear when there is sufficient numerical resolution to separate driving and dissipative scales. At resolutions of 512³ and above, results with a variety of numerical methods show angle-averaged power-law slopes $n = 7/2 - 11/3$ (i.e. $3.5 - 3.67$) for incompressible (i.e. $v/c_s \ll 1$) MHD flows (Müller, Biskamp, & Grappin, 2003; Haugen, Brandenburg, & Dobler, 2004; Müller & Grappin, 2005) and $n = 3.5 - 4.0$ for strongly compressible ($v/c_s \gtrsim 5$) flows both with (Vestuto, Ostriker, & Stone, 2003; Padoan et al., 2007) and without (Kritsuk, Norman, & Padoan, 2006) magnetic fields. Consistent with expectations, spectra are steeper for compressive velocity components than for magnetic fields (and also sheared velocity components if the magnetic field is moderate or strong), and steeper for more supersonic and/or more weakly magnetized models (see also Boldyrev, Nordlund, & Padoan, 2002; Padoan et al., 2004).

When strong mean magnetic fields are present, there is clear anisotropy in the power spectrum, generally consistent with the scaling prediction of Goldreich & Sridhar (1995), for both incompressible and compressible MHD turbulence (Cho & Vishniac, 2000; Maron & Goldreich, 2001; Cho, Lazarian, & Vishniac, 2002a; Vestuto, Ostriker, & Stone, 2003; Cho & Lazarian, 2003).

In order to identify the sources of turbulence in astronomical systems, it is also important to determine the behavior of velocity and magnetic field correlations on spatial scales larger than the driving scale. Numerical simulations, both for incompressible (Maron & Goldreich, 2001; Haugen, Brandenburg, & Dobler, 2004) and compressible MHD turbulence (Vestuto, Ostriker, & Stone, 2003), show that the power spectra below the driving wavenumber scale are nearly flat, $n \approx 0$; that is, “inverse cascade” effects are limited. For spatially-localized forcing (rather than forcing localized in $k$-space), Nakamura & Li (2007) also found a break in the power spectrum, at wavelength comparable to the momentum injection scale. Thus, the forcing scale for internally-driven turbulence in a system can be inferred observationally from the peak or knee of the velocity correlation function. If $v(\ell)$ continues to rise up to $\ell \sim L$, the overall scale of a system, this implies that turbulence is either (i) externally driven, (ii) imposed in the initial conditions when the system is formed, or (iii) driven internally to reach large scales. Note that for systems forced at multiple scales, or both internally and externally, breaks may be evident in the velocity correlation function (or power spectrum).
2.1.2 TURBULENT DISSIPATION TIMESCALES  Recent numerical simulations under quite disparate physical regimes have reached remarkably similar conclusions for the dissipation rates of turbulence. On dimensional grounds, the specific energy dissipation rate should equal $\epsilon U^3/\ell_0$, where $E = U^2/2$ is the total specific kinetic energy, $\ell_0$ is the spatial wavelength of the main energy-containing scale (comparable to the driving scale for forced turbulence; $\ell_0 \lesssim L$), and $\epsilon$ is a dimensionless coefficient. For incompressible turbulence, the largest-scale (4096$^3$ zones) incompressible, unmagnetized, driven-turbulence simulations to date (Kaneda et al., 2003) yield a dimensionless dissipation coefficient $\epsilon = 0.6$. For driven incompressible MHD turbulence (at 1024$^3$ resolution), the measured dimensionless dissipation rate $\epsilon \equiv (1/2)(\dot{E}_{\text{turb}}/E_{\text{turb}})(\ell_0/U)$ also works out to be $\epsilon = 0.6$ (Haugen, Brandenburg, & Dobler, 2004). Quite comparable results also hold for strongly-compressible ($U/c_s = 5$) turbulence at a range of magnetizations $v_A/c_s = 0 - 10$; Stone, Ostriker, & Gammie (1998) found that $\epsilon = 0.6 - 0.7$ for simulations at resolution up to 512$^3$ zones. For decaying compressible MHD turbulence, damping timescales are also comparable to the flow crossing time $\ell_0/v(\ell_0)$ on the energy-containing scale (Stone, Ostriker, & Gammie, 1998; Mac Low et al., 1998; Mac Low, 1999; Padoan & Nordlund, 1999). Thus, although very different physical processes are involved in turbulence dissipation under different circumstances, the overall damping rates summed over all available channels (including shock, reconnection, and shear structures) are nevertheless quite comparable. Defining the turbulent dissipation timescale as $t_{\text{diss}} = E_{\text{turb}}/|\dot{E}_{\text{turb}}|$ and the flow crossing time over the main energy-containing scale as $t_f = \ell_0/U$, $t_{\text{diss}} = t_f/(2\epsilon)$. Since velocities in GMCs increase up to the largest scale, $\ell_0 \rightarrow d$, the cloud diameter. Assuming that on average $U = \sqrt{3}\sigma_{\text{los}}$, the turbulent dissipation time based on numerical results is therefore given by

$$t_{\text{diss}} \approx 0.5 \frac{d}{\sigma_{\text{los}}}.$$  \hspace{1cm} (2)

This result is in fact consistent with the assumption of Mestel & Spitzer (1956) that turbulence in GMCs would decay within a crossing time.

The above results apply to homogenous, isotropic turbulence, but under certain circumstances if special symmetries apply, turbulent damping rates may be lower. One such case is for incompressible turbulence consisting of Alfvén waves all propagating in the same direction along the magnetic field. Note that for the incompressibility condition $\nabla \cdot \mathbf{v} = 0$ to apply, turbulent amplitudes must be quite low ($v \ll c_s$). Since Alfvén waves are exact solutions of the incompressible MHD equations, no nonlinear interactions, and hence no turbulent cascade, can develop if only waves with a single propagation direction are present in this case (see e.g. Chandran, 2004 for a mathematical and physical discussion). A less extreme situation is to have an imbalance in the flux of Alfvén waves propagating “upwards” and “downwards” along a given magnetic field direction. Maron & Goldreich (2001) show that in decaying incompressible MHD turbulence, the power in both upward- and downward-propagating components decreases together until the lesser component is depleted. Cho, Lazarian, & Vishniac (2002a) quantify decay times of imbalanced incompressible turbulence, finding for example that if the initial imbalance is $\approx 50\%$ or $\approx 70\%$, then the time to decay to half the initial energy is increased by a factor 1.5 or 2.3, respectively, compared to the case of no imbalance.

For even moderate-amplitude subsonic velocities, however, Alfvén waves couple
to other wave families and the purely Alfvénic cascade is lost. For strongly supersonic motions, as are present in GMCs, the mode coupling is quite strong. As a consequence, even a single circularly polarized Alfvén wave cannot propagate without losses; a parametric instability known as the “decay instability” (Sagdeev & Galeev, 1969) develops in which three daughter waves (a forward-propagating compressive wave and two oppositely propagating Alfvén waves when $\beta \ll 1$) grow at the expense of the mother wave. The initial growth rate of the instability is $\gamma = (0.1 - 0.3) k v_A$ when $v(k)/c_s = 1 - 3$ and $\beta \equiv 2 c_s^2/v_A^2 = 0.2$, and larger for greater amplitudes and smaller $\beta$ (Goldstein, 1978). The ultimate result is decay into fully-developed turbulence (Ghosh & Goldstein, 1994; Del Zanna, Velli, & Londrillo, 2001). Thus, for conditions that apply within GMCs, even if there were a localized source of purely Alfvénic waves (i.e. initially 100% imbalanced), the power would rapidly be converted to balanced, broad-spectrum turbulence with a short decay time. The conclusion that turbulent damping times within GMCs are expected to be comparable to flow crossing times has important implications for understanding evolution in star-forming regions; these will be discussed in §3.1 and 3.2.2.

2.1.3 PHYSICAL SCALES IN TURBULENT FLOWS In classical incompressible turbulence, the only physical scales that enter are the outer scale $\ell_0$ at which the medium is stirred, and the inner “Reynolds” scale $\ell_r$ at which viscous dissipation occurs. Assuming Kolmogorov scaling $v(\ell) = v(\ell_0)(\ell/\ell_0)^{1/3}$ [for $v(\ell) \sim \sigma_v(\ell) \sim \Delta v(\ell)$], the dissipation scale is

$$\frac{\ell_\nu}{\ell_0} = \left[ \frac{\nu}{\ell_0 v(\ell_0)} \right]^{3/4} \equiv Re^{-3/4}. \quad (3)$$

Here $Re \equiv v(\ell_0)\ell_0/\nu$ is the overall Reynolds number of the flow; if turbulence increases up to the largest scales then $Re = UL/\nu$. With $\nu \sim c_s \lambda_{mfp}$ for $\lambda_{mfp}$ the mean free path for particle collisions, $\ell_\nu/\ell_0 \sim (\lambda_{mfp}/\ell_0)^{3/4}[v(\ell_0)/c_s]^{-3/4}$. In fact, the velocity-size scaling within GMCs has power-law index $q$ closer to 1/2 than 1/3 on large scales, because large-scale velocities are supersonic and therefore the compressible-turbulence results apply. Allowing for a transition from $q = 1/2$ to $q = 1/3$ at an intermediate scale $\ell_s$ where $v(\ell_s) = c_s$ (see below), $\ell_\nu = \ell_s^{1/4}\lambda_{mfp}^{3/4}$. Using typical GMC parameters so that $\lambda_{mfp} \sim 10^{15}$ cm and $\ell_s \sim 0.03$ pc yields $\ell_s \sim 3 \times 10^{-5}$ pc. This is tiny compared to the sizes, $\sim 0.1$ pc, of self-gravitating cores in which individual stars form.

The length $\ell_s$ introduced above marks the scale at which the RMS turbulent velocity is equal to the sound speed. At larger scales, velocities are supersonic and compressions are strong; at smaller scales, velocities are subsonic and compressions are weak. Taking $v(\ell) = v(\ell_0)(\ell/\ell_0)^q$, the sonic scale is $\ell_s = \ell_0[c_s/v(\ell_0)]^{1/q}$, or $\ell_s \approx \ell_0[c_s/v(\ell_0)]^2$ when $q \approx 1/2$. Density perturbations with characteristic scales $\sim \ell_s$ will have order-unity amplitude in an unmagnetized medium. In a magnetized medium, the amplitude of the perturbation imposed by a flow of speed $v$ will depend on the direction of the flow relative to the magnetic field. Flows along the magnetic field will be as for an unmagnetized medium, while flows perpendicular to the magnetic field will create order-unity density perturbations only if $v > (c_s^2 + v_A^2)^{1/2}$. Note that the thermal scale, at which the line-of-sight turbulent velocity dispersion $\sigma_v/\sqrt{3}$ is equal to the one-dimensional thermal speed $\sigma_{th}$, is larger than $\ell_s$ by a factor $\approx 3$.

Another scale that is important for MHD turbulence in fully-ionized gas is the
resistive scale; below this scale Ohmic diffusion would smooth out strong bends in the magnetic field, or would allow folded field lines to reconnect. The resistive scale $\ell_\eta$ is estimated by equating the diffusion term $\sim \eta B(\ell)/(4\pi \ell^2)$ to the flux-dragging term $\sim v(\ell)B(\ell)/\ell$ in the magnetic induction equation. Defining the magnetic Reynolds number as $Rm \equiv v(\ell_0)\ell_0\eta/(\eta B)$, and taking $v(\ell) \sim c_s(\ell/\ell_0)^{1/3}$ at small scales, this yields $\ell_\eta/\ell_\nu = (Re/Rm)^{3/4}$. Since the “magnetic Prandtl number” $Rm/Re$ is very large ($\sim 10^6$), this means that the magnetic field could, for a highly-ionized medium, remain structured at quite small scales (see Cho, Lazarian, & Vishniac, 2002b for discussion of this in the diffuse ISM).

In fact, under the weakly-ionized conditions in star-forming regions, ambipolar diffusion (ion-neutral drift) becomes important well before the resistive (or Ohmic diffusion) scale is reached. Physically, the characteristic ambipolar diffusion scale $\ell_{AD}$ is the smallest scale for which the magnetic field (which is frozen to the ions) is well-coupled to the bulk of the gas, for a partially-ionized medium. An estimate of $\ell_{AD}$ is obtained by equating the ion-neutral drift speed, $\sim B_0\delta B/(m_i + m)/[4\pi \rho_i \rho c_s(\ell)]$, with the turbulent velocity, $\delta v$. Here, $\alpha_{in} = \langle \sigma_{in} | v_i - v_n \rangle \approx 2 \times 10^{-9} \text{ cm}^3 \text{s}^{-1}$ is the ion-neutral collision rate coefficient (Draine, Roberge, & Dalgarno, 1983), and $m_i, m$ and $\rho_i, \rho$ are the ion and neutral mass and density. The resulting ambipolar diffusion scale, assuming $m_i \gg m$ (for either metal or molecular cations) and $\delta v \approx \delta B/\sqrt{4\pi \rho}$, is

$$\ell_{AD} = \frac{v_A}{n_i \alpha_{in}} \approx 0.05 \text{ pc} \left( \frac{v_A}{3 \text{ km s}^{-1}} \right) \left( \frac{n_i}{10^{-3} \text{ cm}^{-3}} \right)^{-1}. \quad (4)$$

Here, $v_A$ is the Alfvén speed associated with the large-scale magnetic field $B_0$.

The ambipolar diffusion scale (4) depends critically on the fractional ionization, which varies greatly within star-forming regions. Regions with moderate $A_V \lesssim 5$ can have relatively large ionization fraction due to UV photoionization, whereas regions with large $A_V$ are ionized primarily by cosmic rays (see §2.3.1). For example, if the electrons are attached to PAHs in dense cores, the ion density is $n_i \approx 10^{-3}(n_H/10^4 \text{cm}^{-3})^{1/2}(\zeta_{CR}/3 \times 10^{-17} \text{s}^{-1})^{1/2}$ (Tiellens, 2005) where $\zeta_{CR}$ is the cosmic ray ionization rate per H atom.. Since $n_i \propto n^{1/2}$, we can express equation (4) in terms of column density and magnetic field strength as $\ell_{AD}/\ell = 0.09(B/10 \mu \text{G})/(N_H/10^{21} \text{ cm}^{-2})$.

For spatial wavelengths $\lambda = 2\pi/k < \pi \ell_{AD}$, MHD waves are unable to propagate in the coupled neutral-ion fluid at all, because the collision frequency of neutrals with ions, $n_i \alpha_{in}$, is less than (half) the wave frequency $\omega = kv_A$. For $\lambda > \pi \ell_{AD}$, MHD waves are damped at a rate $\omega \ell_{AD}/\lambda$ (Kulsrud & Pearce, 1969). Thus, at scales $\ell \lesssim \ell_{AD}$, the magnetic field will be essentially straight and uniform in magnitude, and any further turbulent cascade will be as for an unmagnetized medium. The scale $\ell_{AD}$ is also comparable to the thickness of the $C$-type shocks that are typical under prevailing conditions within GMCs (Draine & McKee, 1993). Further discussion of the interaction between turbulence and ambipolar diffusion is given by Zweibel (2002), Fatuzzo & Adams (2002), and Heitsch et al. (2004).

2.1.4 DENSITY STRUCTURE IMPOSED BY TURBULENCE When turbulent velocities at a given scale are supersonic, they impose density variations within the flow at that scale. For star-forming regions, in which turbulent velocities are increasingly supersonic for scales $\gtrsim 0.1$ pc, the density becomes strongly structured over a wide range of scales. This density structure – which is crucial to the star-formation process – can be characterized statistically in a
The simplest (one-point) statistic is the distribution of mass (or volume) as a function of density, usually referred to as the density PDF (probability density function). For isothermal gas and supersonic turbulence (either forced or decaying), a number of 3D numerical simulations both with (Ostriker, Stone, & Gammie, 2001; Ostriker, 2003; Li et al., 2004) and without (Nordlund & Padoan, 1999; Klessen, 2000) magnetic fields have shown that the density PDF approaches a log-normal distribution when self-gravity is unimportant. This functional form can be understood (Vázquez-Semadeni, 1994; Passot & Vazquez-Semadeni, 1998) to arise as a consequence of multiple, independent dynamical events that alter the density according to

\[ \frac{\rho}{\bar{\rho}} = \Pi_i(1 + \delta_i) \] where \( \delta_i > 0 \) (or \( \delta_i < 0 \)) for compressions (or rarefactions). From the Central Limit Theorem, \( \log(\frac{\rho}{\bar{\rho}}) \) is the sum of independent random variables, and should therefore approach a Gaussian distribution. When the equation of state departs from a simple isothermal form, the density PDF still follows a log-normal over a range of densities, but acquires power-law tails either at high or low density depending on whether the equation of state is softer or stiffer than isothermal (Passot & Vazquez-Semadeni, 1998; Scalo et al., 1998; see also e.g. Wada, 2001).

For a log-normal distribution, the fraction of volume (V) or mass (M) as a function of \( x \equiv \ln(\frac{\rho}{\bar{\rho}}) \) is given by

\[ f_{V,M}(x) = \frac{1}{\sqrt{2\pi}\sigma_x^2} \exp \left[ \frac{-(x \pm |\mu_x|)^2}{2\sigma_x^2} \right] \] (5)

where the mean and dispersion of the distributions are related by \( \mu_x = \sigma_x^2 / 2 \), and the upper and lower signs correspond to volume- and mass-weighting, respectively. For a log-normal distribution, the mass-weighted median density (half of the mass is at densities above and below this value) is \( \rho_{med} = \bar{\rho} \exp(\mu_x) \), whereas the mass-weighted mean density is \( \langle \rho \rangle_M = \bar{\rho} \exp(2\mu_x) \). Based on three-dimensional unmagnetized simulations, Padoan, Jones, & Nordlund (1997) propose that \( \mu_x \approx 0.5 \ln(1 + 0.25M^2) \). Other three-dimensional simulations with magnetic fields (\( \beta = 0.02 - 2 \)) have found \( \mu_x \approx 0.5 - 1 \) for \( M \approx 5 - 10 \) (Ostriker, Stone, & Gammie, 2001). These models confirm that the mean density contrast generally grows as the turbulence level increases, but find no one-to-one relationship between \( \mu_x \) and \( M \) (or the fast magnetosonic Mach number, \( M_F \equiv \sigma_v/(v_s^2 + v_A^2)^{1/2} \)). The large scatter at large \( M \) is because the flow is dominated by a small number of large-amplitude modes (i.e. large “cosmic variance”), some of which are compressive and some of which are shear. With magnetic fields, Ostriker, Stone, & Gammie (2001) found that the lower envelope of the \( \mu_x \) distribution increases with \( M_F \) according to \( \mu_{x,\text{min}} = 0.2 \ln(1 + M_F^2) + 0.5 \) for \( M_F = 0.5 - 2.5 \).

Because the velocity field is spatially correlated, the density distribution will also show spatial correlations over a range of scales. Density correlations can be characterized in terms of the autocorrelation function, the power spectrum, and structure functions of various orders (cf §2.1.1); usually, analyses are applied to \( \delta \rho \equiv \rho - \bar{\rho} \). Using delta-variance techniques, Mac Low & Ossenkopf (2000) show that correlations in density decrease for wavelengths above the velocity driving scale, and that there are relatively modest differences in the density correlations between unmagnetized and magnetized models when all other properties are controlled.
Kim & Ryu (2005) have analyzed the dependence of the spectral index on Mach number for three-dimensional turbulence forced at large spatial scales, using isothermal, unmagnetized simulations at resolution $512^3$. For $\mathcal{M} < 1$, the indices $n_\rho$ or $n'_\rho$ of the density power spectrum $|\delta \rho(k)|^2$ are similar to those of the velocity field in incompressible turbulence – i.e. near $n = 11/3$ or $n' = 5/3$; this is simply because $\delta \rho(k)/\bar{\rho} \sim -k \cdot \mathbf{v}(k)/c_s$ for low-amplitude quasi-sonic compressions (note that even when $\mathcal{M} = 1$, the Mach number for the compressive component of the velocity field is $< 1$). As the Mach number increases, the density power spectrum flattens, reaching $n'_\rho \approx 0.5$ for $\mathcal{M} = 12$. For comparison, a one-dimensional top hat – corresponding to a large clump in 3D – would have $n'_\rho = 2$, whereas a one-dimensional delta function – corresponding to a thin sheet or filament in 3D – would have $n'_\rho = 0$. Note that for the density to take the form of multiple delta functions, the velocity field must generally be a composite of step functions – corresponding to shocks – and has $n' = 2$ for the velocity power spectrum (as discussed above). The low value of $n'_\rho$ at large Mach number implies the density structure becomes dominated by curved sheets and filaments. Curved sheets represent stagnation regions (of the compressive velocity field) where shocked gas from colliding flows settles, and filaments mark the intersections of these curved sheets.

Other statistical descriptions of density structure include fractal dimensions (e.g., Elmegreen & Falgarone, 1996; Stutzki et al., 1998), multifractal spectra (Chappell & Scalo, 2001), and hierarchical structure trees (Houlahan & Scalo, 1992); see Elmegreen & Scalo (2004) for a discussion. The spatial correlation of density can also be characterized in terms of clump mass functions. Clump-finding techniques have been applied to simulations of supersonic turbulent flows by a number of groups; these results will be discussed and compared to observations in §3.3.

2.1.5 OBSERVATIONS OF TURBULENCE For observed astrophysical systems, the intrinsic properties of turbulence cannot be directly obtained, due to line-of-sight projection and the convolution of density and velocity in producing observed emission. A number of different techniques have been developed, calibrated using simulations, and applied to observed data, in order to deduce characteristics of the three-dimensional turbulent flow from the available observations, which include spectral line data cubes (from molecular transitions), continuum emission maps (from dust), maps of extinction (using background stars), and maps of polarization (in extinction and emission from dust). Elmegreen & Scalo (2004) review the extensive literature on observations of turbulence. Here, we will mention just a few results.

The defining property of turbulent motion – in contrast to, for example, the purely random motions of gas particles in a Maxwell-Boltzmann distribution or the highly systematic motions of stars in a rotating system – is the stochastic yet scale-dependent behavior of flow correlations. Larson (1981) was the first to draw attention to the genuine “turbulent” nature of motions internal to star-forming regions, as expressed by an empirical scaling law of the form (1) with $q = 0.38$. Using more homogeneous data, Solomon et al. (1987) obtained a “linewidth-size” scaling index $q \approx 0.5$ for GMCs as a whole. Passot, Pouquet, & Woodward (1988) pointed out that the linewidth-size scaling $\sigma_\ell(\ell) \propto \ell^{1/2}$ observed in star-forming regions is indeed what would be predicted for “Burgers turbulence,” a more appropriate model than Kolmogorov turbulence given the strongly supersonic...
Many subsequent studies have been made of observed scaling behavior of velocities, both for subsystems of a given star-forming region, and for systems that are spatially disjoint. A number of methods have been developed for these investigations, including autocorrelation analysis (Miesch & Bally, 1994) and delta-variance analysis (Ossenkopf et al., 2006) applied to line centroids, the spectral correlation function (Rosolowsky et al., 1999), velocity channel analysis (Lazarian & Pogosyan, 2004; Padoan et al., 2006), and principal component analysis (PCA) (Brunt & Heyer, 2002). Overall, analyses agree in finding power-law linewidth-size relations, with similar coefficients and power-law exponents close to \( q = 0.5 \). The lack of features in velocity correlations at intermediate scales, and more generally the secular increase in velocity dispersion up to sizes comparable to the whole of a GMC, indicates that turbulence is driven on large scales within or external to GMCs (e.g. Ossenkopf & Mac Low, 2002; Brunt, 2003).

Interestingly, turbulence appears to have a “universal” character within most of the molecular gas in the Milky Way, in the sense that the same scaling laws with the same coefficients fit both entire GMCs and moderate-density substructures (observed via CO lines) within them. Using PCA, Heyer & Brunt (2004) find a fit to the amplitudes of line-of-sight velocity components as a function of scale following

\[
\delta v = 0.9 \left( \frac{L_{\text{pca}}}{1 \text{ pc}} \right)^{0.56 \pm 0.02} \text{ km s}^{-1}
\]  

(6)

based on composite data of all PCA components from scales \( L_{\text{pca}} \sim 0.03 - 30 \) pc in a sample of 27 molecular clouds. Using data just within individual clouds, Heyer & Brunt (2004) find a mean scaling exponent that is slightly lower, \( q = 0.49 \pm 0.15 \). Note that the lengths \( L_{\text{pca}} \) entering the relation (6) are the characteristic scales of PCA eigenmodes, and may differ from size scales defined in other ways. For example, the effective GMC cloud diameters as measured by Solomon et al. (1987) are on average about four times the maximum \( L_{\text{pca}} \) found in each cloud. Based on the scaling law (6) the sonic length will be similar, \( \ell_s \sim 0.03 \) pc (allowing for varying definitions of size), in all GMCs. We discuss this empirical result further in §3.1; note that strongly self-gravitating clumps with high densities and surface densities depart from the relation given in equation (6).

For evaluating the density distribution, the most unbiased measurements use dust extinction maps (see Lada, Alves, & Lombardi, 2007 and references therein). A promising new technique for observing the density distribution uses scattered infrared light (Foster & Goodman, 2006), which can probe the structure of molecular clouds for visual extinctions of \( 1 - 20 \) magnitudes at very high spatial resolution (Padoan, Juvela, & Pelkonen, 2006). Consistent with the prediction of numerical simulations (Ostriker, Stone, & Gammie, 2001; Vázquez-Semadeni & García, 2001), distributions of extinction follow log-normal functional forms to an excellent approximation; distributions of integrated intensity from molecular lines, on the other hand, are not log-normal (Ridge et al., 2006), presumably due to a combination of chemistry and/or optical depth effects. Column density distributions of course cannot be directly inverted to obtain volume density distributions. Since the Fourier transform of the column density, \( N(k_x, k_y) \), is equal to \( \delta \rho(k_x, k_y; k_z = 0) \) (up to an overall normalization; here \( z \) is the line-of-sight direction), if statistical isotropy holds then at least the shape of the density power spectrum can be obtained from a well-sampled map of column density. Assum-
ing isotropy, integrated-intensity $^{12}$CO and $^{13}$CO line maps yield density power spectra $|\delta \rho(k)|^2 \propto k^{-n_{\rho}}$ with $n_{\rho} = 2.5 - 2.8$ (Bensch, Stutzki, & Ossenskopf, 2001), which is consistent with the large-$M$ results for density power spectra $n_{\rho}' = n_{\rho} - 2 \approx 0.5$ obtained in the simulations of Kim & Ryu (2005). In principle, features in the density power spectrum should be evident both at the sonic scale, $\ell_s$, and the Jeans scale (these scales are comparable in star-forming regions — Padoan, 1995). A first step toward identifying features in the density power spectrum, using velocity-integrated CO intensity, was taken by Blitz & Williams (1997). It will be very interesting to investigate column density power spectra based on high-resolution extinction maps in both self-gravitating GMCs and unbound molecular clouds, evaluating the slopes and searching for evidence of these breaks.

Other measures of density structure, including the fractal dimension $D \approx 2.3$ empirically measured by Elmegreen & Falgarone (1996), are in agreement with simulations of strongly compressible turbulence (Kritsuk, Norman, & Padoan, 2006). In addition, the typical range of density contrasts obtained for 3D supersonic turbulence (see §2.1.4) is consistent with the compressions required to explain the low effective volume filling factors of gas deduced from CO observations of GMCs (Ostriker, Gammie, & Stone, 1999). If $\langle \ln(n/\bar{n}) \rangle_M = \mu_x$ (eq. 5), then the mass-weighted mean density is $\langle n/\bar{n} \rangle_M = \langle (n/\bar{n})^2 \rangle_V = \exp(2\mu_x)$, so that $\mu_x \approx 1.5$ yields a local density $n = 10^3$ when $\bar{n} = 50$, in agreement with inferred filling factors $\sim 0.1$ (Bally et al., 1987; Williams, Blitz, & Stark, 1995).

Observational estimates of the density and filling factor are often made under the assumption of a constant clump density, however, not the broad distribution of densities expected for a log normal, which may introduce some differences. In more detail, Padoan et al. (1999) have shown that the statistical properties of $^{13}$CO spectra seen in the star-forming Perseus molecular cloud can be well reproduced by synthetic non-LTE spectra created using simulation data cubes from non-self-gravitating supersonic turbulence simulations.

### 2.2 Self-Gravity

The effects of self-gravity on a turbulent cloud can be analyzed with the aid of the virial theorem, which in Lagrangian form (i.e., for a fixed mass) is

$$\frac{1}{2} \ddot{I} = 2(T - T_s) + B + W \quad (7)$$

(Chandrasekhar & Fermi, 1953b; Mestel & Spitzer, 1956), where $I = \int \mathbf{r}^2 dm$ is proportional to the trace of the inertia tensor of the cloud. (It is often assumed that the sign of $\ddot{I}$ determines whether the cloud is expanding or contracting, but in fact it determines the acceleration of the expansion or contraction—Ballesteros-Paredes, 2006.) The term

$$T = \int_{V_{cl}} \left( \frac{3}{2} P_{th} + \frac{1}{2} \rho v^2 \right) \ dV \equiv \frac{3}{2} \rho \sigma^2 V_{cl} \quad (8)$$

is the total kinetic energy in the cloud (thermal plus bulk), where $\sigma^2$ is the 1D mean square velocity (including both thermal $\sigma_{th}^2$ and nonthermal [turbulent] $\sigma_{nt}^2 \equiv \frac{1}{3} \langle v_r \rangle^2$ terms) in the cloud, $V_{cl}$ is the volume of the cloud, and $T_s = \iint_{\partial Q} P_{th} \mathbf{r} \cdot d\mathbf{S}$
is the surface kinetic term. The term

\[ B = \frac{1}{8\pi} \int_{V_{cl}} B^2 dV + \frac{1}{4\pi} \int r \cdot (\mathbf{BB} - \frac{1}{2} B^2 \mathbf{I}) \cdot dS \]  

(9)

is the net magnetic energy, including the effects of the distortion of the field outside the cloud. The volume and surface magnetic terms cancel for a completely uniform magnetic field, since a uniform field exerts no force. Finally,

\[ W = -\int \rho \mathbf{r} \cdot \nabla \Psi \ dV \]  

(10)

is the gravity term, equal to the gravitational self-energy \((1/2) \int \rho \Psi \ dV\) provided the acceleration due to masses outside the system is negligible (as is generally the case for dense clouds embedded in a diffuse turbulent background—Dib et al., 2007).

The virial theorem can also be written in Eulerian form, so that it applies to a fixed volume (McKee & Zweibel, 1992); in that case, the surface term for the kinetic energy includes the dynamic pressure \(\frac{1}{2} \rho \mathbf{v}^2 \cdot dS\), and the theorem itself includes a term \((1/2)(d/dt) \int (\rho \mathbf{v}^2) \cdot dS\) on the left-hand side. This form of the virial theorem is particularly appropriate in a turbulent medium, in which the mass of a cloud is not necessarily fixed. With this form of the equation, clouds that are actively forming or dispersing may have surface kinetic terms comparable to the volume kinetic terms.

Ballesteros-Paredes, Vázquez-Semadeni, & Scalo (1999) examined the various terms in the Eulerian virial theorem for clumps and cores in their turbulent, self-gravitating MHD simulations, and found that the time-dependent terms on the left-hand side are significant, and that the surface terms are generally comparable to the volume terms. Dib et al. (2007) confirmed this and also found that only objects with large density contrasts are virialized. However, in contrast to the other terms in the virial theorem, the time-dependent terms can change sign, so they become less important if the virial theorem is averaged over time for an individual cloud (provided the cloud lives more than a dynamical time) or if it is averaged over an ensemble of clouds (McKee, 1999). In §3.1, we discuss the application of the virial theorem both to observed molecular clouds, clumps, and cores, and to condensations identified within numerical simulations.

The virial parameter is defined to be proportional to the ratio of the total kinetic energy to the gravitational energy (Bertoldi & McKee, 1992; cf. Falgarone, Puget, & Perault, 1992),

\[ \alpha_{vir} \equiv \frac{5\sigma^2 R}{GM}, \]  

(11)

where the numerical coefficient is chosen so that \(\alpha_{vir} = 1\) for a uniform, unmagnetized gas sphere in virial balance \((W = -2T; \text{ but note that such a sphere is not in hydrostatic equilibrium})\). This relation implies that the mean pressure in a cloud varies as the square of the mean surface density, \(\Sigma\),

\[ P_{tot} = \phi_P G \Sigma^2, \]  

(12)

where \(\phi_P \propto \alpha_{vir}\) is a numerical factor of order unity for gravitationally bound objects (McKee & Tan, 2003). The total pressure includes the magnetic pressure; fluctuating magnetic fields have an energy that is about 60% of that of the turbulent kinetic energy (Stone, Ostriker, & Gammie, 1998) and contribute
an effective pressure support that is about 30% of the turbulent kinetic pressure support (McKee & Tan, 2003) (note that $T - T_s$ and $B$ appear in equation 7 with coefficients 2 and 1, respectively). Gravitationally bound objects have $\alpha_{\text{vir}} \sim 1$, which (since $M \sim \rho R^3$) defines gravitational length, time, and mass scales,

$$R_G = \sigma/(G\rho)^{1/2}, \quad t_G = 1/(G\rho)^{1/2}, \quad M_G = \sigma^3/(G^3\rho)^{1/2}.$$  \hspace{1cm} (13)

These scales, derived essentially from dimensional analysis, govern the structure and stability of self-gravitating clouds (the density $\rho$ can be chosen to be the central density, the mean density, or the density at the surface, depending on the application). The gravitational time scale is often expressed in terms of the free-fall time, which is the time for a pressure-free, spherical cloud to collapse to a point due to its self-gravity,

$$t_{\text{ff}} = \left(\frac{3\pi}{32G\rho}\right)^{1/2} = 1.37 \times 10^6 \left(\frac{10^3 \text{ cm}^{-3}}{n_H}\right)^{1/2} \text{ yr},$$

where the numerical value is based on a He abundance of 10% by number.

The simplest case of a self-gravitating cloud is a static isothermal cloud with no magnetic field. For a given surface pressure $P_{\text{th},0} = \rho_0 \sigma_{\text{th}}^2$, the critical mass $M_{\text{cr}}$—i.e., the maximum mass for such a cloud to be in hydrostatic equilibrium (stable or unstable)—is the Bonnor-Ebert mass (Bonnor, 1956; Ebert, 1957),

$$M_{\text{BE}} \equiv 1.182 \frac{\sigma_{\text{th}}^4}{(G^3 P_{\text{th},0})^{1/2}} = 1.182 \frac{\sigma_{\text{th}}^3}{(G^3 \rho_0)^{1/2}}.$$  \hspace{1cm} (15)

For conditions typical of dense clumps within low-mass star-forming regions, this is of order a solar mass: $M_{\text{BE}} = 0.66(T/10 \text{ K})^2/(P_{\text{th},0}/3 \times 10^5 k_B \text{ cm}^{-3} \text{ K})^{1/2} M_\odot$, where $k_B$ is Boltzmann’s constant and the pressure is normalized to the mean kinetic pressure in a typical GMC ($\S 3.1$) (which is similar to the mean thermal pressure in dense clumps). Note that the Bonnor-Ebert mass is very nearly equal to the characteristic gravitational mass $M_G(\rho_0)$, when evaluated at the conditions at the surface of the cloud. The radius of a Bonnor-Ebert sphere is $R_{\text{BE}} = 0.486 \sigma_{\text{th}}/(G\rho_0)^{1/2} = 0.486 R_G(\rho_0)$; this is comparable to the Jeans length (see below).

The importance of the magnetic field to cloud structure is determined by the ratio of the mass to the magnetic critical mass $M_{\Phi}$, which is defined by the condition that the magnetic energy equal the gravitational energy, $\mathcal{B} = |\mathcal{W}|$, for a cold cloud in magnetostatic equilibrium:

$$M_{\Phi} \equiv c_\Phi \frac{\Phi}{G^{1/2}},$$

where $\Phi$ is the magnetic flux threading the cloud (e.g., see the review by McKee et al., 1993). Magnetic fields alone cannot prevent gravitational collapse in magnetically supercritical clouds ($M > M_{\Phi}$), whereas gravitational collapse is not possible in magnetically subcritical clouds ($M < M_{\Phi}$); keep in mind, however, that $M$ can change as the result of flows along the field, and $M_{\Phi}$ can change due to ambipolar diffusion. The numerical coefficient $c_\Phi$ depends on the internal distribution of density and magnetic fields. A cold cloud with a poloidal field and a constant mass-to-flux ratio has $c_\Phi = 0.17$ (Tomisaka, Ikeuchi, & Nakamura, 1988), essentially identical to the critical value of the mass-to-flux ratio for an infinite cold sheet, $G^{1/2}(\Sigma/B)_{\text{cr}} = 1/(2\pi) \simeq 0.16$ (Nakano & Nakamura, 1978). For
clouds with two other distributions of the mass-to-flux ratio, Tomisaka, Ikeuchi, & Nakamura (1988) found that the critical mass-to-flux ratio for the central flux tube corresponds to \( c_\phi \approx 0.17 - 0.18 \). For more complex field geometries, the magnetic flux does not determine the mass that can be supported by magnetic stresses; for example, if the field is poloidal, with half the field pointing one way and half the other, so that the total flux is zero, the mass that could be supported would initially be \( M_\Phi \), but it would go to zero as the field reconnects. For a random field, arguments based on McKee & Holliman (1999) suggest that the mass that can be supported by magnetic fields is comparable to that in equation (16), but with \( \Phi \text{ replaced by } \pi R^2 (B^2)^{1/2} \). Of course, when turbulent magnetic fields are present so are turbulent velocities, which lend their own support to the cloud (see below).

The magnetic critical mass can also be expressed in terms of the mean density and magnetic field in the cloud (Mouschovias & Spitzer, 1976),

\[
\frac{M_B}{M} = \left( \frac{M_\Phi}{M} \right)^3.
\]

For an ellipsoidal cloud of size \( 2Z \) along the axis of symmetry and radius \( R \) normal to the axis, this becomes (Bertoldi & McKee, 1992)

\[
M_B = 79c_\phi^3 \left( \frac{R}{Z} \right)^2 \frac{\bar{v}_A^3}{(G^3 \rho)^{1/2}} = 1020 \left( \frac{R}{Z} \right)^2 \left( \frac{\bar{B}}{30 \mu G} \right)^3 \left( \frac{10^3 \text{ cm}^{-3}}{n_H} \right)^2 \frac{M_\odot}{M},
\]

\[\text{(18)}\]

where the latter expression uses \( c_\phi = 1/(2\pi) \). Note that \( M_B \) has the same form as the gravitational mass \( M_G \), with the velocity dispersion \( \sigma \) replaced by the Alfvén velocity \( \bar{v}_A \equiv B/(4\pi \rho)^{1/2} \). Based on the idea that cores form from sheets that are supported by kinetic pressure along the magnetic field and magnetic tension in the (perpendicular) plane, Shu, Li, & Allen (2004) have introduced another mass scale, \( M_0 = \pi^2 \sigma_A^4/(G^3/2B) \), which yields values \( \sim M_\odot \) when \( \sigma \to \sigma_{\text{th}} \) and \( \bar{B} \to 30 \mu G \).

Just as in the case of stellar structure, it is useful to consider polytropic models of molecular clouds, in which the pressure is a power-law function of the density,

\[
P(r) = K \rho(r)^{\gamma_p},
\]

\[\text{(19)}\]

where \( K \) is constant and \( \gamma_p \) is often written as \( 1 + 1/n \). Here, \( P(r) \) and \( \rho(r) \) represent the total pressure and density averaged over the surface of a sphere of radius \( r \). This approach is based on the microturbulent approximation, in which the turbulent pressure \( \rho \sigma_{\text{nt}}^2 \) is included in the total pressure (Chandrasekhar, 1951a,b); this is equivalent to assuming that the random dynamical motions are isotropic. For a given cloud at a given time, this is reasonable for small-scale motions, but the approximation becomes worse as the scale of the motion becomes comparable to the scale on which the pressure is being evaluated. However, just as in the case of the virial theorem, the microturbulent approximation becomes better – for objects that live more than a dynamical time – if a time average is taken. Polytropes are spherical, so polytropic models apply only to objects with well-defined centers; for such objects, an angular average is also necessary, which improves the accuracy of the microturbulent approximation. Star-forming clumps and cores often appear centrally concentrated and are therefore suitable for modeling with a polytrope, whereas many GMCs do not appear to have well-defined centers and are not very suitable for polytropic models.
For a polytrope, the velocity dispersion obeys $\sigma^2 = P/\rho \propto \rho^{\gamma-1}$. If the mean density decreases with increasing scale (as it does for an object in hydrostatic equilibrium), it follows that the velocity dispersion increases with scale for $\gamma_p < 1$, which is consistent with observations of molecular clouds (Maloney, 1988). Because $n = 1/(\gamma_p - 1)$ is negative in this case, such polytropes are often referred to as “negative-index polytropes.” The stability of a polytrope depends on both $\gamma_p$ and on its adiabatic index $\gamma$, which describes the change in the pressure associated with a given perturbation in density, $\delta \ln P = \gamma \delta \ln \rho$. The value $\gamma = \frac{4}{3}$ is critical for spherical clouds: clouds with $\gamma > \frac{4}{3}$ are gravitationally stable for arbitrarily large masses, whereas those with $\gamma < \frac{4}{3}$ are unstable for sufficiently large masses, or, at fixed mass, for sufficiently high ambient pressures. Correspondingly, the gravitational mass $M_G$ is independent of density for $\gamma = \frac{4}{3}$. Polytropes with $\gamma_p < \frac{6}{5}$ must be confined by an ambient pressure (Chandrasekhar, 1939), and their properties are determined by the pressure of the ambient medium. McKee & Holliman (1999) show that polytropes with $0 < \gamma_p \leq 1$ have masses $\leq 1.182M_G(\rho_0)$; the mean density and pressure of these polytropes are $< 3.8$ times the surface values.

