Numerical Star-Formation Studies—
A Status Report

Ralf S. Klessen1,*, Mark R. Krumholz2, and Fabian Heitsch3

1Zentrum für Astronomie der Universität Heidelberg, Institut für Theoretische Astrophysik, Albert-Ueberle-Str. 2, 69120 Heidelberg, Germany
2Astronomy Department, Interdisciplinary Sciences Building, University of California, Santa Cruz, CA 95064, USA
3Department of Astronomy, University of Michigan, 500 Church St, Ann Arbor, MI 48109, USA

The formation of stars is a key process in astrophysics. Detailed knowledge of the physical mechanisms that govern stellar birth is a prerequisite for understanding the formation and evolution of our galactic home, the Milky Way. A theory of star formation is an essential part of any model for the origin of our solar system and of planets around other stars. Despite this pivotal importance, and despite many decades of research, our understanding of the processes that initiate and regulate star formation is still limited. Stars are born in cold interstellar clouds of molecular hydrogen gas. Star formation in these clouds is governed by the complex interplay between the gravitational attraction in the gas and agents such as turbulence, magnetic fields, radiation and thermal pressure that resist compression. The competition between these processes determines both the locations at which young stars form and how much mass they ultimately accrete. It plays out over many orders of magnitude in space and time, ranging from galactic to stellar scales. In addition, star formation is a highly stochastic process in which rare and hard-to-predict events, such as the formation of very massive stars and the resulting feedback, can play a dominant role in determining the evolution of a star-forming cloud. As a consequence of the wide range of scales and processes that control star formation, analytic models are usually restricted to highly idealized cases. These can yield insight, but the complexity of the problem means that they must be used in concert with large-scale numerical simulations. Here we summarize the state of modern star formation theory and review the recent advances in numerical simulation techniques.

Keywords: Numerical Methods, Astrophysical Hydrodynamics, Radiative Transfer, Star Formation, Stellar Feedback, ISM Evolution, ISM Kinematics and Dynamics, Molecular Clouds.

1. INTRODUCTION

Stars are central to much of modern astronomy and astrophysics. They are the visible building blocks of the cosmic structures around us, and thus are essential for our understanding of the universe and the physical processes that govern its evolution. At optical wavelengths almost all natural light we observe in the sky originates from stars. The Moon and the planets in our solar system reflect the light from our Sun, while virtually every other source of visible light further away is a star or collection of stars. Throughout the millenia, these objects have been the observational targets of traditional astronomy, and define the celestial landscape, the constellations. The most massive stars are very bright, they allow us to reach out to the far ends of the universe. For example, the most distant galaxies in the Hubble Ultra Deep Field are all characterized by vigorous high-mass star formation. Understanding the origin of stars, at present and at early times, therefore is a prerequisite to understanding cosmic history.

Stars are also the primary source of chemical elements heavier than the hydrogen, helium, and lithium that were produced in the Big Bang. The Earth, for example, consists mostly of iron (32%),
oxygen (30%), silicon (15%), magnesium (14%) and other heavy elements. These are produced by nuclear fusion in the interior of stars, and enriching gas to the chemical composition observed today in our solar system must have required many cycles of stellar birth and death.

Today we also know that many stars harbor planetary systems around them, about 300 are known as of fall 2008. The build-up of planets is intimately coupled to the formation of their host stars. Understanding the origin of our solar system and of planets around other stars has a profound impact on how we see our position in the universe. Questions whether we are alone, or whether there is life elsewhere in the cosmos are of broad interest to all of us.

Stars and the planetary systems they may harbor are born in turbulent interstellar clouds of molecular hydrogen with a small fraction of dust mixed in. At optical wavelengths, we see these clouds as dark patches of obscuration along the band of the Milky Way. The dust component blocks the light from stars further away. At far-infrared, sub-millimeter, and radio wavelengths, however, the dust becomes increasingly transparent and we can look into these clouds. These observations reveal extremely complex morphological and kinematic structure, where patches of cold high-density gas are interspersed between regions of low-density warmer material. It is thought that this complicated texture is caused by supersonic turbulence that is generated by large-scale gravitational motions in the galaxy (such as spiral density waves) or by energy and momentum input from stars themselves.

Within molecular clouds, supersonic turbulence and thermal instability lead to a transient, clumpy structure. Some of the resulting density fluctuations exceed the critical mass and density of gravitational stability. These clumps begin to collapse, their central density increases rapidly, and eventually they give birth to new protostars. Clusters of stars form in large regions that become unstable, within which contraction involves multiple collapsing cloud cores. A number of recent reviews have discussed various aspects of stellar birth in these clouds.

Delivered by Ingenta to:
University of California Santa Cruz

Ralf S. Klessen is professor of theoretical astrophysics at Heidelberg University. From 2002 to 2006 he was leading an Emmy Noether research group in theoretical star formation studies at the Astrophysical Institute Potsdam. Before that, he worked as postdoctoral fellow at the University of California at Santa Cruz and at Leiden Observatory. He got his Ph.D. in 1998 from Heidelberg University for his work on star cluster formation with Andreas Burkert at the Max Planck Institute for Astronomy. Ralf Klessen received the Ludwig Biermann Prize of the German Astronomical Society and the Otto Hahn Award from the Max Planck Society. His research interests lie in dynamical phenomena in astronomy and astrophysics on various scales, ranging from stellar dynamical problems in galaxies and individual star clusters, down to hydrodynamical processes in the ISM on parsec scales and a general approach to the theory of supersonic turbulence. In particular he is interested in understanding the physical processes that initiate and regulate the birth of stars at present days as well as in the early Universe. Ralf Klessen is involved in developing a dynamical theory of star formation based on the interplay between interstellar turbulence and self-gravity.

Mark R. Krumholz received his undergraduate degree in physics, with a certificate in applied mathematics, from Princeton University in 1998. He did his graduate work at the University of California, Berkeley, receiving his Ph.D. in physics in 2005, under the supervision of Professors Chris McKee and Richard Klein. From 2005 to 2008 he was a Hubble, Lyman Spitzer, Jr., and Council on Science and Technology Postdoctoral Fellow at Princeton University. Since then he has been an assistant professor of astronomy and astrophysics at the University of California, Santa Cruz. He was recently awarded an Alfred P. Sloan Research Fellowship. His research interests cover a wide range of problems in star formation and the interstellar medium, with particular emphasis on the formation of massive stars and the role of radiation-hydrodynamical effects in this process, regulation of the star-formation rate and efficiency in galaxies, and the formation and evolution of giant molecular clouds. He is also active in developing new numerical techniques in the areas of radiation-hydrodynamics and adaptive mesh refinement techniques.

Fabian Heitsch got his Ph.D. from Heidelberg University in 2001, working with Andreas Burkert and Mordecai-Mark Mac Low on turbulence and fragmentation of magnetized molecular clouds. He received a Feodor-Lynen postdoctoral fellowship from the Alexander-von-Humboldt Foundation in 2001, during which he worked with Ellen Zweibel at the Universities of Boulder-Colorado and Madison-Wisconsin. The Max-Planck-Society awarded him the Otto-Hahn Medal in 2002. In 2003, he took up a research staff position including teaching duties at the University Observatory Munich, until 2006. Currently, he is a research scientist at the Department of Astronomy of the University of Michigan. Heitsch mostly works on plasma dynamics in astrophysical systems, including the development of numerical methods to model these dynamics. The applications range from magnetic reconnection in partially ionized plasmas to the disruption of neutral hydrogen clouds in the Galactic halo. He is currently involved in identifying the initial conditions for star formation by connecting detailed star formation physics with the formation and evolution of the parental molecular clouds.
Stars and their parental clouds are connected via a number of feedback loops. Stars of all ages radiate and will thus heat up the gas in their vicinity. By doing so they influence subsequent star formation. Massive stars emit photons at ultraviolet wavelengths, creating bubbles of hot, ionized gas around them,[11,12] as illustrated in Figure 1. These so-called Hii regions are likely to quench further star formation in their interior, and thus set the star-formation efficiency in the region. The collective action of many Hii regions can destroy entire molecular clouds, and thus have the potential to influence the star-forming properties of galaxies on larger scales. The combined effect of large numbers of supernova explosions is another important mechanism for driving the supersonic turbulence ubiquitously observed in the galactic gas. By the same token, however, these feedback processes may trigger the birth of new stars. The very same processes that terminate star formation in one location may compress gas somewhere else in the galaxy, leading to new star formation.

The density contrast between typical cloud densities and the hydrogen-burning centers of the final stars is enormous, about 24 orders of magnitude, and so is the corresponding spatial range (roughly 8 orders of magnitude). In addition to the large dynamical range, many different physical processes play a role at the various stages of the contraction process. On global scales we need to describe the formation of molecular clouds via large-scale flows of mostly atomic gas in a galactic disk. Internal turbulent compression in the star-forming cloud then sets the initial stage for the protostellar collapse of individual objects. The thermodynamics of the gas, and thus its ability to respond to external compression and consequently to go into collapse, depends on the balance between heating and cooling processes. Magnetic fields and radiative processes also play an important role. Modeling star formation adequately therefore requires the accurate and simultaneous treatment of many different physical processes over many different scales.

In the past, progress has only been achievable by dividing the problem into smaller bits and pieces and by focusing on few physical processes or single scales only. Today, however, algorithmic advances and increasing computational power permit a more integrated approach to star formation. For the first time we are able to combine, for example, magnetohydrodynamics with chemical and radiative processes, and apply these numerical schemes to real astrophysical problems. It is this integrated view of stellar birth that is at the heart of this review. Our goal is to provide an overview of the recent advances in star formation theory with a special focus on the numerical aspects of the problem. We do not aim to be complete, for this we refer the reader to the recent reviews in this field.[4–5,10] Instead we will focus on three selected topics where we think numerical studies have had the largest impact and where we think our understanding of the physical processes that initiate and regulate stellar birth evolves most rapidly. We begin with the large scales and discuss the formation of molecular clouds in galactic disks and the numerical requirements and methodologies needed to do so consistently in Section 2. We then zoom into individual star-forming regions and examine the transition from cloud cores to stars in Section 3. As a third focus point, we discuss in Section 4 the effects of stellar feedback and examine how it alters the star formation process. Finally, in Section 5 we summarize and speculate about the future of numerical star formation research.

2. THE SITES OF STAR FORMATION: MOLECULAR CLOUDS

2.1. Phenomenology of Molecular Clouds

In regions of the interstellar medium (ISM) that are sufficiently dense and well-shielded against the dissociating effects of interstellar ultraviolet radiation, hydrogen atoms bind to form molecules. Star formation appears to occur exclusively within this molecular phase of the ISM. Molecular hydrogen is a homonuclear molecule, so its dipole moment vanishes and it radiates extremely weakly. Direct detection of cold interstellar H$_2$ is generally only possible through ultraviolet absorption studies, such as those made by the the Copernicus[13] and Far Ultraviolet Spectroscopic Explorer (FUSE)[14,15] satellites. Due to atmospheric opacity these studies can only be done from space, and are limited to pencil-beam measurements of the absorption of light from bright stars or active galactic nuclei. Note that rotational and rovibrational emission lines from H$_2$ have also been detected in the infrared, both in the Milky Way and in other galaxies. However this emission comes from gas that has been strongly heated by shocks or radiation, and it traces only a small fraction of the total H$_2$ mass (e.g., Ref. [16]). Due to these limitations, the most common tool for study of the molecular ISM is radio and sub-millimeter emission either from dust grains or from other molecules that tend to be found in the same locations as H$_2$. The most prominent of these is CO, although other tracers such as HCN are beginning to come into wide use.

As of this writing, the Milky Way and a few dozen Local Group galaxies have been mapped in the $J = 1 \rightarrow 0$ or $2 \rightarrow 1$ rotational transitions of CO at a resolution better than 1 kpc,[16–29] and a large number of more distant galaxies have been imaged at lower spatial resolution. The fractions of the ISM within the molecular phase in these galaxies ranges from no more than a few percent in low-surface density dwarfs to near unity in giant...
high-surface density systems. The highest molecular fractions are generally found in the parts of galactic disks with the highest total gas surface densities, but in the most actively star-forming galaxies the molecular fraction can reach ~90% even integrated over the entire galaxy.\(^{30,31}\) In all of the nearby galaxies where high resolution observations are possible the molecular gas is largely organized into giant clouds (the so called giant molecular clouds or GMCs) of mass \(\approx 10^5 - 10^7 M_\odot\) with average densities \(\sim 100\) H\(_2\) molecules per cm\(^3\), separated by a more diffuse intercloud medium. In the Milky Way and galaxies of lower density this medium is mostly atomic or ionized hydrogen, while in the densest nearby galaxies, such as M64,\(^{23}\) it is also molecular.

Molecular clouds across the Local Group all seem to display a number of properties in common. First, when studied with high spatial resolution clouds, they exhibit extremely complex and often filamentary structure, with column densities and corresponding 3-D densities that vary by many orders of magnitude (see Fig. 2 or Table I). Nevertheless, when observed with low resolution, to within factors of a few all molecular clouds seem to have a similar mean surface density of \(\sim 100\) M\(_\odot\) pc\(^{-2}\), corresponding to 0.035 g cm\(^{-2}\).\(^{32,33}\) The constant surface density of molecular clouds is known as one of Larson’s Laws, although there are a number of caveats with these relations and their interpretation.\(^{35}\) Second, the clouds all display linewidths much greater than would be expected from thermal motion, given their inferred temperatures of 10–20 K. The observed linewidth is related to the size of the cloud by

\[
\sigma_{1D} = 0.5 \left(\frac{L}{1.0 \text{ pc}}\right)^{0.5} \text{ km s}^{-1}
\]

(1)

where \(\sigma_{1D}\) is the one-dimensional cloud velocity dispersion and \(0L\) is the cloud size.\(^{19,33,36}\) This is another one of Larson’s Laws. These non-thermal linewidths have been interpreted as indicating the presence of supersonic turbulent motion, since both the low observed star formation rate (see below) and the absence of inverse P-Cygni line profiles indicates that they are not due to large-scale collapse. If one adopts this interpretation, then from these two observed relations one can directly deduce the third of Larson’s Laws, which is that giant molecular clouds have virial parameters\(^{37}\)

\[
\alpha_{vir} \equiv \frac{5 \sigma_{1D}^2 L}{G M} \approx 1
\]

(2)

where \(M\) is the cloud mass. This indicates that these clouds are marginally gravitationally bound, but with enough internal turbulence to at least temporarily prevent global collapse; whether they are truly in virial equilibrium is a topic that we discuss in detail below. (There is also a population of molecular clouds with virial ratios \(\gg 1\), but they have masses \(\ll 10^4 M_\odot\), and do not appear to host star formation.)\(^{38}\) These relations appear to partially or fully break down in starburst galaxies with very high surface densities, where for example the molecular gas velocity dispersion can reach \(\sim 100\) km s\(^{-1}\),\(^{39}\) but it is unknown whether there are analogous relations under these higher density conditions.

