Abstract
The current status of both the observational evidence and the theory of the stellar initial mass function (IMF) is reviewed, with particular attention to the two basic, apparently universal features shown by all observations of nearby stellar systems: (1) a characteristic stellar mass of the order of one solar mass, and (2) a power-law decline of the IMF at large masses similar to the original Salpeter law. Considerable evidence and theoretical work supports the hypothesis that the characteristic stellar mass derives from a characteristic scale of fragmentation in star-forming clouds which is essentially the Jeans scale as calculated from the typical temperature and pressure in molecular clouds. The power-law decline of the IMF at large masses suggests that the most massive stars are built up by scale-free accretion or accumulation processes, and the observed formation of these stars in dense clusters and close multiple systems suggests that interactions between dense prestellar clumps or protostars in forming clusters will play a role. A simple model postulating successive mergers of subsystems in a forming cluster accompanied by the accretion of a fraction of the residual gas by the most massive protostar during each merger predicts an upper IMF of power-law form and reproduces the Salpeter law with a plausible assumed accretion efficiency.

1 Introduction
The stellar initial mass function (IMF), or distribution of masses with which stars are formed, is the most fundamental output function of the star formation process, and it controls nearly all aspects of the evolution of stellar systems. The importance of understanding the origin of the IMF and its possible universality has therefore been a stimulus for much research on star formation, both theoretical and observational, and interest in this subject is of long standing, going back at least to the pioneering study by Nakano (1966) of some of the processes that might be responsible for determining the stellar IMF. In recent years there has been much progress in observational studies relating to the IMF, and somewhat more modest progress in reaching a theoretical understanding of its origin; here I review briefly the current status of both the observational evidence and the theoretical ideas concerning the origin IMF. Other recent reviews of the observations and the theory of the IMF have been given by Scalo (1998), Clarke (1998), Larson (1998, 1999), Elmegreen (1999), and Meyer et al. (2000).

2 Basic Observed Features of the Stellar IMF
Numerous observational studies have been carried out to measure or constrain the IMF in systems with as wide a range in properties as possible in order to establish whether it is universal or whether it varies with place or time, depending for example on parameters such as metallicity. The regions that have been studied with direct star counts so far include the local field star population in our Galaxy and many star clusters of various types and metallicities in both our Galaxy and the Magellanic Clouds. As summarized below, this large body of direct evidence does not yet demonstrate convincingly any variability of the IMF, although the uncertainties are still large. Some indirect evidence based on the photometric properties of more distant and exotic systems suggests that the IMF may vary in extreme circumstances, possibly being more top-heavy in starbursts and high-redshift galaxies (Larson 1998). But this indirect evidence is less secure and will not be discussed further here.

As reviewed by Miller & Scalo (1979), Scalo (1986, 1998), Kroupa (1998), and Meyer et al. (2000), the IMF derived for the field stars in the solar neighborhood exhibits an approximate power-law decline with mass above one solar mass that is consistent with, or somewhat steeper than, the original Salpeter (1955) law; however, below one solar mass the IMF of the field stars clearly flattens, showing a possible broad peak at a few tenths of a solar mass in the number of stars per unit logarithmic mass interval. If the logarithmic slope $x$ of the IMF is defined by $dN/d\log m \propto m^{-x}$, then the slope at large masses is $x \sim 1.5$, while the slope at small masses is $x \sim 0$, the range of values or uncertainty in $x$ being about $\pm0.5$ in each case. The IMF inferred for the local field stars is subject to significant uncertainty, especially in the range around one solar mass, because it depends on the assumed evolutionary history of the local Galactic disk and on assumed stellar lifetimes. In contrast, the IMFs of individual star clusters can be derived with fewer assumptions and should be more reliable, since all of the stars in each cluster have the same age and since, at least in the youngest clusters, all of the stars ever formed are still present and can be directly counted as a function of mass without the need for evolutionary corrections. Much effort has therefore gone into determining IMFs for clusters with a wide range of properties in both our Galaxy and the Magellanic Clouds. As reviewed by von Hippel et al. (1996), Hunter et al. (1997), Massey (1998), and Scalo (1998), the results of these studies are generally consistent with the IMF inferred for the local field stars, and the values found for the slope $x$ of the IMF above one solar mass generally scatter around the Salpeter value $x = 1.35$ (see figure 5 of Scalo 1998). In all cases in which it has been possible to observe low-mass stars, the cluster IMFs also show a flattening below one solar mass. No clear evidence has been found for any systematic dependence of the IMF on any property of the systems studied, and this has led to the current widely held view that the IMF is universal, at least in the local universe.

