THEORY OF STARBURSTS IN NUCLEAR RINGS

Bruce G. Elmegreen¹

RESUMEN

Dos nuevas correlaciones ilustran que los brotes de formación estelar ocurren frecuentemente en regiones con densidades totales altas; una entre la duración y el tamaño de la región de formación estelar y la otra entre dicho tamaño y la luminosidad de la galaxia huésped. Las escalas de tiempo para la formación estelar son más cortas en estas regiones. Se discute el origen de los anillos de formación estelar en espirales barradas y se dan tasas teóricas simplificadas. La tasa de formación estelar al umbral de densidad es proporcional al inverso del cubo del períódo orbital.

ABSTRACT

Two new star formation (SF) correlations, one between duration and size of the SF region, and another between region size and the luminosity of the host galaxy, illustrate that starbursts usually occur where the background galactic virial density is high. The time scales for SF are faster in these regions. The origin of the starburst rings in barred galaxies is discussed, and simple theoretical SF rates are given. The SF rate at the threshold density scales with the inverse cube of the orbit time.

Key words: GALAXIES: STARBURST — STARS: FORMATION

1. OVERVIEW OF GALACTIC STAR FORMATION

Stars form in the dense parts of self-gravitating clouds where the magnetic field diffuses away and the gas can collapse. Observations over the last 10 years have suggested that these clouds are fractal and hierarchically structured, so star formation (SF) in them is likely to be hierarchical as well, forming multiple groups and clusters instead of isolated stars. If the SF efficiency in a cluster ends up high, then the cluster can be self-bound after the gas leaves. In the solar neighborhood, the efficiency is generally low so most clusters in molecular clouds end up unbound. The duration of SF in a region of size D tends to increase with D as

$$t_{SF} \sim 1(D/pc)^{1/2} \text{ My.}$$
 (1)

This relation is shown schematically in Figure 1 and discussed in Elmegreen & Efremov (1996), using the distribution of Cepheid variables in the LMC. It seems to be true for all scales from 0.1 pc to 1000 pc, and so is a general property of SF. The timescale for turbulence is similar to this,

$$t_{turb} \sim 3(D/pc)^{1/2} \text{ My}, \tag{2}$$

from the crossing time, $t_{turb} \sim D/c$ with the observed size-linewidth relation, $c \sim 0.3(D/pc)^{1/2}$ km s⁻¹ for FWHM spectral line velocity c (Blitz 1993). Thus stars form on the turbulent time scale. The implications of this result are vast. SF does not proceed from large to small scales in a cascade of collapse, as formerly thought, but occurs simultaneously on all scales with the smaller regions coming and going many times before the larger regions are finished. Large regions form stars for a long time, so the largest active regions in galaxies like ours contain Cepheid variables and red supergiants in addition to OB stars. SF is hierarchical. This means that star complexes like Gould's Belt, which may span 1 kpc and form stars continuously for 30 My, contain several OB associations at any one time and form several generations of OB associations during their active phase. The OB associations, in turn, contain several active cores or subgroups at any one time, and form several generations of subgroups during their life. Each subgroup should contain several subclusters at any one time and several generations of subclusters over its life. Some of these subclusters mix and become indistinguishable by the time the Sasociation has finished forming; some of the OB associations mix and become indistinguishable by the time the complex has finished forming, and some of the OB associations mix and become indistinguishable by the

¹IBM Research Division, T.J. Watson Research Center, P.O. Box 218, Yorktown Hts NY 10598 USA bge@watson.ibm.com.

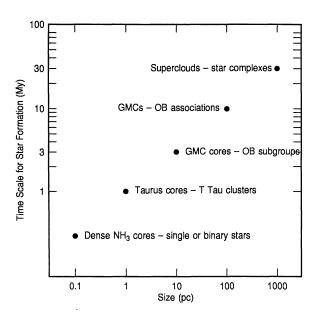


Fig. 1. Schematic diagram showing how the duration of SF in regions of various sizes increases with the square root of the size.