The presence of supersonic turbulence in approximate virial balance with self-gravity indicates that in molecular clouds the turbulent and gravitational energy densities are of the same order of magnitude, and both greatly exceed the thermal energy density. If molecular clouds form in large-scale convergent flows (as we argue below in Section 2.2), then surface terms from ram pressure can also be significant and need to be considered in the virial equations.\(^{32}\) There is also one further energy reservoir that we must mention, magnetic fields. The gas in molecular clouds is a weakly ionized plasma that is tied to magnetic field lines. Observations using Zeeman splitting\(^{43,44}\) and the Chandrasekhar-Fermi effect\(^{45,46}\) indicate that the field strength lies in the range from a few to a few tens of \(\mu G\). The exact value varies from region to region, but in general the magnetic energy density appears comparable to the gravitational and turbulent energy densities. One can describe this state of affairs in terms of magnetic criticality. If the magnetic field threading a cloud is sufficiently strong, then it cannot undergo gravitational collapse no matter what external pressure is applied to it, as long as it is governed by ideal magnetohydrodynamics (MHD). A cloud in this state is called subcritical. In contrast, a weaker magnetic field can delay collapse, but can never prevent it, and a cloud with such a weak field is called supercritical.\(^{47}\) Observations indicate the molecular clouds are close to being, but not quite, magnetically subcritical.\(^{48}\) For further discussions, see Section 3.1.2.

2.2. Cloud Timescales and Cloud Formation

2.2.1. Characteristic Timescales for Molecular Clouds

We can learn a great deal about molecular clouds by considering the timescales that govern their behavior. Because molecular clouds span a huge range of size and density scales, and their evolution times reflect this range, it is convenient to normalize all discussion of timescales to the free-fall time, defined as the time that a pressureless sphere of gas with some initial starting mass density \(\rho\) requires to collapse to infinite density under its own gravity: \(t_f = \sqrt{3\pi/(32G\rho)}\). For a cloud for which the virial parameter \(\alpha_{vir} \approx 1\), this is roughly half the cloud crossing time,\(^{49}\) defined as the ratio of the characteristic size to the velocity dispersion \(t_c = L/\sigma_{1D}\). The timescales \(t_f\) or \(t_c\) define the characteristic timescales on which behavior driven by gravity or limited by the internal gas signal speed can operate. For a molecular cloud detected via CO emission, with a mean number density \(n \approx 100\) cm\(^{-3}\), \(t_f \approx 3 \times 10^4\) yr. We now define some other useful timescales that can be determined from observations, and which yield strong constraints on how molecular clouds must behave. Any complete theory of star formation in molecular clouds must be able to explain each of the three timescales we describe.

The Gas Depletion Time. Perhaps the most fundamental observational timescale for molecular clouds is the gas depletion time \(t_{dep}\), which is defined as the ratio of the mass of a molecular cloud (or population of clouds) to the star formation rate. This defines the time that would be required to convert the cloud completely into stars at its observed star formation rate (SFR), assuming this rate is constant over time. Estimating this number immediately yields a striking conclusion, which is perhaps the most basic problem in star formation. The disk of the Milky Way contains \(\approx 10^9\) M\(_\odot\) of molecular gas, and the observed star formation rate is only a few M\(_\odot\) yr\(^{-1}\), so the gas depletion time \(t_{dep}\) must be a few hundred Myr,\(^{50,51}\) roughly 100 times the free-fall time. (Note, that if we compare this timescale with the age of the Milky Way of close to \(10^9\) yr, we conclude that a continuous inflow of fresh gas is required if the current SFR is at all representative and if we assume we are not living in times when the Milky Way is running out of gas soon. This problem gets worse if we consider proposals that the SFR was higher in the past.)\(^{52}\) One can repeat this exercise for populations of molecular clouds...
Fig. 2. Molecular cloud complex in the constellation Perseus. The image shows the distribution of CO line emission at radio wavelengths. This is a good tracer of total gas mass. Clearly visible is the highly complex and filamentary morphological structure of the cloud. Reprinted with permission from [17], K. Sun et al., Astron. Astrophys. 451, 539 (2006). © 2006, EDP Science.

in both the Milky Way and in other galaxies, using a variety of indicators of the star formation rate, and using a similarly wide variety of techniques to estimate masses of molecular clouds with various densities. Doing so yields the data shown in Figure 3. In this Figure the x-axis indicates the characteristic density to which a particular method of measuring molecular gas is sensitive, and the y-axis shows the ratio \( \frac{t_d}{t_{ff}} \) for that gas. The fact that this ratio is \( \approx 1% \) for low density gas and either remains constant or slowly increases to at most 10% at high densities indicates that the conversion of gas into stars must be inefficient or slow, in the sense that no more than a few percent of the total mass in molecular clouds in a galaxy can be converted into stars per free-fall time. This discrepancy is at the heart of any star formation theory, but before we can address it we must consider some other important timescales.

The Molecular Cloud Lifetime. The gas depletion time tells us how long it would take to convert a molecular cloud into stars completely. The actual lifetime \( t_{ff} \) of the cloud, however, is considerably shorter. Most of the cloud’s mass is never converted into stars. Instead it participates in the perpetual cycle that connects the molecular, atomic, and ionized phases of the ISM.

Table I. Physical properties of molecular cloud and cores.

<table>
<thead>
<tr>
<th>Molecular cloud</th>
<th>Cluster-forming clumps</th>
<th>Protostellar cores</th>
</tr>
</thead>
<tbody>
<tr>
<td>Size (pc)</td>
<td>2–20</td>
<td>0.1–2</td>
</tr>
<tr>
<td>Density ((nH_2)/cm^3)</td>
<td>10^2–10^4</td>
<td>10^2–10^3</td>
</tr>
<tr>
<td>Mass ((M_\odot))</td>
<td>10^2–10^4</td>
<td>10–10^3</td>
</tr>
<tr>
<td>Temperature (K)</td>
<td>10–30</td>
<td>10–20</td>
</tr>
<tr>
<td>Line width (km s^-1)</td>
<td>1–10</td>
<td>0.3–3</td>
</tr>
<tr>
<td>Column density ((g cm^-2))</td>
<td>0.03</td>
<td>0.03–1.0</td>
</tr>
<tr>
<td>Crossing time (Myr)</td>
<td>2–10</td>
<td>1</td>
</tr>
<tr>
<td>Free-fall time (Myr)</td>
<td>0.3–3</td>
<td>0.1–1</td>
</tr>
<tr>
<td>Examples</td>
<td>Taurus, Ophiuchus</td>
<td>L1641, L1709</td>
</tr>
</tbody>
</table>

In external galaxies, estimates of molecular cloud lifetimes are usually obtained from statistical relations between the location of molecular clouds and young star clusters. Star clusters can be age-dated, and determining at what ages they cease to be located preferentially close to molecular clouds gives an estimate of how long molecular clouds live once they form visible star clusters. Correcting for the population of molecular clouds that have not yet produced optically-visible clusters then gives an estimate of \( t_{ff} \). In the LMC this is \( t_{ff} \approx 27 \text{ Myr} \) and in M33 it is \( t_{ff} \approx 20 \text{ Myr} \), so \( t_{ff} \approx 7–10t_{fr} \), with a factor of \( \sim 2 \) uncertainty. Due to the sensitivity limits of the observations, these estimates apply only to GMCS \( \sim 10^5 M_\odot \) or larger.

In star-forming regions within \( \sim 500 \text{ pc} \) of the Sun, we can obtain ages estimates using individual young stars (as opposed to star clusters). Since stellar populations older than about \( 5 \times 10^7 \text{ yr} \) are generally not associated with molecular gas anymore, this technique suggests molecular cloud lifetimes substantially shorter than \( 10^7 \text{ years} \). Placing stars on the Hertzsprung-Russell diagram results in stellar age spread from...
1 Myr up to 3 Myr, with considerable uncertainty but consistent with the above number. Unfortunately these nearby clouds have all masses well below $\sim 10^4 M_\odot$, so there is essentially no overlap between this population and the extragalactic one. Probably in part as a result of this selection effect, these regions are also denser than the giant clouds we can observe in external galaxies, and consequently have lower free-fall times, so $t_{\text{ff}} \sim 1 - 3$ Myr corresponds to $t_{\text{lag}} \sim 1 - 10 t_{\text{ff}}$.

The Lag Time. The third timescale describing molecular clouds is the lag between when they form and when they begin to form stars, which we call the lag time $t_{\text{lag}}$. For molecular clouds in the solar neighborhood (out to 800 pc), the ratio of star-forming clouds over those without clear signs of star formation ranges between 7 and 14. Together with a median age of the young stars in these regions of 1–2 Myr, this entails a lag between cloud formation and onset of star formation of at most $t_{\text{lag}} \approx 1$ Myr, i.e., stars begin to form immediately after (or even during) the formation of the parental cloud. This is consistent with some extragalactic observational evidence that the spatial gap between spiral shocks in $H\alpha$ gas and bright 24 $\mu$m emission downstream of the shock, presumably tracing star formation, corresponds to a lag time $t_{\text{lag}} = 1 - 4$ Myr.

Before proceeding based on this conclusion, however, it is worth mentioning a caution. In both M33 and the LMC, the ratio of molecular clouds that have associated $H\alpha$ regions (detected either via $H\alpha$ or radio continuum emission) to those that do not is significantly smaller than the local ratio of star-forming clouds to non-star-forming ones: 3–4 instead of 7–14. This implies lag times of $\sim 7$ Myr, roughly $2 - 3 t_{\text{ff}}$, between GMC formation and HII region appearance. While the extragalactic clouds used for this measurement are much larger than the solar neighborhood ones and have free-fall times of $\sim 3$ Myr instead of $\sim 1$ Myr, they are comparable in size to the clouds probed by the geometric technique. The discrepancy between $t_{\text{ff}} = 2 - 3 t_{\text{ff}}$ and $t_{\text{lag}} < 1 t_{\text{ff}}$, between these techniques is therefore real. Its origin is unclear. One possibility is that it is simply a result of the different criteria used to measure the onset of star formation: infrared emission versus HII regions. This possibility has yet to be quantitatively evaluated, however. In the absence of an explanation of the discrepancy, we tentatively conclude based on the IR data that $t_{\text{lag}}$ is indeed short. Because of this extremely short timelag, any property the molecular cloud has to allow star formation, i.e., the strong density variations including small filling factors, and the supersonic turbulence, must come from the formation process of the cloud itself.

Strong density variations within molecular clouds are not only observed (Section 3.1.1, Table I), they also are physically mandated. The free-fall time for a spherical cloud of uniform density does not depend on the radius. Thus if one neglects pressure forces, material at the edge of the cloud would arrive at the same time at the center as material close to the center, making it virtually impossible to form isolated stars. “Distributed” star formation can only occur if the cloud acquires high, localized density seeds during its formation process or similarly if pre-existing density fluctuations exist that become strongly amplified, such that the local contraction time is substantially smaller than the global one. The distribution of these high-density regions determines the degree of clustering of the resulting stars, with over-densities that are correlated on small length scales leading to isolated stars or small clusters, while over-densities correlated on large scales produce large clusters. One possible explanation of the presence of small scale density inhomogeneity in newborn molecular clouds is that they form from atomic gas that is thermally bistable (Section 2.2.2). The onset of thermal instability leads to a wide-spread distribution of small-scale non-linear density fluctuations on very short timescales. This could explain how strong density variations occur on small scales even if molecular cloud turbulence is driven on large scales by the assembly of the cloud in the Galactic disk. In addition, GMC’s are highly filamentary and show sheet-like morphology. This means that also boundary effects are important and any pre-existing initial fluctuations are more easily amplified compared to models mentioned above that assume spherical symmetry. We note, however, that the smaller structures in their interior that form star clusters do appear to be more centrally concentrated.

2.2.2. Molecular Cloud Formation

The strict observational limits on $t_{\text{lag}}$, the time between when molecules first appear and when star formation begins, have recently led to a focus on the process of molecular cloud formation. This can be split into three issues, namely the accumulation problem, the chemistry problem, and the issue of rapid fragmentation. We will discuss each of them in turn.

The Accumulation Problem. Building a molecular cloud requires assembling a column density high enough for the dust to shield the cloud against the ambient UV-radiation, and thus to allow CO formation. (H$_2$ forms earlier, because it self-shields very efficiently. However, we “see” the cloud only once it contains CO.) Dust-shielding becomes efficient at column densities of approximately $N = 2 \times 10^{21}$ cm$^{-2}$. If we were to accumulate our prospective cloud at flow parameters typical for the inter-arm ISM, namely at densities 1 cm$^{-3}$ and velocities of 10 km s$^{-1}$, it would take $\sim 60$ Myr to accumulate the shading column density. This timescale is too large, given the maximum lifetimes of GMCs of $\lesssim$ Myr (Section 2.2.1).