Recent studies have provided more information about very faint stars and brown dwarfs, and the IMF estimated for them remains approximately flat or shows only a moderate decline into the brown dwarf regime, consistent with an extrapolation of the IMF of lower main sequence stars and showing no evidence for any abrupt truncation at low masses (Bassiri & Marcy 1997; Martin et al. 1998; Bouvier et al. 1998; Reid 1998). Another area of recent progress has been the determination of IMFs for a number of newly formed star clusters that still contain many pre-main-sequence stars; as reviewed by Meyer et al. (2000), these results again show general consistency with the field star IMF, including a similar flattening below one solar mass and a possible broad peak at a few tenths of a solar mass. Significant numbers of brown dwarf candidates have been found in these young clusters, and although the derivation of an IMF for them is complicated by the need to know their ages accurately, the results again suggest an IMF that is flat or moderately upward-sloping at the low end (Luhman & Rieke 1999; Hillenbrand & Carpenter 1999).

In summary, within the still rather large uncertainties, all of the data that have been described are consistent with a universal IMF that is nearly flat at low masses and that can be approximated by a declining power law with a slope similar to the original Salpeter slope above one solar mass. The fact that the IMF cannot be approximated by a single power law at all masses but flattens below one solar mass means that there is a characteristic stellar mass of the order of one solar mass such that most of the mass that condenses into stars goes into stars with masses of this order. In fact, a more robust statement about the IMF than any claimed functional form is the fact that about 75% of the mass that forms stars goes into stars with masses between 0.1 and 10 $M_\odot$, while about 20% goes into stars more massive than 10 $M_\odot$ and only 5% into stars less massive than 0.1 $M_\odot$. The existence of this characteristic stellar mass is the most fundamental fact needing to be explained by any theoretical understanding of star formation. The second fundamental fact to be explained is that a significant fraction of the mass goes into massive stars in a power-law tail of the IMF extending to masses much larger than the characteristic mass. Possible theoretical explanations of these two basic facts will be discussed in the following sections.

3 The Origin of the Characteristic Stellar Mass
For some years, there have been two contending viewpoints about the origin of the characteristic stellar mass, one holding that it results from a characteristic mass scale for the frag-
mentation of star-forming clouds (e.g., Larson 1985, 1996), and the other holding that it results from the formation of strong outflows at some stage of protostellar accretion (e.g., Adams & Fatuzzo 1996); both effects might in fact play some role, as reviewed by Meyer et al. (2000). The fragmentation hypothesis for the origin of the characteristic mass has recently received support from observations showing that the ρ Ophiuchus cloud contains many small, apparently pre-stellar clumps with masses between 0.05 and 3 M\(_\odot\), whose properties are consistent with their having been formed by the gravitational fragmentation of the cloud, and whose mass spectrum is very similar to the stellar IMF discussed above, including the flattening below one solar mass (Motte, André, & Neri 1998; see also André et al. 1999; André, Ward-Thompson, & Barsony 2000). In particular, the mass spectrum of the clumps in the ρ Oph cloud is quite similar to the mass spectrum of the young stars observed in this cloud (Luhman & Rieke 1999), suggesting that the IMF of the stars derives directly from the mass spectrum of the clumps. A clump mass spectrum consistent with the stellar IMF has also been found in the Serpens cloud by Testi & Sargent (1998), and additional evidence for a possible mass scale of order one solar mass in the structure of molecular clouds has been reviewed by Evans (1999) and Williams, Blitz, & McKee (2000).