2. GALACTIC SCALING LAWS

The duration of SF scales with the square root of the size of the region, but how does the constant of proportionality in this relation vary from place to place, from galaxy to galaxy? We find that there is a continuous change in this constant from large galaxies to small galaxies, spanning 4 orders of magnitude in galaxy luminosity (Elmegreen et al. 1996). This result follows from the observation that the diameter D_c of the largest star complex in a galaxy scales with galaxy luminosity L as

$$D_c(\text{kpc}) = 1.9(L/L*)^{0.55} \text{ (spirals)} \; ; \; D_c(\text{kpc}) = 3.9(L/L*)^{0.55} \text{ (Sm/Im)}.$$
 (3)

Note that $L/L* \sim 0.001 - 0.01$ for an Sm/Im galaxy, so we could write $D_c(\text{kpc}) = 0.31(L/[0.01L*])^{0.55}$ for the Sm/Im case. The relative size of the largest star-forming region, compared to the galaxy size, actually increases with decreasing L, but only by a small amount. Thus small galaxies have small star-forming regions. Figure 2 shows the relevant data from Elmegreen et al. (1996). The point of Figure 2 is that, even though 30 Dor in the LMC may look big, it is still smaller than the largest comparable object (e.g., Gould's Belt) in a galaxy the size of the Milky Way. There is a bias in our impression of SF size when only $H\alpha$ observations are used. $H\alpha$ observations emphasize the regions where OB stars recently formed, and these regions tend to be around 10 My old (the lifetimes of OB stars). This means that they have a certain size given by the product of this lifetime and the turbulent speed (see above). Thus Efremov (1996) finds that all OB associations, in large and small galaxies near us, have a characteristic size of ~ 80 pc. This is not a characteristic scale for SF in general, because there is no characteristic scale for most SF; i.e., SF is apparently hierarchical and self-similar on a wide range of scales. The appearance of a characteristic scale for OB associations comes only because these regions are selected to have a particular age. If we select regions using a different criterion, such as the largest roundish region in a galaxy, regardless of age, then we get the $D_c - L$ relation discussed above. One quantity that does not apparently scale much from galaxy to galaxy is the turbulent speed c of the interstellar gas, which seems to be typically around 10 km s⁻¹. Of course there are regions where this turbulent speed is much larger, as in interacting galaxies or starburst regions (Elmegreen, Kaufman, & Thomasson 1993), but these agitated places can occur in galaxies of all sizes too, depending on the energy input to the ISM. In what we might call normal regions, or "unperturbed" regions, c seems to be about constant to within a factor of 2 (e.g., there are radial variations; Boulanger & Viallefond 1992). Now it follows from the relation $t_{SF} \sim D/c$ that the timescale for SF in the largest region of a galaxy (and presumably scaling to smaller regions as well) decreases with galaxy luminosity, approximately as $L^{1/2}$, like D_c . Thus most SF in dwarf galaxies is burst-like compared to most SF in the main parts of large galaxy disks. Burst-like means that the total duration of SF in the largest regions is very short, say ~ 5 My or less for a dwarf, whereas it is very long, say 50 My, for the same type of region (although larger) in a larger galaxy. For example, Gould's Belt looks quiescent because its large number of stars formed over a long period, 50 My, and 30 Dor looks active because its large number of stars formed over a short period, 5 My, but both are active with different time scales, and the physical process of SF is probably the same in each. Only the constants of proportionality in the (turbulent) scaling relations have changed because of the different galaxy sizes containing these two regions. Why does the constant of proportionality change with galaxy size? The answer to this question comes from the Tully-Fisher (1977) relations between internal galaxy velocity and luminosity L and between galaxy size and L. These relations show that the galaxy size D_{25} (= diameter at 25th mag arcsec⁻¹) decreases with L more rapidly than the internal dynamical velocity, W_D , so the ratio D_{25}/W_D decreases with galaxy luminosity too, approximately as (Elmegreen et al. 1996)

$$D_{25}/W_D = 170 \left(\frac{L}{L*}\right)^{0.375} \text{ pc/km s}^{-1} \text{ (spirals)} ; D_{25}/W_D = 45 \left(\frac{L}{0.01L*}\right)^{0.4} \text{ pc/km s}^{-1} \text{ (Sm/Im)}.$$
 (4)

The quantity D_{25}/W_D for a galaxy is inversely proportional to the square root of the virial density. From the virial theorem,

$$\rho_{vir} \approx \frac{(W_D/D_{25})^2}{G}. (5)$$

Thus the virial density, which is the total density in all forms, gas+stars+dark matter, increases significantly for smaller galaxies. Using equation 4 above, we can write

$$\rho_{vir} \approx 8 \times 10^{-3} \left(\frac{L}{L*}\right)^{-0.75} \text{ M}_{\odot} \text{ pc}^{-3} \text{ (spirals)} ; \quad \rho_{vir} \approx 0.11 \left(\frac{L}{0.01L*}\right)^{-0.8} \text{ M}_{\odot} \text{ pc}^{-3} \text{ (Sm/Im)}.$$
(6)