However, this seemingly compelling argument is not applicable, since it only addresses a one-dimensional situation, resulting
in a sky-filling molecular cloud. Various three-dimensional solutions have been suggested to allow molecular cloud formation within realistically short times. Probably the oldest relies on the Parker instability, 30 a mechanism describing the buoyant rise of evaporated parts of flux tubes in a stratified Galactic disk. The rising flux tube leads to material falling into the valleys, thus accumulating molecular clouds. In combination with a magnetohydrodynamical Rayleigh-Taylor instability to trigger the initial evacuation in spiral shocks, 31 this scenario predicts accumulation timescales of 30 Myr and typical cloud distances of 1 kpc along spiral arms. However, more recent numerical simulations of magnetized disk gas dynamics 78, 79 only found weak signatures of a Parker instability acting along a spiral arm. Instead they identified the Magneto-Jeans instability 60, 80 as a more dominant accumulation mode in a magnetized disk. This instability is driven by in-plane magnetic fields countering the stabilizing effects of the Coriolis force, allowing self-gravitating contractions of overdense regions. Although we know that magnetic fields exist in galactic disks, it is also possible for molecular clouds to form on reasonable timescales in non-magnetized models provided the disk is globally gravitationally unstable. 32–38

On a more local scale, the interstellar medium is filled with flows of various types, many of which result in piling up material through shocks. While it is not always easy to identify specific driving sources in all cases (see, however, 87–89), this is not surprising given the complexity expected as a result of the interaction of neighboring flows. In addition, the presence of massive molecular clouds well out of the Galactic plane (e.g., Orion) clearly suggests the need for some kind of driving. Given the extensive impact that massive stars have on the interstellar medium—HII regions, stellar wind impacts, and ultimately supernova explosions—it is difficult to see how further creation of new star-forming locales by flows with scales of several to tens of pc (or even kpc, in the case of spiral arms) could be avoided. The notion of expanding shells piling up material has led to a “collect & collapse” model, 90 connecting the formation of molecular clouds to energetic events in the ISM.

Thermal instability is at the heart of the rapid fragmentation of the accumulating flows results from the combined action of heating and cooling processes in the ISM. In the diffuse, warm interstellar gas, at densities of \( n \approx 1 \text{ cm}^{-3} \), photo-electric heating by dust grains dominates, while in denser, higher extinction gas, heating by cosmic rays is more significant, becoming dominant deep in the interior of molecular clouds. 

The key to the observationally mandated rapid onset of star formation is to provide the parental cloud with high-amplitude (non-linear) density perturbations during its formation. We will describe a scenario of flow-driven cloud formation and rapid star formation in the context of cloud formation in spiral arms. While differing in details, other environments such as cloud formation in galaxy mergers or in the Galactic molecular ring, are subject to similar physical constraints and may work along similar lines, although this is an issue still to be explored.

Rapid fragmentation of the accumulating flows results from the combined action of heating and cooling processes in the ISM. In the diffuse, warm interstellar gas, at densities of \( n \approx 1 \text{ cm}^{-3} \), photo-electric heating by dust grains dominates, while in denser, higher extinction gas, heating by cosmic rays is more significant, becoming dominant deep in the interior of molecular clouds. In the regime dominated by photo-electric heating, the total heating rate per hydrogen nucleus varies by at most a factor of 10 over a range of densities from \( n \approx 1 \text{ to } 10^3 \text{ cm}^{-3} \). Cooling rates below \( 10^3 K \) depend strongly on the abundances of heavy elements, but have only a weak dependence on temperature at \( T > 100 \text{ K} \). Moreover, in the warm, diffuse ISM, the energy radiated in the dominant cooling lines, such as the hyperfine-structure line of singly ionized carbon at 158 \( \mu \text{m} \), scales as \( n^2 \), while the heating rate depends (roughly) on \( n \). Starting from an equilibrium situation, a small density increase thus leads to a cooling instability. If the size of the density perturbation is small, with a sound-crossing time that is less than its cooling time, then as the gas cools, its density will increase owing to compression from the surrounding warmer medium. If the temperature dependence of the cooling rate is weak, then the increase in the cooling rate produced by the growing density is greater than the decrease caused by the falling temperature. And so the perturbation cools with ever faster growing density. This process will stop when the density dependence of the cooling rate changes, for instance, if the level populations of the dominant coolant reach their local thermodynamic equilibrium values. In this case the cooling rate scales only as \( n \). It will also stop when the temperature dependence of the cooling rate becomes steeper, as will naturally occur at low temperatures. In the ISM, both effects are important, and the thermal instability vanishes once \( n \sim 100 \text{ cm}^{-3} \) and \( T \sim 50 \text{ K} \), resulting in a two-phase structure of interstellar gas, with a warm diffuse phase occupying large volumes, and a cold, dense phase with small filling factors in rough pressure equilibrium. 102, 103

Thermal instability is at the heart of the rapid fragmentation of the accumulating flows. It has been studied in various contexts, such as in generally turbulent media, 104–108 or in cloud formation behind shockfronts 109, 110 or in collisions of gas streams in spiral arm shocks or driven by e.g., expanding supernovae shells. 104, 106, 111, 112 The strength of the thermal instability derives from a combination with dynamical instabilities breaking up coherent shock fronts and shear-flow instabilities, 107, 113, 114 and from the fact that its growth timescales are substantially shorter than those of the hydromagnetic and gravitational instabilities involved. 115

The principal effect of this rapid thermal fragmentation is best gleaned from the evolution of the free-fall times in the forming...
cloud. Figure 4 shows the distribution of free-fall times against time in a molecular cloud being formed by two colliding flows of warm neutral hydrogen gas. Initially, the bulk of the cloud is at free-fall times longer than the simulation duration. At around 7 Myr, a substantial mass fraction of the cloud has dropped to free-fall times as short as 3 Myr. Substantial CO has formed by ≈10 Myr, while star formation sets in at ≈11 Myr, when noticeable mass fractions appear at free-fall times substantially shorter than those in the bulk of the cloud.

One of the key realizations in contrast to earlier models of turbulent fragmentation using periodic boxes is that the finite cloud geometry is crucial not only for the rapid onset of star formation, but also for the rapid formation of CO to overcome the otherwise stringent limitations set by the inflow density and speed. Thus these models bridge a gap between the large-scale simulations of Galactic disk dynamics discussed earlier, and the detailed models of turbulent fragmentation to be discussed in Section 3. The large-scale models do not have sufficient resolution to address the fragmentation and internal dynamics of the resulting molecular clouds, while models on smaller scales generally have to make simplifying assumptions about the boundary conditions.

What’s Missing in Cloud Formation Models? Perhaps the most serious complication with flow-driven cloud formation models is that by themselves they address only one of the fundamental timescales of star forming clouds introduced in Sec. 2.2.1, the lag time, \( t_{\text{lag}} \), between when clouds begin to accumulate and when star formation begins. Taken at face value, they do not explain cloud lifetimes, \( t_{\text{life}} \), nor the low overall star formation rates or equivalently the long gas depletion timescales. \( t_{\text{dep}} \). This is because these models by themselves lack an exit strategy. In the absence of energy sources within the cloud, the accumulated mass will start to collapse globally, the clouds would settle and convert a substantial fraction of their mass into stars, but also the observed limits on the star formation rate. Additional processes, such as feedback, are required to set the two remaining timescales. We come back to this issue in Section 4.

Fig. 4. Distribution of free-fall times against time in a molecular cloud forming in large-scale neutral hydrogen streams. Star formation sets in at ≈11 Myr, when the local free-fall times get substantially shorter than those in the bulk of the cloud. For comparison, substantial CO is visible at ≈10 Myr.

2.3. Modeling Molecular Cloud Formation in Hydrodynamic Simulations with Time-Dependent Chemistry

The chemical composition of the ISM is complex. Over 120 different molecular species have been detected in interstellar space and while many of these are found in detectable amounts only in dense, well-shielded gas, there remain a significant number that have been detected in diffuse, unshielded gas. A full chemical model of the ISM can easily involve several hundred different atomic and molecular species and isotopologues and several thousand different reactions, even if reactions on grain surfaces are neglected, see e.g., the UMIST database.

It is currently impractical to incorporate this amount of chemistry into a 3D hydrodynamical code. The key to constructing time-dependent chemical networks that can run alongside the dynamic evolution of the system therefore is to select a number of chemical species and mutual reaction rates that is small enough so that the chemical network can be solved in a short enough time so that it is tractable to do so during each system timestep and that is large and complete enough so that the overall evolution of the system is still described adequately. In the context of molecular cloud formation it is clearly necessary to be able to follow the formation and destruction of H$_2$ with a reasonable degree of accuracy. Beyond this, the only chemistry that is really required is that which plays a role in determining the thermal balance of the gas. In other words, we need only follow the chemistry of H$_2$ and of a few other major coolants such as C$^+$ or O in low column density gas, or CO in high column density regions. As few as thirty species and two hundred reactions appear to be sufficient for accurately modelling the most important hydrogen, carbon and oxygen chemistry in molecular cloud formation calculations, and a network of this size has been shown to be practical to incorporate into a 3D hydrodynamical code. Synthetic emission maps from turbulent box calculations with such type of chemical network are shown in Figure 5.

Many reaction rates are sensitive to the external radiation field. Molecular hydrogen, for example, can only remain stable in regions where the column density is high enough to significantly attenuate the Galactic radiation field. This means, that any sensible calculation of chemical reaction rates requires knowledge of the local radiation field. Ideally, simulations with time-dependent chemistry running alongside the hydrodynamics should also include full radiative transfer (as discussed in Section 4.2). This is, however, beyond the capabilities of current numerical schemes, and most astrophysical applications treat radiation in a very approximate fashion only, for example, by assuming a constant background field or by computing column densities and optical depths only along the principle axes of the system.

Although it has as yet received only limited attention from computational astrophysicists, efficient coupling between chemical reaction networks and hydrodynamical solvers is an active area of research in fields such as combustion modelling (e.g., Ref. [122]) or atmospheric chemistry (e.g., Ref. [123]). The basic principles are straightforward. One usually uses some form of operator splitting to separate the treatment of the chemistry from the advection and/or diffusion terms. During the chemistry sub-step, one updates the chemical abundances by solving a coupled set of rate equations of the form

\[
\frac{d\eta_j}{dt} = C_j - D_j \eta_j
\]
Fig. 5. Synthetic emission maps from turbulent box calculations with time-dependent chemistry. The top left panel depicts the total H$_2$ column density, while the top right panel shows the integrated optically thin line emission from the tracer molecule CO. For comparison, the lower left image shows the total gas column density. Inspection of the top two images illustrates that CO traces the H$_2$ distribution very well in high column density regions, however, fails to do so for the low-density surface layers that are more strongly exposed to the external radiation field. This is quantified in the bottom right image which shows the ratio between both values. This ratio is related to the so-called “x-factor” that is commonly used to convert CO intensity maps into H$_2$ maps.

Here $n_i$ is the number density of species $i$, and $C_i$ and $D_i$ are chemical creation and destruction terms that generally depend on the temperature $T$ and the chemical abundances of the other reactants in the system. Most chemical reaction networks are stiff—that is, they contain a wide range of different characteristic timescales—and so to ensure stability, it is usually necessary to solve these coupled rate equations with an implicit scheme. The simplest implicit techniques have a computational cost that scales as the cube of the number of species, and so considerable ingenuity has been expended in attempts to reduce this cost, for instance by making use of chemical conservation laws to reduce the number of species that need to be tracked, or by taking advantage of the typically sparse nature of the Jacobian of the coupled set of equations (e.g., Ref. [124]).

The thermal evolution of the gas is usually modelled using a library of cooling functions for each considered species. For example, state-of-the-art calculations include the effects of atomic fine structure cooling (e.g., C, C$^+$, O, Si and Si$^+$), rotational and vibrational cooling of the considered molecules (e.g., H$_2$, HD, CO and H$_2$O), Lyman-$\alpha$ cooling, Compton cooling, and H$^+$ recombination cooling, as well as other processes of lesser importance. The numerical implementation usually adopts the isochoric approximation and computes emission strength using the large-velocity gradient approach, which assumes that the emitted lines are absorbed locally only. During a given hydrodynamic timestep one computes first $\dot{u}_{ad}$, the rate of change of the internal energy due to adiabatic gas physics. Then one has to solve an implicit equation for the new internal energy of the form

$$u^{new} = u^{old} + \left[ \dot{u}_{ad} - \frac{\Lambda (\rho^{new}, u^{new})}{\rho^{new}} \right] \Delta t$$

(4)

where $u^{new}$ and $u^{old}$ are the internal energy per unit mass at the current and old time, respectively, $\rho^{new}$ is the gas density at the current time. It is often necessary to solve this implicit equation simultaneously with the chemical rate equations.
3. FROM CLOUD CORES TO STARS

3.1. Observational Properties of Molecular Cloud Cores

3.1.1. Statistical Properties

Emission line observations and dust extinction maps of molecular clouds reveal extremely complex morphological structure with clumps and filaments on all scales accessible by present day telescopes. Typical parameters of different regions in molecular clouds are listed in Table I. The volume filling factor of dense clumps, even denser subclumps and so on, is rather small, ranging from 10% at densities of \( n \approx 10^3 \text{ cm}^{-3} \) down to 0.1% for \( n > 10^5 \text{ cm}^{-3} \).\(^{51,126-128}\) This hierarchical configuration is often interpreted as being fractal or self-similar,\(^{129-134}\) which however, is still subject to debate.\(^{135}\) It is important to note that star formation always occurs in the densest regions within a cloud, so only a small fraction of molecular cloud matter is actually involved in building up stars, while the bulk of the material remains at lower densities. This is most likely the key to explaining the low star formation efficiencies as discussed above in Section 2.2.

The mass spectrum of clumps in molecular clouds appears to be well described by a power law,

\[
\frac{dN}{dm} \propto m^n
\]

with an exponent usually reported in the range \(-1.3 < \alpha < -1.8\).\(^{136-138}\) The standard cloud decomposition methods, however, are not without pitfalls.\(^{139}\) The observed self-similar behavior indicates that there is no natural mass or size scale between the lower and upper limits of the observations. The smallest observed structures are protostellar cores with masses of a few solar masses or less and sizes of \( \lesssim 0.1 \text{ pc} \). The fact that all studies obtain a similar power law is remarkable, and may be the result of turbulent motions and thermal instability acting on self-gravitating gas. Given the uncertainties in determining the slope, it appears reasonable to conclude that there is a universal mass spectrum for the clumps within a molecular cloud, and that the distribution is a power law within a mass range of three orders of magnitude, i.e., from \( 1 M_\odot \) to about \( 1000 M_\odot \). Hence, it appears plausible that the physical processes that determine the distribution of clump masses are rather similar from cloud to cloud. And vice versa, clouds that show significant deviation from this universal distribution most likely had different dynamical histories.