Although the original analysis of Jeans (1929) showing the existence of critical length and mass scales for the fragmentation of a collapsing cloud was not self-consistent, rigorous stability analyses that yield dimensionally equivalent results can be made for various equilibrium configurations, including spherical, disk-like, and filamentary (Spitzer 1951). In these cases, there is a predicted characteristic mass scale for fragmentation that is a few times \(c_s^2/G\rho\), where \(c_s\) is the isothermal sound speed and \(\rho\) is the surface density of the assumed equilibrium configuration. For a typical molecular cloud temperature of 10 K and a typical surface density of 100 M\(_\odot\) pc\(^{-2}\), this mass scale is about one solar mass, similar to the observed typical stellar mass (Larson 1985). Alternatively, if a collapsing pre-stellar clump forms by the fragmentation of equilibrium configurations but as condensations in a medium with some characteristic ambient pressure \(P\), the minimum mass that can collapse gravitationally is that of a marginally stable ‘Bonnor-Ebert’ sphere with a boundary pressure \(P\), or 1.18 \(c_s^2G^2\rho^{1/2}\) (Spitzer 1968). Since any self-gravitating configuration has an internal pressure \(P \sim \pi G\rho_c^2/2\), this result is dimensionally equivalent to the fragmentation scale given above, and it can be regarded as a different expression for the same physically significant quantity, which can still conveniently be called the ‘Jeans mass’.

It is not yet clear to what extent star-forming molecular clouds or their denser subregions can be regarded as equilibrium configurations, and it may instead be that much of the structure in these clouds consists of transient density fluctuations generated by supersonic turbulence (Larson 1981). Some of the filamentary structure in molecular clouds may be created by violent dynamical phenomena in an active star-forming environment (Bally et al. 1991), and simulations of turbulence in the interstellar medium often show the appearance of transient filamentary features that form where supersonic turbulent flows converge (Vázquez-Semadeni et al. 1995, 2000; Ballesteros-Paredes et al. 1999). In the presence of gravity, some of the densest clumps produced in this way may become self-gravitating and begin to collapse; the initial state for their collapse might then be roughly approximated by a marginally stable Bonnor-Ebert sphere whose boundary pressure is determined by the ram pressure of the turbulent flow. If there is a rough balance between turbulent pressure and gravity in molecular clouds, the turbulent pressure will be approximately equal to the gravitational pressure \(P \sim \pi G\rho_c^2/2\), yielding a pressure \(P \sim 3 \times 10^5\) cm\(^{-3}\) K for a typical surface density \(\rho \sim 100\) M\(_\odot\) pc\(^{-2}\). Alternatively, a typical pressure may be estimated by noting that the correlations among linewidth, size, and density that hold among many molecular clouds (Larson 1981; Myers & Goodman 1983) simplify these clouds all to have similar pressures \(P \rho^2\), for which a typical value is again approximately \(3 \times 10^5\) cm\(^{-3}\) K. For a marginally stable Bonnor-Ebert sphere with a temperature of 10 K bounded by this pressure, the predicted mass and radius are about 0.7 M\(_\odot\) and 0.03 pc respectively (Larson 1991, 1996, 1999). Although factors of 2 may not be very meaningful, these quantities are similar in magnitude to the typical masses and sizes of the pre-stellar clumps observed in molecular clouds (e.g., Motte et al. 1998) and to the characteristic stellar mass noted above. Thus there may indeed be an intrinsic mass scale in the star formation process, and this mass scale may be essentially the Jeans mass as defined above.

There are also some hints that there may be a corresponding size scale for star-forming clumps. Analyses of the spatial distributions of the newly formed T Tauri stars in several regions show the existence of two regimes in a plot of average companion surface density versus separation, namely a binary regime with a steep slope at small separations and a clustering regime with a shallower slope at large separations, with a clear break between them at a separation of about 0.04 pc that may represent the size of a typical collapsing pre-stellar clump (Larson 1995; Simon 1997). Although this interpretation of the observations is not unique and the scale of the break may also depend on superposition effects and on the dynamical evolution of the system (Nakajima et al. 1998; Bate, Clarke, & McCaughrean 1998), the interpretation of the break in terms of a typical clump size may still be valid in low-density regions like Taurus where these effects may not be as important as in denser regions. A similar size scale has been found by Ohashi et al. (1997) in a study of the rotational properties of collapsing pre-stellar clumps, which shows that their specific angular momentum is conserved if \(\rho \geq 0.07\) M\(_\odot\) pc\(^{-2}\), which is smaller than about 0.03 pc; this may represent the characteristic size of a region that collapses rapidly to form a star or binary system (see also Ohashi 1999; Myers, Evans, & Ohashi 2000). Finally, an analysis of the internal kinematics of star-forming cloud cores by Goodman et al. (1998) shows a transition from a turbulent regime on large scales, where the linewidth increases systematically with region size, to a regime of ‘velocity coherence’ on scales smaller than about 0.1 pc, where the linewidth becomes nearly independent of region size. These authors suggest that this change in kinematic behavior is related to the break between the clustering and binary regimes for the T Tauri stars noted above, and they suggest that it has the same basic cause, namely a transition from chaotic dynamics on large scales to more ordered behavior on small scales. Such a transition might be expected because molecular clouds are dominated by turbulent and magnetic pressures on large scales and by thermal pressure on small scales (Larson 1981; Myers 1983), and the transition between the two regimes is in fact what defines the Jeans scale when the latter is calculated by assuming pressure balance between a thermally supported isothermal clump and a turbulent ambient medium. All of the evidence described here is thus consistent with the existence of a scale in the star formation process which is essentially the Jeans scale as derived above.