As discussed in §2.1.5, turbulent regions exhibit a line width–size relation in which the velocity dispersion averaged over a volume increases systematically with size scale, $\sigma_{\text{nt}} \propto r^q$. Observations often show $q \simeq 1/2$, the value expected for Burgers turbulence (see §2.1.1). In general, this line width–size relation reflects the statistical increase in velocity differences with separation between two points, rather than the absolute increase in the local turbulent velocity amplitude with distance from a common center. If the medium is gravitationally stratified, however, the central point has a physical significance, and it is not currently known whether in this situation $q$ varies significantly (locally or globally) from its value in a non-stratified medium. Observations of individual low-mass cores indicate increasing linewidths away from the centers, with $q \simeq \frac{1}{2}$ on large scales (Goodman et al., 1998); similar observations for star-forming clumps or high-mass cores, which are supersonically turbulent, are not yet available. In polytropic models with $\sigma_{\text{nt}} \propto \rho^{(\gamma_p-1)/2} \propto r^q$, the density follows a power-law in radius, $\rho \propto r^{-k_p}$, with $k_p = 2q/(1 - \gamma_p)$. In hydrostatic equilibrium, $k_p = 2/(2 - \gamma_p)$ must hold, so that $q = (1 - \gamma_p)/(2 - \gamma_p)$; the value $q = 1/2$ thus corresponds to $\gamma_p \to 0$. This has motivated the study of equations of state for turbulent gas that include a pressure proportional to the logarithm of the density, so-called “logatropes” (Lizano & Shu, 1989; Gehman et al., 1996). McLaughlin & Pudritz (1996) pointed out a difficulty with previous logatropic models and developed a variant that overcame this problem; however, their model leads to linewidths that actually decrease near the edge of the cloud (McKee & Tan, 2003). An alternative model for clouds in which the inner regions are supported by thermal pressure and the envelopes are supported by turbulent pressure is the “TNT” (thermal/nonthermal) model, in which the density is assumed to be given by the sum of two power laws, one with $k_p = 2$, representing a singular isothermal sphere, and one with $k_p < 2$, representing a turbulent envelope (Myers & Fuller, 1992; Caselli & Myers, 1995). A more rigorous formulation of this type of model is that of a composite polytrope, in which the core and envelope of the cloud have different values of $\gamma_p$ (Curry & McKee, 2000).

Cosmological simulations show that self-gravitating, pressureless matter condenses into filamentary structures (e.g., Springel et al., 2005). This reflects the
nature of evolution of cold, triaxial mass distributions under self-gravity (Lin, Mestel, & Shu, 1965): the first collapse is along the shortest axis, and the second collapse is along the (original) intermediate axis, resulting in a filament aligned along the (original) long axis. Molecular clouds often exhibit filamentary structure as well (Schneider & Elmegreen, 1979; Mizuno et al., 1995; Nagahama et al., 1998; Lada, Alves, & Lada, 1999). This may reflect the effects of self-gravitational evolution, similar to cosmic structure formation. However, it may also reflect the effects of strongly supersonic turbulence. Converging turbulent flows produce curved sheets of shocked gas at stagnation surfaces, and the loci of these sheet intersections are filaments. The morphology of cold, diffuse HI is similar to that in GMCs (e.g. Heiles & Troland 2003; McClure-Griffiths et al. 2006), suggesting that at least some of the filamentary structure in star-forming clouds originates with multi-scale supersonic turbulence; the filaments that are created by turbulent flows may also be (or become) self-gravitating.

Virial balance in filamentary clouds implies $GM/\ell = Gm_\ell \sim \sigma^2$, where $\ell$ is a length along the filament and $m_\ell$ is the mass per unit length. Fiege & Pudritz (2000) have shown that the virial gravity term for a cylindrical cloud with an arbitrary density profile is $W_\ell = -Gm_\ell^2$, which, in contrast to the spherical case, is unchanged by radial compression. They also showed that the critical mass/length is $m_{\ell,cr} = 2\sigma^2/G$. Filaments with $\gamma = \gamma_p \geq 1$ are stable against compression, since the ratio of kinetic energy to gravitational energy does not decrease during compression. Isothermal filaments have a density $\rho = \rho_0/[1 + (r/r_0)^2]^2$, where $r_0 = (2/\pi)^{1/2}R_G(\rho_c)$ (Ostriker, 1964: the properties of filaments with $\gamma_p \geq 1$ are not strongly affected by the ambient medium for $r \gg r_0$, which is why their properties are determined by the central density $\rho_c$.) However, observed filaments often have $\rho \propto 1/r^2$ rather than $1/r^4$ (Lada, Alves, & Lada, 1999). Fiege & Pudritz (2000) have shown that isothermal filaments with helical magnetic fields of the right magnitude can give rise to such a density profile; alternatively, Nakamura & Umemura (2002) have shown that negative index polytropes with $\gamma_p$ slightly less than unity have $\rho \propto 1/r^2$.

In general, masses in excess of the critical mass are subject to fragmentation. In an isothermal, uniform medium, the minimum wavelength for gravitational fragmentation is the Jeans length,

$$\lambda_J = \left(\frac{\pi \sigma_{th}^2}{G\rho_0}\right)^{1/2} = \pi^{1/2}R_G(\rho_0).$$

The corresponding Jeans mass is $M_J \equiv (4\pi/3)(\lambda_J/2)^3\rho = 2.47M_{BE}$, where we have adopted the definition of Binney & Tremaine (1987) (the Jeans mass is elsewhere often defined as $\rho\lambda_J^3 = 6M_J/\pi$, which is even larger than the Bonnor-Ebert mass). For slabs and filaments, there is a fastest growing mode, which will determine the spacing of fragments. For an isothermal slab with surface density $\Sigma$ and with $\rho_c \gg \rho_0$, it is $\lambda_{max,slab} \approx 2^{3/2}\lambda_J(\rho_c) = 4\sigma_{th}^2/(G\Sigma)$, where the Jeans length is defined in terms of the midplane density $\rho_c$. An isothermal filament with $\rho_c \gg \rho_0$ has $m_{\ell} \approx m_{\ell,cr}$ and $\lambda_{max,fil} \approx 1.25 \times 2^{3/2}\lambda_J(\rho_c)$ (Larson, 1985). These estimates assume that the gas is optically thin; if it becomes opaque, fragmentation stops. Low & Lynden-Bell (1976) show that fragmentation ceases for masses $\gtrsim 0.004M_\odot$ (including He—see Whitworth et al. 2007), and that this is relatively insensitive to parameters.

For a thin, rotating disk, rotation stabilizes self-gravitational contraction for
wavelengths greater than the Toomre length $\lambda_T \equiv 4\pi^2 G\Sigma/\kappa^2$, where $\kappa$ is the epicyclic frequency (Toomre, 1964). In order for a rotating gas disk to fragment, the maximum instability scale imposed by angular momentum considerations must exceed the minimum length for fragmentation set by thermal pressure. The Toomre parameter $Q \equiv \kappa \rho_{th}/(\pi G \Sigma) = (\lambda_{\text{max, slab}}/\lambda_T)^{1/2}$, and must be $\gtrsim 0.7 - 1$ for gravitational fragmentation in an isothermal rotating disk, depending on the strength of magnetic fields (Goldreich & Lynden-Bell, 1965; Kim, Ostriker, & Stone, 2002). Allowing for turbulence and for the additional gravity of a stellar disk (for large-scale galactic instabilities), the critical $Q$ is larger (see §3.2.1). Real gases are not strictly isothermal; Gammie (2001) has shown that in a Keplerian disk, the cooling time $t_{\text{cool}}$ and angular velocity $\Omega = \kappa$ must satisfy the condition $t_{\text{cool}} \lesssim 3\Omega^{-1}$ for gravitational runaway to occur (the coefficient 3 is based on 2D simulations with $\gamma_{\text{eff}} = 2$, allowing for dimensional reduction; Rice, Lodato, & Armitage (2005) showed that this coefficient can change by a factor of a few depending on the adopted $\gamma$). Nonlinear instability develops when $Q$ is small even for adiabatic disks, but gravitational collapse of the condensations that form is ultimately halted if $\gamma$ is sufficiently large (Kim & Ostriker, 2001).

### 2.3 Magnetic Fields

The interstellar medium is strongly magnetized, whereas stars are weakly magnetized. How the mass-to-flux ratio increases so dramatically during star formation is one of the classic problems of star formation (Mestel & Spitzer, 1956). We shall characterize this ratio by the ratio of the mass to the magnetic critical mass for poloidal fields, $\mu_\Phi \equiv M/M_\Phi$ (eq. 16, §2.2). Heiles & Troland (2005) found that the median field in the cold H I phase of the ISM (the CNM) is $|B_0| = 6.0 \pm 1.8 \mu$G, and that the CNM is organized into sheets with column densities $2.6 \times 10^{18}$ cm$^{-2}$ $\lesssim N_H$ $\lesssim 2.6 \times 10^{20}$ cm$^{-2}$; the maximum column presumably reflects the transition to molecular hydrogen. It follows that the CNM is magnetically very subcritical, $\mu_\Phi < 0.16$ (throughout this section, we evaluate $M_\Phi$ with $c_\Phi = 1/2\pi$, the value appropriate for sheets). There are thus two parts to the magnetic flux problem: How does the mass-to-flux ratio increase to $\mu_\Phi \gtrsim 2$ so that gravitational collapse can readily occur, and then how does it increase to the very large values ($\sim 10^{5-8}$) characteristic of stars?

Astronomers have two primary methods to measure the strength of magnetic fields in the dense ISM: the Zeeman effect, which measures the line-of-sight component, $B_{\text{los}}$; and the Chandrasekhar-Fermi method (Chandrasekhar & Fermi, 1953a), which measures the component of the field in the plane of the sky, $B_{\text{pos}}$, by comparing the fluctuations in the direction of $\mathbf{B}_{\text{pos}}$ with those in the velocity field (see the reviews by Crutcher, 2005 and Heiles & Crutcher, 2005; note that in the diffuse ISM, magnetic field strengths are also obtained by Faraday rotation and synchrotron observations, with results consistent with Zeeman observations.) The morphology of the field, which is needed for the Chandrasekhar-Fermi method, can be measured from dust polarization and from linear polarization of spectral lines (Goldreich & Kylafis, 1981). The largest compilation of magnetic field strengths in molecular clouds remains that of Crutcher (1999), although it must be noted that the median temperature of the regions with detected fields is 40 K, significantly greater than average. Inferring the intrinsic field strength and column density from measurements of the line-of-sight components is somewhat subtle (Heiles & Troland, 2005); in particular, care must
be exercised in evaluating the average value of $B_\parallel/B$ using logarithmic values, since $(\log B_\parallel/B) < \log(B_\parallel/B)$. However, it is straightforward to infer the median values: $B_{\text{med}} = 2B_{\text{los,med}}$ and, for sheets, $N_{\text{med}} = N_{\text{los,med}}/2$. Most of the structures Crutcher (1999) studied are relatively dense cores, so it is plausible that they are not sheet-like; in that case, the median value of $\mu_\phi$ is 1.65 $\pm$ 0.2 (Heiles & Crutcher, 2005; note that the values in Crutcher 1999 are based on $c_\phi = 0.12$, whereas we are using $c_\phi = 1/2\pi$), the cores are supercritical and the magnetic field is unable to significantly impede gravitational collapse. On the other hand, if the objects in his sample are in fact sheet-like, then the median value of $\mu_\phi$ is reduced to 0.8, and the typical core is about critical. However, it should be noted that none of the cores have observed line-of-sight fields strong enough to ensure that they are subcritical, and for many cases only upper limits on the magnetic field strength are obtained. Subsequent OH Zeeman observations by Bourke et al. (2001) have strengthened these conclusions, although these authors suggest that observations with higher spatial resolution are needed to determine whether the relatively low fields they infer (mostly upper limits) are in part due to variations in the field structure within the telescope beam.

Crutcher (1999) also reached a number of other conclusions on the role of magnetic fields in cores and clumps within molecular clouds from his sample: the observed structures are in approximate virial equilibrium; the kinetic and magnetic energies are in approximate equipartition, as expected theoretically (Zweibel & McKee, 1995); correspondingly, the Alfvén Mach number is $M_A \simeq 1$; the observed motions are highly supersonic, with $M_s \equiv \sqrt{3}\sigma_{\text{nt}}/c_s \simeq 5$; and, to within the errors, $B \propto \rho^{1/2}$, which corresponds to a constant Alfvén velocity (for sources with measured fields, as opposed to upper limits, the average value is $v_A \simeq 2$ km s$^{-1}$, as found previously by Heiles et al., 1993). Basu (2000) showed that the dispersion of the Alfvén Mach number is significantly less than that in the Alfvén velocity in this sample. He argued that a constant value of $M_A$ is to be expected if the clouds are strongly bound, so that the surface pressure is negligible, and if $\mu_\phi$ is about constant. Adopting a median value $M_A = 1.0$ from Crutcher (1999) gives a median value for the magnetic field of

$$B_{\text{med}} \simeq 30 \left( \frac{n_H}{10^3 \text{ cm}^{-3}} \right)^{1/2} \left( \frac{\sigma_{\text{nt}}}{1 \text{ km s}^{-1}} \right) \mu G \quad (n_H \gtrsim 2 \times 10^3 \text{ cm}^{-3}). \quad (21)$$

The value of the density in this relation is $n_H/(4R/3)$, where $R$ is the mean projected radius. Projection effects could cause the actual density to differ from this, but the change is not large for triaxial clouds of the type considered by Basu (2000).

As yet, observations of the mass-to-flux ratio on large scales, up to that of GMCs, are not available. The definition of $\mu_\phi$ implies

$$B = \frac{C^{1/2}\Sigma}{\mu_\phi c_\phi} = 3.80 \left( \frac{N_{H,21}}{\mu_\phi} \right) \frac{A_V}{\delta_{\text{gr}} \mu_\phi} \mu G, \quad (22)$$

where the numerical evaluations are based on $c_\phi = 1/2\pi$, the visual extinction is $A_V = N_H\delta_{\text{gr}}/(2 \times 10^{21} \text{ cm}^{-2})$, and $\delta_{\text{gr}}$ is the dust-to-gas ratio normalized to the local interstellar value. Typical Galactic GMCs have $N_H = 1.5 \times 10^{22}$ cm$^{-2}$ (see §3.1), corresponding to critical magnetic field strength (i.e., such that $\mu_\phi = 1$) of $B_{\text{cr}} = 57\mu G$ for the large-scale mean field. In regions with densities $n_H \approx 2 \times 10^3$ cm$^{-3}$, the lowest for which molecular-line Zeeman observations are available,
Crutcher (1999) reports line-of-sight magnetic field strengths of $\leq 21\mu G$. Allowing for an increase of up to a factor of two for projection effects, and for the fact that the mean magnetic field strength will not increase as the density is reduced by a factor of $\sim 10$ to reach the volume-averaged value in GMCs, we infer that GMCs are supercritical. GMC magnetic fields are not too weak, however: 450 $\mu$m polarimetry of four GMCs shows that the orientation of the field appears to be preserved during the formation of the GMCs and that the energy in the field is comparable to the turbulent energy (Li et al., 2006).

Theoretical arguments are consistent with the empirical evidence that GMCs as well as their sub-parts are supercritical with respect to their mean magnetic fields. Models of self-gravitating, isothermal, magnetized clouds show that large pressure contrasts between the center of the cloud and the edge occur only when the cloud is near its critical mass; furthermore, if the kinetic energy is comparable to the magnetic energy, then large pressure contrasts occur only for $M > M_\phi$. Extending the earlier work of Mouschovias (1976), Tomisaka, Ikeuchi, & Nakamura (1988) found that for the cases they considered with $8\pi P/B^2 = 1$, the central pressure significantly exceeds the surface pressure only when $M$ is quite close to $M_{cr}$, and that in these cases the cloud is magnetically supercritical, $M > M_\phi$. Using this work, McKee (1989) showed that the critical mass is $M_{cr} \simeq M_\phi + M_{BE}$ for quiescent clouds; he assumed that this relation applies to turbulent clouds as well, with $\sigma_{th}$ replaced by the total velocity dispersion $\sigma$, but the validity of this assumption remains to be demonstrated. Since, on large scales, the turbulent magnetic energy is likely comparable to or larger than the mean magnetic energy, and the kinetic energy is at least as large as the magnetic energy (and much greater than the thermal energy), then clouds with $\alpha_{vir} \sim 1$ have $M \gtrsim 2M_\phi$. Nakano (1998) has given a similar, more precise argument that the smaller-scale and less-turbulent cores that form stars are also magnetically supercritical. In both cases, the basic argument is that if gravity is strong enough to overcome both kinetic energy (turbulent plus thermal) and magnetic fields (turbulent and ordered) in order to form a bound object, then it is certainly strong compared to the support from mean magnetic fields alone. Note that this argument does not apply to objects that are not bound, but instead are the result of colliding flows ($\S$3.2.1). It is of great importance to determine observationally the relative importance of magnetic fields and gravity in the large scale structure of molecular clouds.

Most detailed modeling of magnetic fields in non-turbulent clouds is based on the assumption that the field is poloidal (e.g., Mouschovias, 1987); such fields always tend to support clouds against gravity. On the other hand, the toroidal component of a helical field exerts a confining force, and can lead to prolate clouds (Tomisaka, 1991; Fiege & Pudritz, 2000). From a virial analysis of several filamentary clouds, Fiege & Pudritz (2000) find that the self-gravity and the pressure of the ambient medium are inadequate to account for the high mean pressures that are observed in the clouds; they conclude that the data can be explained if these clouds are confined by helical fields. The principal uncertainty in this analysis is that in most cases there is no direct measurement of the ambient pressure.

There are two mechanisms for increasing the mass-to-flux ratio, flows along magnetic fields and ambipolar diffusion. In the part of a bound molecular cloud that is shielded from the interstellar radiation field so that the ionization is due to cosmic rays, the ambipolar diffusion time is about 10 times the free-fall time in
the absence of turbulence (Mouschovias, 1987) and several times faster than this in the presence of turbulence (Zweibel, 2002; Fatuzzo & Adams, 2002; Nakamura & Li, 2005). However, most of the mass of a GMC is ionized primarily by FUV radiation from stars (McKee, 1989), and in this gas the ambipolar diffusion time is much longer. GMCs are very porous, and as a result an even larger fraction of the volume of the cloud is likely to be ionized above the level set by cosmic rays. It follows that flux-freezing is a good approximation on large scales in molecular clouds. Mestel (1985) introduced the concept of the “accumulation length,” \( L_0 \), the size of the region required to achieve a given mass-to-flux ratio when flux-freezing applies, \( \mu_\Phi = M/M_\Phi \propto n_0 L_0/B_0 \). In our notation this yields

\[
L_0 = \left( \frac{c_\Phi B_0}{\mu_\Pi G^{1/2} n_0} \right) \mu_\Phi = 85 \left( \frac{\mu_\Phi B_0 \mu_\Phi^{-6}}{n_0} \right) \text{ pc, \hspace{1cm} (23)}
\]

where \( B_0 \mu_\Phi \equiv B_0/(1 \mu G) \). Using \( n_0 \sim 1 \text{ cm}^{-3} \) for the mean density in the diffuse ISM (since GMC columns are much greater than those of individual CNM clouds) and \( B_0 \mu_\Phi \sim 6 \) for the mean field in the solar vicinity (since the mean field should be similar to the CNM field—Piontek & Ostriker, 2005) gives a large value for this length, \( \sim 1 \text{ kpc} \) if GMCs have \( \mu_\Phi \sim 2 \). In fact, as we shall discuss in §3.2.1, GMC formation from large-scale self-gravitating galactic disk instabilities indeed involves very large accumulation lengths, and yields supercritical clouds.

As discussed above, current observations do not determine whether ambipolar diffusion is necessary for the initiation of gravitational collapse. Theoretical simulations suggest that in the absence of ambipolar diffusion, star formation is strongly suppressed in magnetically subcritical regions, even if \( \mu_\Phi \) is only slightly less than unity (Krasnopolsky & Gammie, 2005). However, similar simulations show that magnetic fields have a relatively small effect in slowing the rate of star formation if the gas is supercritical (Ostriker, Gammie, & Stone, 1999; Heitsch, Mac Low, & Klessen, 2001; Li et al., 2004; Vázquez-Semadeni et al., 2005; Vázquez-Semadeni, Kim, & Ballesteros-Paredes, 2005; Nakamura & Li, 2005). The primary effect of magnetic fields may be to shift the initial collapse to higher masses. Simulations with ambipolar diffusion in weakly ionized plasmas are very challenging. In the strong-coupling approximation, in which the ions are not treated as a separate fluid but the field diffuses relative to the flow, explicit MHD codes have time steps \( \propto \Delta x^2 \), which is prohibitive at high resolution (Mac Low et al., 1995). If the ions are treated as a separate fluid, explicit codes must resolve Alfvén waves in the ions as well as Alfvén waves in the coupled ion-neutral fluid. For ionizations \( \lesssim 10^{-6} \), the Alfvén velocity in the ions can exceed 10\(^4\) km s\(^{-1}\), leading to very small time steps. A potential way around this problem is to increase the ion mass and decrease the ion-neutral coupling constant so that the momentum exchange rate between the ions and neutrals is unchanged (Oishi & Mac Low, 2006; Li, McKee, & Klein, 2006).

One regime in which ambipolar diffusion (or the lack of it) could have a strong effect on the star formation rate is in the outer layers of GMCs, which are dominated by FUV ionization. As remarked above, these regions constitute the bulk of the mass of a GMC, as much as \( \sim 90\% \). FUV photoionization slows ambipolar diffusion, and therefore star formation, when it dominates cosmic-ray ionization, which occurs for visual extinctions \( A_V \lesssim 4 \text{ mag} \) from the surface or \( \sim 8 \text{ mag} \) along a line of sight through the cloud (McKee, 1989). Suppression of star formation in the outer layers of GMCs has been confirmed in the L1630 region of Orion (Li, Evans, & Lada, 1997) and in Taurus (Onishi et al., 1998). To the extent
that ambipolar diffusion is essential for forming molecular cores, the absence or near absence of sub-mm cores in the outer layers of Ophiuchus (Johnstone, Di Francesco, & Kirk, 2004) and Perseus (Enoch et al., 2006; Hatchell et al., 2005) is qualitatively consistent with this prediction. On the other hand, Strom, Strom, & Merrill (1993) find that there is a substantial distributed population of young stars in L1641, although this population is relatively old (5 – 7) Myr.

### 2.3.1 Ionization

The chemistry of molecular clouds is a full subject in its own right. Here we summarize several developments that affect the ionization, which governs the coupling between the gas and the magnetic field. (1) Photodissociation regions (PDRs) are regions of the ISM that are predominantly neutral and in which the chemistry and heating are predominantly due to far-UV radiation (see the review by Hollenbach & Tielens, 1999). Most of the non-stellar infrared radiation and most of the millimeter and submillimeter CO emission in galaxies originates in PDRs. In the typical interstellar radiation field, phot ionization dominates ionization by cosmic rays for extinctions \( A_V < 4 \) mag, which includes most of the molecular gas in the Galaxy (McKee, 1989). (2) Polycyclic aromatic hydrocarbons (PAHs), which contain a few percent of the carbon atoms, often dominate the mid-IR spectrum of star-forming regions and galaxies. It is frequently assumed that PAHs have a low abundance in molecular clouds due to accretion onto dust grains; if this is not the case, they can dominate the ionization balance, since electrons react with them very rapidly (Lepp et al., 1988). (3) \( \text{H}_3^+ \) is a critical ion in initiating ion-molecule reactions in molecular clouds. For many years, the rate of dissociative recombination, \( \text{H}_3^+ + e \rightarrow \text{H}_2 + \text{H} \) or \( \text{H} + \text{H} + \text{H} \), was uncertain, but careful laboratory experiments have shown that the rate coefficient for this reaction is large: a fit to the results of McCall et al. (2003) gives \( \alpha_d(\text{H}_3^+) = 4.0 \times 10^{-7}(T/10 \text{ K})^{-0.52} \text{ cm}^3 \text{ s}^{-1} \). In order to maintain the observed abundance of \( \text{H}_3^+ \) in the face of this high recombination rate, these authors inferred a very high cosmic-ray ionization rate, \( \zeta_{\text{CR}} \approx 6 \times 10^{-16} \text{ s}^{-1} \text{ per H atom (including secondary ionizations), in a diffuse molecular cloud along the line of sight to \( \zeta \) Per. Models involving two gas phases give somewhat lower values of \( \zeta_{\text{CR}} \) (e.g., Dalgarno, 2006), but the correct value is now quite uncertain. In dense clouds, Dalgarno (2006) concludes that the ionization rate is \( \zeta_{\text{CR}} \approx 2.5 - 5 \times 10^{-17} \text{ s}^{-1} \).

(4) Recent observations have established that carbon-bearing molecules freeze out onto dust grains at high densities \( n_H \sim 10^5 \text{ cm}^{-3} \) in low-mass cores, with nitrogen-bearing molecules freezing out at higher densities (di Francesco et al., 2007). This affects the ionization, since it removes abundant ions such as \( \text{HCO}^+ \) from the gas.

Although the chemistry determining the ionization in molecular clouds is complex, simple analytic estimates are possible. In the outer layers of PDRs, carbon is photoionized so that \( n_e \approx n(\text{C}) \). In regions ionized by cosmic rays, the degree of ionization is given by

\[
x_e = \frac{n_e}{n_H} \approx \left( \frac{\zeta_{\text{CR}}}{\alpha n_H} \right)^{1/2}
\]

if the ionization is dominated by molecular ions (including PAHs), where \( \alpha \) is the relevant recombination rate in the chemistry that determines the ionization fraction. If PAHs are depleted, then \( \alpha \approx 10^{-6} \text{ cm}^3 \text{ s}^{-1} \) is the dissociative recombination rate for heavy molecules provided the density is high enough that \( \text{H}_3^+ \) is destroyed primarily by reactions with such molecules; for lower densities, where the ionization is dominated by \( \text{H}_3^+ \), one has \( \alpha = \alpha_d(\text{H}_3^+) \). If PAHs...
are sufficiently abundant that most of the electrons are attached to PAHs, then \( \alpha \simeq 3 \times 10^{-7} \text{ cm}^3 \text{ s}^{-1} \) (Tielens, 2005) and \( n_e \) in equation (24) includes the electrons attached to PAHs. Metal ions can be readily included in the analytic theory (McKee, 1989), but they do not appear to be important in dense cores (Maret, Bergin, & Lada, 2006). Equation (24) is consistent with the results of Padoan et al. (2004) at late times and at high densities for \( \alpha \simeq \alpha_d(\text{HCO}^+) \), which they took to be \( 2.5 \times 10^{-6} \text{ cm}^3 \text{ s}^{-1} \).

3 MACROPHYSICS OF STAR FORMATION

3.1 Physical State of GMCs, Clumps, and Cores

The molecular gas out of which stars form is found in molecular clouds, which occupy a small fraction of the volume of the ISM but, inside the solar circle, comprise a significant fraction of the mass. The terminology for the structure of molecular clouds is not fixed; here we follow the discussion in Williams, Blitz, & McKee (2000). Giant molecular clouds (GMCs) have masses in excess of \( 10^4 M_\odot \) and contain most of the molecular mass. Molecular clouds have a hierarchical structure that extends from the scale of the cloud down to the thermal Jeans mass in the case of gravitationally bound clouds, and down to much smaller masses for unbound structures (Langer et al., 1995; Heithausen et al., 1998). Overdense regions (at a range of scales) within GMCs are termed clumps. Star-forming clumps are the massive clumps out of which stellar clusters form, and they are generally gravitationally bound. Cores are the regions out of which individual stars (or small multiple systems like binaries) form, and are necessarily gravitationally bound. As remarked above, this terminology is not universal; e.g., Ward-Thompson et al. (2007) use “pre-stellar core” to refer to a core, and “cluster-forming core” to refer to a star-forming clump.

A molecular cloud is surrounded by a layer of atomic gas that shields the molecules from the interstellar UV radiation field; in the solar vicinity, this layer is observed to have a column density \( N_H \simeq 2 \times 10^{20} \text{ cm}^{-2} \), corresponding to a visual extinction \( A_V = 0.1 \text{ mag} \) (Bohlin, Savage, & Drake, 1978). A larger column density, \( N_H \simeq 1.4 \times 10^{21} \text{ cm}^{-2} \), is required for CO to form (van Dishoeck & Black, 1988). The layer of gas in which the hydrogen is molecular but the carbon is atomic is difficult to observe, and has been termed “dark gas” (Grenier, Casandjian, & Terrier, 2005).

The mass of a molecular cloud is generally inferred from its luminosity in the \( J = 1 \rightarrow 0 \) line of \(^{12}\text{CO}\) or \(^{13}\text{CO}\). Because \(^{12}\text{CO}\) is optically thick, estimating the column density of \( \text{H}_2 \) molecules from the \(^{12}\text{CO}\) line intensity \( I_{\text{CO}} \) (in units of \( \text{K km s}^{-1} \)) requires multiplication by an “X-factor,” an appropriate name since it is not well understood theoretically; this is defined as \( X = N(\text{H}_2)/I_{\text{CO}} \). Various methods have been used to infer the value of \( X \) in the Galaxy. In one method, observations of \( \gamma \)-rays emitted by cosmic rays interacting with the ISM give the total amount of interstellar matter; the mass of molecular gas follows by subtracting the neutral atomic hydrogen (H I) contribution. With this technique, Strong & Mattox (1996) infer \( X = 1.9 \times 10^{20} \text{ cm}^{-2} (\text{K km s}^{-1})^{-1} \). In another method, subtracting the H I-associated dust emission from the total observed dust emission in the infrared gives a local value \( X = 1.8 \times 10^{20} \text{ cm}^{-2} (\text{K km s}^{-1})^{-1} \) (Dame, Hartmann, & Thaddeus, 2001). Note that both of these methods account for all the molecular hydrogen gas, including the “dark gas.” Allowing for the
atomic shielding layer around a molecular cloud (but not the “dark-gas” layer), Elmegreen (1989) predicted \( X \propto (G_0/Z)^{3/8}/T_b \), where \( G_0 \) is proportional to the intensity of FUV radiation that can photodissociate \( \text{H}_2 \), \( Z \) is the metallicity, and \( T_b \) is the brightness temperature of the line. Maloney & Black (1988) concluded that \( T_b \) should be substantially reduced in regions of low metallicity. Observing the \( ^{13}\text{CO} \) line is advantageous in that it is optically thin in all but the high-density cores. Conversion from \( ^{13}\text{CO} \) intensity to column density involves an assumption of LTE (using temperatures derived from \( ^{12}\text{CO} \)), and a fixed \( \text{H}_2 \) to \( ^{13}\text{CO} \) abundance. Because \( ^{13}\text{CO} \) may be subthermally excited in diffuse regions, however, column densities there will be underestimated. Near-infrared extinction mapping (Lada, Alves, & Lombardi, 2007) offers the prospect of obtaining more accurate masses, at least for nearby molecular clouds.

The observed mass distribution of GMCs is a power law with a relatively sharp cutoff. Let \( dN_c(M) \) be the number of GMCs with masses in the range \( M \) to \( M + dM \). Observations of GMCs inside the solar circle (but excluding the Galactic Center) are consistent with the mass distribution of the form (Williams & McKee, 1997)

\[
\frac{dN_c}{d \ln M} = N_{cu} \left( \frac{M_u}{M} \right)^\alpha \quad (M \leq M_u),
\]

with no GMCs above \( M_u \). Here \( N_{cu} / \alpha \) is equal to the number of clouds eliminated from the distribution by the cutoff at \( M_u \). With \( N_{cu} = 63 \), \( \alpha = 0.6 \), and \( M_u = 6 \times 10^6 \text{M}_\odot \), this cloud mass distribution accounts for all the molecular mass observed inside the solar circle excluding the Galactic Center. An independent analysis by Rosolowsky (2005) finds a similar slope (\( \alpha = 0.5 \pm 0.1 \)), but a somewhat smaller maximum mass (\( M_u = 3 \times 10^6 \text{M}_\odot \), although this does not include the several most massive clouds). These results are necessarily approximate due to the difficulties in identifying clouds from position-velocity data in the inner Galaxy—in particular, blending of clouds along the line of sight is likely to make the true value of the slope steeper (Rosolowsky, 2005). However, the main implications are likely to be robust. First, most of the mass in GMCs is in large clouds—a significant fraction is in clouds with \( M > 10^6 \text{M}_\odot \), and > 80% is in clouds with \( M > 10^5 \text{M}_\odot \) (see also Stark & Lee, 2006). And second, since \( N_{cu} \gg 1 \), the upper limit of the mass distribution, \( M_u \), has a physical significance (McKee & Williams, 1997). If there were no cutoff to the distribution, one would expect about 100 GMCs more massive than \( 6 \times 10^6 \text{M}_\odot \) in the Galaxy, whereas there are none. This upper mass limit may be set by the processes that form GMCs out of diffuse gas (see §3.2.1). It should be noted that the GMCs are embedded in more massive HI “superclouds” (sometimes encompassing multiple GMCs), which also appear to be gravitationally bound (Elmegreen & Elmegreen, 1987). In other Local Group galaxies, GMC mass distributions have similar power laws to that in the Milky Way, with the exception of M33, which has \( dN_c/d \ln M \propto M^{-1.5} \) (Blitz et al., 2007). In more distant galaxies, giant molecular associations (GMAs) with masses up to \( \sim 10^7 \text{M}_\odot \) have been observed (Vogel, Kulkarni, & Scoville, 1988; Sakamoto et al., 1999).

### 3.1.1 Dynamics of GMCs

In a seminal paper, Larson (1981) summarized some of the key dynamical features of GMCs in what are often referred to as “Larson’s laws.” The first result is that GMCs obey a line width–size relation: GMCs are supersonically turbulent with velocity dispersions that increase as a power of the size. For GMCs in the first Galactic quadrant, almost all of which
are inside the solar circle, Solomon et al. (1987) found
\[ \sigma = (0.72 \pm 0.07) \frac{R}{(1 \text{ pc})^{0.5 \pm 0.05}} \text{ km s}^{-1}, \]  
where \( R_{\text{pc}} \equiv R/(1 \text{ pc}) \). (Note that this value for \( \sigma_{\text{pc}} \) is based on a distance to the Galactic Center of 10 kpc; we have not adjusted this value for a more accurate distance since the change is within the errors.) Heyer & Brunt (2004) find that this cloud-to-cloud relation extends to the structure functions within individual GMCs as well (see eq. 6 in §2.1.5), and argue that this demonstrates the universality of the turbulence in moderate-density gas in molecular clouds (see below).

Larson’s second law is that GMCs are gravitationally bound (\( \alpha_{\text{vir}} \simeq 1 \); see eq. 11). (It should be noted that while molecular gas is generally bound in the Galaxy, different physical conditions can lead to substantial amounts of unbound molecular gas or bound atomic gas—Elmegreen, 1993b.) Solomon et al. (1987) determined the masses of clouds in their sample using the virial theorem with \( \alpha_{\text{vir}} = 1.1 \), and then determined the \( X \)-factor. Including He and adjusting the distance to the Galactic Center to 8.5 kpc from 10 kpc, their value of the \( X \)-factor corresponding to a typical GMC with a mass of \( 10^6 M_\odot \) (see above) is \( 1.9 \times 10^{20} \text{ cm}^{-2} (\text{K km s}^{-1})^{-1} \), the same as the value determined from \( \gamma \)-ray observations; hence, the GMCs in their sample are gravitationally bound on average. (Note that this argument is approximate since the \( \gamma \)-ray value for the \( X \)-factor includes the “dark gas,” whereas the value from the virial theorem includes only part of this gas, depending on the morphology of the GMC.) Observations of \( ^{13}\text{CO} \), which is optically thin, permit a direct measurement of the mass, provided the abundance is known. Such observations of a sample of GMCs in the outer Galaxy, where blending of different clouds along the line of sight is negligible, confirm that molecular clouds with \( M \gtrsim 10^4 M_\odot \) (i.e., GMCs) are bound (Heyer, Carpenter, & Snell, 2001). Lower mass clouds become progressively less bound (see also Maloney, 1990), and unbound molecular clouds are for many purposes equivalent to non-self-gravitating clumps within larger GMCs. Based on observations in \( ^{13}\text{CO} \) and other species that are believed to be optically thin, clump mass functions within GMCs follow \( dN/d\ln M \propto M^{-\alpha_{\text{clump}}} \) with \( \alpha_{\text{clump}} = 0.3 - 0.7 \) (Blitz, 1993; Williams, de Geus, & Blitz, 1994). The slope of the clump mass function is similar to that for GMCs as a whole (see eq. 25), possibly because both are determined by turbulent processes within larger, gravitationally-bound systems.

Larson’s third law is that GMCs all have similar column densities. For the Solomon et al. (1987) sample, the mean column density is \( \bar{N}_H = (1.5 \pm 0.3) \times 10^{22} R_{\text{pc}}^{0.0 \pm 0.1} \text{ cm}^{-2} \); this corresponds to an extinction \( A_V = 7.5 \text{ mag} \) with the local dust-to-gas ratio. The corresponding mean surface density of GMCs is \( \Sigma = 170 M_\odot \text{ pc}^{-2} \). However, GMCs in the outer Galaxy are observed to have smaller column densities (Heyer, Carpenter, & Snell, 2001), in part because of the greater sensitivity of these observations.

As Larson pointed out, these three relations are not independent; any two of them imply the third. Indeed, if we express the line width—size relation as \( \sigma = \sigma_{\text{pc}} R_{\text{pc}}^{1/2} \), then

\[ \alpha_{\text{vir}} = \left( \frac{5}{\pi R_{\text{pc}}} \right) \frac{\sigma_{\text{pc}}^2}{G \Sigma} = 3.7 \left( \frac{\sigma_{\text{pc}}}{(1 \text{ km s}^{-1})} \right)^2 \left( \frac{100 M_\odot \text{ pc}^{-2}}{\Sigma} \right) \]  
(27)
relates the three scaling laws. Observations supporting a “universal turbulence law” (eq. 6) in the Galaxy, and thus small differences in $\sigma_{\text{pc}}$ between inner- and outer-Galaxy GMCs (Heyer & Brunt, 2004; Heyer, Williams, & Brunt, 2006), then imply that the value of $\Sigma$ is about the same for all GMCs with $\alpha_{\text{vir}} \sim 1$, regardless of Galactic location. Observational confirmation of this conclusion would be valuable. Provided Larson’s Laws apply, the mean kinetic pressure within GMCs is independent of mass and size, and is given by $P_{\text{kin}} = \frac{\rho \sigma^2}{3} = 3\Sigma \sigma_{\text{pc}}^2/(4 \text{pc})$. For inner-Galaxy GMCs, this is $P_{\text{kin}}/k_B \approx 3 \times 10^9$ K cm$^{-3}$.

These results can also be expressed in terms of the sonic length $\ell_s$ (see §2.1.3), since $\sigma_{\text{pc}} = c_s(2/3 \ell_{s, \text{pc}})^{1/2}$ if $\sigma \approx \sigma_{\text{int}} \gg c_s$, and $\ell = 2R$. Gravitationally bound objects ($\alpha_{\text{vir}} \sim 1$) that obey a line width–size relation with an exponent $\simeq 1/2$ necessarily have surface densities $\Sigma = \left(10/3\pi\alpha_{\text{vir}}\right)c_s^2/(G\ell_s) \sim c_s^2/(G\ell_s)$. The small observed variation in $\Sigma$ for the set of inner-Galaxy GMCs is then equivalent to a small variation in $\ell_s$ in those clouds (since $c_s$ is observed to be about constant). In terms of $\ell_s$, the mean kinetic pressure in GMCs is $P = \Sigma c_s^2/(2\ell_s)$.

Do Larson’s laws apply in other galaxies? Blitz et al. (2007) summarize observations of GMCs in galaxies in the Local Group, in which the metallicity varies over the range (0.1 – 1) solar. They find that the GMCs in most of these galaxies would have luminous masses within a factor two of their virial masses if $X = 4 \times 10^{20}$ cm$^{-2}$ (K km s$^{-1}$)$^{-1}$; alternatively, if $X$ has the same value as in the Galaxy, then the GMCs are only marginally bound ($\alpha_{\text{vir}} \simeq 2$). They conclude that metallicity does not have a significant effect on $X$ since the ratio of the virial mass to the CO luminosity is constant in M33, despite a factor 10 variation in metallicity. (Note, however, that Elmegreen 1989 argues that $X$ depends on the ratio of the metallicity to the intensity of the FUV radiation field, which is not addressed by these results.) Although there is insufficient dynamic range for clear evidence of a relationship between linewidth and size based on current observations, the data are consistent with $\sigma \propto R^{1/2}$ but with values of $\sigma_{\text{pc}}$ (and therefore $\ell_s$) that vary from galaxy to galaxy. Blitz et al. (2007) also find that GMC surface densities have a relatively small range within any given Local group galaxy, while varying from $\sim 50$ M$_\odot$ pc$^{-2}$ for the SMC (L. Blitz, personal communication) to $> 100$ M$_\odot$ pc$^{-2}$ for M33.

There are currently two main conceptual frameworks for interpreting the data on GMC properties. One conception of GMCs is that they are dynamic, transient entities in which the turbulence is driven by large-scale colliding gas flows that create the cloud (e.g. Heitsch et al. 2005; Vázquez-Semadeni et al. 2006; Ballesteros-Paredes et al. 2007). This picture naturally explains why GMCs are turbulent (at least in the initial stages), and why the line width–size relation within clouds has an exponent of 1/2 – simply due to the scaling properties of supersonic turbulence. However, it is less obvious why $\alpha_{\text{vir}} \sim 1$ and why $\Sigma$ has a particular value, since small-scale dense structures may form (and collapse) at stagnation points in a high-velocity compressive flow before sufficient material has collected to create a large-scale GMC. Indeed, based on simulations with a converging flow of $\sim 20$ km s$^{-1}$ with no stellar feedback, Vázquez-Semadeni et al. (2007) find that star formation occurs when the column density is $N_H \approx 10^{21}$ cm$^{-2}$, a factor of ten below the mean observed value for GMCs. They also find that $\alpha_{\text{vir}}$ remains near unity after self-gravity becomes important, although the kinetic energy is primarily due to the gravitational collapse of the cloud, not to internal turbulence.