Most of the objects that enter in the above morphological analyses are not gravitationally bound.\(^{136,140-142}\) It is interesting to note that the distribution changes as one probes smaller and smaller scales and more and more bound objects. When considering prestellar cores, which are thought to be the direct progenitors of individual stars or small multiple systems, then the mass function is well described by a double power law fit \( dN/dm \propto m^{-\alpha} \) following \( \alpha = 2.5 \) above \( \sim 0.5 M_\odot \) and \( \alpha = 1.5 \) below. The first large study of this kind was published by Motle et al.,\(^{143}\) for a population of submillimetre cores in \( \rho \text{ Oph} \). Using data obtained with IRAM, they discovered a total of 58 starless clumps, ranging in mass from \( 0.05 M_\odot \) to \( \sim 3 M_\odot \). Similar results are obtained from the Serpens cloud,\(^{144}\) for Orion B North,\(^{145}\) and Orion B South,\(^{146}\) or for the Pipe Nebula.\(^{147}\) Currently all observational data\(^{143,146-148,153}\) reveal that the mass function of prestellar cores is strikingly similar in shape to the stellar initial mass function, the IMF. To reach complete overlap one is required to introduce a mass scaling or efficiency factor in the range 2 to 10, which differs in different regions. An exciting interpretation of these observations is that we are witnessing the direct formation of the IMF via fragmentation of the parent cloud. However, we note that the observational data also indicate that a considerable fraction of the prestellar cores do not exceed the critical mass for gravitational collapse, much like the clumps on larger scales. The evidence for a one-to-one mapping between prestellar cores and the stellar mass thus is by no means conclusive as we will discuss in more detail in Section 3.3.

3.1.2. Individual Cores

Density Structure. The density structure of prestellar cores is typically estimated through the analysis of dust emission or absorption using near-IR extinction mapping of background starlight, mapping of millimeter/submillimeter dust continuum emission, and mapping of dust absorption against the bright mid-IR background emission.\(^{141}\) A main characteristic of the density profiles derived with the above techniques is that they require a central flattening. The density gradient of a core is flatter than \( r^{-1} \) within radii smaller than \( 2500-5000 \text{ AU} \), and that the typical central density of a core is \( 10^5-10^6 \text{ cm}^{-3} \).\(^{143,144}\) A popular approach to describe these cores as truncated isothermal (Bonnor-Ebert) spheres\(^{155,156}\) that often (but not always) provide a good fit to the data.\(^{157-159}\) These are equilibrium solutions of self-gravitating gas spheres bounded by external pressure. However, such density structure is not unique. Numerical calculations of the dynamical evolution of supersonically turbulent clouds show that transient cores forming at the stagnation points of convergent flows exhibit similar morphology.\(^{141,160}\)

Thermal Structure. The kinetic temperature of the dust and gas components in a core is regulated by the interplay between various heating and cooling processes. At high densities (\( n > 10^3 \text{ cm}^{-3} \)) in the inner part of the cores, the gas and dust have to be coupled thermally via collisions.\(^{162-164}\) At lower densities, which correspond to the outer parts of the cores, the two temperatures are not necessary expected to be the same. Thus, the dust and gas temperature distributions need to be inferred from observations independently. Large-scale studies of dust temperature show that the grains in starless cores are colder than in the surrounding lower-density medium. Far-IR observations toward the vicinity of a number of dense cores provide evidence for flat or decreasing temperature gradients with cloud temperatures of \( 15-20 \text{ K} \) and core values of \( 8-12 \text{ K} \).\(^{165,166}\) These observations are consistent with dust radiative transfer modeling in cores illuminated by the interstellar radiation field,\(^{167-169}\) where the dust temperature is \( \sim 7 \text{ K} \) in the core center and increases up to \( 16 \text{ K} \) in the envelope. The gas temperature in molecular clouds and cores is commonly inferred from the level excitation of simple molecules like CO and NH\(_3\).\(^{170,171}\) One finds gas temperatures of \( 10-15 \text{ K} \), with a possible increase toward the lower density gas near the cloud edges. It is believed that the gas heating in prestellar cores mostly occurs through ionization by cosmic rays, while the cooling is mainly due to line radiation from molecules, especially CO.\(^{162}\) Altogether, the fact that prestellar cores are cold and roughly isothermal with at most a modest increase in temperature from the center to the edge is consistent with numerical models of cores forming from thermal instability.\(^{141,172,173}\) See also Figure 6.
Chemical Structure. Maps of integral line intensity can look very different for different molecular tracers. In particular, the N$_2$H$^+$ and NH$_3$ emission more closely follows the dust emission while the C$^{18}$O and CS emission appears as a "ring-like" structure around the dust emission maximum. For illustration see Figure 7. The common theoretical interpretation of these data is that carbon-bearing species, represented by CO and CS freeze-out on the dust grains from the gas while the abundances of nitrogen–hydrogen bearing molecules, N$_2$H$^+$ and NH$_3$, either stay constant or decay more slowly. At the same time, chemical models of prestellar cores predict that molecules in the core envelope have to be destroyed by interstellar UV field. The chemical stratification significantly complicates the interpretation of molecular line observations and again requires the use of sophisticated chemical models which have to be coupled to the dynamical evolution. From the observational side, the freeze-out of many molecules makes it difficult to use their emission lines for probing the physical conditions in the inner regions of the cores. At the same time, the modeling of the chemical evolution can provide us with important parameters of the cores. For example, the level of CS depletion can be used to constrain the age of the prestellar cores while the deficit of CS in the envelope can indicate the strength of the external UV field. In any case, any physical interpretation of the molecular lines in prestellar cores has to be based on chemical models and should do justice to the underlying density and velocity pattern of the gas.

Kinematic Structure. In contrast to the supersonic velocity fields observed in molecular clouds, dense cores have low internal velocities. Starless cores in clouds like Taurus, Perseus, and Ophiuchus systematically present spectra with close-to-thermal linewidths, even when observed at low angular resolution. This indicates that the gas motions inside the cores are subsonic or at best transsonic, i.e., with Mach numbers $\lesssim 2$. In some cases also inward motions have been detected. They are inferred from the observation of optically thick, self-absorbed lines of species like CS, H$_2$CO, or HCO$^+$, in which low-excitation foreground gas absorbs part of the background emission. Typical inflow velocities are of order of 0.05–0.1 km/s and are observed on scales of 0.05–0.15 pc, comparable to the observed size of the cores. The overall velocity structure of starless cores appears broadly consistent with the structure predicted by models in which protostellar cores form as the stagnation points of convergent flows, but the agreement is not perfect. Simulations of core formation do correctly find that most cores are at most transonic, but the distribution of velocity dispersions has a small tail of highly supersonic cores that is not observed. Clearly more theoretical and numerical work is needed. In particular, the comparison should be based on synthetic line emission maps, which requires to couple a chemical network and radiative transfer to the simulated density profiles as discussed above. In addition, it is also plausible that the discrepancy occurs because the simulations do not include all the necessary physics such as radiative feedback and magnetic fields. Subsonic turbulence contributes less to the energy budget of the cloud than thermal pressure and so cannot provide sufficient support against gravitational collapse. If cores are longer lasting entities there must be other mechanisms to provide stability. Obvious candidates are magnetic fields. However, they are usually not strong enough to provide sufficient support as discussed below. Most observed cores are thus likely to be evolving transient objects that never reach any equilibrium state.

Magnetic Field Structure. Magnetic fields are ubiquitously observed in the interstellar gas on all scales. However, their importance for star formation and for the morphology and evolution of molecular cloud cores remains controversial. A crucial parameter in this debate is the ratio between core mass and magnetic flux. In supercritical cores, this ratio exceeds a critical value and collapse can proceed. In subcritical ones, magnetic fields provide stability. Measurements of the Zeeman splitting of molecular lines in nearby cloud cores indicate mass-to-flux ratios that lie above the critical value, in some cases only by a small margin but very often by factors of many if non-detections are included. The polarization of dust emission offers an alternative pathway to studying the magnetic field structure of molecular cloud cores. MHD simulations of turbulent clouds predict degrees of polarization between 1 and 10%, regardless of whether turbulent energy dominates over the magnetic energy (i.e., the turbulence is super-Alfvénic) or not. However, converting polarization into magnetic field strength is very difficult. Altogether, the current observational findings imply that magnetic fields must be considered when studying stellar birth, but also that they are not the dominant agent that determines when and where stars form within a cloud. Magnetic fields appear too weak to prevent gravitational collapse.
This conclusion means that in many cases and to reasonable approximation purely hydrodynamic calculations are sufficient for star formation simulations. However, when more precise and quantitative predictions are desired, e.g., when attempting to predict star formation timescales or binary properties, it is necessary to perform magnetohydrodynamic (MHD) simulations or even consider non-ideal MHD. The latter means to take ambipolar diffusion (drift between charged and neutral particles) or Ohmic dissipation into account. Recent numerical simulations have shown that even a weak magnetic field can have noticeable dynamical effects. It can alter how cores fragment, change the coupling between stellar feedback processes and their parent clouds, influence the properties of protostellar disks due to magnetic braking, or slow down the overall evolution.

### 3.1.3. Models of Cloud Evolution and Star Formation

There are two main competing models that describe the evolution of the cloud cores. It was proposed in the 1980’s that cores in low-mass star-forming regions evolve quasi-statically in magnetically subcritical clouds. Gravitational contraction is mediated by ambipolar diffusion, causing a redistribution of magnetic flux until the inner regions of the core become supercritical and go into dynamical collapse. This process was originally thought to be slow, because in highly subcritical clouds the ambipolar diffusion timescale is about 10 times larger than the dynamical one. However for cores close to the critical value, as is suggested by observations, both timescales are comparable. Numerical simulations furthermore indicate that the ambipolar diffusion timescale becomes significantly shorter for turbulent velocities similar to the values observed in nearby star-forming regions. The fact that ambipolar diffusion may not be a slow process under realistic cloud conditions, as well as the fact that most cloud cores are magnetically subcritical, has cast significant doubts on any magnetically-dominated quasi-static models of stellar birth.

For this reason, star-formation research has turned into considering supersonic turbulence as being on of the principal physical agents regulating stellar birth. The presence of turbulence, in particular of supersonic turbulence, has important consequences for molecular cloud evolution. On large scales it can support clouds against contraction, while on small scales it can provoke localized collapse. Turbulence establishes a complex network of interacting shocks, where dense cores form at the stagnation points of convergent flows. The density can be large enough for gravitational collapse to set in. However, the fluctuations in turbulent velocity fields are highly transient. The random flow that creates local density enhancements can disperse them again. For local collapse to actually result in the formation of stars, high density fluctuations must collapse on timescales shorter than the typical time interval between two successive shock passages. Only then are they able to ‘decouple’ from the ambient flow and survive subsequent shock interactions. The shorter the time between shock passages, the less likely these fluctuations are to survive. Hence, the timescale and efficiency of protostellar core formation depend strongly on the wavelength and strength of the driving source, and accretion histories of individual protostars are strongly time varying.

Interstellar turbulence is observed to be dominated by large-scale modes. This implies it is very efficient in sweeping up molecular cloud material, thus creating massive coherent structures. The result is a large region in which many Jeans masses of material become unstable to collapse at about the same time, leading to coherent structures in the forming stars. This is a likely explanation for the observed clustering of young stars, as we discuss in the following section.

### 3.2. Spatial Distribution

The advent of sensitive infrared detectors in the last decade has made it possible to perform wide-area surveys. These have led us to recognize that most stars form in clusters and aggregates of various size and mass scales, and that isolated or widely distributed star formation is the exception rather than the rule. The complex hierarchical structure of molecular clouds (Fig. 2) provides a natural explanation for this finding. Star-forming molecular cloud cores vary enormously in size and mass. In small, low-density clouds, stars form with low efficiency, more or less in isolation or scattered around in small groups of up to a few dozen members. Denser and more massive clouds may build up...
stars in associations and clusters of a few hundred members. This appears to be the most common mode of star formation in the solar neighborhood. Examples of star formation in small groups and associations are found in the Taurus-Aurigae molecular cloud. Young stellar groups with a few hundred members form in the Chamaeleon II or ρ-Ophiuchi dark clouds. Each of these clouds is at a distance of about 130 to 160 pc from the Sun. Like most of the nearby young star forming regions they appear to be associated with a ring-like structure in the Galactic disk called Gould’s belt.

The formation of dense rich clusters with thousands of stars is rare. The closest region where this happens is the Orion Nebula Cluster in L1641, which lies at a distance of 410 pc. A rich cluster somewhat further away is associated with the Monoceros R2 cloud at a distance of ~830 pc. The cluster NGC 3603 is roughly ten times more massive than the Orion Nebula Cluster. It lies in the Carina region, at about 7 kpc distance. It contains about a dozen O stars, and is the nearest object analogous to a starburst knot. To find star-forming regions building up hundreds of O stars one has to look towards giant extragalactic H ii regions, the nearest of which is 30 Doradus in the Large Magellanic Cloud, a satellite galaxy of our Milky Way at a distance of 55 kpc. The giant star forming region 30 Doradus is thought to contain up to a hundred thousand young stars, including more than 400 O stars. This sequence as depicted in Figure 8 demonstrates that the star formation process spans many orders of magnitude in scale, ranging from isolated single stars to massive young clusters with several 10^5 stars.

3.3. The Stellar Initial Mass Function and Other Statistical Characteristics of Star Formation

The mass distribution of young stars follows a well-known distribution called the Initial Mass Function (IMF). For stellar masses \( m \geq 1 M_\odot \) it shows a power-law behavior \( dN/dm \propto m^{-\alpha} \), with slope \( \alpha = -2.3 \). Below \( 1 M_\odot \), the IMF flattens, a change in behavior that can be represented either as a lognormal or a change in power law index. At the extreme ends of the stellar mass spectrum, however, our knowledge of both the IMF is limited. Massive stars are very rare and rather short lived. The number of massive stars that are sufficiently near to study in detail and with very high spatial resolution, for example to determine multiplicity, is therefore small. Low-mass stars and brown dwarfs, on the other hand, are faint, so they too are difficult to study in detail. Such studies, however, are in great demand, because secondary indicators such as the fraction of binaries and higher-order multiples as function of mass, or the distribution disks around very young stars or possible signatures of accretion during their formation are probably better suited to distinguish between different star-formation models than just looking at the IMF.