The characteristic stellar mass may thus depend, via the Jeans mass, on the typical temperature and pressure in star-forming clouds, being proportional to \(T^2/P^{1/2}\). The temperatures of molecular clouds are controlled by radiative processes, but their pressures are probably of dynamical origin and result from the cloud formation process, since their internal pressures are much higher than the general pressure of the interstellar medium (Larson 1996). Molecular clouds are probably created by the collisional agglomeration of smaller, mostly atomic clouds in regions where large-scale converging flows assemble the atomic clouds into more complex structures. The resulting cloud collisions produce a ram pressure \(P \rho^2\) which may determine the typical internal pressure of the molecular clouds formed. If the typical density of the colliding clouds is 20 atoms per cm\(^3\) and if they collide with a velocity of 10 km s\(^{-1}\), the ram pressure produced is \(3 \times 10^5\) cm\(^{-3}\) K, similar to the inferred internal pressures of molecular clouds. Thus the typical pressures in molecular clouds can be understood in terms of the structure and dynamics of the atomic component of the interstellar medium. It may further be possible to understand the properties of the atomic clouds in terms of the classical two-phase model of the ISM, which postulates a balance in
thermal pressure between a cool cloud component and a warm intercloud component and predicts cloud densities of a few tens of atoms per cm$^3$ (Field, Goldsmith, & Habing 1969; Wolfire et al. 1995). Thus it may be possible to understand the characteristic temperatures and pressures of molecular clouds, and hence the characteristic stellar mass, in terms of relatively well-studied thermal and dynamical properties of the interstellar medium (Larson 1996).

If the mass scale for star formation depends on the temperature and pressure in star-forming clouds as predicted above, one might expect to see some variability of the IMF between regions with different properties: for example, clouds with higher temperatures might be expected to form stars with a higher characteristic mass (Larson 1985). There is possible evidence for such an effect in extreme cases such as starburst systems and high-redshift galaxies (Larson 1998), but no clear dependence of the IMF on the temperature or other properties of star-forming clouds has been found in local star-forming regions. In fact, clouds with higher temperatures generally also have much higher pressures, so there is a partial cancelation of these effects when the Jeans mass is calculated, and it is not clear that one effect or the other dominates.

Elmegreen (1999) has argued that such an approximate cancelation of effects is to be expected for physical reasons since the cloud temperature depends on radiative heating rates while the overall pressure of the ISM depends on the local column density of matter in a galaxy, both of which increase with the stellar surface density in such a way that $T^2 P^{1/2}$ is approximately constant.

4 The Formation of Massive Stars and the Origin of the Power-Law Upper IMF

The second basic fact about star formation needing to be explained is that the IMF has a power-law tail extending to masses much larger than the characteristic mass, such that about 20% of the total mass goes into stars more massive than $10 M_\odot$. Most of the feedback effects of star formation on the evolution of galaxies depend on energy input from these massive stars, so it is clearly of great importance to understand the origin and possible universality of the upper IMF. At present the formation of massive stars is relatively poorly understood, both observationally and theoretically, so most of what can be said about the origin of the upper IMF remains speculative.