Zuckerman & Palmer (1974) argued many years ago that GMCs cannot be in a state of global collapse without leading to an unrealistically high star formation
rate. Proponents of the transient GMC picture counter by pointing out that most of the gas in GMCs is unbound and never forms stars (e.g., Clark et al., 2005); the global collapse is reversed by feedback from stars that form in the fraction of the gas that is overdense and bound. Indeed, individual star formation proceeds more rapidly than global collapse in essentially all turbulent simulations (see also §3.2.2). However, dominance of global collapse and expansion over large-scale random turbulent motions has not been confirmed from observations.

In the second conceptual framework, GMCs are formed by large-scale self-gravitating instabilities (see §3.2.1), and the turbulence they contain is due to a combination of inheritance from the diffuse ISM, conversion of gravitational energy to turbulent energy during contraction, and energy injection from newly formed stars (§3.2.2); the balance among these terms presumably shifts in time. In the work of Chieze (1987); Elmegreen (1989); Maloney (1988); McKee & Holliman (1999); McKee (1999), GMCs are treated as quasi-equilibrium, self-gravitating objects, so that the virial parameter is near-unity by definition. Whether or not equilibrium holds, the virial parameter is initially of order unity in scenarios involving gravitational instability because GMCs separate from the diffuse ISM as defined structures when they become gravitationally bound. For a quasi-equilibrium, the mean surface density is set by the pressure of the ambient ISM (see §2.2), which in turn is just the weight of the overlying ISM. Elmegreen (1989) has given explicit expressions for how the coefficient in the line width–size relation and the surface density depend on the external pressure, finding results that are comparable with observed values. In particular, the cloud surface density scales with the mean surrounding surface density of the ISM. Even if GMCs are not equilibria, if they are formed due to self-gravitating instabilities in spiral arms (e.g. Kim & Ostriker 2002, 2006) they must initially have surface densities a factor of a few above the mean arm gas density, consistent with observations. Provided that stellar feedback destroys clouds within a few (large-scale) dynamical times before gravitational collapse accelerates, the mean surface density would never greatly exceed the value at the time of formation. Simple models of cloud evolution with stellar feedback (e.g., Krumholz, Matzner, & McKee 2006) suggest that the scenario of slow evolution with $\alpha_{\text{vir}} = 1 - 2$ is self-consistent and yields realistic star formation efficiencies, but more complete studies are needed.

The two approaches to interpreting GMC dynamics correspond to two alternate views on GMC lifetimes. Elmegreen (2000) argued that, over a wide range of scales, star formation occurs in about 1–2 dynamical crossing times of the system, $t_{\text{cross}} \equiv 2R/(\sqrt{3}\sigma)$. Ballesteros-Paredes, Hartmann, & Vázquez-Semadeni (1999) and Hartmann, Ballesteros-Paredes, & Bergin (2001) focused on the particular case of star formation in Taurus, and argued that it occurred in about one dynamical time. The alternate view is that GMCs are gravitationally bound and live at least 2–3, and possibly more, crossing times, $t_{\text{cross}} \simeq 10M_6^{1/4}$ Myr, where $M_6 \equiv M/(10^6 M_\odot)$ and a virial parameter $\alpha_{\text{vir}} \sim 1 - 2$ is assumed. (Note that the crossing time in the nearby star-forming region in Taurus is $\sim 10^6$ yr, whereas in a large GMC it is $\sim 10^7$ yr.) However, GMCs (as opposed to structures within them) cannot be too close to being equilibria, since they do not appear to have a systematic, global density stratification, nor does the line width–size relation within individual clouds differ from that in unstratified clouds. (Note that logarithmic clouds can account for the observed relation $\delta v \propto \ell^{1/2}$ only if they are unbound—see §2.2). Estimates (empirical and theoretical) of GMC lifetimes will
be discussed in §3.2.2. It should be borne in mind that the difference between the two scenarios is only a factor $\sim 2 - 3$ for GMC lifetimes, which makes it difficult to obtain an unambiguous observational resolution purely based on timescales. However, there are major physical distinctions between the limiting cases of the scenarios that are under consideration – e.g. collapse triggered in colliding flows vs. a quasi-steady state supported by internally-driven turbulence. As complete numerical simulations are developed to flesh out the current proposals, it will be possible to distinguish among them using detailed kinematic observations.

### 3.1.2 Clumps and Cores

GMCs are highly clumped, so that a typical molecule is in a region with a density significantly greater than average. Liszt (1993) finds that the typical density of molecular gas in the Galactic plane is $n_H \simeq 3 \times 10^3 \, \text{cm}^{-3}$. Sanders et al. (1993) find a somewhat higher value from a multi-transition study, $n_H \simeq (4 - 12) \times 10^3 \, \text{cm}^{-3}$. However, the mean density in GMCs is considerably less: since $M \propto n_H R^3$ and $N_H \propto n_H R$,

$$\bar{n}_H = \frac{84}{M_6^{1/2}} \left( \frac{\bar{N}_H}{1.5 \times 10^{22} \, \text{cm}^{-2}} \right)^{3/2} \, \text{cm}^{-3}, \quad (28)$$

where we have normalized the column density to the typical value in the Solomon et al. (1987) sample. The effective filling factor of this molecular gas is then

$$f \equiv \frac{\bar{n}_H}{n_H} = \frac{0.028}{M_6^{1/2}} \left( \frac{3000 \, \text{cm}^{-3}}{n_H} \right) \left( \frac{\bar{N}_H}{1.5 \times 10^{22} \, \text{cm}^{-2}} \right)^{3/2}. \quad (29)$$

Note that since $f \leq 1$, clouds with $M \lesssim 10^3 \, M_\odot$ must have column densities less than the Solomon et al. (1987) value if their typical density is $\sim 3000 \, \text{cm}^{-3}$. The small filling factor of molecular gas in GMCs is expected in turbulent clouds (§2.1.4). It should be noted that star-forming clumps are themselves clumpy, and contain the cores that will evolve into stars.

The nature of the interclump medium is uncertain; it is not even known if it is atomic or molecular (Williams, Blitz, & McKee, 2000). Hennebelle & Inutsuka (2006) have suggested that the damping of hydromagnetic waves incident from the ambient ISM could maintain an interclump medium made up of warm H I.

The physical conditions in clumps and cores have been thoroughly reviewed by di Francesco et al. (2007) and Ward-Thompson et al. (2007), and we shall address only a few issues here. First, how well are Larson’s laws obeyed in clumps and cores? Most $^{13}$CO clumps are unbound, and therefore do not obey Larson’s laws (e.g., Carr, 1987); the mass distribution of such clumps can extend in an unbroken power law from several tens of solar masses down to Jupiter masses (Heithausen et al., 1998). On the other hand, Bertoldi & McKee (1992) found that most of the mass in the clouds is concentrated in the most massive clumps, and these appear to be gravitationally bound. The virial parameter for the unbound clumps decreases with increasing mass, in a manner similar to that observed for both the small molecular clouds and clumps within GMCs in the outer Galaxy (Heyer, Carpenter, & Snell, 2001). They found that the velocity dispersion in the unbound clumps is approximately independent of clump mass (or, since the clump density is also about constant in each cloud, of clump size): the unbound clumps do not obey a line width–size relation. Heyer, Carpenter, & Snell (2001) find the same result for clumps in the outer Galaxy. By contrast, Falgarone, Puget, & Perault (1992) found that unbound clumps do obey a line-width size
relation, albeit with considerable scatter. Bertoldi & McKee (1992) showed that
the kinetic pressure in the unbound clumps in their study is comparable to that
in the host molecular cloud, which is $P \approx G \Sigma_{MC}^2$ (eq. 12). Further evidence
on whether clumps or cores are bound is imprinted in their shapes and density
structure and is discussed below.

It is difficult to determine from the data whether there is a line width–size
relation within individual clumps and cores. For low-mass cores, Barranco &
Goodman (1998) (see also Goodman et al. 1998; Tafalla et al. 2004) found that
the nonthermal linewidth decreases and then reaches a minimum “plateau” level
at a finite impact parameter $\sim 0.1$ pc from the center of the core. Because of
projection effects, however, it is not possible to determine whether the observed
turbulence pervades the whole volume interior to that radial impact parameter,
or whether the turbulence is primarily in a shell surrounding a more-quiescent
core; it is also possible that the non-thermal line width is due to coherent oscillations of the cores (Keto et al., 2006). A line width–size relation in the ensemble of different gravitationally bound clumps and cores is expected only if they have
similar surface densities (eq. 27). Jijina, Myers, & Adams (1999) carried out
a comprehensive study of cores and star-forming clumps (in our terminology;
“dense cores” in theirs) observed in NH$_3$ and found that the objects with and
without associated IRAS sources each obeyed a nonthermal line width–size relation with slopes of about 0.5 and 0.8, respectively. When the sample was divided
into objects associated with or without clusters (defined as having at least 30
embedded YSOs), the cluster sample had a weak correlation between line width
and size with a slope of only about 0.2, whereas the non-cluster sample had a
stronger correlation with a slope of about 0.6. However, as remarked above,
Heyer, Carpenter, & Snell (2001) found no evidence for a line width–size relation
for small ($< 10^4$ $M_\odot$) molecular clouds in the outer Galaxy; furthermore, they
did not find evidence for a constant surface density at any mass. Plume et al.
(1997) observed a sample of clumps that are forming high-mass stars, and did
not find a line width–size relation.

The lack of an observed linewidth-size relation in observed unbound clumps
within a given cloud is at first puzzling, since defined volumes should sample
from the overall structure function of the GMC, which follows $\delta v \propto r^{0.5}$ (§2.1.5).
Analysis of turbulence simulations offers a resolution to this puzzle, suggesting
that many apparent clumps in moderate-density tracers such as $^{13}$CO are not in
fact single physical entities. Observationally, clumps are generally identified as
connected overdense peaks in position-velocity data cubes, with the line-of-sight velocity acting as a surrogate for line-of-sight position. Analysis of simulations
shows, however, that many “position-velocity” clumps in fact consist of separate
physical structures superimposed on the sky; correspondingly, many physically-
coherent structures have two or more separate components when observed in
line-of-sight velocity (Pichardo et al., 2000; Ostriker, Stone, & Gammie, 2001;
Ballesteros-Paredes & Mac Low, 2002). Because low-contrast apparent clumps
with any plane-of-sky size may sample from the velocity field along the whole
line of sight, the linewidth varies only very weakly with size. Since a fraction of
the apparent clumps sample velocities from a range of line-of-sight distances no
larger than their transverse extent, however, Ostriker, Stone, & Gammie (2001)
argued that the lower envelope of the “clump” linewidth-size distribution should
follow the scaling defined by the true three-dimensional power spectrum; this is
generally consistent with observations (Stutzki & Guesten, 1990; Williams, de
What about Larson’s third law? Since gravitationally bound clumps have $P \approx G\Sigma_{bd\ clump}^2$, and the mean pressure in a bound clump cannot be much greater than the ambient pressure for a stable structure without an internal energy source, it follows that typically $\Sigma_{bd\ clump}$ is comparable to $\Sigma_{GMC}$ (McKee, 1999). The high-mass, star-forming clumps studied by Plume et al. (1997) violate this conclusion: they have $\Sigma \sim 4800 \ M_\odot \ pc^{-2} \approx 1 \ g \ cm^{-2}$ with considerable dispersion, which is much greater than the typical GMC surface density $\sim 170 \ M_\odot \ pc^{-2}$. There are several possible explanations for this, and it is important to determine which is correct: Are these clumps just the innermost, densest parts of much larger clumps? Do they have much higher pressures than their surroundings but are avoiding gravitational collapse due to energy injection from star formation? Or are they the result of a clump-clump collision that produced unusually high pressures?

The density structure, velocity structure, and shape of cores offer potential means for determining whether they are dynamic objects, with short lifetimes, or quasi-equilibrium, gravitationally bound, objects. The observation of the Bok globule B68 in near-infrared absorption revealed an angle-averaged density profile consistent with that of a Bonnor-Ebert sphere to high accuracy (Alves, Lada, & Lada, 2001). Since then, a number of other isolated globules and cores have been studied with the same technique and fit to Bonnor-Ebert profiles, showing that starless cases are usually close to the critical limit, while cases with stars often match supercritical profiles (Teixeira, Lada, & Alves, 2005; Kandori et al., 2005). Profiles of dense cores have also been obtained using sub-mm dust emission (see di Francesco et al. 2007). A recent study by Kirk, Ward-Thompson, & André (2005) found that “bright” starless cores have density profiles consistent with supercritical Bonnor-Ebert spheres.

Consistency of density profiles with the Bonnor-Ebert profile does not, however, necessarily imply that a core is bound. Analysis of dense concentrations that arise in turbulence simulations show that Bonnor-Ebert profiles often provide a good fit to these structures (provided they are averaged over angles), even when they are transients rather than true bound cores (Ballesteros-Paredes et al., 2007). Even if a cloud with a Bonnor-Ebert profile is bound, however, it need not be in equilibrium: Myers (2005) and Kandori et al. (2005) have shown that density profiles of collapsing cores initiated from near-critical equilibria in fact follow the shapes of static supercritical equilibria very closely. The reason for this is that initially these cores collapse slowly, so that they are approximately in equilibrium; at later times, they evolve via “outside-in” collapse to a state that is marginally Jeans unstable everywhere ($\frac{\rho}{m}$) so that $\rho \propto r^{-2}$ except in a central core; highly supercritical Bonnor-Ebert spheres have profiles with the same characteristic shapes. Thus, not only the density structure but also the level of the internal velocity dispersion and detailed shape of the line profiles must be used in order to distinguish between transient, truly equilibrium, and collapsing objects (Keto & Field, 2005).

Core shapes also provide information on whether cores are transient or are bound, quasi-equilibrium objects. In the absence of a magnetic field, a quasi-equilibrium, bound cloud is approximately spherical. If the cloud is threaded by a magnetic field that tends to support the cloud against gravity, it will be oblate; if the field tends to compress the cloud (as is possible for some helical fields–Wiebe & Pudritz 2000), it will be prolate. If one assumes axisymmetry, the
distribution of observed axis ratios implies dense cores are primarily prolate (Ryden, 1996). However, this conclusion appears to be an artifact of the assumption of axisymmetry: using the method of analysis for triaxial clouds developed by Basu (2000), Jones, Basu, & Dubinski (2001); Jones & Basu (2002) concluded that cores with sizes < 1 pc are in fact oblate. Basu (2000) showed that if the magnetic field is aligned with the minor axis, as in most quasi-equilibrium models, the projection of the field on the plane of the sky will not generally be aligned with the projection of the minor axis, and he argued that the limited polarization data available are consistent with the theoretical expectation that the field in the cloud is aligned with the minor axis of the cloud. Kerton et al. (2003) showed that larger structures, extending up to GMCs, are intermediate between oblate and prolate, and are clearly distinct from the smaller objects. This is consistent with the analyses of clump shapes in turbulence simulations by Gammie et al. (2003); Li et al. (2004), who found that the majority of objects are triaxial. The data on cores and small clumps are thus consistent with (but do not prove) that they are bound, quasi-equilibrium objects. Large clumps and GMCs appear to be farther from equilibrium.

A key feature of the cores that form individual low-mass stars is that they have low nonthermal velocities, whether these cores are found in isolation or clustered with other cores (di Francesco et al., 2007; Ward-Thompson et al., 2007). The mean one-dimensional velocity dispersion in starless cores based on the sample of Benson & Myers (1989) is $0.11 \text{ km s}^{-1}$, such that the three-dimensional velocity dispersion is approximately sonic. This places constraints on theoretical models, and in particular may constrain the nature of turbulent driving. Klessen et al. (2005) compared the results for cores identified in two (unmagnetized) simulations with the same RMS Mach number $\approx 10$, one driven on large scales and the other driven on small scales; the time correlation of the driving force is short for both cases. They found that only with large-scale driving is the mean turbulence level within cores approximately sonic; in the small-scale driving case the preponderance of cores are supersonic. Klessen et al. (2005) also found that the starless cores in their large-scale-driving models are within a factor of a few of kinetic and gravitational energy equipartition. In the 2D MHD simulations of Nakamura & Li (2005) that implement driving by instantaneous injection of radial “wind” momentum when (low-mass) stars are formed, the dense cores that are identified also primarily have subsonic internal motions. Importantly, both types of models show that dense, quiescent cores can form in a turbulent environment; the slow, diffusive formation of quiescent cores central to the older picture of star formation does not seem to be required.

What happens to dense cores once they form? Cores that have sufficient internal turbulence compared to their self-gravity will redisperse within a crossing time. Cores that reach low enough turbulence and magnetization levels (allowing for dissipation) within a few local free-fall times will collapse if $M > M_{\text{cr}}$. Vázquez-Semadeni et al. (2005) found that in globally supercritical 3D simulations with driven turbulence, cores that collapse do so within 3–6 local free-fall times of their formation. Nakamura & Li (2005) found via 2D simulations including ambipolar diffusion that even when the mass in the simulation volume is 20% less than critical, supercritical cores can form; these then either collapse or redisperse within several local free-fall times. In both cases (see also Krasnopolsky & Gammie 2005), only magnetically supercritical cores collapse, as expected. Quiescent cores that are stable against gravitational collapse could in principle
survive for a long time (Lizano & Shu, 1989), undergoing oscillations in response to fluctuations in the ambient medium (Keto & Field, 2005; Keto et al., 2006). Because they are only lightly bound, however, such “failed cores” can also be destroyed relatively easily by the larger-scale, more powerful turbulence in the surrounding GMC. This process is clearly seen in numerical simulations; Vázquez-Semadeni et al. (2005) and Nakamura & Li (2005) found that the bound cores that subsequently disperse do so in $1 - 6 \times t_{\text{ff}}$. Quiescent, magnetically subcritical cores with thermal pressure $\rho_{\text{core}}c_s^2$ exceeding the mean turbulent pressure $\rho_{\text{nt}}c_s^2$ (so that the core would collapse in the absence of magnetic support) cannot easily be destroyed, however, and it is likely that they remain intact until they merge with other cores to become supercritical. Simulations have not yet afforded sufficient statistics to determine the mean time to collapse or dispersal as a function of core properties and cloud turbulence level, or whether there is a threshold density above which ultimate collapse is inevitable.

Observationally, core lifetimes can be estimated by using chemical clocks or from statistical inference. The formation of complex molecules takes $\sim 10^5$ yr at typical core densities, but this “clock” can be reset by events that bring fresh C and C$^+$ into the core, such as turbulence or outflows (Langer et al., 2000). A potentially more robust clock is provided by observations of cold H I in cores: Goldsmith & Li (2005) infer ages of $10^{6.5-7}$ yr for five dark clouds from the low observed values of the H I/H$_2$ ratio. These age estimates would be reduced if clumping is significant and hence the time-averaged molecule formation rate is accelerated, but, as in the case of complex molecules, they would be increased if turbulent mixing were effective in bringing in fresh atomic hydrogen. In simulations of molecule formation in a turbulent (and therefore clumpy) medium, Glover & Mac Low (2007) find that H$_2$ formation is indeed accelerated when compared with the non-turbulent case, although the atomic fractions they found are substantially greater than those observed by Goldsmith & Li (2005). If confirmed, these ages, which are considerably greater than a free-fall time, would suggest that these dark clouds are quasi-equilibrium structures.

Statistical studies of core lifetimes are based on comparing the number of starless cores with the number of cores with embedded YSOs and the number of visible T Tauri stars. The ages of the cores (starless and with embedded YSOs) can then be inferred from the ages of the T Tauri population, provided that most of the observed starless cores will eventually become stars. The results of several such studies have been summarized by Ward-Thompson et al. (2007), who conclude that lifetimes are typically $3 - 5t_{\text{ff}}$ for starless cores with densities $n_{\text{H}_2} = 10^{3.5} - 10^{5.5}$ cm$^{-3}$. This is not consistent with dynamical collapse, nor is it consistent with a long period ($> 5t_{\text{ff}}$) of ambipolar diffusion. It is consistent with the ambipolar diffusion in observed magnetic fields (§2.3), which are approximately magnetically critical. Of course, cores are created with a range of properties, and observational statistics are subject to an evolutionary selection effect: cores that are born or become supercritical evolve rapidly into collapse, and are no longer counted among the starless population. Given a population with a range of intrinsic lifetimes (but similar birth rates), the longest-lived objects will be the best represented. The data rule out the possibility that most cores are born very subcritical and lose their magnetic flux slowly, over $\sim 10t_{\text{ff}}$.

The angular momentum of cores was initially regarded as a bottleneck for star formation, but extensive theoretical analysis led to the conclusion that magnetic fields would provide an effective braking mechanism (e.g., Mestel, 1985;
Mouschovias, 1987). Observations have established that the angular momentum, or equivalently, the rotational energy, of cores is indeed small (e.g., Goodman et al., 1993; Jijina, Myers, & Adams, 1999; Caselli et al., 2002; Pirogov et al., 2003). Goodman et al. (1993) characterized the rotational energy by the parameter

\[ \beta_{\text{rot}} = \frac{1}{3} \left( \frac{v_{\text{rot}}^2}{GM/R} \right), \]  

which is the ratio of the rotational energy to the gravitational binding energy for a uniformly rotating, constant density sphere; they found a median value \( \beta_{\text{rot}} \approx 0.03 \). The specific angular momenta \( j \) in this sample range from \( 6 \times 10^{20} - 4 \times 10^{22} \) cm\(^2\) s\(^{-1}\), and increase with size approximately as \( j \propto R^{3/2} \). Interestingly enough, Burkert & Bodenheimer (2000) showed that rotation arising from sampling turbulent fluctuations with a Burgers power spectrum (and normalization matched to observations) is adequate to account for the observations; in this picture, the role of magnetic braking on small scales is unclear. (On the other hand, magnetic braking appears to be clearly significant in regulating the spin of GMAs and hence GMCs – Rosolowsky et al., 2003; Kim, Ostriker, & Stone, 2003.) Since \( j \propto vR \) and \( v \propto R^{1/2} \) for large-scale turbulence, the observed \( j \propto R^{3/2} \) relation is what would be expected if core turbulence scales similarly. Li et al. (2004) indeed find agreement with this scaling from cores identified in their simulations. Gammie et al. (2003) and Jappsen et al. (2005) both find that the mean specific angular momentum of cores in their models (using grid-based MHD and [unmagnetized] SPH, respectively) is given in terms of the sound speed and large-scale Jeans length by \( \sim 0.1c_s\lambda_J \); Jappsen et al. (2005) show that if the mean density is adjusted so that core masses match those in observed regions, then the mean angular momentum distributions match as well. For cores that collapse, Jappsen et al. (2005) also found that the distributions of \( \beta_{\text{rot}} \) are similar to those obtained by Goodman et al. (1993).

### 3.2 Formation, Evolution, and Destruction of GMCs

#### 3.2.1 CLOUD FORMATION

In principle, GMCs could form either by “bottom-up” or by “top-down” processes. In bottom-up formation, successive inelastic collisions of cold H I clouds would gradually increase the mean cloud size and mass until that of a (self-gravitating) GMC is reached (e.g. Field & Saslaw, 1965; Kwan, 1979). The difficulty with this coagulation scenario, as was early recognized, is that it is very slow; e.g. Kwan (1979) found that the time needed for the peak of the mass distribution to exceed \( 10^5M_\odot \) is more than \( 2 \times 10^8 \) yr. The binary collision time for spherical clouds of radius \( R_{\text{cl}} \) and density \( \rho_{\text{cl}} \) is

\[ t_{\text{collis}} = (\sqrt{\pi}/3)(\rho_{\text{cl}}/\bar{\rho})(R_{\text{cl}}/\sigma), \]

where \( \bar{\rho} \) is the density averaged over large scales and \( \sigma \) is the one-dimensional velocity dispersion over the mean intercloud separation (see Binney & Tremaine 1987, eq. 8-122; note that for considering agglomeration we neglect grazing collisions, choosing a maximum impact parameter \( R_{\text{cl}} \)). Expressed in terms of the cloud “gathering scale” \( R_{\text{gath}} \equiv [3M_{\text{cl}}/(4\pi\bar{\rho})]^{1/3} = 190 \) pc\((M_{\text{cl},6}/n_\text{H})^{1/3}\) in the diffuse ISM, or in terms of the cloud surface density \( \Sigma_{\text{cl}} \equiv M_{\text{cl}}/(\pi R_{\text{cl}}^2) \), the collision time is

\[ t_{\text{collis}} = \frac{\sqrt{\pi}}{3} \left( \frac{\rho_{\text{cl}}}{\bar{\rho}} \right)^{2/3} \frac{R_{\text{gath}}}{\sigma} = \frac{\sqrt{\pi}}{4} \frac{\Sigma_{\text{cl}}}{\bar{\rho}\sigma}. \]
The mean intercloud separation is comparable to $2R_{\text{gath}}$, which exceeds the atomic disk scale height $H \approx 150 \text{ pc}$ (Malhotra, 1995) for $M_{\text{cl},6} \equiv M_{\text{cl}}/10^6 \, M_\odot \gtrsim 0.04$. We can use equation (31) to estimate the collision time if all the diffuse ISM gas were apportioned into equal-mass clouds with equal surface density. Using $\sigma \approx 7 \, \text{ km s}^{-1}$ for the nonthermal velocity dispersion in the diffuse ISM at large ($\gtrsim H$) scales (Heiles & Troland, 2003), $\Sigma_{\text{cl}} \approx 170 \, M_\odot \, \text{pc}^{-2}$ for the mean GMC column (Solomon et al., 1987), and mean density $n_H = 0.6 \, \text{cm}^{-3}$ typical of the diffuse ISM at the Solar circle (Dickey & Lockman, 1990), this yields a collision timescale $> 5 \times 10^8 \, \text{yr}$. Gravitational focusing in principle decreases the cloud-cloud collision time, but in practice this does not help in forming GMCs from atomic clouds since the reduction factor for the collision time, \[1 + \pi G R_{\text{cl}} \Sigma_{\text{cl}} / \sigma^2 \approx 1\], is near unity until the clouds are quite massive ($\gtrsim 10^5 \, M_\odot$). Even if the background density were arbitrarily (and unrealistically) enhanced by a factor 100 to approach $\rho_{\text{cl}}$, the total time of 40 Myr required to build clouds from $10^2 \, M_\odot$ to $5 \times 10^6 \, M_\odot$ (by successive stages of collisions) would still exceed the estimated GMC lifetimes. These lifetimes are set by the time required to destroy clouds by a combination of photodissociation and mechanical unbinding by expanding HII regions (see §§3.2.2). Thus, if coagulation were the only way to build GMCs, the process would be truncated by destructive star formation before achieving the high GMC masses in which most molecular mass is actually found.

Given the timescale problem and other difficulties of bottom-up GMC formation (e.g., Blitz & Shu, 1980), starting in the 1980’s the focus shifted to top-down mechanisms involving large-scale instabilities in the diffuse ISM (e.g., Elmegreen, 1979, 1995). The two basic physical processes that could trigger growth of massive GMCs involve (1) differential vertical buoyancy of varying-density regions along magnetic field lines parallel to the midplane, or (2) differential in-plane self-gravity of regions with varying surface density. The first type of instability is generically termed a Parker instability (Parker, 1966). The second type of instability is generically a Jeans instability, although the simplest form of Jeans instability involving just self-gravity and pressure cannot occur, due to galactic (sheared) rotation (see §2.2). If the background rotational shear is strong, as in the interarm regions of grand design spirals or in flocculent galaxies, there is no true instability but instead a process known as swing amplification (Goldreich & Lynden-Bell, 1965; Toomre, 1981); the dimensionless shear rate must be $d \ln \Omega / d \ln R \lesssim -0.3$ for “swing” to occur (Kim & Ostriker, 2001). If, on the other hand, the mean background dimensionless shear rate is low (as in the inner parts of galaxies where rotation is nearly solid-body, or as in spiral arms), another type of gravitational instability can develop provided magnetic fields are present to transfer angular momentum out of growing condensations (Elmegreen, 1987; Kim & Ostriker, 2001); this is referred to as a magneto-Jeans instability (MJII).

The characteristic azimuthal spatial scale for Parker instabilities is $\lambda_\phi \approx 4 \pi H$ (Shu, 1974). Growth rates are $\propto v_A / H$, which tends to increase in spiral arms; thus these regions have traditionally been considered most favorable for growth of Parker modes (Mouschovias, Shu, & Woodward, 1974). Numerical simulations have shown, however, that Parker instability is not on its own able to create structures resembling GMCs, because the instability is self-limiting and saturates with only order-unity surface density enhancement (Kim et al., 1998; Santillán et al., 2000; Kim, Ryu, & Jones, 2001; Kim, Ostriker, & Stone, 2002). Spiral arms are also the most favorable regions for self-gravitating instabilities (Elmegreen,
1994), since the characteristic (thin-disk) growth rate $\propto G\Sigma_{\text{gal}}/c_s$ is highest there. (Here, $\Sigma_{\text{gal}}$ is the mean gas surface density averaged over large [$\sim$ kpc] scales in the plane of the disk.) Since the spatial wavelengths of Parker and MJI modes are similar, in principle growth of the former could help trigger the latter within spiral arms (Elmegreen, 1982a,b). In fact, it appears that turbulence excited in spiral shocks, together with vertical shear of the horizontal flow, may suppress growth of large-scale Parker modes in arm regions (Kim & Ostriker, 2006). Thus, while the Parker instability is important in removing excess magnetic flux from the disk and in transporting cosmic rays (e.g. Hanasz & Lesch, 2000), it may be of limited importance in the formation of GMCs.

Self-gravitating instabilities, unlike buoyancy instabilities, lead to ever-increasing density contrast if other processes do not intervene. The same is true for the swing amplifier if the (finite) growth is sufficient to precipitate gravitational runaway. The notion that there should be a threshold for star formation depending on the Toomre parameter,

$$ Q \equiv \frac{\kappa c_s}{\pi G \Sigma_{\text{gal}}} = 1.4 \left( \frac{c_s}{7.0 \text{ km s}^{-1}} \right) \left( \frac{\kappa}{36 \text{ km s}^{-1} \text{ kpc}^{-1}} \right) \left( \frac{\Sigma_{\text{gal}}}{13 \ M_\odot \text{ pc}^{-2}} \right)^{-1}, \quad (32) $$

is based on the idea that star-forming clouds can form by large-scale self-gravitating collective effects only if $Q$ is sufficiently low. Here, $\kappa^2 \equiv R^{-3}\partial(\Omega^2 R^4)/\partial R$ is the squared epicyclic frequency, and $c_s$ is the mean sound speed of the gas. Numerical simulations have been used to determine the nonlinear instability criterion, finding that gravitationally bound clouds form provided $Q < Q_{\text{crit}} \approx 1.5$ in model disks that allow for realistic vertical thickness, turbulent magnetic fields, and a “live” stellar component (Kim, Ostriker, & Stone 2003; Li, Mac Low, & Klessen 2005b; Kim & Ostriker 2007; these models do not include global spiral structure in the gas imposed by variations in the stellar gravitational potential – see below.) These results agree with empirical findings for the mean value of $Q$ at the star formation threshold radii in nearby galaxies (e.g. Quirk, 1972; Kennicutt, 1989; Martin & Kennicutt, 2001). The masses of bound clouds formed via self-gravity in galactic disk models where the background gas surface density is relatively uniform are typically a few to ten times the two-dimensional Jeans mass,

$$ M_{J2D} \equiv \frac{c_s^4}{G^2 \Sigma_{\text{gal}}} = 10^7 \ M_\odot \left( \frac{c_s}{7 \text{ km s}^{-1}} \right)^4 \left( \frac{\Sigma_{\text{gal}}}{13 \ M_\odot \text{ pc}^{-2}} \right)^{-1}, \quad (33) $$

depending on the specific ingredients of the model (Kim, Ostriker, & Stone, 2002, 2003; Kim & Ostriker, 2007).

Observations of external galaxies with prominent spiral structure show that most of the molecular gas is concentrated in the spiral arms (e.g. Helfer et al., 2003; Engargiola et al., 2003), and within the Milky Way the most massive clouds that contain most of the mass and forming stars are strongly associated with spiral arms (Solomon, Sanders, & Rivolo, 1985; Solomon & Rivolo, 1989; Heyer & Terebey, 1998; Stark & Lee, 2006). The observed relationship between GMCs and spiral structure suggests that molecular clouds are preferentially born in the high density gas that makes up the arms; this is consistent with theoretical expectations since growth rates for all proposed mechanisms increase with the gas surface density. As noted above, collisional coagulation is expected to be too slow in spiral arms; gravitational instabilities are, however, faster (e.g. Elmegreen, 1990). Taking the ratio of the collision rate $t_{\text{collis}}^{-1}$ to the characteristic growth rate
of the MJI, \( \pi G \Sigma_{\text{gal}}/c_s \), and setting \( \sigma = c_s \) and \( H \approx c_s^2/(\pi G \Sigma_{\text{gal}}) \) (since gas gravity dominates stellar gravity in the arm), the result is \( t_{MJI}/t_{\text{collis}} = 2\Sigma_{\text{gal}}/(\sqrt{\pi} \Sigma_{\text{cl}}) \).

Thus, the collision rate is lower than the self-gravity contraction rate by roughly the surface filling factor of clouds in the arm, \( \sim 0.2 \)–\( 0.5 \) if the arm surface density is enhanced by a factor \( 3 \)–\( 6 \) above the mean value.

To obtain realistic estimates of the masses and other properties of clouds formed via gravitational instabilities, it is necessary to include the effects of spiral structure in detailed numerical models. Diffuse gas entering a spiral arm will in general undergo a shock, significantly increasing the background density (e.g. Roberts, 1969; Shu, Milione, & Roberts, 1973); gas self-gravity enhances the maximum compression factor and also tends to symmetrize the gas density profile across the arm (Lubow, Cowie, & Balbus, 1986). In addition to strong variations in the gas surface density, spiral structure induces corresponding local variations in the gas flow velocity (both compression and streaming). Since the Jeans length is not small compared to the scale of these variations, the background arm “profile” must be taken into account in studying growth of self-gravitating condensations (Balbus, 1988). Kim & Ostriker (2002) performed 2D MHD simulations of this process, showing that bound condensations develop both within the spiral arms themselves, and also downstream from the arms in trailing gaseous “spurs”. Based on 3D local MHD simulations (Kim & Ostriker, 2006) and 2D “thick disk” global simulations (Shetty & Ostriker, 2006), bound gas condensations formed in spiral arms (and arm spurs) have typical masses \( 1 \times 10^7 M_\odot \). This value is \( \sim 10 \times \) the thin-disk 2D Jeans mass using the peak arm density in equation (33), or comparable to the value of the “thick disk” Jeans mass \( M_{J,\text{thick}} = 2\pi c_s^2 H/G \) which is obtained using the gravitational kernel \( \Phi_k = -2\pi G \Sigma_{\text{gal}}/(k + k^2 H) \).

The masses of the bound structures formed via self-gravitating instabilities are comparable to the upper end of the mass function of GMCs in the Milky Way (see eq. 25), allowing for H I envelopes; they are also comparable to the masses of GMAs that have been observed in external spiral galaxies (e.g. Vogel, Kulkarni, & Scoville, 1988; Aalto et al., 1999; Rand, Lord, & Higdon, 1999). In addition, the morphology and spacing of spiral-arm spurs predicted to form via self-gravity effects are consistent with structures that have been observed at a variety of wavelengths (e.g. Elmegreen, 1980; La Vigne, Vogel, & Ostriker, 2006). This spacing (of several times the Jeans length in the arm, \( c_s^2/[G \Sigma_{\text{gal}}, \text{arm}] \)) is similar to that of giant HII regions arranged as “beads on a string” along spiral arms in many grand design spirals (Elmegreen & Elmegreen, 1983), and also similar to the spacings of giant infrared clumps observed with Spitzer along the spiral arms of interacting galaxies IC 2163 and NGC 2207 (Elmegreen et al., 2006). These giant IR clumps typically host a number of individual (optically-observed) H II regions associated with star clusters. While current mm-wavelength observations have insufficient resolution to identify sub-condensations within GMAs in external galaxies, it is expected that their internal turbulence would create density substructure, just as the internal turbulence of GMCs fragments them into clumps. Since the mean density within GMCs is comparable to the typical density of cold clouds in the atomic medium, the pre-existing cloudy structure of the diffuse ISM would contribute to, but not dominate, the internal structure within GMCs. In this “top-down” picture, the more massive, self-gravitating, substructures within GMAs (or analogous atomic “superclouds”) would then become gravitationally bound GMCs.

While large-scale self-gravitating instabilities appear necessary for forming
massive GMCs, and many low-mass GMCs may form via fragmentation of massive GMCs or GMAs, it remains possible that a proportion of the low-mass GMCs form through other mechanisms. Several recent studies have explored the possibility of GMC assembly via colliding supersonic flows (e.g. Chernin, Efremov, & Voinovich, 1995; Vázquez-Semadeni, Passot, & Pouquet, 1995; Ballesteros-Paredes, Hartmann, & Vázquez-Semadeni, 1999; Heitsch et al., 2005; Vázquez-Semadeni et al., 2007); in this scenario the post-shock gas in the stagnation region (which in fact becomes turbulent) represents the nascent GMC. For diffuse ISM gas at mean density \( \bar{\rho} \) with relative (converging) velocity \( v_{\text{rel}} \), a total column of shocked gas \( \Sigma_{\text{cl}} \) builds up over time

\[
\frac{\Sigma_{\text{cl}}}{\bar{\rho} v_{\text{rel}}} = 1.6 \times 10^7 \text{yr} \left( \frac{N_{\text{H}}}{10^{21} \text{ cm}^{-2}} \right) \left( \frac{n_{\text{H}}}{1 \text{ cm}^{-3}} \right)^{-1} \left( \frac{v_{\text{rel}}}{20 \text{ km s}^{-1}} \right)^{-1};
\]

note that modulo order-unity coefficients this is the same as the result in equation (31), with the velocity dispersion of the cloud distribution replaced by the relative velocity of the converging flow. Correlated flows can only be maintained up to the flow timescale over the largest spatial scale of the turbulence, \( \sim 2H \sim 300 \text{ pc} \). With \( v_{\text{rel}} \) equal to the RMS relative velocity \( \sqrt{6} \sigma \) for a Gaussian with 1D velocity dispersion \( \sigma \approx 7 \text{ km s}^{-1} \), this is \( \approx 2 \times 10^7 \text{ yr} \), yielding a column \( \approx 10^{21} \text{ cm}^{-2} \) for \( n \approx 1 \text{ cm}^{-3} \) (note that the shock velocity is about \( v_{\text{rel}}/2 \)). If the interstellar magnetic field does not limit the compression of the shocked gas (an artificial assumption, requiring flow along field lines only), the post-shock gas would have high enough density for significant amounts of \( \text{H}_2 \) to form within the overall accumulation time, and the shielding from the diffuse UV is sufficient for CO to begin to form (Hartmann, Ballesteros-Paredes, & Bergin, 2001; Bergin et al., 2004). However, it should be noted that for a shock velocity of \( 10 \text{ km s}^{-1} \), corresponding to a relative velocity of \( 20 \text{ km s}^{-1} \), less than half the C is in CO after \( 10^8 \text{ yr} \) according to the 1D calculations of Bergin et al. (2004). Turbulence-induced clumping can accelerate molecule formation rates (Elmegreen, 2000; Glover & Mac Low, 2007), alleviating the timescale problem. The molecule formation rate is proportional to the mass-weighted mean density \( \langle n \rangle_M \), which is larger than the volume-weighted mean density \( \bar{n} \equiv \langle n \rangle_V \) in a turbulent flow (see §2.1.4). Since \( \langle n \rangle_M/\bar{n} \sim 10 \) for typical turbulence levels in GMCs, this reduces the typical molecule formation time (Tielens, 2005), \( \sim 2 \times 10^9 \text{ yr} (T/10 \text{ K})^{-1/2}/\langle n \rangle_M \), to 1-2 Myr. Even so, the discussion above shows that the maximum column density produced in the colliding-flow scenario is \( \sim 10^{21} \text{ cm}^{-2} \), which is lower by an order of magnitude than the mean value of the column of molecular gas in Milky Way GMCs; thus, this process can account for at most a small fraction of the molecular gas mass in GMCs. Because gravitational instabilities are suppressed by the high \( Q \) values in interarm regions, on the other hand, the turbulent accumulation mechanism may be more important there. Potentially, this may account for the observed difference (see above) between arm and interarm GMC masses in the Milky Way, as well as for the very low surface densities \( \Sigma_{\text{cl}} \) observed for many of the outer-Galaxy molecular clouds (Heyer, Carpenter, & Snell, 2001).

Finally, we note that the dynamical considerations for gravitationally-bound cloud formation apply whether the diffuse gas is primarily atomic, as is the case in the Solar neighborhood and the outer portions of galaxies more generally, or whether the diffuse gas is primarily molecular, as is true in the inner portions of many galaxies. The time and length scales involved depend on the effective pres-
sure in the diffuse gas, which includes thermal as well as turbulent and magnetic terms. If the diffuse gas is primarily molecular (or cold atomic) by mass, then the mean turbulent and Alfvén speeds will exceed the thermal speed in the dense gas. The thermal sound speed $c_s$ in equations (32) and (33) must then be replaced by an appropriately-defined $c_{\text{eff}}$ that incorporates the effects of turbulent kinetic and magnetic pressures (the form of $c_{\text{eff}}$ would depend on the detailed multiphase structure of the gas). Similarly, the characteristic timescale for self-gravitating cloud formation becomes

$$t_{\text{t2D}} = \frac{c_{\text{eff}}}{G\Sigma_{\text{gal}}} = 3 \times 10^7 \text{yr} \left( \frac{c_{\text{eff}}}{10 \text{ km s}^{-1}} \right) \left( \frac{\Sigma_{\text{gal}}}{100 \text{ M}_\odot \text{ pc}^{-2}} \right)^{-1}. \quad (35)$$

Turbulent velocity dispersions and magnetic field strengths are observed to be similar in the cold and warm diffuse gas in the Solar neighborhood (Heiles & Troland, 2003, 2005), and both observations and simulations (Piontek & Ostriker, 2005) show that magnetic pressure is generally a factor two larger than the thermal pressure.