In contrast to the observational agreement on the IMF, at least above the substellar regime, there is still considerable disagreement on the theoretical side. The origin of the IMF is a major topic of theoretical research which we examine only briefly here to give the necessary theoretical background for our discussion of numerical work. Other reviews provide considerably more detail.

Early models for the origin of the IMF generally relied on statistical arguments, appealing to random processes of collapse in a fractal cloud or to the central limit theorem to explain its characteristic shape. Researchers have also invoked feedback processes that cut off accretion onto individual protostars. Today, however, there are three dominant schools of thought regarding the origin of the IMF, although the boundaries between these pictures are not clearly defined, and numerous hybrid models have been proposed. One model, called core accretion, takes as its starting point the striking similarity between the shape of the observed core mass distribution and the IMF. This model posits that there is a one-to-one relation between the distributions, so that individual cores are the progenitors of individual stars or star systems. The factor of ~3 decrease in mass between cores and stars is the result of feedback processes, mostly protostellar outflows, that eject a fixed fraction of the mass in a core rather than letting it accrete onto the star. This model reduces the problem of the origin of the IMF to the problem of determining the mass spectrum of bound cores, although strictly speaking the idea that the IMF is set by the mass spectrum of cores is independent of any particular model for the origin of that mass spectrum. Arguments to explain the core mass distribution generally rely on the statistical properties of turbulence, which generate structures with a pure powerlaw mass spectrum. The thermal Jeans mass in the cloud then imposes the flattening and turn-down in the observed mass spectrum.

A second model for the origin of the IMF, called competitive accretion, focuses instead in interaction between protostars, and between a protostellar population and the gas cloud around it. In the competitive accretion picture the origin of the peak in the IMF is much the same as in the core accretion model: it is set by the Jeans mass in the prestellar gas cloud. However,
rather than fragmentation in the gas phase producing a spectrum of core masses, each of which collapses down to a single star or star system, in the competitive accretion model all gas fragments down to roughly the Jeans mass. Prompt fragmentation therefore creates a mass function that lacks the powerlaw tail at high masses that we observe in the stellar mass function. This part of the distribution forms via a second phase in which Jeans mass-protostars compete for gas in the center of a dense cluster. The cluster potential channels mass toward the center, so stars that remain in the center grow to large masses, while those that are ejected from the cluster center by N-body interactions remain low mass.271, 272 In this model, the apparent similarity between the core and stellar mass functions is an illusion, because the observed cores do not correspond to gravitationally bound structures that will collapse to stars.273, 274

One potential drawback to both the core accretion and competitive accretion models is that they rely on the Jeans mass to determine the peak of the IMF; but leave unanswered the question of how to compute it. This question is subtle because molecular clouds are nearly isothermal, but they contain a very wide range of densities, and it is unclear which density should be used. A promising idea to resolve this question, which is the basis for a third model of the IMF, focuses on the thermodynamic properties of the gas. The amount of fragmentation occurring during gravitational collapse depends on the compressibility of the gas.275 For polytropic indices $\gamma < 1$, turbulent compressions cause large density enhancements in which the Jeans mass falls substantially, allowing many fragments to collapse. Only a few massive fragments get compressed strongly enough to collapse in less compressible gas though. In real molecular gas, the compressibility varies as the opacity and radiative heating increase. Reference [7] noted that the thermal coupling of the gas to the dust at densities above $n_{\text{crit}} \sim 10^5$ to $10^6$ cm$^{-3}$ leads to a shift from an adiabatic index of $\gamma \sim 0.7$ to 1.1 as the density increases above $n_{\text{crit}}$. The Jeans mass evaluated at the temperature and density where this shift occurs sets a mass scale for the peak of the IMF. The apparent universality of the IMF in the Milky Way and nearby galaxies may be caused by the insensitivity of the dust temperature on the intensity of the interstellar radiation field.276 Not only does this mechanism set the peak mass, but also appears to produce a power-law distribution of masses at the high-mass end comparable to the observed distribution.277

Each of these models has potential problems. In the core accretion picture, hydrodynamic simulations seem to indicate that massive cores should fragment into many stars rather than collapsing monolithically.273, 278, 279 The hydrodynamic simulations almost certainly suffer from over-fragmentation because they do not include radiative feedback from embedded stars,280-286 but no simulation to date has successfully formed a massive core in a turbulent cloud and followed it all the way to the formation of a massive star. In addition, the suggestion of a one-to-one mapping between the observed clumps and the final IMF is subject to strong debate. Many of the prestellar cores discussed in Section 3.1.1 appear to be stable entities,145, 146, 148, 157 and thus are unlikely to be in a state of active star formation. In addition, the simple interpretation that one core forms on average one star, and that all cores contain the same number of thermal Jeans masses, leads to a timescale problem287 that requires differences in the core mass function and the IMF.

The criticism regarding neglect of radiative feedback effects also applies to the gas thermodynamic idea: the cooling curves that these models assume in order to derive the Jeans mass ignore the influence of protostellar radiation on the temperature of the gas, which simulations show can suppress fragmentation in at least some circumstances.281 The competitive accretion picture has also been challenged, on the grounds that the kinematic structure observed in star-forming regions is inconsistent with the idea that protostars have time to interact with one another strongly before they completely accrete their parent cores.183, 288

3.4. Modeling Cloud Fragmentation and Protostellar Collapse

To adequately model the fragmentation of molecular clouds, the formation of dense cloud cores, the collapse of the gravitationally unstable subset of cores, and finally the build-up and mass growth of embedded protostars in their interior is an enormous computational challenge. It requires to follow the evolution of self-gravitating, highly turbulent gas over many order of magnitudes in density and lengthscale. Owing to the stochastic nature of supernova turbulence, it is not known in advance where and when local collapse occurs. One therefore needs highly flexible numerical methods for solving the equations of hydrodynamics, schemes that can provide sufficient degrees of precision and resolution throughout the entire computational domain in an adaptive fashion.

The star formation community is following two highly complementary approaches to reach these goals. One set of methods is based on dividing the computational domain into small volume elements and follow the fluxes of all relevant quantities from one cell to the other. These grid-based methods adopt an Eulerian point of view, because the flow is followed from fixed positions in space. A popular alternative is to split the model cloud into individual parcels of gas and follow their mutual interaction and evolution. Particle-methods therefore correspond to a Lagrangian point of view following the trajectories of individual fluid elements.

3.4.1. Grid-Based Methods

The mathematical formulation of hydrodynamics consists of a set of partial differential equations that relate different flow properties (such as density and velocity) with each other and with thermodynamic quantities (e.g., pressure, temperature or internal energy of the medium). They can be formulated in conservative form corresponding to the conservation of mass, momentum, and energy. As the number of variables is larger than the number of equations in the system, an additional equation is needed to find a unique solution. This closure relation is called equation of state and usually specifies the pressure as function of other thermodynamic variables.289 In a broad sense, the hydrodynamics equations describe how signals propagate through a medium. They specify how local quantities relate to fluxes, e.g., how the density in some control volume depends on the mass flux through its surface. Equations of this kind are called hyperbolic equations.

Numerical solutions to partial differential equations always require discretization of the problem. This means that instead of continuous space and time dimensions we consider a discrete set of points. The computational domain is subdivided into individual volume elements surrounding node points on a grid or unstructured mesh. Finite volume methods are procedures for
representing and evaluating partial differential equations as algebraic equations. They play a key role in computational fluid dynamics. Similar to finite difference schemes, values are calculated at discrete places on a meshed geometry. Volume integrals that contain a divergence term are converted to surface integrals, using the divergence theorem. These terms are then evaluated as fluxes at the surfaces of each volume element. Because the flux entering a given volume is identical to that leaving the adjacent volume, these methods are conservative. Finite volume methods have been in the focus of applied mathematics for decades. They have well defined convergence properties and available code packages have reached a very high degree of maturity and reliability.

Finite volume schemes are most easily formulated for Cartesian grids with fixed cell size. The cell size determines the spatial resolution of the code. Wherever higher resolution is needed, it can be achieved by refining the grid. We speak of adaptive mesh refinement (AMR), when this is done in an automated and locally adjustable way. There are a number of different approaches to AMR in the literature. Most AMR treatments are based on finite-element models on unstructured meshes. They have the advantage to adapt easily to arbitrary complicated boundaries; however, constructing the mesh is time consuming. When using Cartesian grids, one can refine on individual cells or on larger groups of cells, so-called blocks. Cartesian AMR codes nowadays belong to the standard repertoire of numerical star formation studies.

In the following we list a few popular hydrodynamic and magnetohydrodynamic codes that have been developed in the past decade. All but two are freely available, although in some cases registration is needed before being able to download it from the web. ZEUS-MP (http://cosmos.ucsd.edu/lca-www/software/index.html) is a parallel, non-adaptive hydro- and magnetohydrodynamics code with self-gravity and radiation. NIRVANA (http://nirvana-code.aip.de) is an AMR code for non-relativistic, compressible, time-dependent, ideal or nonideal (viscosity, magnetic diffusion, thermal conduc-
tion) MHD. FLASH (http://flash.uchicago.edu/) is a highly modular, parallel adaptive-mesh code initially designed for thermonuclear runaway problems but also capable of a wide variety of astrophysical problems including radiative feedback from stellar sources. ENZO (http://lca.ucsd.edu/projects/enzo) is a hybrid AMR code (hydrodynamics and \(N\)-body) which is designed to do simulations of cosmological structure formation. It has been extended to include magnetic fields, star formation, and ray-tracing radiation transfer. ATHENA (http://www.astro.princeton.edu/jstone/athena) is an MHD code built on a flexible framework that is designed to allow easy and modular extension to include a wide variety of physical processes. The public version is non-adaptive, contains only hydrodynamics and MHD, and uses a fixed Cartesian grid, but extensions exist for non-Cartesian grids, for static and adaptive mesh refinement, for gravity, and for ionizing radiation transfer. Another widely used AMR code for MHD calculations is RAMSES. It is very versatile with applications ranging from the two-phase interstellar medium, to star and planet formation, as well as cosmological structure formation. A grid code used commonly in star formation simulations, but which is not publicly available, is ORION: a parallel hydrodynamics code that includes self-gravity, sink particles coupled to a protostellar evolution model, and diffuse radiative transfer. Last in our short summary is Proteus, a finite-volume method based on a gas-kinetic formulation of the microscopic transport properties. This approach allows the user to fully control the dissipative effects, making the scheme very attractive for e.g., turbulent transport studies. However, adding additional physics is generally more complicated than for other schemes. Proteus is fully parallelized and includes self-gravity, magnetic fields and a two-fluid model for ambipolar drift.

### 3.4.2. Particle-Based Methods

Using a particle based scheme to solve the equations of hydrodynamics was first introduced by Ref. [312] and proposed independently by Ref. [313]. Originally envisioned as a Monte-Carlo approach to calculate the time evolution of a hydrodynamic system, the formalism of smoothed particle hydrodynamics (SPH) is more intuitively understood as a particle interpolation scheme. This provides better estimates for the errors involved and the convergence properties of the method. Excellent overviews of the method and some of its applications provide the reviews of Benz and Monaghan.

In the framework of classical physics, fluids and gases are large ensembles of interacting particles with the state of the system being described by the probability distribution function in phase space. Its time evolution is governed by Boltzmann’s equation. Hydrodynamic quantities can then be obtained in a local averaging process involving scales larger than the local mean-free path. A related approach is facilitated in SPH. The fluid is represented by an ensemble of particles \(i\), each carrying mass, momentum, and hydrodynamical properties. The technique can therefore be seen as an extension to the well known \(N\)-body methods used in stellar dynamics. Besides being characterized by its mass \(m_i\), and velocity \(v_i\), and its location \(r_i\), each particle is associated with a density \(\rho_i\), an internal energy \(e_i\) (equivalent to a temperature \(T_i\)), and a pressure \(p_i\). The time evolution of the fluid is then represented by the time evolution of the SPH particles. Their behavior is governed by the equation of motion, supplemented by further equations to modify the hydrodynamical properties. Thermodynamical observables are obtained by averaging over an appropriate subset of the SPH particles.

Mathematically, the local averaging process for a quantity \(f(r)\) can be performed by convolution with an appropriate smoothing function \(W(r,h)\):

\[
\langle f(r) \rangle \equiv \int f(r') W(r-r',h) d^3 r'
\]

This function \(W(r,h)\) is often referred to as the smoothing kernel. It must be normalized and approach the Dirac delta function in the limit \(h \to 0\). For simplicity, most authors adopt spherical symmetry in the smoothing and averaging process, i.e., the kernel degrades to an isotropic function of the interparticle distances: \(W(r,h) \equiv W(r,h)\) with \(r = |r|\) and \(h = |h|\).

The basic concept of SPH is a particle representation of the fluid. Hence, the spatial integration in the averaging process transforms into a summation over a fixed number of points. For example, the density at the position of particle \(i\) is computed as

\[
\rho(r_i) = \sum m_i W(|r_i - r_j|, h)
\]

In this picture, the mass of each particle is smeared out over the kernel region. The continuous density distribution of the fluid
is then obtained by summing over the local contribution from neighboring elements \( j \). The name “smoothed particle hydrodynamics” derives from this analogy.