Elmegreen et al. (2000) have proposed a more generic model in which stars form by random selection from different subgroups, and if accumulation processes tend to build more massive stars in the more massive subgroups of such a hierarchy, then a power-law upper IMF can be produced (Larson 1992, 1999). Most stars do indeed form in clusters, and in at least some cases there is evidence for hierarchical subclustering (Zinnecker et al. 1993; Gomez et al. 1993; Larson 1995; Elmegreen et al. 2000). Since the larger subgroups in such a hierarchy contain more ‘raw materials’ from which to build massive stars, they will almost certainly produce stars with a mass spectrum extending to a larger maximum mass. If the mass $M_{\text{max}}$ of the most massive star formed in any subgroup increases with a power $n < 1$ of the mass of the subgroup, i.e. if $M_{\text{max}} \propto M_{\text{group}}^n$, and if all stars form in a self-similar hierarchy of such groups, then a power-law IMF is generated whose slope is $x = 1/n$ (Larson 1992). For example, the IMF slope $x = 1.4 \pm 0.4$ suggested by the evidence discussed in Section 2 could be reproduced if $n$ were $0.7 \pm 0.2$.

One hypothesis involving hierarchical structure that has been developed further is that star-forming clouds have fractal structures, and that the universal power-law upper IMF results from a universal fractal cloud structure produced by turbulence (Larson 1992, 1995; Elmegreen 1997, 1999). In the model of Larson (1992), stars are assumed to form by gas accumulation along filaments in a fractal filamentary network, and this resulting IMF slope is scale-free in the sense of the network. Elmegreen (1997) has proposed a more generic model in which stars form by random selection from different...
levels of any fractal hierarchy. However, while there is evidence that molecular clouds have fractal boundary shapes, it is less clear that they have fractal mass distributions, and most of their mass cannot plausibly be fractally distributed but must have a smoother spatial distribution. In any case, even if a fractal picture were correct, the accumulation processes required to form stars in such a model would first form small stars from small cloud substructures before matter could be accumulated from larger regions to form massive stars; the cloud regions that form massive stars would then contain substructure that has already begun to form less massive stars. Such a picture would predict the formation of massive stars only in clusters, as is indeed observed, but the interactions among star-forming clumps and protostars that would necessarily occur during the accumulation of matter to form the more massive stars were not taken into account in the above fractal models. Such interactions would almost certainly play a role in determining the final stellar mass spectrum.

It may in fact be that all of the ideas mentioned above have some merit, and that a more realistic model will involve elements of all of them, namely clump collisions, continuing gas accretion, and hierarchical clustering. In its original form, the clump coagulation model of Nakano (1966) did not take into account the fact that clumps formed by the fragmentation of a contracting cloud will often begin to collapse into stars before colliding and interacting with each other. Many of the colliding clumps will then contain accreting protostars, and the effects of their interactions on the protostellar accretion process and on the structure of the forming system of stars will play an important role in its further development. Since these interactions will generally be dissipative, the star-forming clumps will tend to become bound into progressively larger and denser aggregates (Larson 1990). In this way, star clusters may be built up hierarchically by the merging of smaller subsystems, perhaps basically as in the clump coagulation model of Nakano (1966) but with the clumps replaced here by groups of forming stars. For a brief time, a newly formed cluster of stars may continue to show hierarchical subclustering, but this substructure will soon be erased by dynamical relaxation processes. As smaller systems of forming stars continue to merge into larger ones, the protostars in the most favored central locations may continue to gain mass from larger and larger accretion zones, building up an extended spectrum of stellar masses.

Numerical simulations illustrate the likely importance of interactions for the continuing accretional growth of the more massive stars in such a scenario. Protostars with residual disks can strongly perturb their surrounding disks, causing part of the disk matter to be ejected and part to be accreted by the central star (Heller 1991, 1995); in general, the more massive system tends to gain mass from the less massive one in such interactions. In the simulations of cloud fragmentation and accretion by Larson (1978), the most massive objects gained much of their final mass during episodes of rapid accretion associated with close encounters or mergers between dense clumps. Simulations of accretion processes in forming clusters of stars (Bonnell et al. 1997, 1998; Clarke, Bonnell, & Hillenbrand 2000) show the development of a broad spectrum of masses, the more massive objects tending to form near the cluster center where the accretion and interaction rates are highest; the most massive stars may even gain much of their final mass by mergers between already-formed stars (Bonnell et al. 1998; Stahler et al. 2000).