The transition from having primarily atomic to primarily molecular gas typically occurs where the total gas surface density $\Sigma_{\text{gal}} \approx 12 \text{ M}_\odot \text{ pc}^{-2}$ (Wong & Blitz, 2002; Blitz & Rosolowsky, 2004) and where the mean midplane pressure is inferred to lie in the range $P/k = 10^4 - 10^5 \text{ K cm}^{-3}$ (Blitz & Rosolowsky, 2006). This transition occurs due to a combination of increased self-shielding (hence a lower $\text{H}_2$ dissociation rate) as $\Sigma_{\text{gal}}$ increases, and increased density (hence increased $\text{H}_2$ formation rate) as both $\Sigma_{\text{gal}}$ and the stellar surface density increase toward the center of a galaxy (Elmegreen, 1993a; Blitz & Rosolowsky, 2004).

3.2.2 CLOUD EVOLUTION AND DESTRUCTION  GMCs are born in spiral arms downstream from the large-scale shock fronts, and begin life in a very turbulent state. As they contract under the influence of gravity, this turbulence decays, although the rate of decay is slowed by compression. Mestel & Spitzer (1956) conjectured that turbulence would decay in about a crossing time, and for that reason rejected turbulence as a mechanism for supporting clouds against gravitational collapse (see also Mouschovias, Tassis, & Kunz, 2006, who argue for magnetic support). Indeed, in simulations that do not include energy injection, the contraction eventually evolves into free-fall collapse (cf. Vázquez-Semadeni et al., 2007, who simulated the formation of molecular clouds in colliding flows, as discussed above), which is generally not observed. Furthermore, as discussed in §3.1, turbulence in molecular clouds is observed to be ubiquitous, suggesting that there is some mechanism acting to inject turbulent energy into clouds. How important is energy injection to the structure and evolution of molecular clouds?

Broadly speaking, there are two modes of energy injection, external and internal. External mechanisms tap the turbulence in the diffuse ISM, and because these modes are large scale, external driving would tend to yield a power spectrum that rises all the way to the largest scale in the GMC (see §2.1.1). In terms of total amplitudes, however, external driving may have limited practical importance. For example, while cloud-cloud collisions could drive turbulence, they may be too rare to be important (except possibly within the denser portions of spiral arms; see above). Magnetorotational instability (MRI) in the Galactic disk (Sellwood & Balbus, 1999) connects GMCs to the diffuse ISM by threaded field lines and could therefore help drive GMC turbulence, but it is difficult for this to be effective since GMCs live for only a small fraction of a rotation period (see below), and since MRI is modest compared to other kinetic turbulence drivers in
inner galaxies (Piontek & Ostriker, 2005). Supernovae are the dominant energy source in most of the diffuse interstellar medium (Mac Low & Klessen, 2004), with instabilities in spiral shocks making a significant contribution in spiral arm regions (Kim, Kim, & Ostriker, 2006). It is difficult, however, for these (or other) processes to transmit energy from the diffuse ISM into molecular clouds, which are much denser (Yorke et al., 1989). In fact, the density contrast between molecular clouds and the ambient medium means that energy tends to be reflected from clouds rather than being transmitted into them (e.g., Elmegreen, 1999 and Heitsch & Burkert, 2002 showed this for externally generated Alfvén waves). Thus, “external” energy is primarily limited to the turbulence GMCs inherit from their formation stages.

Internal energy injection is due to feedback from protostars and newly formed stars. This mechanism is observed to be important, but it remains to be shown that it can account for the ubiquity of turbulence given that star formation is intermittent in both space and time. For example, Mooney & Solomon (1988) find that 1/4 of a sample of inner-Galaxy GMCs show no evidence for the presence of O stars. Norman & Silk (1980) were the first to analyze the importance of energy injection by stellar outflows. At the time of their work, bipolar outflows from protostars were unknown, and they focused on winds from T Tauri stars. They suggested that these winds would blow cavities that would govern the structure of the clouds. Franco (1983) and Franco & Cox (1983) considered rotationally driven winds from protostars as well as T Tauri winds. They estimated the star formation rate required to keep swept-up shells colliding at a rate high enough to keep the cloud turbulent, and showed that this was roughly consistent with estimates of the Galactic star formation rate. The effect of stellar energy injection on the surrounding molecular cloud can be described in terms of an energy equation for the cloud (McKee, 1989, 1999; Krumholz, Matzner, & McKee, 2006),

$$\frac{de}{dt} = G - L,$$

where $e$ is the energy per unit mass (including gravitational energy) and $G$ and $L$ represent energy gains and losses, respectively. The injection of energy into a cloud by stars is often accompanied by mass loss from the cloud. Under the simplifying assumption that mass lost from the cloud does not change the energy per unit mass, the specific energy losses are due to the decay of MHD turbulence, which occurs at a rate (see §2.1.2)

$$L = 2.5 \frac{\sigma^3}{R} \left( \frac{2R}{\lambda} \right).$$

(36)

For a cloud of fixed size, the largest driving scale is $\lambda \approx 2R$; for a cloud undergoing global expansion and contraction, the largest scale is $\lambda \approx 4R$. Next, consider the energy gains. Since the shocks associated with the energy injection are radiative, most of the energy injected by outflows is radiated away. The outflow energy available to drive turbulence in the cloud is about $(\sqrt{3}/2)p_w\sigma$, where $p_w = m_w v_w$ is the momentum, $m_w$ the mass, and $v_w$ the velocity of the (proto)stellar outflow; as a result, $G = (\sqrt{3}/2)M_s(p_w\sigma/m_s M)$, where $M_s$ is the star formation rate and $M$ is the cloud (or clump) mass. Balancing the energy gains and losses, $G = L$, then gives the star formation rate necessary to maintain the turbulent motions. Norman & Silk (1980), Franco (1983), and McKee (1989) all estimated outflow momenta $p_w/m_s \approx 50 - 70 \text{ km s}^{-1}$ for a typical stellar mass $m_s = 0.5M_\odot$ and found that the energy injection rate was sufficient to support star-forming clouds against collapse. Li & Nakamura (2006) and Nakamura & Li (2007) have carried out 3D MHD simulations of a forming star cluster, and found that the rate at
which energy is input from protostellar winds maintains the surrounding gas in approximate virial equilibrium. Matzner (2007) has given a general discussion of turbulence driven by protostellar outflows. In particular, he has shown on the basis of dimensional analysis that if outflows occur at a rate per unit volume $S$ and inject an average momentum $I$ into a medium of density $\rho_0$, then on scales small compared to that at which the outflows overlap, the line width—size relation is $\sigma \propto (SIR/p_0)^{1/2}$, where the coefficient is of order unity.

There are both observational and theoretical caveats to this picture. Observationally, Richer et al. (2000) cite $p_w/m_\star \simeq 0.3v_K$, where $v_K$ is the Keplerian velocity at the stellar surface; for $v_K \simeq 200$ km s$^{-1}$, this is $p_w/m_\star \simeq 60$ km s$^{-1}$, in agreement with the values used in these models of energy injection and in agreement with both X-wind and disk models of protostellar outflows. However, some recent observations are in striking disagreement with these estimates: based on observations of CO outflows by Bontemps et al. (1996), Walawender, Bally, & Reipurth (2005) infer a much lower average outflow momentum $p_w = 1.2M_\odot$ km s$^{-1}$ (it should be noted, however, that they infer a much larger momentum, $p_w = 18 M_\odot$ km s$^{-1}$, in a protostellar jet). Quillen et al. (2005) discovered a number of CO cavities in NGC 1333, which they identified as fossil outflows with $p_w \sim 1 M_\odot$ km s$^{-1}$; this estimate of the momentum has been confirmed in numerical simulations of fossil outflows (Cunningham et al., 2006). Like Walawender et al., they concluded that these outflows could support a region of active star formation like NGC 1333 against gravity, but outflows could not support the larger Perseus cloud in which it is embedded. Clearly, more observational work is needed to determine the rate of protostellar energy injection.

The theoretical caveat is that protostellar outflows are unlikely to be effective on the scale of GMCs. Turbulent driving yields a flat power spectrum at scales larger than the input scale of turbulence, regardless of magnetization (see §2.1.1). Because the scale of protostellar jets and outflows is small compared to that of a GMC, driving by low-mass stars is inconsistent with observed turbulence spectra that continue rising up to scales of tens of pc (see §2.1.5); this effect is partly ameliorated by the clustering of stars, since protostellar outflows from a cluster will extend to larger scales than those from individual stars. Protostellar outflows remain a viable energy source for clumps and cores within GMCs, however.

Massive stars can inject more momentum into GMCs through H II regions than do the much more numerous low-mass stars through their outflows (Matzner, 2002); in addition, H II regions can inject energy on the scale of the GMC, overcoming the limitation of protostellar outflows in this regard. The dominant destruction mechanism for GMCs is via photoevaporation by blister H II regions (see below). The momentum given to a GMC by the loss of a mass $\delta M$ in a blister H II region is about $2c_\text{HI} \delta M$, where $c_\text{HI} \simeq 10$ km s$^{-1}$ is the sound speed in the H II region (the factor 2 represents the sum of the thermal and ram pressures in a blister H II region). If the overall star formation efficiency is $\epsilon_{\text{GMC}}$, then a cloud of initial mass $M$ will form a mass of stars $\epsilon_{\text{GMC}} M$ and will lose a mass $(1 - \epsilon_{\text{GMC}})M$ via photoevaporation. The ratio of the momentum given to the cloud by H II regions to that given by protostellar winds is then

$$\frac{p_{\text{HI}}}{p_w} = \frac{(1 - \epsilon_{\text{GMC}})}{\epsilon_{\text{GMC}}} \left( \frac{2c_\text{HI}}{60 \text{ km s}^{-1}} \right) \simeq \frac{1 - \epsilon_{\text{GMC}}}{3\epsilon_{\text{GMC}}},$$

where we used the value $p_w/m_\star \simeq 60$ km s$^{-1}$ from Richer et al. (2000); if protostellar momentum injection is less efficient than this, then H II regions are
correspondingly more important. (Note that if the cloud is disrupted before it
is photoevaporated, as can happen to smaller GMCs—see below—both the mass
loss and the star formation will be reduced, but \( p_{\text{H II}}/p_w \) will be relatively unaf-
fected.) For GMCs, which have star formation efficiencies \( \epsilon_{\text{GMC}} \sim \) a few percent,
\( \text{H II} \) regions dominate the energy injection by an order of magnitude. Large OB
associations do not occur in small molecular clouds, so photoevaporation is less
important in such clouds; Matzner (2002) estimates that protostellar outflows
dominate the energy injection for clouds with masses \( \lesssim 4 \times 10^4 \, M_\odot \).

The same \( \text{H II} \) regions that inject energy to support GMCs also destroy them.
Blitz & Shu (1980) estimated that \( \text{H II} \) regions would inject enough energy to
unbind a GMC in about \( 10^7 \) yr. Using a simple analytic model for the evolution
of blister \( \text{H II} \) regions developed by Whitworth (1979) and confirmed in numerical
simulations by Yorke et al. (1989), Williams & McKee (1997) concluded that an
expanding \( \text{H II} \) region would sweep up more mass than it ionized, so that a very
large \( \text{H II} \) region would disrupt the cloud before ionizing it. The condition for
cloud disruption—that the cloud be engulfed by the \( \text{H II} \) region—is difficult to
meet for massive GMCs (\( \gtrsim 10^6 \, M_\odot \)), but relatively common for small GMCs.
Indeed, the Orion molecular cloud may well have been disrupted by previous
generations of star formation. If one assumes that the specific star formation
rate is independent of GMC mass, then the observed rate of star formation in
the Galaxy implies that GMCs with masses \( \sim 10^6 \, M_\odot \) live for about 30 Myr.
Matzner (2002) assumed that the star formation rate in a cloud is self-regulating
and found that the specific star formation rate increases with cloud mass; as
result, he obtained smaller lifetimes for large GMCs, \( \sim 20 \) Myr. Krumholz,
Matzner, & McKee (2006) obtained a similar value with a more complete model
for GMC evolution (see below).

Observational estimates of GMC lifetimes are difficult, although chemical clocks
can be used to measure the lifetime of clumps and cores within them (see §3.1.2).
For the LMC and M33, Blitz et al. (2007) discuss empirical measures of GMC
lifetimes based on their spatial correlation with \( \text{H II} \) regions and young clusters
and associations. About 25% of the GMCs show no evidence for high-mass star
formation. This can be interpreted as a delay in the onset of star formation
due either to the effects of magnetic fields (McKee, 1989; Tassis & Mouschovias,
2004), or to the time required to create cluster-forming clumps in a turbulent
cloud (from simulations this time is \( \sim 1 - 3 t_H \); e.g. Ostriker, Stone, & Gam-
mie, 2001; Heitsch, Mac Low, & Klessen, 2001; Krasnopolsky & Gammie, 2005;
Vázquez-Semadeni et al., 2005); alternatively, if the star formation rate under-
goes significant fluctuations (e.g., Krumholz, Matzner, & McKee, 2006), it could
simply represent a hysteresis in the star formation rate. In the LMC, about 60% of
star clusters with ages < 10 Myr are within 40 pc of a GMC, while older clusters
have no significant spatial correlation with GMCs; Blitz et al. (2007) therefore
infer that GMCs are destroyed within about 6 Myr of cluster formation. There
are also slightly more than twice as many clouds harboring small \( \text{H II} \) regions as
those containing large \( \text{H II} \) regions and clusters; they infer a lifetime of 14 Myr for
this stage. These statistics imply that the typical time interval between when an
\( \text{H II} \) region turns on and when the cloud is destroyed is \( \sim 20 \) Myr. Including the
GMCs without high-mass stars, they infer a total GMC lifetime of 27 Myr. There
are several caveats to this analysis, however. First, the GMCs are identified only
by their CO emission; since CO is more readily destroyed in the low-metallicity
environment of the LMC than in the Galaxy (van Dishoeck & Black, 1988), some
GMCs could have been missed, both in the early stages of evolution when the density is low and in the late stage when the UV flux is high. Second, the mean mass of the clouds increases in going from the starless sample and the small-H II sample to the cluster sample; in fact, the two largest clouds are in the cluster sample, and overall just 10-15% of massive \((M > 10^{5.5} \, M_\odot)\) GMCs in the LMC and M33 lack (high-mass) star formation. The naive interpretation of this secular trend is that massive GMCs (which contain most of the molecular mass) have a more rapid onset of star formation than do low-mass clouds, although this is not a unique explanation. Next, the associations are unbound and dissolve rapidly; clusters dissolve more slowly, but one still cannot assume that their relative numbers accurately track the relative lifetimes. It is not clear how these latter two issues affect the analysis. Finally, large associations could displace the clouds by the rocket effect rather than destroy them (a process termed “cloud shuffling” by Elmegreen, 1979). Nonetheless, observations of extragalactic GMCs offer the most promising avenue for getting an observational handle on GMC evolution.

How do GMCs and the star-forming clumps within them evolve in the presence of stellar and protostellar feedback? In the quasistatic approximation, a molecular cloud would be in approximate virial equilibrium. Under the additional assumption that the rate-limiting step in star formation is ambipolar diffusion, McKee (1989) found that GMCs contract under the influence of their self-gravity until the contraction is halted by stellar feedback; the equilibrium is stable and the column density is \(A_V \simeq 4 - 8 \, \text{mag}\), comparable to the observed value. Dropping the assumption of quasi-static evolution and using the time-dependent virial theorem for a spherical, homologous cloud, Matzner & McKee (1999) found that a star-forming clump undergoes several bursts of star formation. Krumholz, Matzner, & McKee (2006) extended this work by using a theory for the star formation rate that is consistent with the Kennicutt-Schmidt law (Krumholz & McKee, 2005 — see §3.4) and the full time-dependent virial and energy equations to solve for the evolution of GMCs that are large enough \((\gtrsim 10^6 \, M_\odot)\) that H II regions dominate the energy injection (Matzner, 2002). Their time-dependent integrations assume spherical, homologous evolution for each GMC, and follow the formation and dynamical effects of many individual H II regions, with stars selected from the IMF. Blister H II regions act to destroy the clouds, but they also provide a confining pressure due to the recoil associated with the mass loss. The principal results of the time-dependent models are (1) clouds live for a few crossing times, \(\sim 30 \, \text{Myr}\) for clouds with \(M \gtrsim 10^6 \, M_\odot\), in agreement with observational estimates; (2) clouds are close to equilibrium, with virial parameters \(\alpha_{\text{vir}} \simeq 1 - 2\); (3) the column density of the clouds is \(N_H \simeq 10^{22} \, \text{cm}^{-2}\), in agreement with observation; (4) large clouds are destroyed by photoevaporation, but small clouds \((M \lesssim 2 \times 10^5 \, M_\odot)\) are disrupted before half their mass is photoionized; and (5) GMCs convert \(\sim 5 - 10\%\) of their mass into stars before they are destroyed. H II regions are unable to support GMCs with columns significantly greater than \(10^{22} \, \text{cm}^{-2}\). Krumholz, Matzner, & McKee (2006) conjecture that such clouds can occur in regions in which the mean density is not much less than the density in the GMCs, so that external driving is more efficient; such conditions could occur in starbursts. Testing and extension of these cloud evolution/destruction models via full 3D numerical simulations has not yet been attempted, but development and verification of the necessary computational codes is well underway (Mellema et al., 2006; Mac Low et al., 2006; Krumholz, Stone, & Gardiner, 2006).
3.3 Core Mass Functions and the IMF

3.3.1 OBSERVATIONS OF THE STELLAR IMF AND THE CMF

How is the distribution of stellar masses, or initial mass function (IMF), established? This is one of most basic questions a complete theory of star formation must answer, but also one of the most difficult. Current evidence suggests that the IMF is quite similar in many different locations throughout the Milky Way, with the possible exception of star clusters formed very near the Galactic Center (Scalo, 1998b presents evidence for significant variations in the IMF, but Elmegreen, 1999 argues that much, if not all, of this is consistent with the expected statistical variations). The standard IMF of Kroupa (2001) is a three-part power-law with breaks at $0.08 \, M_\odot$ and $0.5 \, M_\odot$; i.e. $dN_*/d\ln m_* \propto m_*^{-\alpha}$ with $\alpha = 1.3$ for $0.5 < m_*/ M_\odot < 50$, $\alpha = 0.3$ for $0.08 < m_*/ M_\odot < 0.5$, and $\alpha = -0.7$ for $0.01 < m_*/ M_\odot < 0.08$. The slope of the IMF at $m_* > 0.5 M_\odot$ was originally identified by Salpeter (1955), who found $\alpha = 1.35$. Up to $\sim 1 M_\odot$, a log-normal functional form $dN_*/d\ln(m_*) \propto \exp[-(\ln m_* - \ln m_c)^2/(2\sigma^2)]$ provides a smooth fit for the observed mass distribution (Miller & Scalo, 1979), with Chabrier (2005) finding that $m_c \approx 0.2 M_\odot$ and $\sigma \approx 0.55$ applies both for individual stars in the disk and in young clusters; the system IMF (i.e., counting binaries as a single systems) has $m_c = 0.25 M_\odot$. Thus, the main properties of the IMF that any theory must explain are (i) the “Salpeter” power-law slope at high mass, (ii) the break and turnover slightly below $\sim 1 M_\odot$, (iii) the upper limit on stellar masses $\sim 150 M_\odot$ (Elmegreen, 2000; Figer, 2005; Oey & Clarke, 2005), and (iv) the universality of these features over a wide range of star-forming environments, apparently independent of the mean density, turbulence level, magnetic field strength, and to large extent also metallicity. Theory predicts that there should be a lower limit on (sub)stellar masses (Low & Lynden-Bell, 1976), but this has not been confirmed observationally.

Important additional information has been provided by recent mm and submm continuum surveys covering both cluster regions and larger areas in star-forming systems (e.g. Motte, Andre, & Neri 1998; Testi & Sargent 1998; Johnstone et al. 2000, 2001; Motte et al. 2001; Beuther & Schilke 2004; Reid & Wilson 2005, 2006a; Stanke et al. 2006; Enoch et al. 2006; Nutter & Ward-Thompson 2007). Within continuum maps, high-density concentrations representing (starless) cores have been identified in sufficient numbers (and with sufficient resolution) that core mass functions analogous to the IMF can be defined. Similar core mass functions (CMFs) may be derived using extinction data from well-sampled maps (Lada, Alves, & Lombardi, 2007), and molecular line maps in high-density tracers (Onishi et al., 2002). Studies of the CMF using extinction maps are only just beginning, but they promise to be very important given the lower systematic errors that are possible with this method. An excellent recent summary of the statistical properties of observed cores is given by Ward-Thompson et al. (2007).

The CMFs derived from many independent studies and methods are in good agreement with each other, and are remarkably similar in functional form to the stellar IMF. In particular, regardless of the total mass and size of the star-forming cloud, and regardless of whether cores are well separated or highly clustered, the high-end CMF (above $1 M_\odot$) is consistent with a power law. Applying a uniform analysis to data from 11 high- and low-mass star-forming regions, Reid & Wilson (2006b) find $\alpha = 0.8 - 2.1$, with the mean value $\alpha = 1.4$. Observed CMFs for relatively nearby clouds in the references cited above also show a peak and
turnover at low mass in the range $\sim 0.2 - 1 \ M_\odot$. For distant clouds, the peak core mass is larger, but lack of resolution and hence low-mass incompleteness affects these results. Observed well-resolved CMF distributions are thus very similar to the stellar IMF, but shifted to higher mass by a factor of a few. For CMFs derived from mm and submm observations, this factor involves some uncertainty associated with conversion from dust emissivity to total mass. The CMF derived from extinction in the Pipe nebula, which is not subject to this uncertainty, is shifted to higher mass by a factor of 3 with respect to the standard stellar IMF (Alves, Lombardi, & Lada, 2007).

The mirroring of the “universal” IMF by the (possibly also universal) CMF suggests that the stellar mass distribution is imposed early in the star-forming process. The final mass of a star appears to be controlled by the available reservoir of the core from which it forms, rather than, for example, being defined by a termination of accretion due to internal stellar processes. The shift of the observed CMF relative to the IMF nevertheless implies that stellar feedback and other processes in the collapse or post-collapse stage affect stellar masses. Magnetized protostellar disk winds are believed to reduce the stellar mass compared to core mass by a factor of a few (see discussion in §4.2.6). In particular, Matzner & McKee (2000) predicted that the efficiency of a single star formation event in an individual core is $\epsilon_{\text{core}} \approx 0.25 - 0.7$, depending on the degree of flattening, which is comparable to the values implied by the observations cited above. Because the efficiency is not sensitive to the parameters involved, this implies a similar shift from CMF to IMF at all masses. Given the uncertainty in the CMF normalization, the inefficiency of single star formation may account for essentially the whole CMF $\rightarrow$ IMF shift. Some further fragmentation of presently-observed massive cores during their collapse may also occur, but provided that the majority of the mass goes into a single object, this will leave the high-mass end of the CMF relatively unchanged. Since the CMF is already dominated by low-mass cores (by mass as well as by number), the addition of low-mass stars formed as fragments from collapsing high-mass cores would negligibly affect the low-mass end of the IMF.

Molecular line observations of low-mass cores, whether found in isolation (as in Taurus) or in close proximity to other cores in a dense, cluster-forming clump (as in $\rho$ Oph), show that these cores have very low nonthermal internal velocities (Andre et al, 2006). Since weak internal turbulence implies that little density substructure is present within these cores, they are unlikely to undergo subsequent fragmentation during collapse, except to form binaries. The low-mass portion of the CMF should therefore be conserved in mapping to the IMF, modulo mass removal by outflows. Although cores in the high-mass end of the CMF are turbulent and thus in principle subject to further fragmentation, the agreement between the CMF and IMF suggests that this is not a dominant effect.

The environments of observed prestellar cores provide further clues to the processes involved in their formation. Most stars form in clusters (Lada & Lada, 2003; see §4.3.5) and correspondingly, most (starless) molecular cores are part of larger cluster-forming dense clumps. These cluster-forming clumps$^2$, as observed for example in Ophiuchus, Serpens, Perseus, and Orion, have supersonic internal turbulent linewidths (even though the individual cores within them are subthermal). Compared to isolated cores, the cores in clusters tend to be more compact.

$^2$Referred to as “cluster-forming cores” by Ward-Thompson et al. (2007)
in overall size and have higher densities and column densities; they are also lower in mass (Ward-Thompson et al., 2007). The column densities of cluster-forming clumps are generally quite large, and in particular they exceed the mean column densities of the GMCs in which they are formed. In Perseus, where 80% of the mm cores lie in groups and 50% are in clusters (Enoch et al., 2006), 50% of the total cloud mass is at $A_V < 4$ and 80% is at $A_V < 6$, whereas 90% of the mass in prestellar cores is in larger structures that have $A_V > 6$, and 50% is at $A_V > 8$ (Kirk, Johnstone, & Di Francesco, 2006). Similarly in Ophiuchus, the prestellar cores are found in high-column density regions ($A_V > 15$ for > 90% of the core mass), while most of the cloud’s mass has much lower column densities (70% is at $A_V < 7$) (Johnstone, Di Francesco, & Kirk, 2004). The prestellar cores themselves represent only a tiny fraction of the total cloud mass: 5% in Perseus (Enoch et al., 2006), and 3% in Ophiuchus (Johnstone, Di Francesco, & Kirk, 2004); this is comparable to the net observed star formation efficiency over the lifetime of a GMC (see §3.4). On the largest scales, GMCs generally consist of collections of filaments, and both the clusters of cores and most of the individual isolated cores are embedded in these filaments. The structure formation that produces cores, and eventually stars, is therefore clearly a hierarchical process.

3.3.2 THEORETICAL PROPOSALS AND NUMERICAL RESULTS

Many theories have been proposed that aim to explain the IMF or some aspect of it, and more recently to explain the CMF as well (see Elmegreen 2001 and Bonnell, Larson, & Zinnecker 2007 for recent reviews). While none of the proposals to date have won general acceptance, several have introduced elements that are likely to be important in the eventual theory that is developed. Numerical simulations have been valuable in demonstrating that the general characteristics of observed CMFs arise naturally in turbulent, self-gravitating flows, and they have also been useful in testing certain specific hypotheses. However, many features that are seen in the simulations are not yet understood in a fundamental sense, and limited numerical resolution may affect some existing results.


Other simulations have investigated the effects of non-isothermal equations of state. Li, Klessen, & Mac Low (2003) analyzed driven-turbulence SPH simulations that used a range of polytropic indices. Bate, Bonnell, & Bromm (2003) and Bate & Bonnell (2005) analyzed the results of SPH decaying-turbulence simulations with a switch from isothermal to $T \propto \rho^{0.4}$ at density $10^{-13}$ g cm$^{-3}$, to represent the transition from optically-thin to -thick conditions. Jappsen et al.
(2005) and Bonnell, Clarke, & Bate (2006) investigated the result of switching from a weakly cooling to weakly heating barotropic equation of state at a range of densities $n \sim 10^4 - 10^7 \text{cm}^{-3}$, using driven- and decaying-turbulence SPH simulations, respectively.

The distributions obtained by applying clump-finding techniques to simulation “data cubes” share many characteristics, generally showing clump mass functions dominated by the low-mass end and tails at high mass that are (marginally) consistent with a power laws having indices similar to the Salpeter value. In detail, however, the mass functions depend on the adopted clump-finding algorithm and on parameters such as density threshold levels and smoothing scales, as well as on physical properties including the Mach number and the history of a system.

In many simulations that include self-gravity, the high-end slope tends to become shallower over time, as massive objects grow larger. This change in slope may not represent realistic evolution, if massive condensations in fact should undergo fragmentation that the simulations do not follow. Failure to capture fragmentation during collapse could affect results from either grid-based or SPH simulations. Fragmentation is seeded by turbulence, which imposes fluctuations in the density (Sasao, 1973). These fluctuations grow as a condensation collapses, and in principle could ultimately result in fragmentation if they become locally Jeans unstable (Hunter, 1962; Lynden-Bell, 1973). Fluctuations at smaller and smaller mass scales would grow to be highly nonlinear if collapse were to proceed unchecked, so in order to capture this numerically, turbulence at scales below the sonic scale would have to be resolved before collapse commences in a given region. However, it is likely that in reality fragmentation of collapsing high mass condensations is prevented by real physical effects, rather than numerical effects: accretion onto stars formed early in the collapse process heats the surrounding gas significantly, which helps limit further fragmentation (Krumholz, 2006), and for condensations in cluster-forming regions, outflows from nearby stars inject small-scale turbulence that may provide support sufficient to prevent rapid localized collapse (Tan, Krumholz, & McKee 2006; Nakamura & Li 2007; see also Klessen 2001). Until physical processes enter to limit further breakup of massive condensations during their evolution, self-similarity implies they would fragment due to the same turbulent processes that produced the massive condensations in the first place. This is presumably why the IMF in clusters is the same as that in distributed star formation. Observed dense clumps break up into individual small cores when imaged at high resolution, suggesting that much of the fragmentation is in place prior to collapse. Indeed, the fact that the turnover of the IMF is similar to the Jeans or Bonnor-Ebert mass evaluated at the mean turbulent kinetic pressure within GMCs (see §3.1), $P_{\text{kin}}/k_B \sim 3 \times 10^5 \text{ K cm}^{-3}$, suggests that the ambient pressure that sets the “typical” star’s mass is not significantly increased above this level by collapse prior to fragmentation. However, note that the correspondence between the turnover in the IMF and the Bonnor-Ebert mass evaluated at the mean turbulent pressure appears to break down in star clusters like the Orion Nebula Cluster (Hillenbrand & Hartmann, 1998) and globular clusters (Paresce & De Marchi, 2000), which are believed to have formed at substantially greater pressures (McKee & Tan, 2003); the reason for this is not clear.

A common numerical “shortcut” to studying cluster formation is to focus on just a single cluster, rather than the whole hierarchical system; this allows the collapse and fragmentation to be better resolved. Models of this kind initiate
a simulation at high density with comparable internal turbulent and gravitational energy. However, this approach misses an aspect of the real situation which may be quite important: self-gravitating massive condensations develop out of non-self-gravitating gas in which perturbations have already been imposed by turbulence. In simulations where the initial kinetic energy does not exceed the gravitational energy, collapse occurs before the turbulence is able to imprint a realistic density structure on the system, such that the subsequent fragmentation may also be unrealistic. In particular, this may lead to massive fragments continuing to grow over time as they capture low-turbulence unstructured gas from their surroundings (“competitive accretion” — see §4.1.2). To obtain a reliable measure of the high-end CMF from numerical models, it will be necessary to perform simulations that include large scales as well as cluster scales, and adequately resolve massive condensations both prior to and during collapse. In addition, physical processes representing the feedback from star formation must be properly included in order to impose realistic limits on fragmentation, coalescence, and accretion after collapse begins.

Another feature of numerical simulations that is at least qualitatively in accord with observations is the presence of a resolved peak and turnover in the CMF. Exactly how the location of this peak depends on model parameters, however, is not well yet determined. In some simulations, the CMF peak is found to be at masses comparable to the initial Jeans mass of the system (these are primarily low-Mach-number simulations), while in other simulations the peak is at much lower mass (these are primarily at high Mach number). The turbulent power spectrum can also affect the position of the CMF peak, and in some simulations the peak is seen to move to larger mass over time. In fact, the position of the peak for an isothermal simulation with a fixed turbulence scaling law must be a function of two dimensionless parameters, the ratio of the total mass in the system to the initial Jeans mass, and the turbulent Mach number on the largest scale. For magnetized simulations, an additional parameter is the ratio of mass to the magnetic critical mass. Although limited dependence on parameters has been explored, a comprehensive and controlled study has not yet been performed. Note that the mass-weighted density in a turbulent system increases as the turbulent Mach number increases (see §2.1.4), so that the Jeans mass at the “typical” (mass-weighted) cloud density decreases as the turbulence level increases, for a given mean (volume-weighted) density and Jeans mass. This probably accounts for why the peak of the CMF was found to be far below the mean Jeans mass in studies with high $M$, and close to the mean Jeans mass in studies with lower $M$.

A recurrent theme in star formation theory is that the characteristic mass – defined by the peak of the IMF – is the Jeans mass at some preferred density. An upper limit on the preferred density, and hence a lower limit on the fragment mass, is the value at which which the optical depth is unity over a Jeans length; this yields a minimum fragment mass $\approx 0.007 M_\odot$ (Low & Lynden-Bell, 1976). More recently, Larson (2005) has argued that the thermal coupling of gas to dust at densities above $n_H = n_c \approx 10^6 \text{ cm}^{-3}$ results in a shift from weakly-decreasing to weakly-increasing temperature as a function of density ($T \propto \rho^{-0.27}$ changes to $T \propto \rho^{0.07}$ at $T_{\text{min}} \approx 5 \text{ K}$), and that the Jeans mass $\rho \approx 0.3 M_\odot$ at this inflection point sets the preferred mass scale in the IMF. Part of Larson’s argument is that if structure is filamentary, then the filaments will contract radially while $\gamma < 1$; fragmentation into protostellar cores would occur when the filament’s central density reaches $n_c$ and $\gamma$ exceeds unity. This argument does not take into
account, however, that the mass per unit length of a filament may be determined primarily by the turbulence which originally creates it. In this case, the density $n_c$ defines a Jeans length (see eq. 20 in §2.2), so that the mass scale that emerges would be set by this (fixed) length scale $\sim \lambda_J(n_c)$ times the (variable) filament mass per unit length. The simulations of Jappsen et al. (2005), which vary the density $n_c$ at which the temperature minimum occurs, provide qualitative support for Larson’s proposal in that the peak of the CMF moves to lower mass as $n_c$ increases. The scaling of peak mass with $n_c$ in the simulations is not consistent with the predicted $m_{\text{peak}} \propto n_c^{-0.95}$ scaling, however. In addition, these models did not test dependence on other parameters that may be important, such as the Mach number of the turbulence or the total mass of the system.

A recent comprehensive proposal to explain the CMF and IMF has been advanced by Padoan & Nordlund (2002, 2004) (hereafter “PN”). They argue that because the strength of any given compression (in a shock) is related to its corresponding (pre-shock) spatial scale $\ell$, a power-law turbulence spectrum $|v(\ell)| \propto \ell^q$ will result in a distribution of clump masses that itself follows a power law. In particular, they propose that the clump mass function produced by turbulence in a magnetized medium will obey $dN(m)/d\ln m \propto m^{-3(3-2q)}$. They further propose that at a given mass $m$, the fraction of clumps created by turbulence that collapse is obtained by integrating the density PDF down to the density at which that mass would be Jeans unstable, i.e. $\rho_{\text{min}} = \pi^3\sigma_{\text{th}}^6/(36m^2G^3)$. With this prescription, at high mass the limit of the integral $\rho_{\text{min}} \rightarrow 0$ and PN find $\alpha = 3/(3-2q) \approx 1.4$. The position of the CMF peak would depend on the properties of the density PDF; for a log-normal PDF ($f_M$; see eq. 5 in §2.1.4) with $\mu_x = 0.5 - 2$, the peak mass would be between $0.8 - 0.1$ times the Jeans mass at the mean (volume-weighted) density in the cloud.

The proposal of PN is attractive in its overall thrust, and analysis of numerical simulations (Padoan et al., 2007) shows promising consistency with some of the model predictions, such as a steepening of the CMF (larger $\alpha$) with steeper velocity power spectrum (larger $q$). The PN proposal, however, also suffers from missing links in its theoretical underpinnings: (1) The effective value of $v_A$ is defined by PN such that the typical compression $\rho'/\rho$ in a shock moving at $v$ is a factor $v/v_A$ (in fact, compression factors depend on the magnetic field direction as well as strength). This effective $v_A$ is assumed to be independent of scale, and for numerical comparisons with data they adopt a value small compared to the typical value in a GMC of $\sim 2$ km s$^{-1}$. (2) The argument used to obtain $\alpha = 3/(3-2q)$ for turbulent clumps assumes that each pre-shock volume $\ell^3$ maps to a number of post-shock volumes of mass $m$ that is independent of $\ell$; i.e. $\ell^3N(m)/L^3 = \text{const}$. While this is plausible, other arguments can be made that draw on the scale-free nature of turbulence, yet yield different results. Fleck (1996) and Elmegreen & Falgarone (1996) have argued that in non-self gravitating turbulence one obtains $mN(m) = \text{const}$. This scaling corresponds to converting a constant fraction of the mass or volume behind each shock into clumps, $\ell^3\rho N(m)/(L^3\rho) = \text{const.}$, where $\ell' = \ell v_A/v = \ell p/p'$ is the post-shock scale. One might also propose that the filling factor of post-shock clumps within the whole volume should be constant, i.e. $\ell^3 N(m)/L^3 = \text{const}$. This leads to $N(m) \propto m^{-(3-3q)/(3-2q)}$, or $\alpha = 0.75$ for $q = 1/2$. While the assumption $\ell^3 N(m)/L^3 = \text{const}$ in the PN formulation yields results that are in agreement with measured CMFs, a physical argument is needed to explain why this is the correct choice among several plausible alternatives. In particular, since PN’s argument for the slope
\( \alpha = 3/(3 - 2q) \) involves only turbulence, why does this value of \( \alpha \) disagree with the distinctly-shallower empirical mass spectrum (\( \alpha \sim 0.5 \); see §3.1) of non-self-gravitating clumps in GMCs? (3) The argument PN use to obtain a formula for the mass function does not appear to take account of substructure within clumps at a given mass scale although the presence of substructure is implicit in their picture. In particular, they assume that any region that is unstable by the thermal Jeans criterion will collapse. An implicit requirement for this is that at each density, a contiguous volume containing a mass in excess of the Jeans mass is present. More generally, since hierarchical density structures are clearly important in nature (most cores and stars are clustered), any fundamental theory should identify how this this comes about. Given these difficulties, it appears premature to accept the PN proposal in its current form, although it is promising as a basis for future development.

### 3.4 The Large-Scale Rate of Star Formation

Much of this review focuses on the detailed physical processes of star formation at and below GMC scales. To understand the structure of a given galaxy, however, or the evolution of a population of galaxies over cosmological timescales, often only a very gross characterization of the star formation processes – such as the overall star formation rate (SFR) and the resulting IMF – is adequate. Many empirical studies of disk galaxies characterize the SFR in terms of the number of stars formed per unit time per unit area \( \Sigma_\ast \); this is usually reported using either averages over the whole of a galaxy within some outer radius \( R \), or using azimuthal averages over an annulus of width \( dR \) to give \( \Sigma_\ast(R) \). Both of these methods average over regions that may have widely varying SFRs, and the results must be carefully interpreted as strong nonlinearities are involved. Fortunately, with the data becoming available from large-scale galactic mapping surveys (e.g. SONG and SINGS; Helfer et al., 2003; Kennicutt et al., 2003), it will soon be possible to characterize SFRs on scales large compared to individual GMCs but small enough to separately measure, e.g., SFRs for arm and interarm regions.

More fundamental than \( \Sigma_\ast \) is the star formation or gas consumption timescale. This is defined by \( t_{g\ast} \equiv \Sigma_g/\Sigma_\ast = M_g/M_\ast \), where \( \Sigma_g \) is the gas surface density; the second equality assumes that the same area average is used for the total gas mass \( M_g \) and star formation rate \( M_\ast \). The resulting timescale depends on the gas tracer(s) chosen, which determines the range of gas densities included in \( \Sigma_g \). For a chemical species tracing gas in a class of structures denoted by \( S \) that have mean internal gas density \( \langle \rho \rangle_V = \rho_S \), and total mass \( M_S \), a convenient fiducial time for comparison to \( t_{S\ast} \equiv M_S/M_\ast \) is the free-fall time obtained by using \( \rho_S \) in equation (14). The star formation or gas consumption rate is then

\[
\dot{M}_\ast \equiv \varepsilon_{f,H,S} \frac{M_S}{t_{ft,H,S}},
\]

where the efficiency over a free-fall time is \( \varepsilon_{f,H,S} = t_{ft,H,S}/t_{S\ast} = \Delta M_\ast(t_{ft,H,S})/M_S \) (see Krumholz & McKee, 2005 and Krumholz & Tan, 2007; note that they denote this quantity by \( \text{SFR}_{f,H,S} \)). Note that the star formation efficiency \( \varepsilon_{f,H,S} \) over the free-fall time at the mean local density of structures \( S \) differs from the star formation efficiency \( \varepsilon_S \) over the mean lifetime of individual structures in class \( S \), which is discussed below. The definition in equation (38) is particularly useful for describing star formation on local scales within GMCs, in which different
molecular transitions trace a relatively limited range of densities, and in which densities can be obtained for individual structures that are spatially resolved and have mass measurements from molecular lines, dust continuum or extinction observations. Since most of the molecular gas, and essentially all star formation, is found within GMCs, the SFR in a region with local surface density in GMCs, $\Sigma_{\text{GMC}}$, is given by $\dot{M}_s = \epsilon_{\text{ff,GMC}} \Sigma_{\text{GMC}} / t_{\text{ff,GMC}}$. Here, $t_{\text{ff,GMC}}$ is calculated using the free-fall time within GMCs in a given region. Typical mean densities within GMCs are $n_H \sim 100 \text{ cm}^{-3}$, implying $t_{\text{ff,GMC}} \sim 4 \text{ Myr}$, but this may vary due to the effects of spiral arms, for example. In clumps or cores within GMCs, $\rho_S$ can be larger by orders of magnitude compared to $\rho_{\text{GMC}}$, yielding a corresponding decrease in $t_{\text{ff}}$.