In star formation studies SPH is popular because it is intrinsically Lagrangian. As opposed to mesh-based methods, it does not require a fixed grid to represent fluid properties and calculate spatial derivatives.\(^{325}\) The fluid particles are free to move and—in analogy—constitute their own grid. The method is therefore able to resolve very high density contrasts, by increasing the particle concentration where needed. This it most effective, if the smoothing length is adaptable.\(^{317}\) There is no need for the complex and time-consuming issue of adaptive grid-refinement. However, the method has its weaknesses compared to grid-based methods. For example, its convergence behavior is mathematically difficult to assess and the method has problems reproducing certain types of dynamical instabilities.\(^{310}\) The algorithmic simplicity of the method and the high flexibility due to its Lagrangian nature, however, usually outweigh these drawbacks and SPH remains one of the numerical workhorses of current star-formation studies. There are various implementations of the method. Examples of very popular codes are GADGET,\(^ {125, 320}\) GASOLINE,\(^ {321}\) Magma,\(^ {322}\) and the various descendents of the SPH program originally developed by Benz.\(^ {315, 323, 324}\)

### 3.4.3. Sink Particles as Subgrid-Scale Models for Protostars
A fundamental problem for modeling protostellar collapse and star formation are the enormous density contrasts that need to be covered. Regions of high density require small cell sizes in grid-based methods,\(^ {325}\) or equivalently, small-particle masses in SPH calculations.\(^ {324, 326}\) In order to guarantee stability, every numerical scheme must resolve the traversal of sound waves across the minimum resolution element, i.e., either across one cell or across the smoothing kernel of individual SPH particles. This is the so called Courant Friedrich Lewy criterion. It causes the time integration stepsize to get smaller and smaller as the density increases. As a consequence, a computation virtually grinds to a halt during gravitational collapse.

When modeling the build-up of entire clusters of stars, or even following the accretion of the bulk of a core’s mass onto a single star, this problem clearly needs to be overcome. One way out is to introduce sink particles. Once the very center of a collapsing cloud core exceeds a certain density threshold (usually several thousand times the mean density, or using a threshold based on the Jeans mass) it is replaced by one single particle which inherits the combined masses, linear and angular momentum of the volume it replaces and which has the ability to accrete further gas from the infalling envelope. This permits to follow the dynamical evolution of the system over many global free-fall times, however, at the cost of not being able to resolve the evolution at densities above the threshold value. In a sense, sink particles introduce “inner boundaries” to the computational domain. They have been successfully introduced to grid-based\(^ {161, 306}\) as well as particle-based methods.\(^ {277, 321}\)

Each sink particle defines a control volume with a fixed radius. It lies typically between a few and a few hundred astronomical units, AU, depending on the specific goals of the calculation. For comparison, the radius of Earth’s orbit per definition is exactly 1 AU. In most cases the sink radius is chosen such that the Jeans scale below the threshold density is sufficiently resolved.\(^ {324, 325}\) There are, however, implementations where sink particles have radii equivalent to one cell only. Because the interior of the control volume is not accessible, the physical interpretation is often very difficult and subject to debate. Usually sink particles are thought to represent individual protostars or dense binary systems. This is supported by detailed one-dimensional implicit radiation hydrodynamic calculations which demonstrate that a protostar will build up in the very center of the control volume about 10\(^ 3\) yr after sink creation\(^ {327}\) which will swallow most of the infalling material.

Protostellar collapse is accompanied by a substantial loss of specific angular momentum, even in the absence of magnetic braking.\(^ {328, 329}\) Still, most of the matter that falls in will assemble in a protostellar disk. It is then transported inward by torques from magnetorotational and possibly gravitational instabilities.\(^ {330–338}\) With typical disk sizes of order of several hundred AU in simulations of the formation of star clusters, the control volume fully encloses both star and disk. Even in higher resolution calculations that focus on single cores, the control volume contains the inner part of the accretion disk. If low angular momentum material is accreted, the disk is stable and most of the material ends up in the central star. In this case, the disk simply acts as a buffer and smooths eventual accretion spikes. It will not delay or prevent the mass growth of the central star by much. However, if material that falls into the control volume carries large specific angular momentum, then the mass load onto the disk is likely to be faster than inward transport. The disk grows large and may become gravitationally unstable and fragment. This may lead to the formation of a binary or higher-order multiple.\(^ {330, 336}\) Indeed an initial binary fraction of almost 100% is consistent with observations of star clusters.\(^ {340}\) To some degree this can be taken into account by introducing an appropriate scaling factor. In a cluster environment the protostellar disk may be truncated by tidal interactions and loose matter.\(^ {341, 342}\) The importance of this effect depends strongly on the stellar density of the cluster and its dynamical evolution. Further uncertainty stems from the possible formation of O or B stars in the stellar cluster. Their intense UV radiation will trigger evaporation and gas removal, again limiting the fraction of sink particle mass that turns into stars. Similar holds for stellar winds and outflows. These feedback processes are discussed in Section 4 below.

### 4. THE IMPORTANCE OF FEEDBACK

#### 4.1. Feedback Processes

Most star formation simulations to date have neglected the effects of feedback on the star formation process. While this is computationally simpler, it is clearly not physically correct, and its omission leads to a number of obvious differences between simulations and observations. For example, simulations without feedback produce star formation that is too rapid and efficient\(^ {281, 284}\) and significantly overproduce brown dwarfs compared to observations.\(^ {186, 280, 285}\) To make progress the next generation of simulations will have to remedy this omission.

#### 4.1.1. Radiation Feedback

We begin our consideration of feedback processes by examining the effects of radiation from young stars. It is convenient to distinguish three distinct types of radiative feedback on star formation. The dominant sources of radiation in forming star clusters...
are the young massive stars, which begin their lives accreting rapidly, producing high accretion luminosities from the initial infall onto their surfaces. Later on in their formation these stars radiate prodigiously via Kelvin-Helmholtz contraction and then nuclear burning. The first effect of the radiation they produce is on their immediate environs. Their starlight is absorbed by dust grains suspended in the circumstellar gas, exerting a pressure that opposes gravity. Second, as the radiation diffuses out of the dusty gas clouds around a massive star it heats the gas. This affects the process of fragmentation, and thus plays a role in determining the stellar mass function for all stars born in strongly irradiated environments. Third, once massive stars contract onto the main sequence, they become significant sources of ultraviolet radiation, which can dissociate molecules, ionize atoms, and drive strong shocks throughout the star-forming cloud. These processes can both inhibit and promote star formation.

**Radiation Pressure in Massive Star Formation.** The first effect is perhaps the best studied, and has been the subject of several recent reviews.\(^9\)–\(^11\),\(^343\) so we skip over it relatively quickly. As early as the 1970s researchers considering the formation of massive stars realized a fundamental problem. The largest stars in nearby galaxies have masses \(\sim 100–150 \, M_\odot\)\(^344\)–\(^346\) and for these stars radiation pressure on electrons within the star is the dominant support mechanism. In effect, these stars are at their internal Eddington limit. However, the Thompson cross-section is smaller than the cross-section of dusty gas to stellar radiation reprocessed into the infrared by an order of magnitude. Thus if a massive star is at its Eddington limit with respect to the ionized, dust-free gas in its interior, it must exceed the Eddington limit by an order of magnitude with respect to the dusty gas found in molecular clouds. How then is it possible for dusty gas to accrete and form a massive star, since the outward radiation force on the accreting material should be significantly stronger than the pull of gravity\(^345\)–\(^349\).

Analytic treatments of the problem suggest that the solution lies in the non-sphericity of the accretion process: if the dusty gas around a protostar is sufficiently opaque, it can collimate the radiation, reducing the radiation force over some fraction of the solid angle to the point where gravity is stronger and accretion can occur.\(^350\)–\(^353\) Simulations appear to bear out this solution, at least preliminarily. Hydrodynamic simulations in two dimensions using a flux-limited diffusion approach (see below) are able to form stars up to \(40 \, M_\odot\) before radiation pressure reverses infall,\(^354\) while three-dimensional simulations show no signs of a limit on the upper masses of stars imposed by radiation pressure.\(^355\),\(^356\) In both the 2-D and 3-D cases, radiation is strongly beamed toward the poles of an accretion disk, allowing gas to accrete through parts of the equatorial plane shielded by the disk. In 3-D, this self-shielding is further enhanced by the organization of the gas into opaque, dense filaments, while radiation escapes through optically thin channels. This effect appears to allow the formation of stars with no clear upper mass limit.

**Radiation Heating and the IMF.** The second radiative effect is heating of the gas, with the resulting modification of the initial mass function. Increasing the gas temperature suppresses fragmentation, and the observed overproduction of brown dwarfs in isothermal simulations\(^267\) is at least in part due to their omission of radiative feedback.\(^357\) Similarly, radiative feedback is a strong candidate solution to another mystery about massive star formation: why would \(\sim 100 \, M_\odot\) of gas, a mass that represents
tens to hundreds of Jeans masses at the typical densities and temperatures of a molecular cloud that has not yet begun to collapse, ever collapse coherently rather than fragmenting into many objects\(^269\),\(^278\). A possible answer is that the accretion luminosity produced by the collapse of a dense core in a massive star-forming region is sufficient to suppress a high level of fragmentation, converting a collapse that might have produces \(\sim 100\) small stars into one that produces only a few massive ones.\(^280\),\(^282\) However, the overall importance of this process and its details are subject to ongoing debate, and clearly more work is required on this important subject.\(^358\)

The simulations of how radiative heating affects fragmentation that have been published to date\(^284\)–\(^285\),\(^356\) all confirm the analytically-predicted outcome. Radiative heating reduces the amount of fragmentation that occurs during the collapse of massive pre-stellar cores. Figure 9 shows an example of this effect, comparing two simulations that are identical in every respect except that one is done with radiative transfer and one is done without it. However, none of the simulations published to date have formed enough objects to produce a well-sampled mass function. Radiative feedback does clearly alleviate the problem of overproduction of brown dwarfs,\(^283\),\(^285\) but it is not yet known whether simulations with radiative feedback will naturally produce the observed IMF, or if further physical processes must be included. However, it does seem clear that any results derived from simulations using the isothermal or optically-thin cooling approximations must be regarded as likely subject to over-fragmentation.

**High Energy Radiation and Star Formation Efficiency.** The third form of radiative feedback from massive stars is high energy radiation that is capable of dissociating hydrogen molecules (photon energies above 11 eV) and ionizing hydrogen atoms (photon energies above 13.6 eV). The former creates a photodissociation region (PDR), a volume of mixed atomic and molecular gas at temperatures of hundreds of Kelvin, too warm to form stars. The latter rapidly heats the gas around a massive star to \(\sim 10^4\) K and raises its sound speed to \(\sim 10\) km s\(^{-1}\), forming a structure known as an H\(_\text{ii}\) region. Except in the case of stars with very weak ionizing fluxes, or in environments where the magnetic field strongly confines the ionized region, the shock front generated by an expanding H\(_\text{ii}\) region generally overruns the PDR created by dissociating radiation and traps it between the ionization front and the shock front. For the purpose of molecular cloud dynamics, therefore, ionizing radiation is usually the more important effect.\(^359\)

Unlike radiative heating and radiation pressure, dissociating and ionizing feedback do not become significant until fairly late in the star formation process. Early on rapid accretion swells massive stars to radii of \(\sim 100 \, R_\odot\)\(^359\),\(^360\) and these large radii lead to low surface temperatures, reducing the fraction of a massive star’s power that emerges at energies above 13.6 eV. Moreover, even the full main sequence ionizing luminosity from a massive star will not escape from the stellar vicinity if accretion onto the star is sufficiently rapid and covers enough of the stellar surface.\(^286\),\(^298\),\(^361–364\) In this quasi-spherical case, the H\(_\text{ii}\) region is kept from expanding and the Strömgren radius is small. However, the situation may change once protostellar outflows are taken into account (see next Section 4.1.2). These effectively remove high-density material along the rotational axis of the system. This may lead to an H\(_\text{ii}\) region that escapes along the outflow axis while
remaining confined in the equatorial direction, or it may allow the H\textsc{i} region the break free entirely from its parent core.

Once ionization does begin to break out of a massive protostellar core, it is likely to be the most significant of the three types of radiative feedback. Since 10 km s\textsuperscript{-1} is much greater than the escape speed from a molecular cloud under Milky Way conditions, ionized gas escapes from star-forming clouds into the ISM, reducing the amount of mass available for star formation and unbinding molecular clouds. Furthermore, since 10 km s\textsuperscript{-1} is much larger than the sound speed in the non-ionized molecular gas, once they form H\textsc{i} regions expand dynamically, driving shocks into the neutral material. Analytic models suggest that this can both promote star formation, by sweeping up gas into sheets that subsequently fragment by gravitational instability, and inhibit it, by driving turbulent motions. Clearly more work is required to determine which effect dominates.

Simulations of these processes are still quite primitive, and most have focused on small molecular clouds that are already in a process of free-fall collapse when the simulation begins. Within this limited context simulations have produced a number of qualitative conclusions. First, single ionizing sources at the molecular cloud centers do not easily unbind those clouds, even if they deposit an amount of energy larger than the cloud binding energy. This is because in a cloud with a pre-existing density structures, most of the energy is deposited in low-density gas that freely escapes from the cloud, while the higher density material is largely unaffected. Thus, in this context the effects of ionization on reducing the star formation efficiency are modest. Second, ionization can drive significant velocity dispersions in neutral gas, possibly generating turbulence. Third, ionization does sweep up material and promote collapse, but numerical simulations indicate that this effect may again be modest. Much of the swept up gas in these calculations was already on its way to star formation due to gravity alone, and the compression produced by an H\textsc{i} region shock only modifies this slightly.

This work is only a beginning, and many questions remain. First, none of the simulations to date have included multiple ionizing sources that are simultaneously active, so that interactions between expanding H\textsc{i} region shells can promote both star formation and turbulence. Since massive stars form in clusters, however, multiple sources should be the rule rather than the exception. Second, the simulations have for the most part focused on small, tightly-bound proto-cluster gas clouds being ionized by rather small ionizing luminosities corresponding to single stars, rather than larger, lower density, more loosely bound molecular clouds subjected to the ionizing flux of an entire star cluster. The effects of ionizing radiation may be greater in the latter case than in the former. Finally, only two simulations of ionizing radiation feedback to date have included magnetic fields and only then in a very idealized context. Since magnetic fields can tie together high- and low-density regions of a cloud, they may significantly increase the effects of ionization feedback.