The simple Bondi-Hoyle accretion model of Zinnecker (1982) assumes a protostellar accretion rate that increases with mass in qualitatively the expected way, and it predicts a rapidly increasing spread in protostellar masses and the growth of a power-law tail on the IMF that is qualitatively similar to what is observed. However, it also has the unrealistic feature that it predicts the unlimited runaway growth in mass of the most massive protostar because it is assumed to accrete matter from a region of unlimited size. More realistically, each protostar in a forming cluster will have an accretion zone of finite size associated with the substructures (Larson 1978), and the total amount of gas available to form massive stars will be limited by the size of the cluster. Since

the gas supply is depleted as accreting protostars continue to gain mass from it, a decreasing amount of mass is available to build stars of higher and higher mass, resulting in an IMF with $x > 1$ in which there is less and less mass in stars of increasing mass, as is observed. The amount of mass accreted by each protostar may then be determined by the amount of gas in the subsystem in which it forms, and by the effects of continuing interactions and mergers among the subsystems in a forming cluster; each such interaction or merger is likely to cause additional gas to be accreted by the most massive protostar present. One can easily construct simple interaction and accretion schemes based on these ideas that generate a power-law IMF. The only essential requirement is that the accretion processes involved are basically scale-free, that is, they do not depend on any new mass scale larger than the Jeans mass. This would be the case if, for example, each interaction or merger between two subsystems causes a constant fraction of the remaining gas to be accreted by the most massive protostar present. If we assume, in the simplest formulation of such a model, that the mass of the most massive protostar increases by a constant factor $f$ when the mass of the system to which it belongs increases by another constant factor $g$ because of a merger with another system (for example, $g = 2$ for equal-mass mergers), then the mass of the most massive star formed in a cluster built up by a sequence of such mergers increases as a power $n$ of the cluster mass, where $n = \log f / \log g$. If all stars more massive than the Jeans mass are formed in a self-similar hierarchy of such merging subsystems, then the assumptions of the hierarchical clustering model of Larson (1991, 1992) are satisfied and a power-law upper IMF is produced that has a slope $x = 1/n$. The Salpeter slope is recovered if, for example, $g = 2$ and $f = 5/3$; then $n = \log(5/3)/\log 2 = 0.74$ and $x = 1/n = 1.36$. If the most massive protostar grows by accreting residual gas, then it can be shown that in this simple example, 1/6 of the remaining gas is accreted during each merger. Conversely, if it is assumed that 1/6 of the remaining gas is accreted by the most massive protostar during each merger, then a Salpeter IMF is produced. If the fraction of the gas accreted in each merger varies between 1/10 and 1/4, then the predicted value of $x$ varies between 1.18 and 1.71. These results are not very sensitive to the assumption of equal-mass mergers; for example, if the mass ratio of the interacting subsystems is not 1 but 3, a typical value for clump coalescence models, and if again 1/6 of the remaining gas is accreted during each merger, the resulting IMFs of these assumptions do not seem especially implausible in the light of the observational evidence and the theoretical results noted above, and they all result in IMF slopes that are consistent with the observations, within the uncertainties.

While it would be easy to construct more elaborate and perhaps more realistic accretion models that also yield power-law IMFs, what is really needed to advance our understanding of the origin of the upper IMF is better physical input regarding the processes involved in the accretional growth of massive stars, and estimates of the efficiency of these processes, for example the fraction of residual gas accreted during each interaction or merger between subsystems. The processes involved can now be studied in some detail with numerical simulations, which as noted above have already begun to simulate some of the processes likely to be important. At present these simulations do not provide sufficient quantitative information to test in any detail the kind of model that has been proposed. However, it is more detailed simulations of the kind of interaction/accretion picture suggested above, and if the accretion processes involved are indeed approximately scale-free and characterized by similar efficiencies, important progress will have been made toward understanding the formation of massive stars and the origin of the upper IMF. Ultimately such simulations will have to reproduce not only the IMF of the massive stars but also the clustering and binary properties of these stars as well, and this test will place stronger constraints on the models. Whatever processes may be involved, the formation of massive stars cannot be understood without explaining the striking facts that they form only in