The values of $\epsilon_{\text{ff}}$ are generally low ($\lesssim 0.01$) over a wide range of density tracers (see below), and vary only weakly with $\rho_S$. Krumholz & Tan (2007) point out that this suggests that turbulence is limiting star formation, although magnetic regulation is also possible (but probably not on GMC scales, since they appear to be magnetically supercritical – §2.3). In the turbulence-regulation picture, the low overall efficiency of star formation on GMC scales (over their own free-fall times) is dictated by the low fraction of gas that concentrates into structures that are sufficiently dense to collapse before being dispersed by turbulence. The weak variation of $\epsilon_{\text{ff}}$ with density follows naturally if the density obeys a log-normal distribution, which is consistent both with numerical simulations of supersonic turbulent flows and with observations of extinction statistics (§§2.1.4,2.1.5). For a log-normal distribution defined by equation (5), let $M_S$ be the mass with densities in the range $\delta x_S$ surrounding $x_S \equiv \ln(\rho_S/\bar{\rho})$, where $\bar{\rho} \equiv \langle \rho \rangle_V$ is obtained by dividing total GMC mass by total GMC volume. Then since the star formation rate $\dot{M}_s$ is independent of tracer,

$$\frac{\epsilon_{\text{ff}}}{\epsilon_{\text{ff,GMC}}} = \frac{\sqrt{4\pi \mu_x}}{\delta x_S} \exp \left[ \frac{(x_S - 2\mu_x)^2}{4\mu_x} - \frac{3}{4} \frac{\mu_x}{x_S} \right],$$

where $\mu_x \equiv \langle \ln(\rho/\bar{\rho}) \rangle_M$ is the (mass-weighted) mean within a GMC. With $\mu_x \sim 1.5$ [corresponding to mass-weighted mean density $\langle \rho/\bar{\rho} \rangle_M = \exp(2\mu_x) \sim 20$, typical of GMCs] and $\delta x_S \sim 1$, $\epsilon_{\text{ff}}/\epsilon_{\text{ff,GMC}}$ is between 1.4 and 3.4 for $1 < \rho_S/\langle \rho \rangle_M < 10$; larger values of $\mu_x$ keep $\epsilon_{\text{ff}} \sim \epsilon_{\text{ff,GMC}}$ over a larger range of densities. Approximate constancy of $\epsilon_{\text{ff}}$ over a range of densities implies the approximate relation $M_S \propto t_{\text{ff}} \propto \rho_S^{-1/2}$ from equation (38). Physically, this is because the equilibrium fraction of mass in a GMC in structures at densities significantly above $\bar{\rho}$ is equal to the ratio of the destruction time to the formation time of those structures, $t_{\text{dest},S}/t_{\text{form},S}$. In a turbulent medium, the destruction time is of order the dynamical time of the structure, which decreases with increasing density and decreasing size, whereas the formation time is of order the dynamical time of the GMC for all structures, since the large-scale flow dominates. Note, however, that this discussion does not apply to regions in which self-gravity is strong but turbulence is weak, as occurs in low-mass pre-stellar cores. In such cores, $\epsilon_{\text{ff}}$ rises by an order of magnitude or more to $\sim 0.1$. Quiescent cores have individual lifetimes of a few $t_{\text{ff}}$ (see §3.1.2), and net efficiency of star formation in each core $\sim 1/3$ (see §3.3.1 and §4.2.6). These structures have evolved to have internal densities (and hence self gravity) high enough that they can resist destruction by the ambient turbulence. In regions such as forming clusters, where self-gravity causes strong departures from the overall log-normal density distribution in GMCs and high
gravity is offset by locally-driven turbulence, the relation (39) would also not be expected to apply.

Even within a given density regime, there may be significant cloud-to-cloud variations in local conditions such that \( t_S \) need not be a universal quantity even for structures observed in a given tracer. Indeed, Mooney & Solomon (1988) showed that for Milky Way GMCs with virial masses (traced in CO) \( M_{\text{CO}} = 10^4 - 5 \times 10^6 \, M_\odot \) and infrared luminosities \( L_{\text{IR}} \propto M_* \), the ratio \( t_{\text{GMC},*} \propto M_{\text{CO}}/L_{\text{IR}} \) varies over two orders of magnitude and is not correlated with \( M_{\text{CO}} \). Williams & McKee (1997) came to a similar conclusion from their analysis of OB associations and GMCs in the Galaxy. With a total GMC mass \( \sim 10^9 M_\odot \) in the Galaxy and a star formation rate of several \( M_\odot \, \text{yr}^{-1} \), the mean value of \( t_{\text{GMC},*} \approx 3 \times 10^8 \, \text{yr} \), which translates to \( \epsilon_{\text{GMC}} \sim 0.01 \) if \( n_H \sim 100 \, \text{cm}^{-3} \) in GMCs. For dense gas clumps in GMCs, however, it appears that conditions are more uniform, such that there is less scatter in \( t_S \) for dense gas tracers. In particular, Wu et al. (2005) show that the ratio \( L_{\text{HCN}}/M_{\text{dense clumps}}/M_* \) measured in Milky Way star-forming regions agrees with the same values measured in high-redshift galaxies (Gao & Solomon, 2004), for which there is only one order of magnitude scatter. Wu et al. (2005) estimate a corresponding star formation timescale of \( t_{\text{HCN},*} = 8 \times 10^7 \, \text{yr} \). If the typical density of HCN-emitting gas is \( \sim 10^5 \, \text{cm}^{-3} \), the corresponding efficiency per free-fall time is \( \epsilon_{\text{HCN}} \sim 0.002 \). Krumholz & Tan (2007) apply slightly different factors to convert total HCN and IR luminosities to gas masses and star formation rates, and obtain \( \epsilon_{\text{HCN}} \sim 0.006 \). These values of \( \epsilon_{\text{G}} \) are small compared to those for individual cores (\( \sim 0.1 \)), which in clustered regions (where most stars form) have densities \( \sim 10^7 \, \text{cm}^{-3} \) (Ward-Thompson et al., 2007) that are large compared to the densities traced by HCN.

In spite of the large scatter in \( t_{\text{G}} \) from one local region to another (in various density tracers), empirical studies have shown that when averaged over large scales, \( t_{\text{G}} \) is correlated with the global properties of gas in a galaxy. The early studies of Schmidt (1959, 1963) sought to characterize the star formation rate as a power law (with index \( > 1 \)) in the mean gas density (both volume and surface density); this would then translate to \( t_{\text{G}} \) (or \( t_{\text{G}}/\epsilon_{\text{G}} \)) that varies as a negative power of gas density. More recently, following Kennicutt (1989), a number of empirical studies of disk galaxies have identified and explored “Kennicutt-Schmidt” (or KS) laws of the form \( \Sigma_* \propto \Sigma_g^{p+1} \), for which \( t_{\text{G}} \propto \Sigma_g^{-p} \). The original study of Kennicutt investigated correlations of \( \Sigma_*(R) \) (based on \( H\alpha \)) with the total \( \Sigma_g(R) \) (including both atomic and molecular gas); he found an index \( p = 0.3 \) for \( \Sigma_g(R) \) above a threshold level. Kennicutt (1998) studied correlations of global averages of \( \Sigma_\ast \) with \( \Sigma_g \) (again combining atomic and molecular gas). For the whole sample including normal galaxies, the centers of normal galaxies, and starbursts, the fitted index is \( p = 0.4 \); the index is slightly larger for just normal spirals. Recent years have seen a number of other studies of the \( \Sigma_g - \Sigma_* \) relationship based on annular averages in galaxies, using \( H\alpha \), radio continuum, or far-IR to measure star formation, and using either the total gas surface density or just the molecular gas contribution from CO observations (Wong & Blitz, 2002; Murgia et al., 2002; Boissier et al., 2003; Heyer et al., 2004; Komugi et al., 2005; Schuster et al., 2007). Most of these studies have found \( p \) in the range \( 0.3 - 0.4 \), although larger values of \( p \) have been obtained in some analyses that include both atomic and molecular gas. For dense gas as traced by HCN, Gao & Solomon (2004) found a linear relationship between the integrated star formation rate and the total mass
of dense gas, i.e. $p = 0$, based on a sample including both normal galaxies and luminous/ultraluminous IR galaxies. For the same sample, the fitted SFR-gas mass index is $p = 0.7$ for less-dense molecular gas observed in CO lines. All of these fits involve (at least) an order of magnitude scatter about the mean relation. Taken together, these results imply that the amount of dense gas available for star formation increases nonlinearly with the global amount of lower-density gas, but that the star formation rate in this dense gas is independent of global galactic properties.

A second approach to characterizing global SFRs introduces the global timescale associated with the galaxy, the orbital period $t_{\text{orb}} = 2\pi/\Omega$. For grand design spirals, the SFR is expected to be proportional to the rate at which gas passes through spiral arms, since GMCs are expected (and observed) to form rapidly in the high-surface-density gas behind the spiral shock (e.g. Roberts 1969; Kim & Ostriker 2002, 2006; Shetty & Ostriker 2006). Shu (1973) appears to have been the first to propose this idea, and showed that it is roughly consistent with observations of star formation in the Galaxy. Wyse (1986) proposed that GMCs, and hence stars, are the result of atomic cloud-cloud collisions at a rate $\propto \Sigma_H^2(\Omega - \Omega_p)$, where $\Omega_p$ is the pattern speed. More generally, Wyse & Silk (1989) suggested that the star formation rate should scale as $\dot{\Sigma}_g \propto \Sigma_g \Omega$. This has been confirmed by Kennicutt (1998); the resulting two forms for the KS law are

$$\dot{\Sigma}_g = 0.017\Sigma_g \Omega \simeq (2.5 \pm 0.7) \times 10^{-4} \left(\frac{\Sigma_g}{1 \, M_\odot \, \text{pc}^{-2}}\right)^{1.4 \pm 0.15} M_\odot \, \text{yr}^{-1} \, \text{kpc}^{-2}. \quad (40)$$

The fact that there are two forms of the star formation law implies that there is a correlation between $\Sigma_g$ and $\Omega$; Krumholz & McKee (2005) found $\Omega \propto \Sigma_g^{0.49}$ for a sample comprised of both normal and starburst galaxies (Kennicutt, 1998; Downes & Solomon, 1998). The reason for this correlation is not known at present, but may be related to an overall tendency for velocity dispersions to increase at large surface densities (see below). The corresponding gas consumption time is $t_{g*/\text{t}_{\text{orb}}} \approx 10$ with $t_{g*}$ evaluated for the entire galaxy and $t_{\text{orb}}$ evaluated at the outer edge of the star formation. Subsequent observations have found $t_{\text{mol}*/\text{t}_{\text{orb}}} \sim 10 - 100$ when considering the molecular gas alone (Wong & Blitz, 2002; Murgia et al., 2002).

Since most star formation is observed to take place within bound GMCs, it is useful to introduce $f_{\text{GMC}} \equiv \Sigma_{\text{GMC}}/\Sigma_g$, i.e. the fraction of gas that is found in GMCs. The surface densities must be averaged over a region containing a large number of GMCs, since the specific star formation rate has very large fluctuations; the average can be over a local patch of a galaxy, an azimuthal ring, or an entire galaxy. Equation (38) implies

$$\dot{\Sigma}_g = \epsilon_{\text{eff, GMC}} \frac{\Sigma_g f_{\text{GMC}}}{t_{\text{eff, GMC}}}. \quad (41)$$

This form of the star formation law is particularly useful if most of the gas is in GMCs, $f_{\text{GMC}} \simeq 1$. Since the gas density in the midplane $\rho_g \propto \Omega^2/GQ^2$ in terms of the Toomre $Q$ parameter, and since $t_{\text{eff, GMC}}^{-1} \propto \rho_{\text{GMC}}^{1/2} \propto \rho_g^{1/2}$, it follows that $\dot{\Sigma}_g \propto \epsilon_{\text{eff, GMC}} \Sigma_g f_{\text{GMC}}/(Q t_{\text{orb}})$ (Krumholz & McKee, 2005; see below), which is similar to the “orbital time” form of the KS law.
An alternative expression for the SFR follows by noting that if GMCs form from diffuse gas at a rate $M_{\text{diffuse}}/t_{\text{diffuse}}$ and are destroyed by star formation at a rate $M_{\text{GMC}}/t_{\text{GMC}}$, then $f_{\text{GMC}} = t_{\text{GMC}}/t_{\text{lc}}$, where $t_{\text{lc}} \equiv t_{\text{diffuse}} + t_{\text{GMC}}$ is the life-cycle time for gas in the galaxy. The mean efficiency of star formation in any GMC over its lifetime $t_{\text{GMC}}$ is $\epsilon_{\text{GMC}} = \epsilon_{\text{eff,GMC}}(t_{\text{GMC}}/t_{\text{GMC}})$; in the Milky Way, the observed average value of $\epsilon_{\text{GMC}}$ is about 0.05 (e.g., Williams & McKee, 1997), corresponding to $t_{\text{GMC}}/t_{\text{GMC}} \approx 5$. The star formation rate is then

$$\dot{\Sigma}_* = \epsilon_{\text{GMC}} \frac{\Sigma_g}{t_{\text{lc}}}.$$  

(42)

This equation leads to a simple interpretation of the empirical result $\dot{\Sigma}_* \propto \Sigma_g \Omega$ for grand design spirals. In this situation the mean lifecycle time of the gas should be equal to the timescale between successive encounters with spiral arms, $t_{\text{lc}} = (2\pi/m)\Omega(\Omega - \Omega_p)^{-1}$ for an $m$-armed spiral. If $\epsilon_{\text{GMC}}$ varies only modestly with radius, then well inside corotation (which is most of the star-forming disk) the overall star formation rate should obey $\dot{\Sigma}_* \propto \Sigma_g \Omega$. More generally, consider an arbitrary disk galaxy in which the gas is primarily diffuse, so that the GMC formation time $t_{\text{diffuse}}$ is much greater than the GMC destruction time (or lifetime) $t_{\text{GMC}}$, and as a result the life-cycle time $t_{\text{lc}} \simeq t_{\text{diffuse}}$. The characteristic timescale $t_{\text{diffuse}}$ for formation of self-gravitating structures in a disk with surface density $\Sigma_g$ is the two-dimensional Jeans time $t_{\text{J,2D}} = c_{\text{eff}}/(G\Sigma_g)$ (eq. 35). (Actual GMC formation timescales differ from $t_{\text{J,2D}}$ due to rotation and disk-thickness effects – see references in §3.2.1). Using the definition of the Toomre $Q$ parameter (eq. 32), $t_{\text{orb}} = (t_{\text{J,2D}}/Q)(2\kappa/\Omega)$, so that from equation (42), $\dot{\Sigma}_* \sim 3\epsilon_{\text{GMC}} \Sigma_g (Q t_{\text{orb}})^{-1}$. Thus, both the diffuse-dominated and GMC-dominated cases yield $\dot{\Sigma}_* \propto \Sigma_g/(Q t_{\text{orb}})$, assuming that the efficiency factors are comparable in different regions of a galaxy and from one galaxy to another.

Since star formation tends to deplete the gas in any region until $Q$ is near the critical value $Q_{\text{crit}} \simeq 1.5$ (theory: Quirk, 1972; observation: Martin & Kennicutt, 2001; Wong & Blitz, 2002; Murgia et al., 2002; Boissier et al., 2003), this yields the “orbital time” form of the KS law (including the normalization) when $\epsilon_{\text{GMC}} \approx 0.05$. GMC formation on a timescale $\sim t_{\text{J,2D}}$, implying a star formation rate $\dot{\Sigma}_* \propto \Sigma_g^2/c_{\text{eff}}$ if $\epsilon_{\text{GMC}}$ is weakly-varying, also yields the other form of the KS law if the effective velocity dispersion increases with surface density according to $c_{\text{eff}} \propto \Sigma_g^{1-p}$. A significant increase in the gas velocity dispersion at small radii, where $\Sigma_g$ is generally larger, has been noted in several galaxies (Kenney, Carlstrom, & Young, 1993; Sakamoto, 1996; Walsh et al., 2002; Lundgren et al., 2004; Schuster et al., 2007), although there is no quantitative theory for this increase.

A quantitative theory for the galactic star formation rate must determine the star formation efficiency (e.g., $\epsilon_{\text{eff,GMC}}$ or $\epsilon_{\text{GMC}}$) as well as the corresponding overall rate. (An exception is the theory of Silk, 1997, who notes that the porosity of hot gas in a galaxy is determined by the SFR; the SFR can thus be expressed in terms of the porosity, but this remains uncertain.) Tan (2000) proposed that the overall star formation rate is determined by cloud-cloud collisions, but he set the efficiency based on comparison with observations. Elmegreen (2002, 2003) suggested that the star formation rate per unit volume is $\dot{\rho}_* \simeq \epsilon_{\text{core}} f_{\text{core}} (G\rho_{\text{core}})^{1/2}/\rho$, where $f_{\text{core}}$ is the fraction of gas in dense cores; the value of this was set by comparison with observation. Simulations by Kravtsov (2003), by Li, Mac Low, & Klessen (2005a,b, 2006), and by Tasker & Bryan (2006) show that the fraction of
high-density gas scales as $\Sigma_g^{1.4}$, but the definition of “high-density” is arbitrary and the dependence of the SFR on this definition is not known.

A theory for the star formation efficiency per free-fall time has been given by Krumholz & McKee (2005); for galaxies in which $f_{\text{GMC}} \approx 1$, this is proposed as a complete theory of the KS law. The first three assumptions underlying the theory have been discussed above: they assume that star formation occurs primarily in GMCs, so that the star formation rate is described by equation (41), that the density PDF in GMCs is log normal, as inferred from simulations of turbulence in gas that is approximately isothermal (§2.1.4), and that the IMF has the standard form. The final assumption is that low-mass stars form in all gas dense enough that the sonic length in the surrounding turbulent gas exceeds the Jeans length ($\lambda_J < \ell_s$) (Padoan, 1995), with an efficiency $\epsilon_{\text{core}} \approx 1/2$ from the theoretical estimate of Matzner & McKee (2000). The condition that $\lambda_J < \ell_s$ ensures that critical Bonnor-Ebert spheres are not torn apart by turbulence; the corresponding critical density implies that the thermal pressure in the cores matches the turbulent pressure in the environment, $\rho_{\text{core}}/\rho \approx M^2$. An important part of this last assumption is that the regions that are dense enough to satisfy $\lambda_J < \ell_s$ have masses large enough to collapse. The normalization for the star formation rate is based on the simulations of Vázquez-Semadeni, Ballesteros-Paredes, & Klessen (2003). These assumptions lead to a star formation efficiency $\epsilon_{\text{eff},\text{GMC}} \approx 0.017\alpha_{\text{vir}}^{-0.68}(M/100)^{-0.32}$. Since the virial parameter in GMCs is of order unity and the Mach numbers are somewhat smaller than 100 in regions where they have been observed (and $\lesssim$ a few 100 even in unresolved starbursts), this corresponds to a typical $\epsilon_{\text{eff},\text{GMC}} \approx 0.02$. For galaxies in which the gas is not fully molecular, Krumholz and McKee adopt the phenomenological result for $f_{\text{GMC}}$ obtained by Blitz & Rosolowsky (2006). They show that the resulting star formation rate agrees well with Kennicutt’s observed relations (eq. 40).

Finally, we remark that controversy continues to surround the question of what physical process defines the observed outer-disk thresholds $R_{\text{th}}$ for active star formation. Kennicutt (1989) and Martin & Kennicutt (2001) argue that disk thresholds are set by gravitational stability considerations in shearing, rotating disks, and find a mean value of $Q \approx 1.4$ when they adopt a constant value $c_{\text{eff}} = 6 \text{ km s}^{-1}$ for their sample. Numerical simulations of isothermal gas disks, including both disk thickness effects and the gravity from an active stellar disk, quantitatively support this conclusion (Li, Mac Low, & Klessen, 2005a; Kim & Ostriker, 2007). On the other hand, Schaye (2004), building on the suggestion of Elmegreen & Parravano (1994), argues that star formation thresholds are defined by the condition that the pressure is high enough that a cold component of the atomic ISM can exist. This transition point depends on the UV intensity and metallicity (e.g., Wolfire et al., 2003), but typically corresponds to threshold surface density $\approx 3 - 10 M_\odot \text{ pc}^{-2}$. Schaye essentially reverses the argument for a $Q$ threshold: he argues that when a significant fraction of the ISM becomes cold, $Q$ drops significantly and gravitational instability ensues. An advantage of this proposal is that it can naturally account for the isolated patches of star formation that occur outside $R_{\text{th}}$ (Ferguson et al., 1998; Boissier et al., 2006), since star formation occurs wherever the pressure of the gas exceeds the critical value. However, while the model gives a necessary condition for star formation, it does not give a sufficient condition: some of the galaxies in the Martin & Kennicutt (2001) sample have $R_{\text{th}}$ inside the radius at which the gas becomes molecular, which in turn is inside the radius at which cold atomic gas first appears.
(see also de Blok & Walter, 2006, who find evidence of a cold atomic component even in non-star-forming regions). In these cases, it is possible that MRI-driven turbulence in the outer disk maintains the effective $Q$ greater than the critical value even when some of the gas is cold (Piontek & Ostriker, 2007).

4 MICROPHYSICS OF STAR FORMATION

4.1 Low-Mass Star Formation

Star formation is traditionally divided into two parts: Low-mass stars form in a time short compared to the Kelvin-Helmholz time, $t_{KH} = Gm^2/RL$, whereas high-mass stars form in a time $\gtrsim t_{KH}$ (Kahn, 1974). This distinction between low-mass and high-mass protostars is not fully satisfactory, however, since for a sufficiently high accretion rate any protostar would be classified as “low-mass.” We somewhat arbitrarily divide low and high-mass stars at a mass of $8 \, M_\odot$. Protostars that will form stars with masses significantly below this value have luminosities dominated by accretion, and they form from cores that have masses of order the thermal Jeans mass. Protostars above this mass have luminosities that are dominated by nuclear burning unless the accretion rate is very high, and if they form from molecular cores, those cores are significantly above the thermal Jeans mass. Low-mass stars undergo extensive pre-main sequence evolution in the Hertzsprung-Russell diagram, from the point on the “birthline,” where they cease accreting and are revealed (Stahler, 1983; see also Larson, 1972), to the main sequence. Here we briefly review the current understanding of how such stars form.

4.1.1 Theory of core collapse and protostellar infall

As discussed above, low-mass stars appear to form from gravitationally bound cores. The time scale for the collapse of these cores determines both the time scale for the formation of a star and the accretion luminosity. Note that the rate of infall onto the star-disk system, $\dot{m}_\text{in}$, can differ from the rate of accretion onto the protostar, $\dot{m}_*$, since some of the infalling gas can be temporarily stored in the disk. The collapse of such cores and the growth of the resulting protostars has been reviewed by Larson (2003), and we draw on this work here. At the outset of theoretical studies of star formation, it was realized that isothermal cores undergoing gravitational collapse become very centrally concentrated, with a density profile that becomes approximately $\rho \propto r^{-2}$ (Bodenheimer & Sweigart, 1968; Larson, 1969). Prior to the formation of the protostar, there is a central, thermally supported region of size $r \simeq \lambda_f$. Collapse of a marginally unstable core begins near the outer radius. The $r^{-2}$ density gradient is created as the wave of collapse propagates inward, leaving every scale marginally unstable as the collapse accelerates (cf. Larson, 2003). That is, since $\lambda_f \sim c_s/(G\rho)^{1/2}$ (eq. 20), a sphere that is marginally unstable at each scale, $r \sim \lambda_f$, will have $\rho \sim c_s^2/(Gr^2)$ when the protostar is first formed; the corresponding infall rate is

$$\dot{m}_\text{in} \sim \frac{M_G}{t_G} = \frac{c_s^2}{G} \Rightarrow \dot{m}_\text{in} = \phi_\text{in} \frac{c_s^3}{G},$$

(43)

where the gravitational mass and radius are defined in equation (13) and $\phi_\text{in}$ is a numerical factor that is typically $\gtrsim 1$. (When the effect of protostellar outflows is included, the infall rate is reduced by a factor $\epsilon_\text{core} < 1$.) Although this result was first derived for an isothermal sphere, Stahler, Shu, & Taam (1980)
emphasize that it should apply approximately to the collapse of any cloud that is initially in approximate hydrostatic equilibrium, with \( c_s^2 \approx c_{\text{eff}}^2 = c_s^2 + v_{\text{turb}}^2 \) including the effects of magnetic fields and turbulence as well as thermal pressure; Shu, Adams, & Lizano (1987) suggest that \( c_{\text{eff}} \lesssim 2c_s \), however. This infall rate explicitly depends only on the sound speed, but it implicitly depends on the density of the core: since the core was assumed to be initially in hydrostatic equilibrium, equation (43) is equivalent to
\[
\dot{m}_{\text{in}} \sim \frac{M_{\text{core}}}{t_G} \propto M_{\text{core}} P^{1/2}.
\]

There are two limiting cases for the gravitational collapse of an isothermal sphere. In the first case, originally considered by Larson (1969) and Penston (1969) and extended by Hunter (1977), one begins with a static cloud of constant density and follows the formation of the \( r^{-2} \) density profile. At the time when the protostar first forms (i.e., when the central density reaches infinity in this idealized calculation), the collapse is highly dynamic, with an infall velocity of \( 3.3c_s \). The infall rate onto the star is large, rapidly increasing from \( \dot{m}_{\text{in}} = 29c_s^3/G \) at the moment of protostar formation to \( \dot{m}_{\text{in}} = 47c_s^3/G \). In the physically unrealistic case of an infinite, uniform medium, the accretion rate would remain at this high value; in practice, the accretion rate rapidly declines after the formation of a point mass (see below). In the opposite case, considered by Shu (1977), one assumes that the evolution to the \( r^{-2} \) density profile is quasi-static (most likely due to the effects of magnetic fields—see below), so that the infall velocities are negligible at the moment of protostar formation. The resulting initial configuration is the singular isothermal sphere (SIS), which is an unstable hydrostatic equilibrium. The collapse is initiated at the center, and the point at which the gas begins to fall inward propagates outward at the sound speed (the “expansion wave”): \( R_{\text{ew}} = c_s t \). This solution is therefore termed an “inside-out” collapse. For \( r \geq R_{\text{ew}} \), the density is that of a SIS, \( \rho = c_s^2/(2\pi G r^2) \); for \( r < R_{\text{ew}} \), the gas accelerates until it reaches free fall, with
\[
v = - \left(2Gm_*/r^2\right)^{1/2} \text{ and } \rho_0 \propto r^{-3/2}.
\]
The generalized post-core-formation solutions of Hunter (1977) share the same density and velocity scalings at small radii. The infall rate for Shu’s expansion wave solution is constant in time,
\[
\dot{m}_{\text{in}} = 0.975c_s^3/G = 1.54 \times 10^{-6}(T/10 \text{ K})^{3/2} \ M_\odot \text{ yr}^{-1}.
\]

The total mass inside the expansion wave at time \( t \) is \( 2\dot{m}_{\text{in}} t \), so that about half this mass is in the protostar (i.e., \( f_{\text{ew}} \equiv m_*/m_{\text{ew}} \approx 1/2 \)). Larson (2003) describes the Larson-Penston-Hunter (LPH) and Shu solutions as “fast” and “slow” collapse, respectively, and suggests that reality is somewhere in between. A general discussion of the family of self-similar, isothermal collapse solutions has been given by Whitworth & Summers (1985).

Observations suggest that the cores that form low-mass stars initially have density profiles that approximate those of Bonnor-Ebert spheres (§3.1). Foster & Chevalier (1993) used time-dependent simulations to follow the collapse of such spheres under the assumption that support is by thermal pressure alone. They found that the collapse of the innermost, nearly uniform, part of a critical Bonnor-Ebert sphere (i.e., one with a center-to-edge density contrast of 14.1) approaches, but does not reach, the Larson-Penston (LP) solution prior to and at the time of core formation. Shortly thereafter, the infall rate begins to decline; there is no phase of constant infall for a critical Bonnor-Ebert sphere. On the other hand, a sphere that is initially in an unstable equilibrium with a larger center-to-edge density contrast has an extended (outer) region in which the density scales as \( r^{-2} \). In this case the infall rate starts off as in the critical Bonnor-Ebert case.
and then declines to the constant value for an SIS. The infall rate eventually decreases below the SIS value when a rarefaction wave from the boundary of the cloud reaches the origin (see Vorobyov & Basu, 2005a). Numerical simulations of gravitational collapse in an unmagnetized, turbulent medium show that the initial spike and subsequent decline of the infall rate are typical (Schmeja & Klessen, 2004).

Most of the theoretical work (except for simulations) on low-mass star formation has neglected the role of turbulence in the core. This is generally a valid approximation for low-mass cores, but it becomes increasingly inaccurate as the core mass increases. When turbulence is included, it is generally in the microturbulent approximation. Bonazzola et al. (1987, 1992) treated the turbulent pressure as being scale-dependent, and suggested that turbulence could stabilize GMCs while allowing smaller scales to undergo gravitational collapse. Lizano & Shu (1989) introduced a phenomenological model for turbulence, a logatropic equation of state (§2.2), to treat the contraction of the core. Myers & Fuller (1992) and Caselli & Myers (1995) modeled cores that are supported in part by turbulent motions with a density distribution that is the sum of an $r^{-2}$ power law for a thermal core and a flatter power law for the turbulent envelope. Turbulent cores also can be approximately modeled as polytropes with $\gamma_p < 1$ (§2.2), and when cores collapse, the adiabatic index $\gamma$ can exceed unity (McKee & Zweibel, 1995; Vázquez-Semadeni, Canto, & Lizano, 1998). Ogino, Tomisaka, & Nakamura (1999) generalized the Foster & Chevalier (1993) calculation and found that both the peak infall rate and the rate of decline of the infall rate are increased for $\gamma = \gamma_p > 1$. McLaughlin & Pudritz (1997) generalized the Shu solution to singular polytropic and singular logatropic spheres. They showed that the expansion wave accelerates in time as $R_{\text{ew}} \propto t^{2-\gamma_p}$ and the infall rate increases as $\dot{m}_{\text{in}} \propto t^{3(1-\gamma_p)}$ (logatropes correspond to $\gamma_p \rightarrow 0$). The ratio of the mass in the protostar to that engulfed by the expansion wave is $f_{\text{ew}} \simeq 1/2, 1/6, 1/33$ for $\gamma_p = 1, 2/3$ and a logatrope, respectively. The infall rate for a singular polytropic sphere can also be expressed as $\dot{m}_{\text{in}} = \phi_s m_*/t_{\text{ff}}$, where $t_{\text{ff}}$ is the free-fall time measured at the initial density of the gas just accreting onto the star and $\phi_s \simeq 1.62 - 0.96/(2 - \gamma_p)$ (McKee & Tan, 2002; this is valid for $0 < \gamma_p < 1.2$). Inside-out collapse solutions for clouds that are initially contracting and that have $\gamma \neq \gamma_p$ have been developed by Fatuzzo, Adams, & Myers (2004). For clouds that are supported in part by turbulence, the decay of the turbulence can initiate the collapse of the core (Myers & Lazarian, 1998).

In the innermost regions of the collapsing core, the opacity eventually becomes large enough that the gas switches from approximately isothermal behavior to adiabatic behavior. The initial calculations were carried out by Larson (1969), and recent calculations include those by Masunaga, Miyama, & Inutsuka (1998), Masunaga & Inutsuka (2000), and Wuchterl & Tscharnuter (2003); all assume spherical symmetry. The gas begins to become adiabatic at a density $\rho \sim 10^{-13}$ g cm$^{-3}$. The “first core” forms when the gas becomes hot enough to stop the collapse, and an accretion shock forms at a radius $\sim 5$ AU and with an enclosed mass $\sim 0.05 M_\odot$. Once the gas is hot enough to dissociate the molecular hydrogen, a second collapse ensues and the protostar is formed. When opacity effects are included, the maximum infall rate is limited to about $13 c_s^2 / G$ (Larson, 2003), and the average infall rate over the time required to assemble 80% of the final stellar mass is about 1.5-3 times the SIS value (Wuchterl & Tscharnuter, 2003).
Effects of Rotation  

The two classical problems of star formation are the angular momentum problem and the magnetic flux problem: a star has far less angular momentum and magnetic flux than an equivalent mass in the interstellar medium. Magnetic fields effectively remove angular momentum so long as the contraction of the core is sub-Alfvenic and the neutral and ionized components of the infalling gas are reasonably well coupled (e.g., Mestel, 1985; Mouschovias, 1987). Once either of these conditions breaks down, the gas will collapse with (near-)constant specific angular momentum, $j = \varpi v_\varpi$, where $\varpi$ is the cylindrical radius, provided the transport of angular momentum by turbulence and gravitational torques is unimportant. For collapse at constant $j$, a disk will form with a radius

$$R_d = \frac{(R_d v_{\text{Kep}})^2}{R_d v_{\text{Kep}}^2} = \frac{j^2}{R_d v_{\text{Kep}}^2} = \frac{\Omega_0^2 \varpi_0^4}{G m_{*d}} = 3 \varpi_0 \beta_{\text{rot}}(\varpi_0),$$

(45)

where $v_{\text{Kep}} = (G m_{*d}/R_d)^{1/2}$ is the Keplerian velocity, $m_{*d}$ is the mass of the star and disk (assumed to be equal to the initial mass $M(\varpi_0)$), $\beta_{\text{rot}}$ is the rotational energy parameter defined in equation (30), and $\varpi_0$ and $\Omega_0$ are the initial cylindrical radius and angular velocity. In the collapse to a disk, the radius shrinks by a factor $3 \beta_{\text{rot}}$. Note that if the rotational velocity is proportional to the velocity dispersion, as might be expected for a cloud supported by turbulent motions (Burkert & Bodenheimer, 2000), then $\beta_{\text{rot}}(\varpi_0)$ is constant (Goodman et al., 1993), and the disk radius is a fixed fraction of the initial radius. On the other hand, clouds supported primarily by thermal pressure are generally assumed to be in uniform rotation. Recall that Goodman et al. (1993) found that cores typically have $\beta_{\text{rot}}(R_{\text{core}}) \sim 0.02$.

As in the non-rotating case, two limits for rotating collapse have received the greatest attention. These studies have generally assumed isothermality and have focused on inviscid, axisymmetric flow, although the latter conditions are likely to be violated in real disks, as discussed in §4.2 below. If the core initially has constant density and is rotating slowly, then it collapses to a disk that evolves to a configuration with a singular surface density profile, $\Sigma \propto \varpi^{-1}$ (Norman, Wilson, & Barton, 1980; Narita, Hayashi, & Miyama, 1984). The self-similar solution for the collapse of a rotating disk has been obtained by Saigo & Hanawa (1998), who pointed out that this solution is the analog of the Larson-Penston-Hunter (LPH) solution for non-rotating collapse (i.e., it includes the time after the formation of the central singularity in $\Sigma$). A quasi-equilibrium disk with a radius $R_d \simeq j^2/G m_{*d}$ (eq. 45) grows after formation of the central singularity. Since both $M(\varpi)$ and $j$ scale as $\varpi^2$ in the inner part of the initial spherical cloud, it follows that $j \propto M(\varpi_0)$. Angular momentum is conserved during disk formation, so when a mass $M(\varpi_0) = m_{*d}$ has collapsed into the disk, $R_d \propto j^2/m_{*d} \propto m_{*d}$, which grows linearly in time in the isothermal case. Note also that since $\Sigma \propto \varpi^{-1}$ in the disk, it follows that after disk formation $M$ and therefore $j \propto \varpi$; as a result, the rotational velocity in the disk is constant, independent of $\varpi$. The infall rate into the central disk is about $(3 - 11) c_s^3/G$, depending on the angular momentum; this is significantly less than that for the non-rotating LPH solution (Saigo & Hanawa, 1998). In this solution the gas outside the equilibrium disk is dynamically contracting and is assumed to be itself in a thin disk. Numerical calculations indicate that relaxation of the thin disk approximation increases the accretion rate by about a factor 2 for the case they considered.

Alternatively, if the core settles into a centrally concentrated, spherical, quasi-
equilibrium state prior to collapse, a slow, inside-out collapse ensues. The density
distribution of the supersonically infalling gas in the vicinity of the disk has
been determined by Ulrich (1976) and by Cassen & Moosman (1981) under
the assumptions that the mass is dominated by the central protostar and that the gas
is spherically symmetric far from the protostar. The outer radius of the disk is
given by equation (45), but the disk is far from Keplerian—there is a large inward
velocity that leads to a dynamically contracting outer disk and quasi-equilibrium
inner disk (Stahler et al., 1994). This solution for the inner part of the infall can
be joined smoothly to the solution for the collapse of an SIS (Terebey, Shu, &
Cassen, 1984). More generally, if the cloud initially has a power-law density profile
\( \rho \propto r^{-k_\rho} \) with \( k_\rho > 1 \), then it is straightforward to show that, for \( \Omega_0 = \text{const} \),
the disk radius is \( R_d \propto m_{sd}^{(k_\rho+1)/(3-k_\rho)} \). For the collapse of a slowly rotating SIS
\( (k_\rho = 2) \), this equation implies that the disk radius varies as \( R_d \propto m_{sd}^3 \propto t^3 \)
(Cassen & Moosman, 1981). This rapid increase in disk radius with time is based
on the assumption that the cloud can evolve to a rigidly rotating SIS. Subsequent
work (described below) shows that even when magnetic fields are included, this
condition is difficult to realize, and \( R_d \) tends to increase only linearly with time.

Effects of Magnetic Fields  Poloidal magnetic fields prevent gravitational
collapse when they are sufficiently strong (subcritical cores), and they inhibit
contraction otherwise (supercritical cores), as discussed in §2.2. Magnetic ten-
sion acts to dilute the force of gravity. In nonrotating disks, this effect can be
modeled approximately by adopting an effective gravitational constant (Basu,
1997; Nakamura & Hanawa, 1997; Shu & Li, 1997), \( G_{\text{eff}} = (1 - \mu_\Phi^2)G \), where
\( \mu_\Phi \equiv M/M_\Phi \) (§2.2) is assumed to be independent of \( r \) and must be greater than
unity for gravitationally bound clouds; this equation is exact if the disk is also
thin (Shu & Li, 1997). As a result, the mass of a thin, nonrotating isothermal,
magnetically supercritical disk in equilibrium is \( M(r) = [1/(1 - \mu_\Phi^{-2})]2c^2r/G \) (Li
& Shu, 1996), which can be much larger than in the absence of magnetic support.
If such a disk is initially in static equilibrium (which is difficult to arrange), the
infall rate resulting from an inside-out collapse has \( \phi_{\text{in}} = 1/(1 - \mu_\Phi^{-2}) \) in equation
(43) to within about 5% (Li & Shu, 1997; Allen, Shu, & Li, 2003).

Mouschovias and his students have carried out an extensive set of calculations
of the evolution of a magnetized disk assuming that the disk is thin and axisym-
metric, which reduces the calculation to one spatial dimension. They followed
the evolution from a subcritical initial state to supercritical collapse under the
influence of ambipolar diffusion (Fiedler & Mouschovias, 1993), including the ef-
ects of charged grains (Ciolek & Mouschovias, 1994, 1996, 1998) and rotation
(Basu & Mouschovias, 1994, 1995a,b). In these calculations, the magnetic field
has a characteristic hour-glass shape in which the field is normal to the disk
and flares above and below it; observations that are consistent with this geom-
etry have been obtained recently at a resolution of several hundred AU (Girart,
Rao, & Marrone, 2006). Galli & Shu (1993a,b) show that even if the magneto-
tized core began with a spherical shape, it would collapse to a disk, which they
term a “pseudodisk” since it is not rotationally supported. The calculations of
Mouschovias and his students cited above typically began in a very subcritical
state, with \( \mu_\Phi \simeq 0.25 \), and stopped when the central density reached \( 10^{9.5} \text{ cm}^{-3} \),
which is about the point at which the central regions are expected to become
opaque and non-isothermal. They found that thermal pressure exerts an out-
ward force \(\approx 30\%\) of that due to gravity, whereas centrifugal accelerations are negligible in this phase of evolution. The central part of the core undergoes an extended phase of evolution until it becomes supercritical, at which point the contraction accelerates and mass-to-flux ratio increases more slowly.

Basu (1997) has given a semi-analytic treatment of these results. He showed that the surface density profile in the inner core is \(\Sigma(r, t) \approx \Sigma_c(t)/(1+(r/R)^2)^{1/2}\), where \(\Sigma_c\) is the central surface density and \(R(t) = 2c_s^2/G\Sigma_c(t)\) is the radius of the region in which thermal pressure is sufficiently strong to maintain an approximately constant density (Narita, Hayashi, & Miyama, 1984 found a similar result for rotating collapse without a magnetic field). The supercritical core has a radius \(R_{\text{crit}}\) obtained by setting the central surface density equal to twice the critical value \([\Sigma_c = \mu_B B_0/(2\pi G^{1/2})\) with \(\mu_B = 2\), where \(B_0\) is the initial field strength in the core]. He showed that the slow increase in the mass-to-flux ratio, \(\mu_B \propto \Sigma^{0.05}\), results in a significant reduction in magnetic support at high densities. The density profile in the inner part of the disk \((r \ll R_{\text{crit}})\) has \(k_\rho \approx 2\), but the increasing relative importance of magnetic forces in the outer regions cause it to flatten out so that the mean value in the entire core is \(k_\rho \approx 1.6\). Basu found that the rotational velocity is independent of \(r\), just as Saigo & Hanawa (1998) did for the non-magnetic case. Extending the problem to include the time after protostar formation, Contopoulos, Ciolek, & Königl (1998) obtained a similarity solution to the non-rotating collapse problem with ambipolar diffusion and found an infall rate \(\dot{m}_{\text{in}} = 5.9c_s^3/G\) at the time of protostar formation. Two-dimensional numerical calculations have confirmed this result and have shown that the infall rate subsequently drops by somewhat less than a factor 2 (Ciolek & Königl, 1998).