4.1.2. Protostellar Outflow Feedback

Outflows from young stars provide another significant source of feedback on local scales in star-forming regions. During the process of accretion onto young stars about 10% of the gas that reaches the inner accretion disk is ejected into a collimated wind that is launched at a speed comparable to the Keplerian speed close to the stellar surface. Theoretical predictions and observational data agree very well on this value. Another quantity that is well constrained by observations is the net momentum flux of the material entrained in the outflow. It is typically \( p \sim 0.3 M v_K \), where \( M \) is the accretion rate onto the star plus disk and \( v_K \) is the Keplerian velocity at the stellar surface. Outflow momentum flux correlates well with source luminosity across a very wide range in luminosity, suggesting that all protostars show a common wind launching mechanism independent of mass. Since the wind momentum flux is much greater than \( L/c \), this mechanism is almost certainly hydromagnetic rather than radiative in nature. The two primary
theories for this are the x-wind\cite{78} and the disk wind.\cite{379,380,382} Common to both models is the idea that matter gets loaded onto magnetic field lines and then accelerated outwards by centrifugal forces. When considering feedback on scales large compared to the accretion disk around the source, the details of how the wind is launched matter little. All magneto-centrifugal winds approach the same distribution of momentum flux per unit angle at large distances from the launching region.\cite{386} On larger scales, on which the outflow interacts with ambient material in the core, the opening angle varies depending on mass and age of the protostar. Outflows from low-mass stars appear quite well collimated, and remain so up to roughly B stars\cite{384,387,388}. The opening angles of O star outflows are wider, but it is unclear if this widening is an inherent property of the outflow or a result of the interaction between the outflow and the ambient gas.

One important difference between outflow and radiation feedback is that outflow feedback is more democratic. The most massive stars in a cluster dominate its radiative output, because (except at the very highest stellar masses) luminosity is a very strong function of mass, and ionizing luminosity an even stronger one.\cite{389} The dependence of luminosity on mass is strong enough to overcome the relative dearth of massive stars compared to low-mass ones. For outflows the reverse is true. The total mass accretion rate onto all the stars in a cluster is necessarily dominated by the low-mass stars, since they comprise the bulk of the stellar mass once star formation is complete. The Keplerian velocity at a star’s surface varies as $\sqrt{M/R}$, where $M$ and $R$ are the star’s mass and radius, and this ratio is only a very weak function of mass for main sequence stars. Thus, we might expect low-mass stars near the peak of the IMF to dominate outflow feedback. This simple analysis neglects the effect that more massive stars have shorter Kelvin-Helmholtz times and thus reach smaller radii more rapidly than low-mass stars, giving them larger Keplarian velocities at earlier times. Including this effect in a more careful analysis suggests that each logarithmic bin in mass contributes roughly equally to the total amount of momentum injected into a cluster.\cite{390} This has two important consequences: first, it means that outflow feedback can be important even in small clumps that do not form massive stars. Second, it means that simulations of outflow feedback cannot focus exclusively on the most massive stars, but must instead consider all stars as sources.

Outflows can influence their immediate surroundings as well as the cluster in which they form. On small scales, they reduce the star formation efficiency by removing mass from a collapsing core, both directly and via material that the outflow entrains as it escapes the core. Analytic estimates suggest that this process removes 25–75% of the mass in a core,\cite{261} but this is highly uncertain since no simulations of the collapse of individual cores with outflows that are capable of evaluating this estimate have been published.

Outflows could also modify the star formation process by removing mass from protocluster gas clumps and possibly by driving turbulence within them.\cite{261,391,392} However, it must be noted that this process cannot be the main driver of turbulence on global molecular cloud scales as outflows typically have short length scales. Instead this turbulence is probably driven on scales comparable to or larger than typical cloud sizes.\cite{223,225,394,395} One possible candidate for the origin of this large-scale turbulence is convergent flows in the galactic disk (see Section 2.2.2) either driven by gravitational instability in the disk,\cite{85,84} by collisions between molecular clouds,\cite{86} or caused by supernova explosions.\cite{396,397} Another candidate is giant H\textsc{ii} regions created by clustered star formation within clouds, which have size scales comparable to entire GMCs.\cite{54,134,370} In addition, there seems to be no difference between the measured turbulence content of cloud clumps that are still in the so-called dark phase, i.e., before star formation has set in, and cloud clumps that are already actively building up stars in their interior.\cite{32} This indicates that, at least at birth, star-forming regions must have turbulent motions that were imprinted as part of the formation process. Conversely, however, observations show that high column density star-forming clumps within GMCs lie above the linewidth-size relation observed for GMCs as a whole.\cite{32,49,71} This suggests that their turbulence cannot be supplied from large scales motions within the parent GMCs. Either these regions are powered by gravitational collapse, in a scaled-down version of the scenario described in Section 2.2.2, or they are driven by internal sources.

Outflows are a natural candidate for this, and the deviation from a simple powerlaw linewidth-size relation predicted by analytic models appears to be consistent with what is observed.\cite{393}

Simulations of protostellar outflows to date fall into two categories. Local simulations focus on the interaction of a single outflow with an ambient medium at high resolution, while larger-scale simulations follow an entire gas clump and star cluster including multiple outflows, but at significantly lower resolution. Local simulations attempt to understand the driving of turbulence by single outflows in detail. However, interpretation of these results is difficult, since there is no simple way to separate “turbulence” from the coherent motion caused by a single outflow. Different authors analyzing simulations in different ways have come to opposite conclusions, with some arguing that outflows cannot drive supersonic turbulence,\cite{398} while others conclude that it can.\cite{399,400}

Global simulations of outflow feedback generally find that it has strong effects on the cluster formation process. In the absence of energy input, simulations of cluster-forming gas clumps find that any turbulence initially present decays rapidly, leading to a global collapse in which an appreciable fraction of the mass is converted into stars within a few dynamical times.\cite{267,271,401,403} Simulations that include outflow feedback found that outflows can change this picture. They eject mass from the densest and most actively star-forming parts of a cluster, reducing the star formation rate, while at the same time injecting enough energy to slow down overall collapse and maintain a constant level of turbulent motions.\cite{206,404,405} As a result, the star formation rate drops to < 10% of the mass being converted into stars per free-fall time, and rather than undergoing a runaway collapse the clump reaches a slowly evolving quasi-equilibrium state. Figure 10 illustrates this effect in a simulation of the formation of a star cluster including protostellar outflow feedback. At the start of the calculation the kinetic energy falls as the initial turbulent velocity field decays. Consequently the cloud contracts and at about half of a free-fall time stars start to form and drive outflows. This energy input changes the subsequent evolution, and the cloud’s global contraction is halted or at least significantly retarded. A definite answer would require to follow the dynamics over a longer period in time.

All the simulations published to date have significant limits. With only one exception they treat the wind as an instantaneous explosion, rather than a continuous beam injected over $\sim 10^3$ yr as we observe. The individual explosion spikes are clearly visible in Figure 10. The current studies are also characterized by
low numerical resolution (128$^3$ cells), which makes the energy injected more space filling, which is one of the main characteristics of interstellar turbulence. In addition, the calculations are performed in a periodic box, so energy cannot leave the star-forming cloud. In reality some very strong pencil-beam outflows escape their parent clumps and cover distances of a few parsec.\cite{408} The one simulation published thus far that does include time history of accretion and better resolution only considers the effects of outflows from stars larger than 10 $M_\odot$, thereby neglecting the majority of the outflow power. Clearly the problem of outflow feedback and how it affects star formation is in need of further study.

4.1.3. Other Types of Feedback

Although radiation and protostellar outflows are thought to be the dominant feedback processes in star formation, two other mechanisms are worthy of brief discussion: supernovae, and winds ejected by stars on the main sequence and post-main sequence. In terms of sheer energetics, it might seem odd to ignore supernovae as a major source of feedback. However, two compensating effects reduce their role in regulating star forming clouds. The first is timescales. Even the most massive stars do not explode as supernovae until 3–4 Myr after formation,\cite{409} and this is comparable to or longer than the formation time of star clusters. Thus supernovae come too late to affect the formation of individual star clusters, although they may be able to affect their parent giant molecular clouds, which have longer lifetimes.

A second effect, however, mitigates the impact of supernovae on GMC scales as well. Supernovae occur only after H ii regions and stellar winds have carved large cavities of hot, ionized gas around the massive stars that produce them. If a supernova occurs while this bubble is still embedded within its parent cloud, much of its energy is radiated away while the blast wave is still confined to the bubble. Simulations find that, as a result, the mass that is removed from the cloud by a supernova plus ionization is typically only $\sim$10% larger than that removed by ionization alone.\cite{54,370,410,411} If, on the other hand, the bubble of hot gas created by ionization has broken out of a massive star’s parent cloud by the time the star explodes, both distance and an impedance mismatch make it difficult to deliver much of the supernova energy to the cloud. In a few Myr the expanding H ii region around a massive star cluster can push back the parent GMC by $\sim$10 pc or more from the site of the supernova, which then occurs in an ionized medium whose density is 2–4 orders of magnitude lower than that of the cloud. Both the distance and the large density jump serve to shield a molecular cloud from the effects of a supernova, so that very little of the supernova energy is deposited in the molecular cloud. This effect means that, even on GMC scales, supernovae are unlikely to be the dominant feedback mechanism. It is important to note, however, that the energy that is not deposited in the molecular cloud itself does nevertheless affect the remainder of the ISM on galactic scales. Consequently supernovae are likely to dominate the energetics of the ISM on large scales.\cite{4}

Main sequence winds are also thought to be subdominant as feedback mechanisms due to the effects of ionization.\cite{370,412} Stellar winds initially expand into a bubble of ionized gas created by the ionization from a massive star, and they create a radiative shock within that bubble where much of the wind energy is dissipated. Only after the stellar wind shock reaches up to the ionization-created shock can stellar winds begin to provide feedback to the parent cloud. Even then the increase in total kinetic energy in the shock is modest for stars up to at least 35 $M_\odot$\cite{413} and even for 60 $M_\odot$ stars is only of order unity.\cite{414}

4.2. Modeling Feedback

4.2.1. Numerical Methods for Non-Ionizing Radiation Feedback

Stars emit the bulk of their radiation in the visible part of the spectrum, but the dusty clouds in which stars form are generally very opaque to visible light until late in the star formation process, when most of the gas has already been accreted or dispersed. As a result, direct stellar radiation tends to be absorbed by dust and reprocessed into the infrared close to the star that emits it, and modeling the resulting diffuse infrared radiation field is the primary goal of most numerical methods for simulating non-ionizing radiation feedback.

Focusing on the diffuse infrared radiation field simplifies the radiative transfer problem considerably, since the primary opacity source at infrared wavelengths is dust rather than atomic or molecular lines, and because in the IR scattering is negligible compared to absorption.\cite{415} Even with these simplifications, though, it is possible to solve the full equations of (magneto)-hydrodynamics plus the equation of radiative transfer for this problem only in one dimension.\cite{416,418} Such an approach is unfortunately too computationally expensive to be feasible in three or even two dimensions. Instead, one must simplify the problem even further.

One approach is simply to modify the standard optically-thin cooling curve used in simulations without feedback by using an approximation to estimate the optical depth and reduce the cooling rate appropriately.\cite{419,420} In this case one need not to solve a radiative transfer problem at all. This is an advantage, since it means that the radiation step has nearly zero computational cost, but it is also a limitation. Because it lacks a treatment of radiative transfer, this approach allows gas to heat up due to adiabatic
compression, but not because it is being illuminated by an external radiation source. In particular, in this approach there is no way for stars to heat gas. Since stellar radiation provides significantly more energy than gravitational compression once the first collapsed objects form,\textsuperscript{201, 206–208, 351} this technique is only suitable for simulating star formation up to the point when the first parcels of gas collapse to stars.

The most common and simplest approach that can go past first collapse and follow either accretion or subsequent star formation, and the only one used so far in "production" simulations,\textsuperscript{201, 351, 352, 421, 422} is the flux-limited diffusion approximation.\textsuperscript{221, 424} The underlying physical idea is simple: in an optically thick environment like the dusty clouds in which stars form, radiation diffuses through the gas like heat, and the radiative flux $\mathbf{F}$ obeys Fick’s Law: $\mathbf{F} = -\kappa E/(3\kappa \rho)$, where $E$ is the radiation energy density, $\rho$ is the gas density, and $\kappa$ is the specific opacity, with units of area divided by mass. This is the standard diffusion approximation, and it can be made either in a gray form by integrating $E$ and $\mathbf{F}$ over all frequencies, or in a multi-group form in which one divides the spectrum into some number of intervals in frequency and computes a separate energy density and flux for each interval.\textsuperscript{10, 354, 425}

The pure diffusion approximation encounters a problem when the opacity is low, since if $\kappa \rho$ is sufficiently small the flux can exceed $c E$, violating the constraints of special relativity. The flux-limiting approach is to solve this problem by modifying the law for the flux to $\mathbf{F} = -\lambda \kappa E/(\kappa \rho)$, where $\lambda$ is the flux limiter, a dimensionless function of $E$ and $\kappa \rho$ that has the properties $\lambda \to 1/3$ when the gas is optically thick, and $\lambda \to \kappa E/|\mathbf{F}|$ when it is optically thin. This limiting behavior ensures that flux approaches the correct Fick’s Law value when the optical depth is high, and correctly reaches a maximum magnitude of $c E$ at low optical depth. Many functional forms are possible for $\lambda$. The most commonly-used one is the Levermore & Pomraning limiter $\lambda = R^{-1} \coth(R - R^{-1})$, with $R = |\mathbf{E}|/(\kappa \rho E)^{224, 426}$.

Given a formula for computing the radiation flux in terms of the radiation energy density, it is possible to drop all moments of the equation of radiative transfer except the zeroth one, so that the set of equations to be solved consists of the standard equations of HD or MHD, with some added terms describing the interaction of radiation with the gas, plus one additional equation for the radiation energy density. One treats feedback from stars in this formulation simply by adding it as a source term or a boundary condition in the radiation energy equation. The resulting set of equations may be written using either a comoving\textsuperscript{205, 427–431} or a mixed-frame\textsuperscript{307, 425, 432} formulation. The former approach is more suited to implementation in a code that is either Lagrangian, such as SPH, or based on van Leer advection,\textsuperscript{433} but has the disadvantage that the equations are not explicitly conservative, and so the resulting codes cannot precisely conserve energy. The mixed-frame equations, on the other hand, are explicitly conservative, which makes them preferable for codes based on a conservative update, particularly those involving adaptive mesh refinement. In either form the equations are still significantly more expensive to solve than the corresponding non-radiative ones, since codes must handle radiative diffusion implicitly in order to avoid severe constraints on the time step, but solutions are within the reach of modern supercomputer simulations.