Shu, Li, & Allen (2004) have considered very different initial conditions, in which a uniform field threads a singular isothermal sphere. Since the flux-to-mass ratio is zero at the origin and increases outward, magnetic effects are negligible at the center and become important only at a characteristic length scale \(R_{\text{ch}} = \pi c_s^2/G^{1/2} B\), corresponding to the condition that \(M_G \sim M_B\) (a similar length scale arises in studies of collapse with ambipolar diffusion, as may be inferred from Basu & Mouschovias, 1995b). They follow the collapse under the assumption that the flux is frozen to the gas and find that the infall rate declines after the expansion wave reaches \(R_{\text{ch}}\); this is analogous to the result of Ciolek & Mouschovias (1995), who found a similar result for the case of collapse with ambipolar diffusion when the expansion wave reached a point at which the ambipolar diffusion was inhibited by photoionization. The initial conditions assumed by Shu, Li, & Allen (2004) lead to a poloidal field that decreases inward, whereas calculations that start from non-singular initial conditions and include ambipolar diffusion find that the poloidal field strongly increases inward.

A full similarity solution for the evolution of the collapsing core after it has fallen into a thin disk and a protostar has formed at the center, including rotation, magnetic fields and ambipolar diffusion, has been obtained by Krasnopolsky & Königl (2002). They assumed that the gas is in a thin disk with a constant rotational velocity (see above); how this assumption would be affected by turbulence, which would thicken the disk and transport angular momentum, is unclear. The infalling gas goes through two shocks, a C-shock (which has a structure dominated by ambipolar diffusion—e.g., Draine & McKee, 1993) and a shock at the outer edge of the centrifugally supported disk. When the protostar reaches \(1M_\odot\), the C-shock is at about \(10^3\) AU, and the centrifugal shock is at about \(10^2\) AU, consistent
with data on T Tauri systems (see §4.2.1 below). Within the self-similar framework, they find that magnetic braking can be adequate to maintain accretion onto the central protostar; in this case there would be no need for internal disk stresses to drive accretion. The infall rate in their fiducial case is $4.7c^3/G$; for a gas at 10 K, this corresponds to a star formation time $t_{\text{sf}} = 1.3 \times 10^5 (m_s/M_\odot)$ yr. Their solution does not include an outflow, but they show how one might be included and estimate that this could reduce the accretion rate by a factor $\lesssim 3$. In sum, based on the theoretical work to date, it is clear that the infall rate is proportional to $c^3_{\text{eff}}/G$, where $c_{\text{eff}}$ is an effective sound speed (Stahler, Shu, & Taam, 1980), but the value of the coefficient and its time dependence have yet to be determined in realistic cases.

The magnetic flux problem in star formation is that stars have very large values of the mass-to-flux ratio ($\mu_\Phi \sim 10^4 - 10^5$ in magnetic stars, $\sim 10^8$ in the Sun—Nakano, 1983), whereas they form from gas with $\mu_\Phi \sim 1$. This problem does not have an adequate solution yet, but it appears that it must be resolved in part on scales $\lesssim 1000$ AU and in part on smaller ($\sim$ AU) scales. Detailed calculations of the ionization state of the infalling and accreting gas show that the ionization becomes low enough that the field decouples from the gas at densities of order $10^{10.5} - 10^{11.5}$ cm$^{-3}$ (Nishi, Nakano, & Umebayashi, 1991; Desch & Mouschovias, 2001; Nakano, Nishi, & Umebayashi, 2002); decoupling occurs at a somewhat lower density after the formation of the central protostar, due to the stronger gravitational force (Ciolek & Königl, 1998). Li & McKee (1996) showed that once the field decouples from the gas, magnetic flux accumulates in the accretion disk as the gas flows through the field and onto the protostar. The pressure associated with this field is strong enough to drive a C-shock (which has a structure dominated by ambipolar diffusion) into the infalling gas. The radius of the shock is predicted to be several thousand AU at the end of the infall phase of a 1 $M_\odot$ star; inside the shock, the field is approximately uniform (except close to the star) and the gas settles into an infalling, dense disk that they identified with the “outer disk” observed in HL Tau (Hayashi, Ohashi, & Miyama, 1993).

These results have been confirmed and improved upon by Contopoulos, Ciolek, & Königl (1998), Ciolek & Königl (1998) and Krasnopolsky & Königl (2002). Tassis & Mouschovias (2005) have carried out 2D axisymmetric calculations with careful attention to the evolution of the ionization and find that the location of the shock oscillates, leading to fluctuations in the accretion rate; it is important to determine if this effect persists in a full 3D simulation. Tassis & Mouschovias (2007a,b,c) find that the magnetic field in the central region ($r \lesssim 10$ AU) is about 0.1 G at the end of their calculation, when the central star has a mass $\sim 0.01 M_\odot$; this is at the low end of the fields inferred in the early solar nebula from meteorites, which are in the range 0.1 – 10 G (Morfill, Spruit, & Levy, 1993). They show that ohmic dissipation becomes as important as ambipolar diffusion at densities $\gtrsim 10^{12.5}$ cm$^{-3}$, but it does not affect the total magnetic flux. However, even though these processes significantly reduce the field within a few AU of the protostar, they are not sufficient to reduce the magnetic flux in the protostar to the observed value (Nakano & Umebayashi, 1986; Li & McKee, 1996; Li, 1998; Ciolek & Königl, 1998; Tassis & Mouschovias, 2005). It is possible that turbulent diffusion (Li & McKee, 1996) or magnetic reconnection (Mestel & Strittmatter, 1967; Galli & Shu, 1993b) plays a role in further reducing the magnetic flux. Reconnection alters the topology of the field and can displace the region in which the flux crosses the forming or accreting disk. However,
reconnection cannot actually destroy flux (a common misconception), since at sufficient distance from the protostar the plasma is a good conductor and the total flux inside this conductor must be conserved. At the present time, the solution to the magnetic flux problem remains incomplete.

As remarked above (§4.1.1), magnetic fields are thought to play a critical role in solving the classical angular momentum problem by means of magnetic braking (Mestel, 1985; Mouschovias, 1987). Indeed, magnetic braking when the field is frozen to the matter is so effective that Allen, Li, & Shu (2003) and Galli et al. (2006) have argued that magnetic reconnection is required to reduce the field and therefore the braking enough that a Keplerian disk can form. The infall solution of Krasnopolsky & Königl (2002), which includes ambipolar diffusion, and the numerical simulations of Hujerat et al. (2000), which include both turbulence and ambipolar diffusion, suggest that Keplerian disks can form without reconnection, but nonetheless indicate that predicting the evolution of the specific angular momentum of the infalling gas is a complex problem. However, it is not clear that any of these theoretical models are consistent with the observations of Ohashi et al. (1997), which show that the specific angular momentum in gas associated with several protostars in Taurus is constant for $10^{-3} \text{ pc} < r < 0.03 \text{ pc}$.

Numerical simulations, as opposed to numerical integration of the underlying partial differential equations, are required to study core collapse in 2D or 3D without additional assumptions (such as self-similarity or a thin-disk condition). A critical review of numerical simulations of low-mass star formation is given by Klein et al. (2007). To date, such simulations have not included ambipolar diffusion, nor have they simultaneously included radiative transfer and magnetic fields; most simulations have also stopped prior to the formation of the protostar. A prediction of these simulations is that a slow ($v \sim c_s \sim 0.2 \text{ km s}^{-1}$) outflow should occur at large radii ($\sim 10^3 \text{ AU}$; Tomisaka, 1998, 2002; Banerjee & Pudritz, 2006). These authors suggest that this outflow is related to the observed bipolar outflows, but Allen, Li, & Shu (2003) disagree. In any case, this large scale outflow could be important in setting the outer boundary conditions for the jets and higher-velocity outflows that are observed (see §§4.2.4, 4.2.5).

4.1.2 Bondi-Hoyle Accretion Once a protostar star has formed by gravitational collapse of a core, it can continue to grow by gravitational accretion from the ambient medium. Most treatments of this process do not distinguish between the gas that accretes directly onto the star and the gas that first falls onto the disk. Hoyle & Lyttleton (1939) first developed the theory of accretion by a moving point mass, and Bondi (1952) extended the theory to the case in which the star is at rest in a medium of finite temperature. Today, gravitational accretion by a stationary object is generally referred to as Bondi accretion, whereas that by a moving object is referred to as Bondi-Hoyle accretion. If the density and sound speed far from the star are $\rho$ and $c_s$, respectively, and the star is moving at a velocity $v_0 = M_0 c_s$ through the ambient medium, then the characteristic radius from which the star accretes is

$$R_{BH} = \frac{Gm_*}{(1 + M_0^2 c_s^2)},$$  \hspace{1cm} (46)$$

The accretion rate is

$$\dot{M}_{BH} = 4\pi\dot{\phi}_{BH}R_{BH}^2\rho c_s(1 + M_0^2)^{1/2} = \frac{4\pi\dot{\phi}_{BH}G^2m_*^2}{(1 + M_0^2)^{3/2}c_s^2},$$  \hspace{1cm} (47)$$
where $\phi_{BH}$ is a number of order unity that fluctuates somewhat due to instabilities in the flow (Ruffert & Arnett, 1994).

There are a number of assumptions that go into this result: (1) The mass inside $R_{BH}$ is dominated by the mass of the star—i.e., the self-gravity of the accreting gas is negligible. One can show that

$$\frac{M(R_{BH})}{m_*} = \frac{5.85}{(1 + M_0^2)^{3/2}} \left( \frac{m_*}{M_{BE}} \right)^2,$$

so that this condition is equivalent to requiring that the stellar mass be small compared to the Bonnor-Ebert mass (eq. 15) in the ambient medium (for $M_0 \lesssim 1$). (2) The tidal gravitational field is negligible; when it is not, $R_{BH}$ is replaced by the tidal radius (Bonnell et al., 2001a). (3) The magnetic field is negligible. Based on dimensional scalings, a rough approximation for the effect of a magnetic field on the accretion rate would be to make the replacement $v_0 \rightarrow (v_0^2 + v_A^2)^{1/2}$. (4) The ambient gas is moving at a uniform velocity. In fact, gas in molecular clouds is generally supersonically turbulent, so that an accreting star experiences large fluctuations in both the density and velocity of the accreting material. Krumholz, McKee, & Klein (2006) showed that the mean accretion rate in a turbulent medium is given by equation (47) with $\rho$ equal to the mean density, $M_0$ replaced by $M_{turb}$, and $\phi_{BH} \approx 3.5 \ln(0.70 M_{turb})$, provided that the 3D Mach number $M_{turb}$ of the turbulence is large compared to $M_0$ and compared to unity (they verified this result for $M_0 = 0$ and $3 \leq M_{turb} \leq 10$). This result was derived from simulations of isothermal gas, but it should be approximately valid for other equations of state also. The median accretion rate is significantly less than the mean, however.

The dominant paradigm for star formation is gravitational collapse, but an alternative is that stars (or at least relatively massive stars) are formed primarily by the capture and subsequent accretion of matter that is initially unbound to the star (Zinnecker, 1982; Bonnell et al., 1997). Since protostars compete for gas from a common reservoir, this process is termed “competitive accretion.” The simulations of Bonnell et al. (1997) show that a few of the fragments gain most of the mass; these are the ones that reside primarily in the central regions of the clump and have the highest accretion rates. Since gravitational accretion scales as $m_*^2$, initial differences in protostellar masses are amplified. This process has the potential of producing the initial mass function, and it also naturally leads to massive stars being centrally concentrated in clusters, as observed (Bonnell et al., 2001b). A key issue for competitive accretion is, what is the level of turbulence in the ambient medium? There is general agreement that competitive accretion is ineffective if the medium has turbulent energy comparable to gravitational energy, with $\alpha_{vir}$ order unity (see eq. 11), whereas it is effective if the turbulence is sufficiently weak, $\alpha_{vir} \ll 1$ (Bonnell et al., 2001a; Krumholz, McKee, & Klein, 2005a; magnetic fields, which tend to suppress accretion, have not been considered yet). Bonnell and his collaborators (Bonnell & Bate, 2006 and references therein) argue that the gas throughout star-forming clumps has a very low turbulent velocity so that protostars in clusters can accrete efficiently. On the other hand, Krumholz, McKee, & Klein (2005a) argue that stellar feedback and the cascade of turbulence from larger scales ensure that the star-forming clumps are sufficiently turbulent to be approximately virialized and to therefore have negligible competitive accretion. Analysis of data from several star-forming clumps shows that stars in these clumps could grow by only $(0.1 - 1)\%$ in a dynamical
time, far too small to be significant. The timescale for the formation of star clusters is an important discriminant between these models: Star clusters form in about $2t_g$ if turbulence is allowed to decay, whereas it can take significantly longer if turbulence is maintained (Bonnell, Bate, & Vine, 2003). The observational evidence discussed by Tan, Krumholz, & McKee (2006) and Krumholz & Tan (2007) favors the longer formation time. This controversy can be resolved through more detailed observations of gas motions in star-forming clumps and through more realistic simulations that allow for the evolution of the turbulent density fluctuations as the clump forms and evolves to a star-forming state, and that incorporate stellar feedback.

4.1.3 Observations of low-mass star formation

The growth of protostars can be inferred through observations of the mass distribution surrounding the protostar, the velocity distribution of this circumstellar gas, and the non-stellar radiative flux. The mass and/or temperature distribution on both small and large spatial scales can be inferred by modeling the spectral energy distribution (SED) of the continuum. Protostellar SEDs are conventionally divided into four classes, which are believed to represent an evolutionary progression (Myers et al., 1987 divided sources into two classes; Lada, 1987 introduced Classes I-III; Adams, Lada, & Shu, 1987 discussed a similar classification; and Andre, Ward-Thompson, & Barsony, 1993 introduced Class 0). Andre, Ward-Thompson, & Barsony (2000) have summarized the classification scheme:

**Class 0**: sources with a central protostar that are extremely faint in the optical and near IR (i.e., undetectable at $\lambda < 10 \mu$m with the technology of the 1990’s) and that have a significant submillimeter luminosity, $L_{\text{submm}}/L_{\text{bol}} > 0.5\%$. Sources with these properties have $M_{\text{envelope}} \gtrsim m_*$. Protostars are believed to acquire a significant fraction, if not most, of their mass in this embedded phase.

**Class I**: sources with $\alpha_{\text{IR}} > 0$, where $\alpha_{\text{IR}} \equiv d \log \lambda F_\lambda / d \log \lambda$ is the slope of the SED over the wavelength range between 2.2 $\mu$m and 10–25 $\mu$m. Such sources are believed to be relatively evolved protostars with both circumstellar disks and envelopes.

**Class II**: sources with $-1.5 < \alpha_{\text{IR}} < 0$ are believed to be pre-main sequence stars with significant circumstellar disks (classical T Tauri stars).

**Class III**: sources with $\alpha_{\text{IR}} < -1.5$ are pre-main sequence stars that are no longer accreting significant amounts of matter (weak-lined T Tauri stars).

These classes can also be defined in terms of the “bolometric temperature,” which is the temperature of a blackbody with the same mean frequency as the SED of the YSO (Myers & Ladd, 1993).

Unfortunately, the geometry of the source can confound this classification scheme (e.g., Masunaga & Inutsuka, 2000). It is well-recognized that a given source can appear as a Class II source at small or moderate inclination angles (so that the central star is visible) and as a Class I source at large inclination angles (so that the central source is obscured by the disk). A similar ambiguity can involve Class 0 sources if the protostellar envelope is flattened due to the presence of a large-scale magnetic field or contains cavities created by protostellar jets. White et al. (2007) summarize the observational evidence that many of the properties of Class I and Class II sources are similar, which is consistent with inclination effects confusing the evolutionary interpretation of the SEDs. This
ambiguity can be alleviated by radio or sub-millimeter observations of the envelopes, which yield masses that are independent of inclination; Motte & André (2001) find that about 40% of the sources in Taurus that are classified as Class I on the basis of their SEDs have envelope masses < 0.1 M⊙ and are thus unlikely to be true protostars. More sophisticated modeling of the SEDs can also clarify the evolutionary sequence of young stellar objects; for example, Robitaille et al. (2006) have calculated 2 × 10^5 model SEDs, including the effects of outflow cavities, that can be automatically compared with observed SEDs to infer the properties of the source. Counts of sources at different evolutionary stages together with an estimate for the age for one of the stages allows one to infer the lifetimes for all the stages. Typical estimates for the ages are 1 – 2 × 10^5 yr for Class I sources and 1 – 3 × 10^4 yr for Class 0 sources (Andre, Ward-Thompson, & Barsony, 2000).

There is a significant discrepancy between the protostellar accretion rates that are observed and those that are expected, resulting in the so-called “luminosity problem” (Kenyon et al., 1990). The luminosity due to accretion onto the star is

\[ L_{\text{acc}} = f_{\text{acc}} \frac{G m_* \dot{m}_*}{R_*} = 3.1 f_{\text{acc}} \left( \frac{m_*}{0.25 M_\odot} \right) \left( \frac{\dot{m}_*}{10^{-6} M_\odot \text{ yr}^{-1}} \right) \left( \frac{2.5 R_\odot}{R_*} \right) L_\odot, \]

(49)

where \( f_{\text{acc}} \) is the fraction of the gravitational potential energy released by accretion, the rest being carried off in a wind or absorbed by the star (e.g., Ostriker & Shu, 1995), 0.25 M⊙ is the typical mass of a protostar (i.e., half the mass of a typical star), and 2.5 R⊙ is the corresponding radius (Stahler, 1988). There are two main ways to estimate the expected average accretion rate, \( \langle \dot{m}_* \rangle \). (1) Since both the mass ejected from the disk and the mass stored in the disk are generally a small fraction of the stellar mass (see §4.2), the average accretion rate \( \langle \dot{m}_* \rangle \) should be comparable to the infall rate \( \dot{m}_{\text{in}} \). Theoretically, for \( T = 10 \) K, this is \( \dot{m}_{\text{in}} = 1.5 \times 10^{-6} \epsilon_{\text{core}} \phi_{\text{in}} M_\odot \text{ yr}^{-1} \) (eqs. 43, 44). If the fraction of the core mass that goes into the star is \( \epsilon_{\text{core}} \simeq 1/3 \) (Matzner & McKee, 2000; Alves, Lombardi, & Lada, 2007), and if the envelope infall rate is that expected from rotating, magnetized collapse (\( \phi_{\text{in}} \simeq 5 \)—Krasnopolsky & Königl, 2002), then the infall rate is \( \dot{m}_{\text{in}} \simeq 2.5 \times 10^{-6} M_\odot \text{ yr}^{-1} \). Observationally, the properties of the envelopes around Class I objects inferred from the SEDs give similar infall rates (Kenyon, Calvet, & Hartmann, 1993), so this estimate for \( \epsilon_{\text{core}} \phi_{\text{in}} \) cannot be too far off. (2) A direct estimate of the average accretion rate is that forming an 0.5 M⊙ star in 2 × 10^5 yr, the estimated upper limit on the duration of the embedded stage, requires \( \langle \dot{m}_* \rangle = 2.5 \times 10^{-6} M_\odot \text{ yr}^{-1} \), comparable to the estimated infall rate. The average luminosity corresponding to this accretion rate is \( \langle L \rangle \simeq 8 L_\odot \). The problem is that the observed median luminosity of the bona fide Class I sources (i.e., those with significant molecular envelopes) in Taurus is about 0.5 – 1 L⊙ (White & Hillenbrand, 2004 and Motte & André, 2001, respectively), almost an order of magnitude smaller. The problem is significantly worse than this, however, since only a small fraction of the luminosity is due to accretion (Muzerolle, Hartmann, & Calvet, 1998); White & Hillenbrand (2004) find that the fraction is about 25%. The clearest statement of the luminosity problem is as an accretion rate problem: The observed accretion rates in Class I protostars are 1 – 2 orders of magnitude smaller than those needed to form a star during the lifetime of a Class I object.
Kenyon et al. (1990) suggested two solutions to this problem. One solution is that significant accretion continues into the T Tauri stage, but this appears to be ruled out by the fact that such stars accrete very slowly \((10^{-8} \, M_\odot \, \text{yr}^{-1})\), and there is not a significant disk or envelope mass reservoir that they can draw on for episodic accretion. The other solution is that most of the accretion occurs in the embedded stage, but it is episodic, so that the median accretion rate is much smaller than the mean. They suggested that the high accretion-rate stage of protostellar accretion could correspond to FU Orionis objects, which are very luminous \((\text{typically} \, 200 - 800 \, L_\odot)\) — Hartmann & Kenyon, 1996 and have have accretion rates \(\sim 10^{-4} \, M_\odot \, \text{yr}^{-1}\). Such outbursts could be due to thermal instability (Bell & Lin, 1994; Bell et al., 1995), although this would only affect the inner disk, and hence the outburst would be limited in duration and total mass accreted; or to gravitational instability in the disk (see §4.2.2 and Vorobyov & Basu 2005b, 2006). However, Hartmann & Kenyon (1996) estimate that the observed FU Ori objects can account for only about \(5 - 20\%\) of the mass of stars forming in the solar vicinity; this discrepancy has not disappeared in the intervening decade. Recent observational studies of the central stars in Class I sources differ on their evolutionary status: White & Hillenbrand (2004) argue that the Class I protostars are similar to T Tauri stars and are thus past the main protostellar accretion phase, whereas Doppmann et al. (2005) come to the opposite conclusion. If protostars are close to their final mass by the time they become Class I sources, then they must gain most of their mass in the Class 0 stage. In this case the luminosity problem remains, albeit in a milder form: The mean accretion rate required to form an \(0.5 \, M_\odot\) star in \(3 \times 10^4\ yr\) (the estimated upper limit on the lifetime of a Class 0 source—Andre, Ward-Thompson, & Barsony, 2000) is \(1.7 \times 10^{-5} \, M_\odot \, \text{yr}^{-1}\), corresponding to \(\langle L \rangle \simeq 50 \, f_{\text{acc}}(2.5 \, R_\odot/R_*) \, L_\odot\); by contrast, the median luminosity of the Class 0 sources listed by Andre, Ward-Thompson, & Barsony (2000) is about \(10 \, L_\odot\). This luminosity problem could be alleviated if a significant fraction of the accretion energy is carried off by the powerful protostellar outflows that accompany these sources, so that \(f_{\text{acc}} \lesssim 1/2\). However, accounting for a value of the infall rate as high as the inferred accretion rate onto the star remains a theoretical challenge: The magnetized collapse models discussed above give infall rates of a few \(\times 10^{-6} \, M_\odot \, \text{yr}^{-1}\) (for \(\epsilon_{\text{core}} \sim 1/3\), significantly less than required. A resolution of the luminosity problem thus remains elusive.

Spectroscopic observations using molecular transitions can give both the mass and velocity distributions in collapsing cores (Myers, Evans, & Ohashi, 2000), but to date the spatial resolution of these data is generally \(\gtrsim 100\ AU\). Evidence for infall in unresolved cores is provided by the “infall asymmetry” (Lucas, 1976; Leung & Brown, 1977; Myers et al., 1996): optically thick, infalling gas in which the excitation temperature rises toward the center produces a characteristic line profile in which the blue wing is stronger than the red wing. Observations of samples of starless cores (Lee, Myers, & Tafalla, 1999, 2001), Class 0 sources (Gregersen et al., 1997), and Class I sources (Gregersen et al., 2000) show a “blue excess” \([(\text{blue asymmetries} - \text{red asymmetries})/\text{(number of sources)}]\) of about \(0.25 - 0.35\), indicating that many of these sources are undergoing collapse (Myers, Evans, & Ohashi, 2000). Unfortunately, it has proved difficult to carry out unambiguous observational tests of the theoretical models for protostellar accretion. Furuya, Kitamura, & Shimaga (2006) mapped the infall in a young Class 0 source and found reasonably good agreement with the LPH solution (\(\phi_{\text{in}} \simeq 20\) in eq. 43);
this source appears to be very young, since for $r > 100$ AU there is no evidence for the $\rho \propto r^{-3/2}$ density profile expected for accretion onto a protostar of significant mass. Tafalla et al. (1998) and Lee, Myers, & Tafalla (2001) found that infall is more extended than expected in inside-out collapse models, although this infall may reflect the formation of small clusters rather than individual stars. They also found that the infall velocity is faster than expected in standard ambipolar diffusion models; however, the velocities are consistent with the collapse of magnetically supercritical cores (Ciolek & Basu, 2000). A potentially important result is that Ohashi et al. (1997) found that cores in Taurus are in solid body rotation on scales $\gtrsim 0.03$ pc but conserve angular momentum on smaller scales. The physical significance of this length scale could be inferred by determining its value in other molecular clouds.

**Brown Dwarfs** Since brown dwarfs represent the low-mass extreme of star formation, they can shed light on the earliest stages of star formation. As a result of a great deal of observational work over the past decade, it has been established that most brown dwarfs form by the same mechanism as most stars (Luhman et al., 2007; Whitworth et al., 2007): the initial mass function, velocity and spatial distributions at birth, multiplicity, accretion rates, circumstellar disks, and outflows are all continuous extensions of those for hydrogen-burning stars. This is to be expected, since stars near the H-burning limit at $0.075 M_\odot$ reach their final mass long before hydrogen burning commences. Following Whitworth et al. (2007) and Chabrier et al. (2007), we shall assume that brown dwarfs form by gravitational instability on a dynamical timescale, and that their composition reflects that of the ambient interstellar medium. By contrast, planets are believed to form in circumstellar disks and to have an elemental composition with an excess of heavy elements. With these definitions, the observational distinction between giant planets with masses $\gtrsim M_J$ and small brown dwarfs is somewhat indistinct, but should eventually be amenable to spectroscopic determination (Chabrier et al., 2007). The lower limit to the mass of a brown dwarf is set by the condition that the star become opaque to the radiation it emits while undergoing gravitational collapse (Low & Lynden-Bell, 1976); including helium, this is about $4 \times 10^{-3} M_\odot \approx 4 M_J$ (Whitworth et al., 2007). The smallest brown dwarfs detected to date have masses $\sim (0.01 - 0.02) M_\odot$ (Luhman et al., 2007).

In order for a brown dwarf to form, its mass must exceed the Bonnor-Ebert mass, even if it forms via shock compression (Elmegreen & Elmegreen, 1978); the pressure at the surface of the core that forms the brown dwarf must therefore be $P/k_B \gtrsim 10^9 (T/10 \text{ K})^4 (10^{-2} M_\odot/m_{BD})^2$ K cm$^{-3}$ at the surface of the brown dwarf. Assuming that brown dwarfs form by turbulent fragmentation, Padoan & Nordlund (2004) show that such pressures can be reached in a large enough fraction of the mass of the cluster IC348 to account for the brown dwarfs observed there. This model is based on the assumption that the gas is isothermal and that as a result the density PDF is a log-normal (eq. 5). The remaining mystery is, why is the IMF of brown dwarfs relatively constant (at least to within a factor 2) when their numbers are exponentially sensitive to the mean of the log-normal, which depends on the Mach number at the largest scale in the cloud? This mystery is part of the larger mystery as to why the IMF appears to be universal, but in the brown dwarf regime the exponential sensitivity to ambient conditions potentially offers an opportunity to determine how ambient conditions affect the
IMF. Whitworth et al. (2007) review a number of other mechanisms for brown dwarf formation that could contribute in some cases: (1) Hierarchical fragmentation—Protostellar cores can fragment as they collapse, and indeed this process is believed to lead to the formation of binary and multiple star systems (§4.1.3). Simulations are as yet inadequate to evaluate the effectiveness of this process in producing brown dwarfs; in particular, it is important to include proper treatment of the radiative transfer (Boss et al., 2000; Whitehouse & Bate, 2006). (2) Disk fragmentation—Several analyses have shown that fragmentation in disks around low-mass stars is suppressed for \( r \lesssim 100 \) AU (Matzner & Levin, 2005; Rafikov, 2005; Whitworth et al., 2007), but again accurate treatment of radiative transfer is essential for quantifying this further. Goodwin & Whitworth (2007) have suggested that brown dwarfs form in disks beyond 100 AU and that the resulting binaries are disrupted by passing stars. (3) Premature ejection of protostellar embryos (Reipurth & Clarke, 2001) —This is a variant of the hierarchical fragmentation scenario, in which low-mass protostars that form via fragmentation are ejected before they can accrete enough matter to reach the hydrogen-burning limit. Reipurth & Clarke (2001) suggested that the ejected brown dwarfs would have a higher velocity dispersion, a more extended spatial distribution and smaller disks than their more massive cousins. This has not been observed (Luhman et al., 2007), but SPH simulations suggest that in fact these differences between brown dwarfs and hydrogen-burning stars are relatively small (Goodwin et al., 2005); more accurate calculations are needed to determine the magnitude of the differences. A difficulty with the ejection model is that the cluster simulations that support it produce too many single stars to be consistent with observation (Goodwin & Kroupa, 2005). Observations of BD-BD binaries can provide a strong test of models for brown dwarf formation, particularly the ejection model (Burgasser et al., 2007). (4) Photoevaporation—Cores that are close to an O star can undergo a radiation-driven implosion (Klein, Sandford, & Whitaker, 1980; Bertoldi, 1989); the subsequent equilibrium photoevaporation produces very high pressures (Bertoldi & McKee, 1990). This process may produce brown dwarfs (e.g., Whitworth & Zinnecker, 2004), but more work is needed to determine if the number of such brown dwarfs is significant.

**Binaries** Stars are roughly evenly divided between those that are in multiple systems (mainly binaries) and those that are single. For stars of mass \( \sim 1 \, M_{\odot} \), Duquennoy & Mayor (1991) found that the fraction of stellar systems that are multiple—i.e., the ratio of the total number of binaries, triples, etc. divided by the total number of systems, including single stars—is \( f_{\text{mult}} = 0.58 \). Multiplicity declines for smaller masses, and Reid & Gizis (1997) find \( f_{\text{mult}} \approx 0.3 \) for stars in the solar neighborhood, which are primarily M stars. Lada (2006) finds a similar result when one averages over the entire IMF: the majority of stellar systems (as opposed to the majority of stars) are single. The fraction of singles is smaller at birth, however: Higher-order multiples (triples, etc.) are often dynamically unstable, and in dense environments collisions among stellar systems can disrupt wide binaries (Kroupa, Petr, & McCaughrean, 1999). Observations summarized by Duchêne et al. (2007) show that the multiplicity among T Tauri stars in low-density associations such as Taurus-Auriga and Ophiuchus and among Class I sources in high-density regions such as L1641 in Orion is about twice as high as
among field stars. In dense star-forming regions like Orion, however, this excess multiplicity is soon erased—for example, the multiplicity in the Orion Nebula Cluster (which is older than L1641) is the same as in the field. This introduces a complication in comparing the core mass function with the IMF (§3.3): on average, cores produce more than one star, and the properties of the resulting stellar systems evolve with time. This is not a major complication, however, since each core typically produces only 2-3 stars (Goodwin & Kroupa, 2005), and the distribution of secondary masses in typical binaries appears to follow the field star IMF (Goodwin et al., 2007). It should be noted that the multiplicity of the stars is imprinted on their spatial distribution: Larson (1995) found a clear break in the density of companions in Taurus at about 0.04 pc, separating binaries and multiple stars from larger scale clusters.

Binaries raise two important issues in the theory of star formation: What is the role of binaries in reducing the angular momentum inherited by protostars? What determines how molecular cores fragment? Binaries do not appear to be effective in taking up the angular momentum of the initial core: Fisher (2004) has shown that turbulent molecular cores must lose 99%-99.9% of their initial angular momentum in order to qualitatively account for a number of features of the binary population with periods \( \gtrsim 10^3 \) day. This angular momentum loss is generally assumed to be due to magnetic braking (§4.1.1), but Jappsen & Klessen (2004) suggest that gravitational torques can also contribute. Binaries can remove angular momentum on small scales by ejecting a companion, which hardens the remaining binary. However, Goodwin & Kroupa (2005) point out two limitations on this process: (1) it tends to create equal mass binaries, which are not common for typical stars, and (2) it would create a population of single stars significantly larger than observed.

There is an extensive literature on the fragmentation of protostellar cores into binary and multiple protostars that is reviewed by Bodenheimer et al. (2000), Duchêne et al. (2007) and Goodwin et al. (2007). Simulations of fragmentation are very challenging because of the enormous range of scales involved, and it does not appear that any of the simulations carried out to date have enough resolution and enough physics (i.e., including MHD and radiative transfer) to adequately address the problem (Klein et al., 2007). In particular, a number of simulations produce 5-10 fragments per core, whereas observations show that most cores produce only 2-3 fragments (Goodwin & Kroupa, 2005).

4.2 Disks and Winds

4.2.1 OBSERVATIONS OF DISKS Because protostellar cores are rotating, collapse with conservation of angular momentum results in the formation of a centrifugally-supported disk (§4.1.1). Observed sizes and rotation parameters for low-mass dense cores predict disk sizes \( \lesssim 1000 \) AU, consistent with high-resolution submillimeter continuum observations that indicate average (dust) disk sizes around T Tauri stars of \( \approx 200 \) AU (Andrews & Williams, 2007); similar results are obtained using mm interferometry (Kitamura et al., 2002). The disks around T Tauri disks extend inward to \( \sim 0.04 \) AU based on modeling of observed CO vibrational emission lines (Najita et al., 2006); these inner radii are smaller than the inner disk radii inferred for dust disks, presumably because the dust sublimes. The initial sizes of circumstellar disks are more difficult to determine, because protostellar systems in the earliest stages (prior to the T Tauri
stage) are still enshrouded in dusty envelopes that emit at similar wavelengths to the disk; in a few cases where the inner envelope emission can be spatially separated out, disk sizes appear similar (e.g. Jørgensen et al. 2005).

Masses of protostellar disks are estimated using the continuum flux in millimeter and submillimeter wavelengths. The first disk mass estimates for T Tauri stars were obtained from observations of the total flux at a single millimeter wavelength under the assumption of optically thin emission (e.g., Beckwith et al. 1990), but subject to an uncertainty in the overall normalization due to the uncertainty in the dust opacity coefficient (since \( M_{\text{disk}} \propto F_{\nu} / (\nu^2 \kappa_{\nu}) \)). Multiwavelength (sub-mm to cm) observations suggest that the dust opacity law \( \kappa_{\nu} \propto \nu^\beta \) has a distinctly shallower slope than holds for dust in the diffuse ISM, presumably due to grain growth (e.g., Beckwith & Sargent 1991; Beckwith, Henning, & Nakagawa 2000). Interpretation of the multiwavelength flux data as implying a change in \( \beta \) is complicated by the fact that some of the short-wavelength emission can be optically thick (which would yield \( F_{\nu} \propto \nu^2 \) independent of \( \beta \) for a disk that is optically thick at all \( \nu \)). However, spatially resolved observations can be combined with modeling to correct for optically-thick contributions varying with \( \nu \) and \( R \), with the resulting median \( \beta \lesssim 1 \) (e.g. Natta et al. 2007; Andrews & Williams 2007; Lommen et al. 2007), suggesting that the largest “grains” are in fact cm-size pebbles (e.g. Wilner et al. 2005; Rodmann et al. 2006). Total disk masses for T Tauri systems are estimated to be in the range \( \sim 10^{-3} - 10^{-1} \) \( M_\odot \), with a median near 0.005 \( M_\odot \) from submillimeter observations (e.g. Andrews & Williams 2005), but these may severely underestimate the true masses if a large fraction of the grains have grown to mm or cm sizes and thus emit only weakly in the submm (Hartmann et al., 2006; Natta et al., 2007). Determining the distribution of mass within disks is difficult because submm emission is likely optically thick in the inner regions, while at longer wavelengths there is insufficient resolution to probe the inner-disk regions (Andrews & Williams, 2007). Finally, we note that disk mass determinations assume a cosmic ratio of gas to dust; at late evolutionary stages, photoevaporation may preferentially remove gas and planet formation may preferentially remove dust.

The thermal structure of protostellar disks is likely quite complex. Disks can be heated both externally via irradiation from the central star, and internally from dissipation and thermalization of orbital kinetic energy as the gas accretes (Calvet et al., 1991; Chiang & Goldreich, 1997). As a consequence, the vertical temperature distribution depends on details of the system and can have a local minimum at intermediate altitude (D’Alessio et al., 1998). The vertical temperature distribution together with its dependence on radius must be self-consistently calculated, since the flaring of the disk surface affects the amount of radiation intercepted from the central star (see discussion and references in Dullemond et al., 2007). In addition, gas and dust temperatures may differ in the upper atmospheres where the densities are low and stellar X rays strongly heat the gas (Najita et al., 2006). Quite sophisticated radial-vertical radiative models (including grain growth and settling) have been developed that agree well with observed spectral energy distributions from \( \mu \)m to mm wavelengths (see e.g. Dullemond & Dominik, 2004, D’Alessio et al., 2006 and references therein). The IR emission signatures, including PAH features at 3 – 13 \( \mu \)m and edge-on silhouette images, as well as scattered-light/polarization observations in optical and near-IR, indicate that although some grains have grown to large sizes, small grains still remain in disk atmospheres (see references and discussion in Dullemond et al., 2007 and
Disk lifetimes are inferred based on stellar ages combined with IR and mm/sub-mm emission signatures, which are sensitive to warm dust. Multiwavelength Spitzer observations of the nearby star-forming cluster IC 348 (Lada et al., 2006) show that for ~70% of stars, disks have become optically-thin in the IR (implying inner disks $R < 20$ AU have been removed) within the 2-3 Myr age of the system; disk fractions are slightly higher (~50%) for Solar-type stars than in those of higher or lower mass. Observations of other clusters are consistent with these results (Sicilia-Aguilar et al., 2006a). $L$-band observations of disk frequencies in clusters spanning a range of ages (Haisch, Lada, & Lada, 2001) suggest that overall disk lifetimes are $\approx 6$ Myr. Even in the 10 Myr old cluster NCG 7160, however, a few percent of stars still show IR signatures of disks (Sicilia-Aguilar et al., 2006a), and disk lifetimes appear to be inversely correlated with the mass of the star (Hernandez et al., 2007). Signatures (or their absence) of dusty disk emission are also well correlated with evidence (or lack) of accretion in gaseous emission line profiles (see below) in systems at range of ages, indicating that gas and dust disks have similar lifetimes (Jayawardhana et al., 2006; Sicilia-Aguilar et al., 2006b). Andrews & Williams (2005) found, for a large sample of YSOs in Taurus-Auriga, that in general those systems with near-IR signatures of inner disks also have sub-mm signatures of outer disks, and vice versa; they conclude that inner and outer disk lifetimes agree within $10^5$ yr.

Accretion in YSO systems is studied using a variety of diagnostics (see e.g. Calvet, Hartmann, & Strom (2000)), including continuum “veiling” of photospheric absorption lines and optical emission lines, which are respectively believed to arise from hot (shocked) gas on the stellar surface and from gas that is falling onto the star along magnetic flux tubes. Gullbring et al. (1998) measured a median accretion rate for million-year-old T Tauri stars of $\sim 10^{-8}$ $M_\odot$ yr$^{-1}$, and White & Ghez (2001) found similar accretion rates for the primaries in T Tauri binary systems. A recent compilation of observations (White & Basri, 2003; Muzerolle et al., 2003; Calvet et al., 2004) shows an approximate dependence of the accretion rate on stellar mass $\dot{M}_{\text{disk}} \propto m_2^2$, although with considerable scatter (Muzerolle et al., 2005). This scaling of the accretion rate with stellar mass is potentially explained by Bondi-Hoyle accretion from the ambient molecular cloud (Padoan et al., 2005). However, such a model accounts only for the infall rate onto the star-disk system, not the disk accretion rate; these need not agree. In addition, it does not account for the accretion seen in T Tauri stars outside molecular clouds (Hartmann et al., 2006). During their embedded stages (a few $\times 10^5$ yr), low-mass stars have typical disk accretion rates similar to or slightly larger than those of TTSs (White et al., 2007). As discussed in §4.1.3, the infall rates from protostellar envelopes typically exceed disk accretion rates by a factor 10-100, so it is possible that mass is stored in the disk and released intermittently, in brief but prodigious accretion events similar to FU Ori outbursts (Kenyon et al., 1990; Hartmann & Kenyon, 1996).

For high-mass protostars, observations suggest that there are at least two classes of disks (Cesaroni et al., 2007). In moderate-luminosity sources corresponding to B stars ($L \lesssim \text{few} \times 10^4 L_\odot$), the disks appear to be Keplerian, with masses significantly less than the stellar mass and time scales for mass transfer $\sim 10^5$ yr. In luminous sources ($L \gtrsim 10^5 L_\odot$), the disks are large ($4 - 30 \times 10^3$ AU) and massive ($60 - 500 M_\odot$). Consistent with the discussion in §§4.1.1 and 4.1.1, the disks are observed to be non-Keplerian on these
large scales. To distinguish these structures from the disks observed around B stars, Cesaroni (2005) terms them “toroids.” The inferred infall rates in these disks are of order $2 \times 10^{-3} - 2 \times 10^{-2} \, M_\odot \, yr^{-1}$, corresponding to mass transfer time scales of order $10^4 \, yr$ (Zhang, 2005). In view of their large size and mass, they may be circumcluster structures rather than circumstellar ones. Indeed, one of the best studied luminous sources, G10.8-0.4, is inferred to have an embedded cluster of stars with a total mass $\sim 300 \, M_\odot$ (Sollins et al., 2005). Simulations of the formation of an individual massive star in a turbulent medium give a disk size $\sim 10^3 \, AU$, significantly smaller than the size of the toroids (Krumholz, Klein, & McKee, 2005). To date, no disks have been observed in the luminous sources on scales $\lesssim 10^3 \, AU$. Most likely, this is because of the observational difficulties in observing such disks; it should be borne in mind, however, that there is no direct evidence that these sources are in fact protostellar. Including disks around both B stars and the toroids around luminous sources, Zhang (2005) finds that the mass infall rate in the disks scales as $M_{\text{disk}} \propto m_{\text{star}}^{2/2}$, although there are substantial uncertainties in the data for the luminous sources.

4.2.2 ACCRETION MECHANISMS

The most fundamental theoretical question about YSO disks is what makes them accrete; while many mechanisms have been investigated, the problem is still open. In large part this is because the accretion process depends on a complicated interplay of MHD, radiative transfer, chemistry, and even solid state physics. The MHD is itself non-ideal, since the medium is partially ionized, and in addition self-gravity is important in many circumstances. Self-gravity effects and the level of electrical conductivity are very sensitive to thermal and ionization properties, which in turn are determined by chemistry and radiative transfer (including X-rays and cosmic rays), and the latter are strongly affected by grain properties that evolve in time due to sticking and fragmentation. Compounding the difficulty imposed by the interactions among the physical processes involved is the lack of exact knowledge of initial and boundary conditions: how does collapse of the rotating protostellar core shape the distribution of mass in the disk, starting from the initial disk-building stage and continuing (although at a reduced rate) with later infall? Finally, there is the difficulty imposed by the huge dynamic range in space and time; disks themselves span a range of $\sim 10^4$ in radius and $10^6$ in orbital period, while the small aspect ratio $H/R \ll 1$ (where $H$ is the scale height of the disk) implies a further extension in dynamic range is required for numerical models that resolve the disk interior.

Processes proposed to transport angular momentum in YSO disks generally fall into one of three categories: purely hydrodynamic mechanisms, MHD mechanisms, and self-gravitating mechanisms (e.g., see the reviews of Stone et al. 2000 and Gammie & Johnson 2005). Within the last decade, it has become possible to investigate mechanisms in each class using high-resolution time-dependent numerical simulations in two and three dimensions, in which the stresses that produce transport are explicitly obtained as spatial correlations of component velocities, magnetic fields, and the density and pressure for a self-consistent flow. Prior to the computational revolution that made these investigations possible, and continuing into the present for modeling in which large radial domains and long-term evolution is required, many studies have made use of the so-called “alpha prescription” for angular momentum transport. In this approach (Shakura & Sunyaev, 1973; Lynden-Bell & Pringle, 1974; Pringle, 1981), a stress tensor is
defined that yields an effective viscous torque between adjacent rings in a differentially rotating disk. On dimensional grounds, and using the fact that the shear stress should be zero for solid-body rotation, this stress can be written as \( T_{R,\phi} = -\alpha P d\ln\Omega/d\ln R \); i.e. the effective kinematic viscosity is taken to obey \( \nu = \alpha \sigma_{\text{th}}^2 / \Omega = \alpha \sigma_{\text{th}} H \). This effective viscosity Ansatz makes it possible to study disk evolution with a purely hydrodynamic, one-dimensional model. While the “\( \alpha \)-model” approach has been essential to progress on modeling disk observables, it is limited in its ability to capture realistic dynamics since the coefficient is arbitrary (and usually taken as spatially constant) and the adopted functional form for \( T_{R,\phi} \), while dimensionally correct, may not reproduce the true behavior of non-linear, time-dependent, three-dimensional flows (e.g., see Ogilvie 2003; Pessah, Chan, & Psaltis 2006). For a Keplerian disk, \(-d\ln\Omega/d\ln R = 3/2\) and in steady state the mass accretion rate is \( \dot{M}_{\text{disk}} = 3\pi \Sigma \nu = 3\pi \Sigma \alpha \sigma_{\text{th}}^2 / \Omega \); i.e. the ratio of radial inflow speed to orbital speed is \( (v_R/v_\phi) = (3/2)\alpha (\sigma_{\text{th}} / v_\phi) = (3/2)\alpha (H/R)^2 \). Observed accretion rates of TTSs require \( \alpha \sim 10^{-2} \) (Hartmann et al., 1998). Since the effective viscosity is equal to a characteristic length scale for angular momentum transport times a characteristic transport speed, the empirically-determined viscosity corresponds to a few percent of the value that would obtain if transport occurred at sonic speeds over distances comparable to the scale height of the disk.

Using the infall rate scaling of equation (43), the ratio of the disk accretion rate to the infall rate is

\[
\frac{\dot{M}_{\text{disk}}}{m_{\text{in}}} \sim \frac{\alpha}{\phi_{\text{in}}} \left( \frac{M_{\text{disk}}}{m_*} \right) \left( \frac{R}{H} \right) \left( \frac{T_{\text{disk}}}{T_{\text{core}}} \right)^{3/2},
\]

where we have assumed that the gravitational potential is dominated by the star. The outer-disk temperature is not much larger than the temperature in the core, and \( R/H \sim 10 \) for the outer disk, so the disk accretion rate is much lower than the infall rate unless \( M_{\text{disk}}/m_* \) or \( \alpha/\phi_{\text{in}} \) exceeds \( \sim 0.1 \). This is not the case for TTSs, but during the embedded stages the disk masses may be larger, and (possibly as a consequence of larger \( M_{\text{disk}} \) and self-gravity; see below) the values of \( \alpha \) may be larger as well.

**Hydrodynamic Mechanisms** The simplest transport mechanisms would be purely hydrodynamic. Turbulence generated either through convection (due to vertical or radial entropy gradients), through shear-driven hydrodynamic instabilities, or through external agents (such as time-dependent, clumpy infall) could in principle develop velocity field correlations \( \langle \rho \delta v_R \delta v_\phi \rangle \) of the correct sign \( (> 0) \) to transport angular momentum outward. Ryu & Goodman (1992) showed, however, that convective modes tend to transport angular momentum inward, rather than outward, and Stone & Balbus (1996) confirmed from three dimensional numerical simulations with turbulence driven by convection that angular momentum transport is inward. Convection driven by radial entropy gradients also transports angular momentum inward, and is generally stabilized by differential rotation (Johnson & Gammie, 2006).

Several analytic studies have shown that purely hydrodynamic disturbances in Keplerian-shear disks are able to experience large transient growth (Chagelishvili et al., 2003; Klahr, 2004; Umurhan & Regev, 2004; Johnson & Gammie, 2005a; Afshordi, Mukhopadhyay, & Narayan, 2005), especially for the case of two-dimensional (i.e. \( z \)-independent) columnar structures. Conceivably, transient
growth of sheared waves could lead to self-sustained turbulence with outward transport of angular momentum, if new leading wavelets could be continually reseeded in the flow via nonlinear interactions (Lithwick, 2007). While transient growth is indeed seen in two-dimensional \((R - \phi)\) numerical simulations, it is subject to secondary Kelvin-Helmholtz instability that limits the growth when \(|kR\delta v_R|/\Omega \gtrsim 1\) (Shen, Stone, & Gardiner, 2006). The turbulence that results also appears to decay without creating leading wavelets to complete the feedback loop, but this may be due to limited numerical resolution. Other numerical evidence, together with analytic arguments, suggest that nonlinear shear-driven hydrodynamic instabilities are unable to maintain turbulence for Rayleigh-stable rotational profiles (in which angular momentum increases outward, i.e. \(\kappa^2/\Omega^2 = 2d\ln(\Omega R^2)/d\ln R > 0\)) (Balbus, Hawley, & Stone, 1996; Hawley, Balbus, & Winters, 1999). Since simulations using the same numerical methods show that analogous Cartesian shear flows do exhibit nonlinear instability, rotating systems are presumably stabilized by Coriolis forces and the epicyclic motion that results. One potential concern is that the effective Reynolds numbers of numerical experiments are too low to realize nonlinear shear-driven instabilities and self-sustained turbulence. Very recently, however, Ji et al. (2006) reported from laboratory experiments at Reynolds numbers up to millions that hydrodynamic flows with Keplerian-like rotation profiles in fact show extremely low levels of angular momentum transport, corresponding to \(\alpha < 10^{-6}\).

Although it may be difficult to grow perturbations from instabilities in uniform Keplerian disks, it is still possible that disks are born with large internal perturbations, and that ongoing infall at all radii can continually resupply them. Simulations have shown that two-dimensional disks with non-uniform vorticity tend to develop large-scale, persistent vortices that are able to transport angular momentum outwards (Umurhan & Regev, 2004; Johnson & Gammie, 2005b). Three-dimensional simulations, however, show that vortex columns tend to be destroyed (Barranco & Marcus, 2005; Shen, Stone, & Gardiner, 2006). While off-midplane vortices can be long-lived (Barranco & Marcus, 2005), the angular momentum transport in three-dimensional simulations is an order of magnitude lower than for the two-dimensional case (Shen, Stone, & Gardiner, 2006), and secularly decays. Further investigation of this process is needed, and it is particularly important to assess whether vorticity can be injected at a high enough rate to maintain the effective levels of \(\alpha \sim 10^{-2}\) needed to explain observed TTS accretion.

**MHD Mechanisms** The introduction of magnetic fields considerably alters the dynamics of circumstellar disks. The realization by Balbus & Hawley (1991) that weakly-or-moderately magnetized, differentially rotating disks are subject to a powerful local instability – now generically referred to as the magnetorotational instability (MRI) – revolutionized the theory of accreting systems. Early axisymmetric numerical simulations showed robust growth and development of the so-called channel solution (Hawley & Balbus, 1991), while three-dimensional numerical simulations showed emergence of quasi-steady state saturated turbulence (Hawley, Gammie, & Balbus, 1995; Matsumoto & Tajima, 1995; Brandenburg et al., 1995) in which the angular momentum transport is outward, and is dominated by the magnetic stresses, \(\langle -B_R B_\phi/(4\pi) \rangle\). Much effort has been devoted to exploring the MRI as a basic mechanism driving accretion in a variety
of systems; Balbus & Hawley (1998) and Balbus (2003) summarize many of the
these developments. The effective value of $\alpha$ depends on the mean vertical mag-
netic flux, which presumably evolves over long timescales, and can easily exceed
0.1 (e.g., Hawley, Gammie, & Balbus 1996; Stone et al. 1996; Sano et al. 2004).
While MRI almost certainly plays an important role in driving accretion in
YSO systems, it is not a magic bullet. The difficulty is that substantial portions
of these disks may have ionization too low for MRI to be effective (Jin, 1996;
Gammie, 1996; Glassgold, Najita, & Igea, 1997; D’Alessio et al., 1998; Igea &
Glassgold, 1999), creating a “dead zone”. A critical review of the requirements for
MRI to develop in partially-ionized disks is given in Gammie & Johnson (2005);
Ohmic diffusion appears to be the main limiting effect, with the saturated-state
value of $\alpha$ dropping when $v^2_{A,\sigma}/\eta \Omega \lesssim 1$, where $\eta$ is the resistivity (Sano & Stone,
2002; Turner, Sano, & Dziourkevitch, 2007). In the very innermost parts of
YSO disks ($R \lesssim 0.1$ AU), alkali metals are collisionally ionized where the stellar
irradiation maintains the temperature above $\sim 2000$K, so MRI can operate. In
the outer disk (beyond several AU), and in the mid-disk’s surface layers, column
densities are low enough ($\Sigma < \Sigma_\alpha \sim 100$ g cm$^{-2}$) that X rays or cosmic rays
can penetrate the disk to ionize it. (For comparison, the surface density in the
minimum Solar nebula is $\Sigma = 1700(R/AU)^{-1.5}$ g cm$^{-2}$ – Hayashi, Nakazawa,
& Nakagawa, 1985.) Unfortunately, the extent of the MRI-active region in the
outer disk is very sensitive to the presence and size distribution of dust particles;
if small grains are present and well-mixed, the active region is quite limited,
while it can become very large if all the dust is incorporated in large particles or
settles to the midplane (Sano et al., 2000; Fromang, Terquem, & Balbus, 2002;
Desch, 2004; Salmeron & Wardle, 2005). Even if ionizing radiation is limited
to the surface layers by high total disk columns, if small grains are absent (an
extreme assumption) the gas-phase recombination rate is low enough such that
turbulence with rapid vertical mixing can maintain non-negligible ionization in
the interior. Turner, Sano, & Dziourkevitch (2007) have shown, using direct
numerical simulations, that the dead zone can be effectively eliminated in this
(optimistic) scenario; while the very center of the disk at 1 AU is not unstable to
MRI, the interior is still conductive enough that magnetic fields generated nearer
the surface can induce accretion in the midplane.

One of the possible consequences of spatially-varying conductivity in disks is
that the accretion rate will, in general, vary with radius. If only a surface layer
$\Sigma_\alpha$ is “active”, in the sense of being sufficiently conductive to support MRI with
effective viscosity coefficient $\alpha_\alpha$, then the accretion rate in that layer will be
$M_a = 3\pi\Sigma_\alpha \alpha_\alpha \sigma_{th}^2 / \Omega$. Since $\Sigma_\alpha$ varies slowly with radius (for the case of external
ionization) while the combination $\sigma_{th}^2 / \Omega$ tends to decrease inward, dropout
from the accretion flow can accumulate within the “dead zone” that is sandwiched
between active layers (Gammie, 1996). If the “dead zone” remains completely
inactive, then matter will build up until it becomes dynamically unstable and
begins to transport angular momentum by gravitational stresses (see below),
potentially leading to transient bursts of accretion (Gammie, 1996; Armitage,
Livio, & Pringle, 2001).

Finally, we note that MHD winds (see §4.2.5) may remove angular momentum
from disks, driving the matter remaining in the disk to accrete in order to main-
tain centrifugal balance. The angular momentum deficit is transferred to the disk
by magnetic stresses, so that only the matter that is well-coupled to magnetic
fields will be affected. Thus, the above considerations regarding ionization also apply to wind-driven accretion.

**Self-Gravitational Mechanisms** Accretion disks that have sufficiently small values of the Toomre parameter $Q = \kappa \sigma_{th}/(\pi G \Sigma) \sim (H/R)(m_*/M_{\text{disk}})$ are subject to nonlinear growth of density perturbations via the swing amplifier (see §§2.2, 3.2.1). Then, in addition to hydrodynamic Reynolds stresses $\langle \rho \nu_x \nu_y \rangle$ and MHD Maxwell stresses $\langle B_x B_y/(4\pi) \rangle$, gravitational “Newton stresses” $\langle B_x g_{xy} \rangle/(4\pi G)$ (where $g = -\nabla \Phi$) also contribute to the radial transport of angular momentum. Gammie (2001) showed that if the disk is in equilibrium such that cooling removes the energy dissipated by mass accretion at a rate per unit area $\Sigma \kappa \sigma_{th}^2/[(\gamma - 1) t_{\text{cool}}]$, then $\alpha^{-1} = (9/4)\gamma(\gamma - 1) t_{\text{cool}}$, where $\gamma$ is the effective (two-dimensional) adiabatic index (which takes into account vertical degrees of freedom, and depends on the three-dimensional index and degree of self-gravity). Numerical simulations with simple cooling prescriptions (constant $t_{\text{cool}} \Omega$) show that the disk can settle into a self-regulated state with $Q$ near unity (Gammie, 2001; Lodato & Rice, 2004; Rice, Lodato, & Armitage, 2005; Mejía et al., 2005), provided that $t_{\text{cool}} \Omega$ is not too small (in which case the disk fragments). For disks that are not externally illuminated, Johnson & Gammie (2003) performed two-dimensional simulations with realistic opacities (and a “one-zone” vertical radiative transfer approximation for cooling), and found that the transition between fragmentation and non-fragmentation lies in the range $t_{\text{cool}} \Omega = 1 - 10$. The corresponding $\Sigma$ at the transition point increases with $\Omega$, such that outer disks are the most active regions gravitationally. Values of $\alpha$ up to 0.5 are possible, with the equilibrium condition prediction satisfied down to $t_{\text{cool}} \Omega \approx 3$ and $\alpha \approx 0.1$. Using a three-dimensional model of an 0.07 $M_\odot$ disk with realistic cooling, Boley et al. (2006) find a value of $\alpha \sim 0.01$ over a large range of radii $> 20$ AU.

In view of the limitations on $\alpha$, Cesaroni et al. (2007) argue that accretion rates are limited to values substantially smaller than inferred for the formation of high-mass stars (§4.3.1). On the other hand, Krumholz, Klein, & McKee (2007) find that disks around high-mass protostars can transfer mass inward at the same rate that it falls in. They carried out simulations of high-mass star formation in a turbulent medium and included radiative transfer rather than prescribing the heating and cooling rates. They found that large amplitude $m = 1$ modes develop that give effective values of $\alpha$ of order unity, in qualitative agreement with the isothermal disk results of Laughlin & Bodenheimer (1994).

Disks that are illuminated sufficiently strongly will have the temperature set by the external radiation field, rather than internal dissipation of energy. In that case, whether self-gravity is important or not depends essentially on the amount of matter present in a given region. Where the surface density is high enough so that $Q$ is near but not below the critical value $\approx 1.4$, self-gravitational stresses will be appreciable but not so large as to cause fragmentation. Analytic estimates assuming steady state and accretion heating as well as irradiation (Matzner & Levin, 2005; Rafikov, 2005) indicate that fragmentation is only possible in the outer portions of disks, although more massive disks, around more massive stars, are more subject to fragmentation (Kratter & Matzner, 2006). At temperatures comparable to those in observed systems, disks with masses $\gtrsim 0.1$ $M_\odot$ are candidates for having significant mass transport due to self-gravitating torques (Mayer et al., 2004). Thus, self-gravity is likely to be particularly important during the
embedded stage of disk evolution, when disk masses are the largest. Vorobyov & Basu (2005b, 2006) propose, based on results of two-dimensional simulations, that recurrent “bursts” of accretion due to self-gravity are likely to develop during the early stages of protostellar evolution. A number of other results from models of self-gravitating disk evolution (with an emphasis on criteria for planet formation through fragmentation) are presented in the review of Durisen et al. (2007).

4.2.3 DISK CLEARING While a large proportion of the mass in the disk ultimately accretes onto the star, conservation of angular momentum requires that some of the matter be left behind. MHD winds during the main lifetime of the disk remove some of this material (see §4.2.5). What remains is either incorporated into planets, or removed by photoevaporation. Although planet formation is inextricably coupled to disk evolution, recent developments in this exciting – and rapidly expanding – field are too extensive to summarize here. A number of excellent recent reviews appear in Protostars and Planets V.

Disks can be irradiated by UV and X ray photons originating either in their own central stars, or in other nearby, luminous stars (see e.g. reviews of Hollenbach, Yorke, & Johnstone 2000 and Dullemond et al. 2007). EUV radiation penetrates only the surface layer of the disk, where it heats the gas to $\sim 10^4$ K (the ionization and heating depth is determined by the Strömgren condition); FUV penetrates deeper into the disk (where densities are higher), but heats gas to only a few 100 K (Hollenbach et al., 1994; Johnstone, Hollenbach, & Bally, 1998). The characteristic radial scale in the disk for a thermally-driven wind is the gravitational radius $r_g = Gm/(kT)$, where $T$ is the temperature at the base of the flow. Pressure gradients enable flows to emerge down to $(0.1 - 0.2)r_g$ (Begelman, McKee, & Shields, 1983; Font et al., 2004; Adams et al., 2004). EUV-driven winds are most important in the inner disk, since the gravitational potential there is too deep for FUV-heated regions at modest temperatures to escape.

Observations discussed above (see also Simon & Prato 1995 and Wolk & Walter 1996) indicate that the inner and outer disks surrounding YSOs disperse nearly simultaneously and on a very short ($\sim 10^5$ yr) timescale, based on the small number of transition objects between classical and weak T Tauri systems and the typical CTT lifetimes of a few to several Myr. Since the accretion time of the outer disk itself determines the system lifetime, rapid removal of the outer disk must be accomplished by other means; photoevaporation is the most natural candidate. Models of photoevaporation that also include viscous disk evolution (which allow spreading both inward and outward) have very recently shown that rapid and near-simultaneous removal of the whole disk indeed occurs (Clarke, Gendrin, & Sotomayor, 2001; Alexander, Clarke, & Pringle, 2006a,b). In this process, the accretion rate declines slowly over time until the photoevaporative mass loss rate at some location in the inner disk exceeds the rate at which mass is supplied from larger radii. The inner disk, which is no longer resupplied from outside, then drains rapidly into the star. At the same time, the radiative flux onto the outer disk grows as it is no longer attenuated by the inner disk’s atmosphere; the photoevaporation rate in the outer disk climbs dramatically, and it is removed as well.

4.2.4 OBSERVATIONS OF YSO JETS AND OUTFLOWS Young stellar systems drive very powerful winds. The clearest observable manifestations of YSO winds are the central “Herbig-Haro” jets consisting of knots of ionized gas
(v \gtrsim 100 \text{ km s}^{-1}), and the larger-scale bipolar outflows consisting of expanding lobes of molecular gas (v \sim 10 \text{ km s}^{-1}). “Jet-like” outflows (i.e. high-v, narrow molecular structures) are also observed in some circumstances (see below). The high velocities of jets indicate that they represent (a part of) the primary wind from the inner part of the star-disk system, while the low velocities and large masses of (broad) molecular outflows indicate that they are made of gas from the star’s environment that has been accelerated by an interaction with the wind. In addition to these observed signatures, there may be significant gas in a large-scale primary wind surrounding the jet, which remains undetected due to lower excitation conditions (low density, temperature, and/or ionization fraction).

Outflows are ubiquitous in high-mass star formation as well as in low-mass star formation (Shepherd & Churchwell, 1996). Outflows from high-mass protostellar objects with L < 10^5 L_\odot (corresponding to m_* < 25 M_\odot—Arnett, 1996) are collimated (Beuther et al., 2002b), but somewhat less so than those in low-mass protostars (Wu et al., 2004). In some cases, jets are observed with the outflows, and in these cases the momentum of the jet is generally large enough to drive the observed outflow (Shepherd, 2005). No well-collimated flow has been observed in a source with L > 10^5 L_\odot; as remarked above, disks that are clearly circumstellar have not been observed in such sources either. Beuther & Shepherd (2005) have proposed an evolutionary sequence that is consistent with much of these data: A protostar that eventually will become an O star first passes through the HMPO stage with no H II region and with a well-collimated jet. When the star becomes sufficiently massive and close to the main sequence that it produces an H II region, the outflow becomes less collimated. The collimation systematically decreases as the star grows in mass and the H II region evolves from hypercompact to ultracompact (see §4.3.4). The remainder of this section focuses on winds and outflows from low-mass stars, which have been observed in much greater detail than their high-mass counterparts.

Recent reviews focusing on the observational properties of jets include those of Eisloeffel et al. (2000), Reipurth & Bally (2001), and Ray et al. (2007). Jets are most commonly observed at high resolution in optical forbidden lines of O, S, and N, as well as Halpha, but recent observations have also included work in the near-IR and near-UV. For CTTs, which are YSOs that are themselves optically revealed, observed optical jets are strongly collimated (aspect ratio at least 10:1, and sometimes 100:1), and in several cases extend up to distances more than a parsec from the central source (Bally, Reipurth, & Davis, 2007). The jets contain both individual bright knots with bow-shock morphology, and more diffuse emission between these knots.

The emission diagnostics from bright knots are generally consistent with heating by shocks of a few tens of km s^{-1} (Hartigan, Raymond, & Hartmann, 1987; Hartigan, Morse, & Raymond, 1994), producing post-shock temperatures T_e \approx 10^4 \text{ K}. The electron density n_e, ionization fraction x_e = n_e/n_{H}, and temperature T_e can be estimated using line ratios (Bacciotti & Eisloffel, 1999). Analyses of spectra from a number of jets yields a range of parameters n_e = (50–3 \times 10^3) \text{ cm}^{-3} and x_e = 0.03 – 0.6 so that n = (10^3 – 10^5) \text{ cm}^{-3} (Podio et al., 2006). The total mass loss rate in jets \dot{M}_{\text{jet}}, and hence the total jet momentum flux, \dot{M}_{\text{jet}}v_{\text{jet}}, can be estimated using jet densities and velocities together with an emission filling factor, yielding \dot{M}_{\text{jet}} = (10^{-8} – 10^{-7}) \text{ M}_\odot \text{ yr}^{-1} for CTTs (Podio et al., 2006). For Class 0 sources, which are much more luminous and have much higher accretion rates, estimated mass-loss rates in jets based on O I emission (from shocked
gas) extend up to $\dot{M}_{\text{jet}} \sim 10^{-6} \, M_\odot \, \text{yr}^{-1}$ (Ceccarelli et al., 1997). Inferred values of $\dot{M}_{\text{jet}}$ are generally correlated with estimates of $\dot{M}_{\text{disk}}$ from veiling (Hartigan, Edwards, & Ghandour, 1995), with the ratio in the range 0.05 – 0.1 (Ray et al., 2007).

The densities and temperatures obtained from jet diagnostics indicate internal pressures $P/k_B = (10^7 - 10^9) \, \text{K cm}^{-3}$ in the jet, exceeding the ambient pressure in the surrounding core and GMC by a factor $10^2 - 10^4$. In principle, infalling envelope gas could provide a “nozzle” to collimate an emerging wind, but simulations indicate that only relatively weak winds can be so confined as to produce a narrow jet (Delamarter, Frank, & Hartmann, 2000). This implies that observed jets must be contained within a broader wind, with collimation likely produced by magnetic hoop stresses (see below). Emission line analyses in fact indicate that a lower-velocity ($\sim (10 - 50) \, \text{km s}^{-1}$) wind component is present near the source, surrounding the high-velocity flow of a few 100 km s$^{-1}$ that emerges as the large-scale jet (Hartigan, Edwards, & Ghandour, 1995; Hirth, Mundt, & Solf, 1997; Bacciotti et al., 2000; Pyo et al., 2005). Since velocities of MHD winds scale with the Keplerian rotation speed of the footpoint (see below), the presence of both high- and low-velocity components suggests that winds are driven from a range of radii in the disk. Recent high-resolution observations have detected signatures of differential rotation in jets, using near-UV, optical, and near-IR lines (Bacciotti et al., 2002; Ray et al., 2007); these also indicate a range of wind launch points.

Recent reviews of the observational properties of molecular outflows include those of Bachiller & Tafalla (1999), Richer et al. (2000), and Arce et al. (2007). Like jets, classical molecular outflows can extend to distances 0.1 – 1 pc from the central star, but they have much lower velocities (up to a few tens of km s$^{-1}$) and collimation (aspect ratio $\sim 3 - 10$). In a few very young, embedded sources, molecular jets with much higher velocities and aspect ratios have been observed in H$_2$, CO, and SiO lines (e.g., Gueth & Guilloteau 1999; Beuther et al. 2002a; Lee et al. 2007). The total momentum flux carried in CO outflows is correlated with the bolometric luminosity of the source and is discussed in §3.2.2. For embedded sources with $L_{\text{bol}} = 1-10^5 L_\odot$, the momentum flux is $10^{-4} - 10^{-1} \, M_\odot \, \text{km s}^{-1} \, \text{yr}^{-1}$ (Richer et al., 2000); in optically-revealed sources, this declines considerably (e.g. Bontemps et al. 1996).

Detailed spectroscopic and morphological analysis of outflows enable intercomparisons with theoretical models. Mapping of outflows reveals both simple expanding shells, and more complex features such as multiple cavities and bow shock structures that are suggestive of episodic ejection events (Lee et al., 2002); outflow lobes become broader and more irregular over time (Arce & Sargent, 2006). Channel maps and position-velocity diagrams in some sources show parabolic structures that are consistent with driving by wide-angle winds, and in other sources show spur structures that are consistent with jet driving (Lee et al., 2000, 2001). The strongly curved morphology of internal bow shocks (as seen in both molecular and atomic tracers) indicates that the wind must have velocities that decrease away from the poles, since a time-variable wind with latitudinally-constant velocity produces nearly flat internal shocks (Lee et al., 2001). This implies, in turn, that the wind is driven from a range of radii in the disk, rather than arising from only a narrow region.
4.2.5 DRIVING MHD WINDS AND JETS

It was recognized very early on that jets and outflows contain more momentum than could possibly be driven by radiation pressure (Lada, 1985), whereas the high efficiencies and velocities found by Blandford & Payne (1982) for MHD winds driven from accretion disks in near-Keplerian rotation suggested that the same magnetocentrifugal mechanism could drive winds in YSO systems (Pudritz & Norman, 1983). The main requirement for these winds to develop is for the disk to be threaded by magnetic fields of sufficient strength. The mathematical theory of MHD winds and jets is presented in e.g. Spruit (1996) and Pudritz (2004).

Over the years, two main types of MHD wind models for YSO systems have been explored. One, the “x-wind” model (see Shu et al. 1994 and references in Shu et al. 2000 and Shang et al. 2007), focuses on the interaction region between the stellar magnetosphere and the inner accretion disk as the source of the wind. In this model, a large portion of the stellar dipole flux is taken to be concentrated into a small range of radii near the point where the magnetosphere and disk corotate. Since YSOs are rapid rotators, the corotation point is close to the star, and the wind that would be launched could have terminal speed of a few 100 km s\(^{-1}\), as is observed in jets. The second class of MHD wind models assumes that a much larger region of the disk is threaded by open field lines, such that there would be a range of terminal wind speeds, reflecting the range of rotation speeds at the magnetic field’s footpoints in the disk (see references in Konigl & Pudritz 2000 and Pudritz et al. 2007). For disk winds, the poloidal magnetic flux could in part be generated locally (e.g. by an MRI dynamo), in part be advected inward with the collapse of the prestellar core, and in part originate in the star and diffuse outward into the disk. Since one type of wind would not exclude the other, it is likely that both x-winds and disk winds are present at some level. This might help, for example, explain particular features of jets such as their strong central density concentration as well as the apparent decrease in velocity from inside to outside.

The observed rotation velocities in jets can be used to infer the launch point in the disk (Anderson et al., 2003). From the Bernoulli equation for a cold flow along a streamline which rotates with angular velocity \(\Omega_0\), the quantity \(E = \frac{1}{2} |v|^2 + \Phi_g - v_\phi \Omega_0 R\) is constant, where \(\Phi_g = -Gm_*/r\) is the gravitational potential and in this section \(R\) denotes the cylindrical radius. At \(R_{\text{obs}}\), where the wind is observed (sufficiently beyond the Alfvén transition), the dominant terms in the \(E\) equation are the first and the last terms on the RHS. For the cases of interest, \(v_{\phi,\text{obs}} \ll v_{p,\text{obs}}\), where \(v_p\) is the poloidal velocity, and \(|E|^{1/2} = (3/2)^{1/2} \Omega_0 R_0 \ll v_{p,\text{obs}}\), so that \(\Omega_0 \approx v_{p,\text{obs}}^2/(2v_{\phi,\text{obs}} R_{\text{obs}})\). The specific angular momentum \(j = R(v_\phi - B_\phi B_p/[4\pi \rho p])\) is also conserved along streamlines. One can show that this is equal to \(\Omega_0 R_A^2\), where \(R_A\) is the Alfvén radius of the wind. Since \(j\) is dominated by the kinetic term at large distance (where the wind is super-fast-magnetosonic), observations can be used to infer the ratio \(R_A/R_{\text{obs}} \approx \sqrt{2v_{\phi,\text{obs}}/v_{p,\text{obs}}}\). For the low velocity component of DG Tau, Anderson et al. (2003) find from calculating \(\Omega_0\) as above that the wind launch point radii are \(\sim 0.3 - 4\)AU, implying a disk wind. The high velocity component could originate as either an x-wind or a disk wind from smaller radii. For DG Tau, the inferred ratio \(R_A/R_0 \approx 2 - 3\) is also consistent with numerical solutions that have been obtained for disk winds (see Pudritz et al. 2007 for a summary). This implies that the angular momentum carried by the wind, \(\dot{M}_{\text{wind}} \Omega_0 R_A^2\), which equals the
angular momentum lost by the disk, $\dot{M}_{\text{disk}} \Omega_0 R_0^2$, can drive accretion at a rate $\dot{M}_{\text{disk}}/\dot{M}_{\text{wind}} = (R_A/R_0)^2 \sim 4 - 9$.

The acceleration of MHD winds is provided by a combination of the centrifugal “flinging” effect produced by rigid poloidal fields, and gradients in the toroidal magnetic pressure in the poloidal direction (e.g. Spruit (1996)). Beyond the Alfvén surface, magnetic hoop stresses will tend to bend streamlines toward the poles. Full cylindrical streamline collimation, in the sense of $v_p \parallel z$ asymptotically, can only occur if $B_0 R$ is finite for $R \to \infty$ (Heyvaerts & Norman, 1989). Using solutions in which all velocities scale as $v, v_A \propto r^{-1/2}$ and the density and magnetic field respectively scale as $\rho \propto r^{-q}$ and $B \propto r^{-(1+q)/2}$, Ostriker (1997) showed, however, that cylindrically-collimated disk winds are slow, in the sense that the asymptotic value of $v_p/\Omega_0 R_0$ is at most a few tenths. Since observed jets are fast, they must either have their streamlines collimated by a slower external wind, or else be collimated primarily in density rather than velocity. Time-dependent simulations have also shown that the degree of collimation in the flow depends on the distribution of magnetic flux in the disk; cases with steeper distributions of $B$ with $R$ tend to be less collimated in terms of streamline shapes (Fendt, 2006; Pudritz, Rogers, & Ouyed, 2006).

The idea that nearly radially-flowing wide-angle MHD winds may produce a “jetlike” core, with density stratified on cylinders, was first introduced by Shu et al. (1995) in the context of x-winds. This effect holds more generally, however, as can be seen both analytically (Matzner & McKee, 1999) and in simulations (see below). Asymptotically, the density approaches $\rho \to |B_0| R k/(\Omega_0 R^2)$ where $k$ is the (conserved) mass flux-to-magnetic flux ratio (also termed the mass-loading parameter). Since nearly radially-flowing winds must be nearly force-free, $|B_0| R$ varies weakly with $R$, such that if the range of $k/\Omega_0$ over footpoints is smaller than the range of $R$ over which the solution applies (which is generally very large), the wind density will vary as $R^{-2}$. The $R^{-2}$ dependence cannot continue to the origin; Matzner & McKee (1999) suggested that precession, internal shocks due to fluctuating wind velocity, or magnetic instabilities would result in a flattening of the density close to the axis so that the momentum flux in the wind $\rho v_w \propto (1 + \theta_0^2 - \cos^2 \theta)^{-1}$, where $\theta$ is the angle of the flow relative to the axis and $\theta_0 \ll 1$ measures the size of the flattened region. This distribution gives approximately equal amounts of momentum in each logarithmic interval of angle for $\theta > \theta_0$. Several time-dependent numerical MHD simulations have demonstrated this density collimation effect for wide-angle winds (Gardiner, Frank, & Hartmann, 2003; Krasnopolsky, Li, & Blandford, 2003; Anderson et al., 2005).

Magnetized winds are subject to a variety of instabilities (e.g. Kim & Ostriker 2000; Hardee 2004), which may contribute to enhancing the confinement of the jet, structuring the jet column (yielding wanderers, twists, and clumps), and mixing with the ambient medium at interfaces. Since jets are likely surrounded by wider winds, they are to some extent protected from the development of Kelvin-Helmholtz and helical modes that that disrupt jets propagating through ambient gas, although development of axisymmetric pinch modes may still contribute to the formation of HH knots (Hardee & Rosen, 2002). In addition, lightly-loaded poloidal flux within the central core of the wind/jet may help suppress the growth of large-scale pinch and kink instabilities (Ostriker & Shu, 1995; Anderson et al., 2006). Time-dependent simulations focusing on the portion of the wind flow above the disk show that while steady winds are possible in certain ranges of the
mass-loading parameter for a given distribution of magnetic flux, in other ranges no steady solution is possible (Ouyed & Pudritz, 1999; Anderson et al., 2005). Since the spectral diagnostics of HH objects indicate shock speeds of a few tens of km s\(^{-1}\), it is plausible that they form due to nonlinear steepening and shocking of wind instabilities.

### 4.2.6 ORIGINS AND EFFECTS OF OUTFLOWS

Overall, the structure and kinematics of molecular outflows suggest that they are driven by winds that originate from a range of radii in the disk, with a dense central core (seen as a jet) surrounded by a lower-density, lower-velocity wide-angle wind. Jet driving and wind driving of outflows have traditionally been explored separately, although in practice they would operate in tandem.

Jets drive outflows as bow shocks, with ambient material swept into a thin shell and carried away from the body of the jet as the shock overtakes and entrains it (Raga & Cabrit, 1993; Masson & Chernin, 1993), mixing newly shocked material with material that is already flowing outward (Smith, Suttner, & Yorke, 1997). The leading bow shock is itself created due to pressure forces at the circumferential boundary of the working surface at the head of the jet, which drive transverse flows. Jets with internal shocks can create analogous bow shocks, with the difference that internal bow shocks would propagate into the wind that surrounds the jet, whereas the leading bow shock would propagate into the ambient medium. Leading bow shocks tend to be fairly narrow, because the cooling of shocked gas in the working surface limits the transverse thrust that can be applied to the shell (Downes & Ray, 1999). As a consequence, the width of the shell increases only as the cube root of the distance from the head of the jet (Masson & Chernin, 1993; Ostriker et al., 2001). Thus, bow shocks have difficulty explaining broad outflows. On the other hand, the “convex spur” velocity features seen in some systems agree well with the predictions of bow shock models (Lee et al., 2001; Ostriker et al., 2001).

For wide-angle winds, the momentum flux contained in the transverse bulk motion of the wind is large compared to the thrust that could be provided by pressure forces in the shell of shocked (strongly cooling) gas, so that a momentum-conserving “snowplow” flow is a good approximation. Shu et al. (1991) developed the “wind-swept shell” model of outflows based on this concept, which was able to explain the large opening angles seen in most outflows. Li & Shu (1996) and Matzner & McKee (1999) extended the wind-swept shell analytic model to incorporate the characteristic \(R^{-2}\) density stratification and logarithmic collimation of streamlines of asymptotic wide-angle MHD winds, also allowing for latitudinal density stratification in the surrounding core. The mass-velocity and position-velocity relations for these analytic models agree well with those in observed outflows. Numerical simulations of outflows swept up by wide-angle winds (Lee et al., 2001; Shang et al., 2006) are in good agreement with the results of analytic models.

Outflows affect both the immediate environment of the forming star (removing mass from the core before it can collapse into a disk), the clump in which the core forms (also removing mass, and injecting energy), and the larger-scale cloud (injecting energy). The effects of energy injection on clumps forming clusters of stars, and on GMCs as a whole, are discussed in §4.3.5 and §3.2.2, respectively. Mass removal by winds is related to the star formation efficiency, as we next discuss.
The star-formation efficiency $\epsilon$ can be defined for individual cores, for star-forming clumps, or for GMCs. The correspondence between the Core Mass Function and the IMF has been discussed in §3.3; they are related by the core star-formation efficiency, $\epsilon_{\text{core}} \equiv m_*/M_{\text{core}}$. (The individual-star IMF must also take into account the multiplicity of the stars formed in a given core.) Nakano, Hasegawa, & Norman (1995) showed that outflows from protostars could reverse the infall and determine $\epsilon_{\text{core}}$; they assumed spherical winds and found $\epsilon_{\text{core}} \sim$ a few percent. Matzner & McKee (2000) calculated the dynamics of the outflows including collimation and obtained $\epsilon_{\text{core}} \sim 0.25 - 0.75$, depending on the degree of flattening of the core due to magnetic support. Subsequent observations suggest $\epsilon_{\text{core}} \approx 1/5 - 1/3$ (§3.3.1), at the low end of this range. They also evaluated the star-formation efficiency for a star-forming clump, and found typical values somewhat less than 0.5. The predicted values of $\epsilon_{\text{core}}$ are inversely proportional to the momentum per unit mass in the outflow, $p_w/m_*$; they are consistent with observation for $p_w/m_* \sim 40 \, \text{km s}^{-1}$ as assumed, but not if $p_w/m_*$ is much smaller (see §3.2.2 for a discussion of the values of $p_w$ inferred from observation). Both Nakano, Hasegawa, & Norman (1995) and Matzner & McKee (2000) found that $\epsilon_{\text{core}}$ is only weakly dependent on the core mass, so that the CMF and the IMF should be similar in shape, as observed (§3.3).

4.3 High-Mass Star Formation

High-mass protostars are characterized by Kelvin-Helmholtz times less than the accretion time, so that they undergo nuclear burning while still accreting (§4.1). This leads to two powerful feedback effects that do not apply to low-mass protostars, radiation pressure and photoionization (Larson & Starrfield, 1971). Furthermore, high-mass protostars tend to form in dense clusters, so that interaction with other protostars and newly formed stars may be important in their evolution. Drawing on the review of Beuther et al. (2007) we first summarize work on infall onto high-mass protostars and then discuss the feedback effects.

4.3.1 Protostellar Infall

High-mass star formation is generally taken to be a scaled up version of low-mass star formation: The accretion rate is $\dot{m}_* \sim \epsilon_{\text{eff}}^3 / G$, where the effective sound speed $c_{\text{eff}}$ includes the effects of thermal gas pressure, magnetic pressure, and turbulence (Stahler et al. 1980, although they did not address the issue of high-mass star formation). As discussed in §4.1, there may be a numerical factor of a few in front of the $\epsilon_{\text{eff}}^3 / G$. Wolfire & Cassinelli (1987) found that accretion rates of order $10^{-3} \, M_\odot \, \text{yr}^{-1}$ are needed to overcome the effects of radiation pressure for the highest stellar masses, and attributed this to the high values of $c_{\text{eff}}$ in high-mass star forming regions. Myers & Fuller (1992) used their “TNT” model (§2.2) to infer formation times for $(10 - 30) \, M_\odot$ stars of $(6 - 10) \times 10^5 \, \text{yr}$; the turbulent envelopes allow equilibrium cores to have greater densities and shorter collapse times than those supported by thermal pressure alone. Caselli & Myers (1995) extended this to more massive stars and found formation times $> 10^6 \, \text{yr}$ for stars of $100 \, M_\odot$, a significant fraction of the main sequence lifetime. On the other hand, by modeling the spectral energy distributions (SEDs) of high-mass protostars, Osorio, Lizano, & D’Alessio (1999) inferred that high-mass stars form in somewhat less than $10^5 \, \text{yr}$, and favored a logatropic model for the density distribution of the core. Nakano et al. (2000) inferred an accretion rate of $10^{-2} \, M_\odot \, \text{yr}^{-1}$ (corresponding to a formation time of a few thousand years) for the source IRc2 in Orion based on the assumption that...
the accretion rate is \( \sim 10 \epsilon_{\text{eff}}^3 / G \), with the effective sound speed \( c_{\text{eff}} \) determined from the observed line width.