Although it is the tool of choice for star formation simulations at present, the pure flux-limited diffusion approximation does have some important limitations. Diffusion methods do not correctly represent shadowing effects which appear in systems that are optically thin or nearly so, nor can they model direct stellar radiation before it is absorbed and re-radiated isotropically. Pure diffusion also assumes that the gas and dust are thermally well-coupled. While this approximation is a good one at densities $\sim 10^4$ cm$^{-3}$ or more, it may fail at lower densities. Diffusion also neglects cooling via molecular line emission, which can also be important at lower densities.\textsuperscript{354}

The literature contains a variety of numerical techniques to address these shortfalls. One can handle imperfect dust-gas coupling by explicitly including it in the iterative radiative transfer update.\textsuperscript{431} To handle direct stellar radiation or molecular cooling as well as the diffuse IR field, one can use a hybrid approach that combines a diffusion step with a ray-tracing step\textsuperscript{435} or an optically-thin cooling step. To correctly model shadowing, one can use a more sophisticated radiative transfer method than diffusion, such as Monte Carlo, ray-tracing,\textsuperscript{436} variable tensor Eddington factor (VTEF),\textsuperscript{437} or $S_n$ transport.\textsuperscript{438} However, with the exception of the dust-gas coupling method, none of these techniques have thus far been used in any "production" simulations of star formation. In some cases this is simply a matter of the necessary techniques not yet having been implemented into the codes most commonly used for star formation studies. These techniques, such as two-step approaches for the diffuse IR field and direct stellar fields and line radiation, are likely to appear in production simulations in the next few years. In other cases, however, the limitation is one of computational expense. For example the VTEF and $S_n$ methods have thus far only been used in two-dimensional calculations, simply because in three dimensions they have proven prohibitively expensive. Remedyng these problems will require significant advances in radiative transfer methodology to solve.

### 4.2.2. Numerical Methods for Ionizing Radiation

In comparison to non-ionizing radiation, handling ionizing radiation is conceptually more straightforward. Rather than a diffuse field arising from the repeated reprocessing of stellar radiation by dust grains, the ionizing radiation field in a star-forming region consists mostly of photons directly emitted from a stellar surface. Only in the outer parts of low-density ionized regions with sharp density gradients does reprocessed radiation make up a significant part of the photon field.\textsuperscript{439} The dominance of a relatively small number of point sources of radiation translates into a much simpler computational problem. By far the most common approach for solving it is to adopt the on-the-spot approximation,\textsuperscript{440} in which one assumes that recombinations of ionized atoms into the ground state yield ionizing photons that are re-absorbed immediately near the point of emission. One therefore ignores photons emitted by recombinining atoms entirely, and one solves the transfer equation along rays from the emitting stellar source, balancing recombinations into excited states against ionizations along each ray.

Within this over-arching framework, there are a variety of subtleties about how one draws the rays and updates the gas state. For example, the ray-drawing procedure can range in complexity from grids of rays restricted to radial paths originating at the center of a spherical grid,\textsuperscript{388} up to a variety of schemes for handling casting rays either with a fixed\textsuperscript{172, 173, 422, 446} or adaptive\textsuperscript{307, 444–446} ray grid, or through a field of SPH particles.\textsuperscript{447–449} Similarly,
there are a variety of possible time-stepping strategies for handling the interaction of radiation heating with (magneto-) hydrodynamics. The simplest are Strömgren volume methods, in which one assumes that the gas reaches radiation and thermal equilibrium instantaneously. Solving time-dependent equations for the thermal and chemical structure but not resolving the relevant timescales hydrodynamically represents a middle ground, while the most complex option is to restrict the hydrodynamic time step to resolve gas heating and cooling times. As always in numerics, there is a trade-off between speed and quality of solution. Fully resolving the ionization heating time produces measurably more accurate solutions, but is of course significantly more expensive than resolving it only marginally or assuming instantaneous equilibration.

4.2.3. Numerical Methods for Protostellar Outflows

Protostellar outflows are a natural result of the process of collapse and accretion that produces stars, and simulations of star formation that treat MHD and gravity with sufficient resolution do not need to include any additional physics to produce outflows. However, sufficient resolution here means that the simulation must resolve the outflowing region, which is typically no more than ~10 stellar radii. Achieving such high resolution is prohibitively expensive for simulations that span more than a tiny fraction of the total formation time of a star, let alone a whole star cluster, so the most common procedure in such simulations is to use radiative feedback. Replace the collapsing region with a sink of some sort, and model an outflow emerging from that sink via a subgrid model that sits on top of whatever sink model the computation uses. Such a model for outflows must specify the amount of mass and momentum contained, as well as the angular distribution of these quantities. Of these quantities the momentum is the best constrained by observations, since it remains unchanged even as the outflow gas entrains the material it encounters in the protostellar core. The mass flux and the angular distribution of the outflow are less well constrained, since these are altered as the outflow ages and interacts with its environment. The correct value to use in a given simulation probably depends on the length scale that simulation resolves, since the mass and opening angle both increase as outflowing gas moves away from the star and interacts with its environment. Most simulations assume the standard value of 10% for the ratio of infalling to outflowing mass. A good approximation for the opening angle is to assume the momentum of the outflowing gas is distributed with a profile \( p \propto 1/(r \sin \theta)^2 \), where \( r \) is the distance from the star and \( \theta \) is the angle relative to the star’s axis of rotation. The opening angle adopted in numerical simulations, however, varies enormously all the way from 0° to 90°.

Once one has chosen a physical model for the outflow mass, momentum, and angular distribution in a given simulation, there remains the question of how to implement it numerically. As noted above, all stars contribute significant amounts of momentum, so any realistic approach must include contributions from any star that forms in a simulation. In grid codes this problem is generally straightforward; one simply adds mass and momentum with the desired angular distribution to the computational cells in or immediately around the sink region for each star. In particle-based codes the problem is more complex, since one must take care to avoid artificial clumping due to the discrete nature of the particles, and to ensure the momentum deposition in the region surrounding the sink is not altered by numerical interpenetration of the SPH particles. A variety of strategies are available to solve these problems, but they are computationally expensive, which presents a potential problem for simulations with large numbers of sources.

Once the subgrid model is in place, outflows are much easier to simulate than radiation feedback, because outflow evolution is governed solely by hydrodynamics or MHD plus gravity, the physical mechanisms that are already included in any code used to simulate star formation. Beyond those involved in computing the subgrid model and injecting the outflow, the only additional computational cost that outflows impose on a code comes from the fact that outflow velocities can reach hundreds of km s\(^{-1}\), significantly greater than the \( \lesssim 10 \) km s\(^{-1}\) turbulent or infall speeds typically found in simulations that omit outflows. The higher speeds require smaller simulation time steps, and a corresponding increase in computational cost.

5. SUMMARY AND OUTLOOK

In this review we presented an overview of the current state of numerical star-formation studies. We have restricted ourselves to the early phases of stellar birth, from the formation of molecular clouds through the build-up of stars and star clusters in their interior. We have left out the problem of accretion disks and protostellar evolution and point to other reviews in this context.

We hope we have illustrated that the question of stellar birth in our Galaxy and elsewhere in the universe is far from being solved. Instead the field is rapidly evolving and has gone through a significant transformation in the last few years. In numerical star formation studies, we notice a general trend away from solely considering isolated processes and phenomena towards a more integrated multi-scale and multi-physics approach in today’s computer simulations. In part this is triggered by the growing awareness that many physical processes contribute more or less equally to the formation of stars, such that it is not possible to single out individual effects. Reliable and quantitative predictions can only be made on the basis of taking all relevant physical phenomena into account. Another reason for this development is the tremendous increase in computational capability provided by the advent of (relatively) easy-to-handle massively-parallel supercomputers, coupled with new and more efficient numerical algorithms for these machines.

If we examine the past and current state of the art, then it is evident that most studies so far have focussed on a small number of physical processes only. Typically, one had a single question in mind, such as what happens if we include one particular physical process? How does it affect the system? How does it modify possible equilibrium states? And how does it influence the dynamical evolution if we apply perturbations? The processes and phenomena about which these questions have been asked include hydrodynamics, turbulence, gravitational dynamics, magnetic fields, nonequilibrium chemistry, and the interaction of radiation with matter. Typically only one or two of them have been included in any given simulation. More sophisticated approaches include larger numbers of processes, but no simulation so far has considered all of them. The challenge in the past was mainly to do justice to the inherent multidimensionality.
of the considered problems. For example, stellar birth in turbulent interstellar gas clouds with highly complex spatial and kinematical structure is an intrinsically three-dimensional problem with one- or two-dimensional approaches at best providing order-of-magnitude estimates. Including multiple physical processes gave way before the challenge of simulating in three dimensions.

This era is coming to an end. Many of today’s most challenging problems are multi-physics, in the sense that they require the combination of many (if not all) of the above-mentioned processes, and multi-scale, in the sense that unresolved microscopic processes can feed back onto macroscopic scales. This is true not only for star formation studies, but applies to virtually all fields of modern astrophysics. For example, the coagulation of dust species to larger particles or the interaction of dust with the radiation field from the central stars will eventually feedback into the dynamical behavior of the gas in protostellar accretion disks and hence has severe consequences for the formation and mass growths of planetary systems. Similarly, star formation and baryonic feedback are crucial ingredients of understanding galaxy formation and evolution in cosmological models. In a realistic description of cosmic phenomena, one is faced with the highly non-linear coupling between quite different kinds of interactions on a variety of scales. Star formation is no exception.

This is not only a challenge, it is also a chance, because it may open up new pathways to successful collaborations across astrophysical disciplines. It also reaches out to scientists in neighboring fields, such as applied mathematics or computer science. For example, only few groups around the world are able to fully benefit from the massively parallel computing architectures that are currently being developed. Peak performances with $\sim 100$ teraflops will only be attainable on thousands of CPUs, sustained petaflop computing may require as many as $10^5$ CPUs. This asks for a completely new approach to parallel algorithm design, a field where modern computer science is far ahead of the schemes currently used in astrophysics and star-formation studies. Regular methodological exchange with applied mathematicians and possibly numerical fluid dynamicists thus holds the promise of both transferring new methods into astrophysics and raising the awareness of mathematicians about numerical challenges in astrophysics.

Towards the end, we want to speculate about a few of the what we think are the most interesting and pressing open problems in modern star formation theory and where the current advances in computational power and algorithmic sophistication are likely to have a major impact.

**What drives interstellar turbulence?** Observations show that turbulence in molecular clouds is ubiquitous, and that, with the exception of the dense cores discussed above, it seems to follow a universal relationship between velocity dispersion and size. Even extragalactic molecular clouds appear to obey similar scalings. There are no variations in the turbulent properties between GMCs with little and much star formation, which might seem to argue for galaxy-scale driving, but there is also no systematic variation in GMC properties within a galaxy or between galaxies, which would seem to argue that internal processes must be important. So what is the relative importance of internal and external forcing mechanisms in driving ISM turbulence? Does the answer depend on the length scales that one examines, or on the place where one looks?

**How does the multi-phase nature of the ISM influence stellar birth?** Star formation appears to follow fairly universal scaling laws in galaxies that range from mildly H$_2$-dominated (such as the Milky Way) to galaxies that are strongly H$_2$-dominated (such as local starbursts). Does the presence or absence of a significant atomic phase play an important role in regulating star formation, either directly (e.g., by limiting the amount of molecular gas "eligible" for star formation) or indirectly (e.g., by driving turbulent motions via thermal instability)? How does the star formation process change, if at all, in galaxies such as dwarfs that contain very little molecular gas?

**How does stellar feedback influence star formation?** Stars produce different types of feedback: outflows, main sequence winds, ionizing and non-ionizing radiation, and supernovae. Which, if any of these, are responsible for controlling the rate and efficiency of star formation? Does the answer to this question change in different galactic environments, i.e., are there different processes acting in the denser molecular clouds found in circum-nuclear starbursts than in the tenuous outer regions of the galaxy?

**What determines the statistical properties of a stellar population, and are these properties universal?** On the observational side, is the stellar IMF and binary distribution at present days different in different galactic environments, or is it truly universal? Especially in rich clusters our observational basis still needs to be extended. The same holds for variations with metallicity as can be traced in the Local Group. Is the IMF in the Large Magellanic Cloud (with metal abundances of $\sim 1/3$ of the solar value) and the Small Magellanic Cloud (with $\sim 1/10$ of that value) really similar to the Milky Way? On the theoretical side, what processes are responsible for the (non-)variation of the IMF? The critical mass for gravitational collapse can vary enormously between different environments. Yet the IMF in globular clusters, for example, appears to be the same as in regions of distributed star formation as in Taurus. Hence, there must be additional physical processes that influence the fragmentation behavior of the interstellar gas and determine the resulting stellar mass spectrum.

**Acknowledgments:** We thank all our collaborators for many stimulating, encouraging and sometimes controversial discussions. In particular, we thank Javier Ballesteros-Paredes, Robi Banerjee, Ian A. Bonnell, Paul C. Clark, Bruce G. Elmegreen, Christoph Federrath, Simon C. O. Glover, Lee W. Hartmann, Patrick Hennebelle, Anna-Katharina Jappsen, Richard I. Klein, Kai-Diin M. Kratter, Mordecai-Mark Mac Low, Christopher D. Matzner, Christopher F. McKee, Stella S. R. Offner, Wolfram Schmidt, Jonathan C. Tan, Todd A. Thompson, and Enrique Vázquez-Semadeni. RSK thanks for support from the Germany Science Foundation, DFG, via the Emmy Noether grant KL 1358/1, grants KL 1358/4 and KL 1358/5, and the priority program SFB 439 Galaxies in the Early Universe. MRK acknowledges support from National Science Foundation grant AST-0807739 and from the National Aeronautics and Space Administration through the Spitzer Space Telescope Theoretical Research Program, through a contract issued by the Jet Propulsion Laboratory, California Institute of Technology. FH acknowledges support from National Science Foundation grant AST-0807305 and from the National Aeronautics and Space Administration through the Herschel Cycle-0 Theoretical and Laboratory Astrophysics Research program, NHSC 1008.