The turbulent core model for high-mass star formation (McKee & Tan, 2002, 2003) follows from the assumption that such stars form in turbulent, gravitationally bound cores (virial parameter \( \alpha_{\text{vir}} \sim 1 \)). The turbulence is self-similar on all scales above the Bonnor-Ebert scale, where thermal pressure dominates. The star-forming clump and the protostellar cores within it are assumed to be centrally concentrated so that the pressure and density have a power-law dependence on radius, \( P \propto r^{-k_p}, \rho \propto r^{-k_p} \). It follows that the cores are polytropes (\S 2.2), and since the Bonnor-Ebert scale is small, the cores are approximately singular. The protostellar infall rate is determined by the surface density of protostellar core, which in turn is comparable to that of the clump in which it is embedded. The regions of high-mass star formation studied by Plume et al. (1997) have surface densities \( \Sigma_{\text{cl}} \sim 1 \) g cm\(^{-2} \), corresponding to visual extinctions \( A_V \sim 200 \) mag; these values are similar to those for observed star clusters in the Galaxy (e.g., \( \sim 0.2 \) g cm\(^{-2} \) in the Orion Nebula Cluster, 0.8 g cm\(^{-2} \) for the median globular cluster and \( \sim 4 \) g cm\(^{-2} \) in the Arches Cluster). By contrast, regions of low-mass star formation have \( \Sigma \sim 0.03 \) g cm\(^{-2} \), corresponding to \( A_V \sim 7 \) mag (Onishi et al., 1996).

The radius of a protostellar core is

\[
R_{\text{core}} = \left( \frac{M_{\text{core}}}{\pi \Sigma_{\text{core}}} \right)^{1/2} \approx 0.06 \left( \frac{m_{sf}}{60 \epsilon_{\text{core}} M_{\odot}} \right)^{1/2} \frac{1}{\Sigma_{\text{cl}}^{1/2}} \text{ pc}, \tag{51}
\]

where \( m_{sf} \) is the final stellar mass. The second expression is based on the result that the surface density of a typical core is comparable to that of the clump in which it is embedded; cores near the center of a clump have higher surface densities, and the sizes are correspondingly smaller. Using the results of McLaughlin & Pudritz (1997) for the inside-out collapse of a singular polytrope and adopting \( k_p = \frac{3}{2} \), a typical density power law from Plume et al. (1997), McKee & Tan (2003) found that the typical infall rate and the corresponding time to form a star of mass \( m_{sf} \) are

\[
m_{sf} \simeq 0.5 \times 10^{-3} \left( \frac{m_{sf}}{60 \epsilon_{\text{core}} M_{\odot}} \right)^{3/4} \Sigma_{\text{cl}}^{3/4} \left( \frac{m_s}{m_{sf}} \right)^{0.5} M_{\odot} \text{ yr}^{-1}, \tag{52}
\]

\[
t_{sf} \simeq 1.3 \times 10^5 \left( \frac{m_{sf}}{60 \epsilon_{\text{core}} M_{\odot}} \right)^{1/4} \Sigma_{\text{cl}}^{-3/4} \text{ yr}, \tag{53}
\]

where \( \Sigma_{\text{cl}} \) is the surface density (in g cm\(^{-2} \)) of the several thousand \( M_{\odot} \) clump in which the star is forming. For typical values of \( \Sigma_{\text{cl}} \sim 1 \) g cm\(^{-2} \), the star formation time is of order \( 10^5 \) yr and the infall rate is of order \( 10^{-3} M_{\odot} \) yr\(^{-1} \). This infall rate is large enough to overcome the effects of radiation pressure at the dust destruction front, thereby addressing one of the key theoretical difficulties for models of high-mass star formation (see below). The mean infall rate could be somewhat larger than given in equation (52) if the core was initially overdense or contracting, and turbulence in the core could generate large fluctuations in the infall rate. However, the infall rate given above is only a few times greater than the free-fall value and is unlikely to be much larger.

The key assumptions in this model are that stars form from pre-assembled cores (although since the cores are turbulent, there will be significant mass exchange with the ambient medium); that the cores and the clumps in which they
are embedded are in approximate virial equilibrium; and that they are magnetically supercritical, so that the magnetic field does not significantly limit the rate of accretion. Evidence in support of the first assumption has been obtained by Beuther, Sridharan, & Saito (2005) and Sridharan et al. (2005); the remaining assumptions are also subject to observational test. The model is necessarily approximate, since it treats the turbulence as a local pressure (the microturbulent approximation), and since it incorporates all the feedback effects due to radiation pressure and photoevaporation in the core star formation efficiency, $\epsilon_{\text{core}}$, which was assumed to be of order $1/2$. Some of the large density fluctuations in the supersonically turbulent cores will form low-mass stars, but most of the mass of the core is assumed to go into one or two massive stars. Dobbs, Bonnell, & Clark (2005) have criticized the model on the ground that the massive cores would fragment and form many low-mass stars rather than a single massive star, but radiative heating by the rapidly accreting high-mass protostar strongly suppresses fragmentation (Krumholz, 2006; Krumholz, Klein, & McKee, 2007). The turbulent core model is consistent with the correspondence between the core mass function and the IMF (§3.3), and it naturally allows for the disks and winds associated with high-mass stars (see §4.2) since it is an extrapolation of low-mass star formation theory. The cores are predicted to be denser than the clump in which they are embedded by about $(M_{\text{clump}}/M_{\text{core}})^{1/2}$, which is much greater than unity for stellar mass cores embedded in clumps with $M > 10^3 M_\odot$; this naturally overcomes the crowding problem.

An alternative class of gravitational collapse models involves rapidly accelerating accretion ($m_\ast \propto m_q^q$ with $q > 1$, so that $m_\ast \to \infty$ in a finite time in the absence of other effects). Building on the work of Norberg & Maeder (2000), Behrend & Maeder (2001) assumed that the accretion rates are proportional to the mass outflow rates observed in protostellar outflows; since the outflows are swept-up material, the justification for this assumption is unclear. They found $t_{\text{sf}} \sim 3 \times 10^5$ yr for massive stars, with most of the growth occurring in the last 10% of this time. In the competitive accretion model (Bonnell et al., 1997; see §4.1.2), massive stars form via Bondi-Hoyle-Lyttleton accretion ($m_\ast \propto m_\odot^2$). Keto (2002, 2003) has studied this model further, focusing on the associated H II regions. For a 10 $M_\odot$ star in a typical high-mass star-forming clump observed by Plume et al. (1997), which has a mass $\sim 4000 M_\odot$ and a virial parameter of order unity, the Bondi-Hoyle accretion rate is much smaller than that expected in the turbulent core model (McKee & Tan, 2003), even after allowing for the turbulent enhancement factor $\phi_{\text{BH}}$ (eq. 47). The rate of Bondi-Hoyle accretion increases if the virial parameter is small, if the infall occurs onto a cluster of stars that is much more massive than a single star (as Keto comments), or if the infall occurs from a significantly less massive clump. In the latter two cases the assumptions underlying Bondi-Hoyle accretion begin to break down, and further study is needed to determine the infall rate. Edgar & Clarke (2004) have shown that radiation pressure halts Bondi-Hoyle accretion when the star is moving supersonically relative to the gas for $m_\ast > 10 M_\odot$, since the luminosity is large enough that radiation pressure deflects gas away from the star.

In view of the challenges facing conventional theories of high-mass star formation, Bonnell, Bate, & Zinnecker (1998) made the radical suggestion that high-mass stars form via stellar collisions. This model requires stellar densities $\sim 10^8$ stars pc$^{-3}$ during the brief period in which the stars grow by merging. This coalescence model produces an IMF that is in qualitative agreement with obser-
vations, although no feedback effects were included in the calculations (Bonnell et al., 2001b). This model faces a number of challenges: (1) The required stellar density is far greater than has been observed in any Galactic star cluster. For example, W3 IRS5 is one of the densest clusters observed to date, with 5 proto OB stars in a sphere of radius 0.015 pc (Megeath, Wilson, & Corbin, 2005); the corresponding stellar density $\sim 4 \times 10^5$ pc$^{-3}$ is lower than required by the coalescence model by more than 2 orders of magnitude, although it must be borne in mind that the number of lower mass stars in that volume is currently unknown. (2) For large OB protoclusters, the hypothesized ultra-dense state would produce a very luminous, compact source, yet this has never been observed. (3) The mass loss that is hypothesized to reduce the cluster density to observed values must be finely tuned in order to leave the cluster marginally bound. (4) Finally, it is difficult to see how the model could account for observed associated disks and outflows (§§4.2.1, 4.2.4) above. Bally & Zinnecker (2005) discuss a number of observational tests of the coalescence model, and suggest that the wide-angle outflow from OMC-1 in the Orion molecular cloud could be due to the merger of two protostars that released $10^{48} - 10^{49}$ erg. Two variants of the coalescence model have been suggested: Stahler, Palla, & Ho (2000) proposed that gas bound to the protostars could increase the cross section for collisions, although they did not explain why this would result in stellar coalescence rather than the formation of a binary. Bonnell & Bate (2005) have proposed an explanation for this: assuming that the gas has negligible angular momentum (which is plausible if the turbulence is weak, as assumed in the competitive accretion model), then accretion drives the stars in the binary to closer separations and ultimately to a merger. The stellar density required for the binary coalescence model is $\sim 3 \times 10^6$ stars pc$^{-3}$, substantially smaller than in the direct coalescence model but higher than observed nonetheless. On the other hand, Krumholz & Thompson (2007) argue that pre-main sequence evolution of tight, high-mass protostellar binaries can lead to equal mass binaries, as often observed, rather than to mergers.

4.3.2 Observations of high-mass protostars

Beuther et al. (2007) have summarized the current state of observations of high-mass star formation. They divide the formation of individual high-mass stars into four stages:

1. High-Mass Starless Cores (HMSCs)
2. High-Mass Cores harboring accreting Low/Intermediate-Mass Protostar(s) destined to become high-mass star(s)
3. High-Mass Protostellar Objects (HMPOs), with $m_\star \gtrsim 8 M_\odot$
4. Final Stars

The earliest stages of high-mass star formation may occur in the Infrared Dark Clouds (IRDCs—Egan et al., 1998), which have properties consistent with being the dense clumps out of which clusters eventually form (Simon et al., 2006). To date, few true HMSCs have been detected—high-mass cores often appear to have some signatures of star formation. In the turbulent core model, this could be because the central densities in the cores are much greater than the mean densities (in contrast to the case for low-mass cores), and the time scale for gravitational collapse is correspondingly shorter. The lack of true HMSCs is also consistent with the competitive accretion model or the coalescence model, since in these models HMSCs do not exist. Evidence for High-Mass Cores harboring low/intermediate mass protostars, or possibly relatively low-mass HMPOs, has
McKee & Ostriker

been obtained only recently (Beuther, Sridharan, & Saito, 2005; Sridharan et al., 2005). HMPOs are often (but not always) associated with Hot Molecular Cores (HMCs), which have a rich chemistry (van der Tak, 2005). HMPOs are often associated with H$_2$O and Class II CH$_3$OH maser emission, although the interpretation of this emission remains ambiguous. HMPOs are also associated with H II regions (see §4.3.4); many should have hypercompact H II regions and some should be associated with ultra-compact H II regions, but most ultra-compact H II regions are associated with the final stars.

Observational tests of infall models for high-mass protostars are difficult due to their large distances (typically $\gtrsim 2$ kpc), crowding, large extinctions, and confounding effects of H II regions. Several tests are possible: if confirmed, the correspondence between the core mass function and the IMF (§3.3) would be consistent with the turbulent core model; the properties of disks and winds associated with HMPOs can provide important clues (§4.2.1); the spectral energy distributions (SEDs) of embedded sources provide information on the distribution of circumstellar matter on scales smaller than can be directly resolved (Osorio, Lizano, & D’Alessio, 1999; Chakrabarti & McKee, 2005; Whitney et al., 2005); and chemical clocks can provide direct measures of the time scale for the growth of HMPOs. To this end, Doty, van Dishoeck, & Tan (2006) have developed the first model for the chemical evolution of an HMPO, including the evolution of the central source, infall, and adsorption and desorption of ices from grains. They find that the time scale for the warm chemistry is set by the time it takes for matter to flow through the warm region, and that the total age of the HMPO they study (AFGL 2591) is $(0.3 - 1) \times 10^5$ yr.

4.3.3 Forming stars in the presence of radiation pressure

One measure of the importance of radiation pressure is to compare the stellar luminosity with the luminosity at which the force due to radiation pressure balances gravity. Since dust provides the dominant opacity to non-ionizing radiation in the interstellar medium, this generalized Eddington luminosity is

$$L_{E,d} = \frac{4\pi c G m_\ast}{\kappa_d},$$

where $\kappa_d$ is the dust opacity per unit mass, and $c$ is the speed of light. The dust in the infalling gas sublimates when it reaches the dust destruction front at $r = R_{dd} \simeq 1.2 \times 10^{15} (L/10^5 L_\odot)^{1/2}$ cm (Wolff & Cassinelli, 1987). We approximate the radiation field outside $R_{dd}$ as a blackbody with a temperature $T$ that declines with radius. As an example, consider the Pollack et al. (1994) dust model: $\kappa_d(T)$ first rises with temperature as the average frequency increases, but then declines for $T \gtrsim 600$ K as some of the grain species sublimate. The maximum opacity is $\kappa \simeq 8$ cm$^2$ g$^{-1}$, which leads to $L_{E,d} \simeq 1600 (m_\ast/M_\odot) L_\odot$. Since main sequence stars have luminosities $L \simeq 10 (m_\ast/M_\odot)^3 L_\odot$ for $7 M_\odot \lesssim m_\ast \lesssim 20 M_\odot$ (inferred from Arnett, 1996), the infalling gas and dust pass through a region in which the force due to the infrared radiation exceeds that due to gravity if $m_\ast \gtrsim 13 M_\odot$. For somewhat larger masses, the net force is outward over a sufficiently large region that the infall is stopped. At the dust destruction front, the gas and dust are exposed to the stellar UV radiation, for which $\kappa \approx 200$ cm$^2$ g$^{-1}$. However, this radiation interacts with the matter only once with this opacity, since it is emitted in the infrared after absorption; as a result, the condition for the infall to persist is that its momentum exceed that of the radiation, $\dot{M}_\text{in} v_\text{in} > L/c$ (Larson & Starrfield, 1971; Kahn, 1974; Wolfire & Cassinelli, 1987). High infall rates
Several mechanisms have been proposed to permit the formation of massive stars in the face of radiation pressure:

**Reduced dust opacity**: Based on 1D, multifluid calculations of steady flows with both graphite and silicate grains with a range of sizes, Wolfire & Cassinelli (1987) found that a reduction in the dust-to-gas ratio of at least a factor 4 is needed in order for accretion to proceed for stars with $m_* \geq 60 M_\odot$.

**Rotation**: Nakano (1989) showed that the higher ram pressure associated with disk accretion helps overcome the UV radiation pressure problem, and escape of the infrared radiation from the disk alleviates the infrared radiation pressure problem. He found that accretion could continue onto a 100 $M_\odot$ star with an accretion rate as small as $10^{-4} M_\odot$ yr$^{-1}$, 50 times smaller than for spherical accretion. Jijina & Adams (1996) showed that there are a range of conditions for which infrared radiation pressure cannot halt infall prior to the formation of a disk in the context of the Terebey, Shu, & Cassen (1984) model for rotating collapse, even for quite massive stars. Yorke & Sonnhalter (2002) have carried out the most detailed axisymmetric numerical simulations to date. Using frequency-dependent radiative transfer, they found that radiation pressure limited the maximum stellar mass that could be formed from a 120 $M_\odot$ core to 43 $M_\odot$; they point out that this is an upper limit, since it did not allow for fragmentation or for the effects of outflows.

**Rapid infall**: Edgar & Clarke (2003) have shown that radiation pressure becomes moot if the protostellar core is sufficiently dense that the protostellar mass is inside the dust destruction front. The models they considered to achieve this condition were far from virial equilibrium and had constant density, which led to very large accretion rates $\sim 10^{-2} M_\odot$ yr$^{-1}$. However, models with centrally concentrated initial density profiles ($\rho \propto r^{-1}$) and normal dust could not produce stars with $m_* > 16 M_\odot$.

**Beaming**: Nakano (1989) pointed out that disks redirect the infrared radiation toward the poles, reducing the radiative force in the plane (Yorke & Bodenheimer, 1999 termed this the “flashlight effect” and emphasized its observational importance). Krumholz, McKee, & Klein (2005b) showed that the cavities produced by outflows from massive stars would allow the infrared radiation to escape, reducing the the radiation pressure in the infalling gas and permitting infall over a substantial range of solid angle.

**3D effects**: 3D simulations with flux-limited, gray radiative transfer show that the accreting gas is subject to radiation-driven Rayleigh-Taylor instabilities, which facilitate the escape of the radiation in low column regions and the accretion of the gas in high column regions (Krumholz, Klein, & McKee, 2005). There is no evidence that radiation pressure halts the accretion up to $m_* = 35 M_\odot$, a substantially higher mass than was found in axisymmetric simulations with gray transfer (Yorke & Sonnhalter, 2002). Turner, Quataert, & Yorke (2007) have shown that the dusty envelopes of HMPOs are subject to the photon bubble instability, which further promotes infall.

### 4.3.4 Photoionization feedback: H II regions

The H II regions associated with HMPOs provide strong feedback on infall and accretion, and may play a role in defining the maximum stellar mass. They are classified into two types: *Ultra-compact* H II (UCHII) regions have diameters (0.01 – 0.1) pc, densities $\geq 10^4$ cm$^{-3}$, and emission measures $\int n_e^2 dl \geq 10^7$ pc cm$^{-6}$ (Wood & Church-
well, 1989). *Hypercompact* H II (HCHII) regions have diameters < 0.01 pc with emission measures $\geq 10^8$ pc cm$^{-6}$ (Beuther et al., 2007; for a slightly different definition and a review of both types of H II region, see Hoare et al., 2007). HCHII regions often appear in tight groups in high-mass star-forming regions, and they often have broad radio recombination lines with widths that can exceed 100 km s$^{-1}$.

The high accretion rates characteristic of HMPOs delay the point at which the stars reach the main sequence (McKee & Tan, 2003; Krumholz & Thompson, 2007), thereby delaying the time at which the photosphere is hot enough to produce an H II region. High accretion rates also quench the emission of ionizing photons once the star has reached the main sequence (Walmsley, 1995). Close to the star—i.e., inside the gravitational radius $r_g = Gm/c_i^2 = 3.2 \times 10^{15}(m_*/30 M_\odot)$ cm, where $c_i \approx 10$ km s$^{-1}$ is the isothermal sound speed of the ionized gas—spherically accreting gas is in free fall, with $\rho \propto r^{-3/2}$. For an ionizing photon luminosity $S$, the radius of the HCHII region is

$$R_{\text{HCHII}} = R_* \exp(S/S_{cr}),$$

(55)

where

$$S_{cr} = \frac{\alpha^{(2)} m_e^2}{8 \pi \mu_H^2 G m_*} = 5.6 \times 10^{50} \left(\frac{m_*}{10^{-3} M_\odot \text{ yr}^{-1}}\right)^2 \left(\frac{100 M_\odot}{m_*}\right) \text{ s}^{-1},$$

(56)

(Omukai & Inutsuka, 2002), where $\alpha^{(2)}$ is the recombination rate to excited states of hydrogen and where we have replaced the proton mass in their expression with $\mu_H = 2.34 \times 10^{-24}$ g, the mass per hydrogen nucleus. Provided the accretion is spherical, the H II region is quenched for $S \lesssim S_{cr}$. If $S/S_{cr}$ is not too large ($\lesssim 7$), $R_{\text{HCHII}}$ is less than $r_g/2$ and the infall velocity at the Stromgren radius exceeds $2c_i$, the minimum velocity of an R-critical ionization front; as a result there is no shock in the accretion flow and the H II region cannot undergo the classical pressure-driven expansion (Keto, 2002). If the accretion is via a disk, as is generally expected, then the ionizing photons can escape out of the plane of the disk, and the H II region will not be trapped (Keto & Wood, 2006; Keto, 2007). Disk accretion is often associated with the production of winds, and Tan & McKee (2003) have suggested that such winds confine HCHII regions: the winds clear the gas along the axis, and the ionizing radiation then illuminates the inner surfaces of the winds. If correct, this offers the possibility of a powerful diagnostic for determining the nature of disk winds associated with massive stars. van der Tak & Menten (2005) found very compact radio emission aligned with the outflows in two high-mass protostellar sources, consistent with this picture. When the ionizing luminosity becomes large enough, however, the wind will become ionized and the H II region will evolve to a UCHII state. The ionizing photons will photoevaporate the surface of the disk at a rate of order $\dot{m}_{\text{evap}} \sim \text{few} \times 10^{-5}(S/10^{49} \text{ s}^{-1})^{1/2} M_\odot \text{ yr}^{-1}$; absorption of ionizing photons by dust can significantly affect this (Hollenbach et al., 1994; Richling & Yorke, 1997). This mass-loss rate is too small to be important in setting the maximum mass of the star (although it can be important in primordial star formation—McKee & Tan, in preparation). Absorption of ionizing photons by dust must also be taken into account when inferring the ionizing luminosity of the central star from the properties of the H II region (Dopita et al. 2006 and references therein).
4.3.5 Star Formation in Clusters  Most stars are born in clusters (e.g., Lada & Lada, 2003; Allen et al., 2007), and this is particularly true of high-mass stars. The mass distribution of clusters appears to obey a universal power law, $dN_{\text{cluster}}/d\ln M \propto M^{-\alpha}$, with $\alpha \approx 1$. With this distribution, $MdN_{\text{cluster}}/d\ln M = \text{const}$; taken together, clusters in each decade of mass have the same total number of stars. Lada & Lada (2003) find that very young clusters within 2 kpc of the Sun that are still embedded in their natal molecular clouds obey this power law for $M \gtrsim 50 ~M_\odot$; the upper limit of the observed distribution is set by the largest cluster expected in the area they surveyed. The mass distribution of OB associations in the Galaxy also has a power-law distribution with $\alpha \approx 1$ (McKee & Williams, 1997); they inferred that the distribution extended from $\sim 50 ~M_\odot$ to $2 \times 10^5 ~M_\odot$ and could account for all the stars formed in the Galaxy. Kennicutt, Edgar, & Hodge (1989) found that the luminosity distribution of H II regions in disk galaxies obeys $dN/d\ln L \propto L^{-1.5}$, which is consistent with an $M^{-1}$ distribution since the luminosity is proportional to mass for associations that are large enough to fully sample the IMF. The distribution of OB associations in the SMC has $\alpha = 1$ from the largest associations down to associations with a single OB star (Oey, King, & Parker, 2004). The star clusters in the “Antennae” galaxies show $\alpha = 1$ over the mass range $10^4 ~M_\odot < M < 10^6 ~M_\odot$ (Zhang & Fall, 1999); this is one of the best determined cluster mass functions, and has an error, including systematic errors, estimated as $\pm 0.1$. The mass distributions of open clusters and globular clusters are also consistent with an $M^{-1}$ distribution at birth (Elmegreen & Efremov, 1997). Dowell, Buckalew, & Tan (2006) found $\alpha \approx 0.9$ for clusters in irregular galaxies and $\alpha \approx 0.75$ in disk galaxies, but comment that this result could be affected by the low spatial resolution of the data. The $M^{-1}$ mass distribution of clusters is intermediate between the high-mass part of the IMF ($M^{-1.35}$) on the one hand, and the observed mass distribution of GMCs ($M^{-0.6}$) and the clumps within them ($M^{-0.3}$ to $M^{-0.7}$; §3.1) on the other. It is important to understand the origin of the differences among these power laws, which appear to be real.

The structure of star clusters contains clues to their formation. High-mass stars in Galactic clusters that are massive enough to contain a number of such stars are observed or inferred to be segregated toward the center of the cluster. The large fraction of O stars that are runaways can be naturally explained if they originate in dense, mass-segregated clusters and undergo dynamical interactions (Clarke & Pringle, 1992). Hillenbrand & Hartmann (1998) analyzed the spatial distribution of stars in the Orion Nebula Cluster (ONC) and concluded that the high-mass stars were born preferentially near the center. Using N-body simulations, Bonnell & Davies (1998) showed that it takes a relaxation time, $t_{\text{relax}} \simeq 0.1(N_c/\ln N_c)t_{\text{cross}}$ ($\simeq 14t_{\text{cross}}$ for $N_c = 1000$), for high-mass stars to collect near the center of a cluster due to dynamical interactions. In the case of the ONC, they argued that significant dynamical mass segregation has occurred, but not enough to account for the observed central concentration of OB stars; they concluded that therefore the observed mass segregation is primordial. Tan, Krumholz, & McKee (2006) suggested both a greater age and a longer crossing time for the ONC, but the basic conclusion does not change. In NGC 3603, the most luminous Galactic star cluster that is not heavily obscured, Stolte et al. (2006) found that the maximum mass of the stars decreases away from the center of the cluster, with all the most massive stars being quite close to the center, and again concluded that the segregation is primordial. These arguments
for primordial mass segregation have been weakened by the realization that sub-clustering in the initial cluster significantly accelerates the rate of dynamical mass segregation (McMillan, Vesperini, & Portegies Zwart, 2007). In addition, most estimates do not account for dynamical friction between the stars and the surrounding gas, which can considerably reduce the mass segregation timescale (Ostriker, 1999). However, recent observations have reinforced the argument for primordial mass segregation: Megeath, Wilson, & Corbin (2005) have found a Trapezium-like cluster in W3 IRS5 that is deeply embedded in molecular gas and is only half the radius of the ONC. In ρ Oph, Stanke et al. (2006) found direct evidence for primordial mass segregation by showing that the the core mass function exhibits mass segregation as well; this also supports the correspondence between the core mass function and the IMF discussed in §3.3. The cluster R136 in the LMC appears to be an exception to the rule that high-mass stars are centrally concentrated, since half the massive stars are located outside the central core (Stolte et al., 2006).

Only a small fraction of clusters survive as bound clusters to an age of $10^8$ yr; Lada & Lada (2003) estimate this fraction as $4 - 7\%$. In order for a cluster to remain bound, its natal clump must have a high star formation efficiency. Analytic estimates suggest that if the gas in the clump is removed suddenly, such as by an H II region, one requires $\epsilon_{\text{clump}} > 0.5$ in order for the cluster to remain bound, whereas if the gas is removed gradually, the cluster will expand adiabatically and lower values suffice (Hills, 1980; Mathieu, 1983). Numerical calculations show that a fraction of the cluster survives even if the mass ejection is abrupt (Lada, Margulis, & Dearborn, 1984). Kroupa, Aarseth, & Hurley (2001) modeled the evolution of the ONC with $\epsilon_{\text{clump}} = 0.3$ and a sudden mass ejection; they concluded that 30% of the mass of the ONC would remain bound, and that it would evolve into a cluster like the Pleiades. Star formation efficiencies in the embedded clusters in the solar neighborhood are observed to be $\simeq 0.1 - 0.3$ (Lada & Lada, 2003); since star formation is ongoing in these clusters, the final value of the star formation efficiency, $\epsilon_{\text{clump}}$, is near the upper limit of this range. Matzner & McKee (2000) calculated the star formation efficiency for clumps in which the mass loss is dominated by protostellar outflows ($M \lesssim 1 - 3 \times 10^3 M_\odot$). (Note that the clump star formation efficiency, $\epsilon_{\text{clump}}$, which is the fraction of the mass of a clump that goes into a cluster of stars, is distinct from the core star formation efficiency, $\epsilon_{\text{core}}$, which is the fraction of a core mass that goes into a single or binary star.) They estimated $\epsilon_{\text{clump}} \simeq 0.4$ for clumps with escape velocities $v_{\text{esc}} \simeq 2$ km s$^{-1}$, comparable to the observed value. The star formation efficiency is predicted to rise with $v_{\text{esc}}$—i.e., with increasing mass and/or density of the clump.

The star formation efficiency for larger clusters, ranging up to globular clusters and super star clusters (SSCs; e.g., Ho & Filippenko, 1996) is most likely determined by the H II regions that form in the clusters. Since globular clusters are much more centrally concentrated than open clusters that form in the disk of the Galaxy today, it is likely that their star formation efficiency, $\epsilon_{\text{clump}}$, was higher. (Note that a high star-formation efficiency over the life of the clump, $\epsilon_{\text{clump}}$, is consistent with a low value of the star-formation efficiency per free-fall time, $\epsilon_{\text{ff},\text{clump}}$, only if clusters form over a number of free-fall times, and conversely, cluster formation in $1 - 2$ dynamical times requires a relatively high value of $\epsilon_{\text{ff},\text{clump}}$—see §3.4.) Using a simple phenomenological model, Elmegreen & Efremov (1997) showed how $\epsilon_{\text{clump}}$ should increase with both the mass of the
natal clump and its pressure, $P \propto \Sigma_d^2$. They point out that high pressures are naturally produced in merging galaxies, accounting for the large number of super star clusters seen in such systems. The high surface densities of globular clusters implies that they necessarily formed in high pressure environments (see also McKee & Tan, 2003). High star formation efficiencies are possible in a clump with embedded H II regions since their destructive effect is significantly reduced in a clump composed of dense cores (Tan & McKee, 2001; Dale et al., 2005). For sufficiently massive and concentrated clusters, the escape velocity exceeds the sound speed of ionized gas, and this can further increase the star formation efficiency (Kroupa & Boily, 2002; Matzner, 2002; Tan & McKee, 2004).

How do stars form in clusters? The natal clumps of embedded clusters in the solar neighborhood have densities $\sim 10^{4-5} \text{ cm}^{-3}$ and masses $\sim 10^{2-3} M_\odot$ (Lada & Lada, 2003), and more broadly distributed high-mass star-forming clumps in the Galaxy have densities $\sim 10^5 \text{ cm}^{-3}$ and masses $\sim 10^{3.5} M_\odot$ (Plume et al., 1997). These extreme conditions have led to suggestions that the process of high-mass star formation is qualitatively different from that observed in regions of low-mass star formation (§4.3.1) or that it is triggered by an external effect. In a review of triggered star formation, Elmegreen (1992) pointed out that triggering generally does not affect the star formation efficiency by more than a factor 2. In any case, triggered star formation loses much of its meaning in a theory of star formation based on turbulence, since in most cases the triggering event is just a manifestation of the intermittency of the turbulence. The observed correspondence between the core mass function and the IMF (§3.3) and the constancy of the star formation efficiency per free-fall time (§3.4) suggest a more unified picture in which stars form via gravitational collapse in a turbulent medium over most, if not all, the range of observed clustering.

5 OVERVIEW OF THE STAR FORMATION PROCESS

Key goals of a theory of star formation are to predict the rate of star formation and the distribution of stellar masses on the macroscopic scale, and to predict the properties of individual stars from the initial conditions on the microscopic scale. In the past decade, there has been a paradigm shift in the theory from star formation in a quasistatic medium to star formation that occurs in a super-sonically turbulent one, and this has led to significant progress on both fronts. Based on our current understanding, the narrative of star formation contains the following elements:

- The road to star formation in a disk galaxy like the Milky Way begins when massive ($\sim 10^7 M_\odot$) bound structures condense out of the diffuse ISM as a result of gravitational instabilities, frequently initiated within spiral arms.

- The most massive structures (GMAs or HI superclouds) inherit high levels of internal turbulence from the diffuse ISM, and this combines with self-gravity to cause fragmentation into GMCs of a range of masses, as well as clumps within the GMCs.

- The turbulence within GMCs is highly supersonic and approximately Alfvénic. It imposes a log-normal distribution of densities, and creates a spectrum of gas condensations over a wide range of spatial scales and masses. This structure is hierarchical.
• This turbulence damps in about one crossing time, and as yet it is not understood exactly how, and for how long, the highly intermittent sources of energy in the interstellar medium (including within GMCs themselves) can maintain the observed universal level of turbulence in GMCs.

• Spatially-defined structures within GMCs tend to have internal velocity dispersions that increase with size as $\sigma \propto \ell^{0.5}$, which is understood to reflect the underlying power spectrum scaling expected for supersonic turbulence.

• Some of the densest regions created by turbulence become self-gravitating cores with masses that are typically of order the Bonnor-Ebert mass. The distribution of core masses appears to be similar to the initial mass function (IMF) for stars, and turbulence appears to be important in defining this distribution.

• These cores are frequently clustered, due to the dominance of large scales in the turbulent flow. Forming cores sample from the local vorticity of the turbulence to determine their spins. The rate of core formation can be estimated based on the turbulent properties of a GMC.

• Dense cores that begin or become magnetically supercritical undergo collapse, first becoming strongly stratified internally. Observations show that magnetic fields in cores are roughly critical, and this is consistent with inferred core lifetimes.

• Continued accretion after the collapse of a core can occur if the surrounding ambient medium has a sufficiently low level of turbulence, but it is not yet known how much this can increase the masses of stars.

• The collapse of a core leads to the formation of a rotating disk interior to an accretion shock; significant magnetic flux is lost in this collapse process, although based on current results this is not enough to account for the small fluxes observed in stars.

• Disks accrete due to a combination of processes that transport angular momentum outward; these transport mechanisms include gravitational stresses when the surface density is high enough, and magnetic stresses when the ionization is high enough.

• Powerful winds are magnetocentrifugally driven from the surface of circumstellar disks at a range of radii. The inner portion of the wind, which arises nearest the central star, becomes collimated into a jet-like flow due to magnetic hoop stresses.

• The impact of a wide-angle, stratified disk wind on the protostellar core sweeps up much of the ambient gas into a massive molecular outflow. This reduces the net efficiency of star formation to $\sim 1/3$. The combined action of many outflows also helps to energize dense, star-cluster-forming clumps.

• Massive stars form from cores that are considerably more massive than a Bonnor-Ebert mass, and are most likely highly turbulent. Radiation pressure strongly affects the dynamics of massive star formation, but can be overcome by the combined action of disk formation, protostellar outflows and radiation-hydrodynamic instabilities in the accreting gas. It is not clear whether protostellar feedback determines the maximum mass of the stars that form.
• Massive, luminous stars ionize their surroundings into HII regions. The expansion of these regions into ambient gas at \( \sim 10 \text{ km s}^{-1} \) energizes GMCs, contributing to the large-scale turbulent power. However, this process is difficult to regulate, and can unbind GMCs within a few dynamical crossing times. By the time they are finally destroyed, GMCs may have lost much of their original mass by photoevaporation.

• The destruction of GMCs returns almost all of the gas they contain to the diffuse phase of the ISM, with a mean star formation efficiency over the cloud lifetime of \( \sim 5\% \). This low efficiency can be understood as a consequence of the small fraction of mass that is compressed into clumps dense enough that turbulence does not destroy them before they collapse.

• The return of GMC gas to the diffuse ISM completes the cycle of star formation, which then begins anew.

The coming decade will test and revise this narrative of star formation, particularly with the advent of ALMA and JWST and the continued advances in numerical simulation. Turning this narrative into a quantitative, predictive theory will provide a foundation for addressing many of the outstanding questions in astrophysics today, ranging from the formation of planets to the evolution of galaxies and the origin of the elements.

Acknowledgements We are grateful to our expert readers, J. Bally, G. Basri, S. Basu, E. Bergin, L. Blitz, R. Crutcher, B. Elmegreen, C. Gammie, L. Hartmann, M. Heyer, R. Kennicutt, S. Kenyon, M. Krumholz, C. Lada, Z.-Y. Li, C. Matzner, S. Offner, P. Padoan, J. Tan, E. Vazquez-Semadeni, and E. Zweibel, for their insightful comments on draft sections of the manuscript, and to our editor, E. van Dishoeck, for her comments on the entire manuscript. We are also grateful to C.-F. Lee for his help producing the figure of HH111 and to Nathan Smith for his help with the figure of the Carina Nebula. The work of CFM and ECO was supported by the National Science Foundation under grants AST 0606831 and AST 0507315, respectively. In preparing this review, we have relied upon the search and archive facilities provided by NASA’s Astrophysics Data System Bibliographic Services.

References

Balbus, S. A. 2003, ARAA, 41, 555
Balbus, S. A., & Hawley, J. F. 1998, Reviews of Modern Physics, 70, 1
Ballesteros-Paredes, J., Klessen, R. S., Mac Low, M.-M., & Vazquez-Semadeni, E. 2007, Protostars and Planets V, 63
Beckwith, S. V. W., Henning, T., & Nakagawa, Y. 2000, Protostars and Planets IV, 533
Beuther, H., & Schilke, P. 2004, Science, 303, 1167
Beuther, H., & Shepherd, D. 2005, Cores to Clusters: Star Formation with Next Generation Telescopes, 105
Blitz, L. 1993, Protostars and Planets III, 125
Bodenheimer, P., Burkert, A., Klein, R. I., & Boss, A. P. 2000, Protostars and Planets IV, 675
Theory of Star Formation

Bromm, V., & Larson, R. B. 2004, ARAA, 42, 79
Calvet, N., Hartmann, L., & Strom, S. E. 2000, Protostars and Planets IV, 377
Cesaroni, R., Galli, D., Lodato, G., Walmsley, C. M., & Zhang, Q. 2007, Protostars and Planets V, 197
Chandrasekhar, S. 1939, An Introduction to the Study of Stellar Structure, (Chicago, Ill., The University of Chicago press) [1939]
Doppmann, G. W., Greene, T. P., Covey, K. R., & Lada, C. J. 2005, AJ, 130, 1145
Draine, B. T., & McKee, C. F. 1993, ARAA, 31, 373
Dubrulle, B. 1994, Physical Review Letters, 73, 959
Duchêne, G., Delgado-Donate, E., Haisch, K. E., Jr., Loinard, L., & Rodríguez, L. F. 2007, Protostars and Planets V, 379
Ebert, R. 1957, Zeitschrift fur Astrophysik, 42, 263
Elmegreen, B. G. 1992, Star Formation in Stellar Systems, 381
Elmegreen, B. G. 1993b, Protostars and Planets III, 97
Elmegreen, B. G. 1995, Molecular Clouds and Star Formation, 149
Elmegreen, B. G., & Scalo, J. 2004, ARAA, 42, 211
Ghosh, S., & Goldstein, M. L. 1994, JGR, 99, 13351
Hartmann, L., & Kenyon, S. J. 1996, ARAA, 34, 207
Hayashi, C., Nakazawa, K., & Nakagawa, Y. 1985, Protostars and Planets II, 1100
Hollenbach, D. J., & Tielens, A. G. G. M. 1999, Reviews of Modern Physics, 71, 173
Hujeirat, A., Myers, P., Camenzind, M., & Burkert, A. 2000, New Astronomy, 4, 601
Klein, R. I., Sandford, M. T., II, & Whitaker, R. W. 1980, Space Science Reviews, 27, 275
Kolmogorov, A.N. 1941, C. R. Acad. Sci. USSR 30, 301
Konigl, A., & Pudritz, R. E. 2000, Protostars and Planets IV, 759
Krumholz, M. R., Klein, R. I., & McKee, C. F. 2005c, IAU Symposium, 227, 231
Lada, C. J. 1985, ARAA, 23, 267
Lada, C. J., & Lada, E. A. 2003, ARAA, 41, 57
Larson, R.B. 2003, Rept. Prog. Phys. 66, 1651
Mac Low, M.-M., & Klessen, R. S. 2004, Reviews of Modern Physics, 76, 125
Mac Low, M.-M., Klessen, R. S., Burkert, A., & Smith, M. D. 1998, Physical Review Letters, 80, 2754
Mestel, L. 1985, Protostars and Planets II, 320
Müller, W.-C., Biskamp, D., & Grappin, R. 2003, Phys Rev E, 67, 066302
Müller, W.-C., & Grappin, R. 2005, Physical Review Letters, 95, 114502
Narita, S., Hayashi, C., & Miyama, S. M. 1984, Progress of Theoretical Physics, 72, 1118
Nordlund, Å. K., & Padoan, P. 1999, Interstellar Turbulence, 218
Passot, T., & Vázquez-Semadeni, E. 1998, Phys Rev E, 58, 4501
Pringle, J. E. 1981, ARAA, 19, 137
Pudritz, R. E., Ouyed, R., Fendt, C., & Brandenburg, A. 2007, Protostars and Planets V, 277
Reipurth, B., & Bally, J. 2001, ARAA, 39, 403
Sasao, T. 1973, PASJ, 25, 1
Scalo, J., & Elmegreen, B. G. 2004, ARAA, 42, 275
Shepherd, D. 2005, IAU Symposium, 227, 237
Shu, F. H. 1973, IAU Symp. 52: Interstellar Dust and Related Topics, 52, 257
Stahler, S. W., Palla, F., & Ho, P. T. P. 2000, Protostars and Planets IV, 327
Tan, J. C., & McKee, C. F. 2001, Starburst Galaxies: Near and Far, 188
Tan, J. C., & McKee, C. F. 2003, IAU Symposium, 221, 274P
Toomre, A. 1981, Structure and Evolution of Normal Galaxies, 111
van der Tak, F. F. S. 2005, IAU Symposium, 227, 70
Walmsley, M. 1995, Revista Mexicana de Astronomia y Astrofisica Conference Series, 1, 137
Williams, J. P., Blitz, L., & McKee, C. F. 2000, Protostars and Planets IV, 97
Yorke, H. W., & Richling, S. 2002, Revista Mexicana de Astronomia y Astrofisica Conference Series, 12, 92
Zuckerman, B., & Palmer, P. 1974, ARAA, 12, 279
Zhang, Q. 2005, IAU Symposium, 227, 135