This is the average virial density for all the material inside the radius at 25th mag $arcsec^{-2}$; the virial density can be higher than this inside smaller radii, scaling approximately as the inverse square of radius for a flat rotation curve. The increase in virial density with smaller galaxy size implies that the dynamical time scale decreases for small galaxies. This includes the orbit time at the edge D_{25} , and it includes the duration of SF in the largest regions. This latter variation follows from the $D_c(L)$ correlation for complex size. If we divide the $D_c(L)$ correlation by the $D_{25}/W_D(L)$ correlation, we get

$$\frac{D_c}{D_{25}} = \frac{11 \text{km s}^{-1}}{W_D} \left(\frac{L}{L*}\right)^{0.175} \text{ (spirals)} \; ; \; \frac{D_c}{D_{25}} = \frac{6.8 \text{km s}^{-1}}{W_D} \left(\frac{L}{0.01L*}\right)^{0.15} \text{ (Sm/Im)}.$$
 (7)

These relations imply that D_c equals approximately half the Jeans length in a medium with a 1D Gaussian velocity dispersion of $c_G \sim 6-11$ km s⁻¹ and a density comparable to the virial density,

$$D_c \approx \frac{0.5c_G}{(G\rho_{vir}/\pi)^{1/2}},\tag{8}$$

for nearly constant c_G in all galaxies. We may conclude from this that in essentially all galaxies, star formation operates on the gravitational length scale whenever the gas density exceeds the local virial density. There is a threshold density for SF given by

$$\rho > \rho_{vir},$$
(9)

averaged over the thickness of the gas layer. This gives a threshold column density $\rho_{vir}D_c$, which is similar to the Kennicutt (1989) critical column density for spirals, but also applicable to dwarfs where Q and the usual disk instability analysis are not valid (see discussion in Elmegreen et al. 1996). The time scale for the duration of SF in a region can now be related to the local orbit time. Since $t_{SF} \sim D_c/c$ on the largest scale for FWHM $c \sim 2c_G$, we have $t_{SF} \sim 0.5D_{25}/W_D$. The 1D virial velocity in a galaxy is proportional to the rotation speed: for a thin disk, $W_D \sim 2V_{rot}/3^{1/2}$. This means that t_{SF} scales with the orbit time, $\sim 0.14t_{orbit}$, from these numbers. There are easily factor-of-two uncertainties in the size-line width relation for molecular clouds and in the size scaling of the SF time, so it is best to evaluate the constant of proportionality in the $t_{SF} - t_{orbit}$ relation from observations on a large scale. A better estimate is $t_{SF} \sim 0.2t_{orbit}$, which gives about the age of Gould's Belt, 50 My, in a galaxy the size of ours. In terms of galaxy luminosity, this becomes

$$t_{SF} \approx 50 \left(\frac{L}{L*}\right)^{0.4} \text{ My.} \tag{10}$$

168

ELMEGREEN

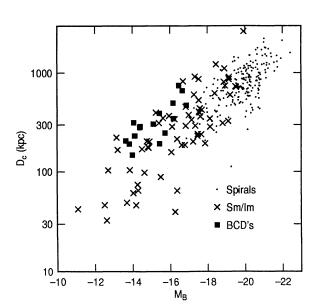


Fig. 2. The diameters of the largest star complexes in spiral galaxies (small dots), Sm/Im galaxies (x symbols and BCD galaxies (squares) are shown as a function of the absolute galaxy magnitude.

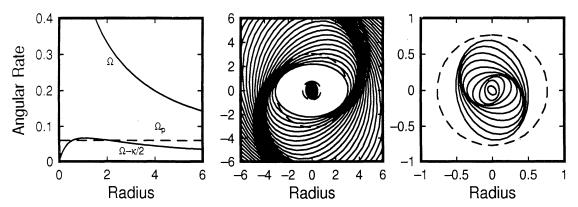


Fig. 3. Rotation properties and gas orbits when there are two inner Lindblad resonances.

Thus the time scale for SF in the largest patches decreases with galaxy size, from 50 My in large galaxies to 5 My or less in small galaxies because D_c/c , or, equivalently, $1/(G\rho_{vir})^{1/2}$, decreases with galaxy size. As a result, SF is more intense, concentrated, and burst-like in small galaxies than in large galaxies.

3. APPLICATIONS TO NUCLEAR STARBURST RINGS

3.1. Formation of the Ring

The above discussion of dwarf and other normal galaxies illustrates the essential property of SF: it begins when the average gas density exceeds approximately the local virial density from the background galaxy. It starburst regions, this virial density is always very high, so SF is always intense and burstlike. In the main disks of spiral galaxies, the gas may exceed this critical value for a large fraction of the Hubble time because the SF rate is low and the gas consumption time long ($\sim 5t_{orbit}/\epsilon$ for efficiency $\epsilon \sim 0.1$, see below). In region with high virial densities, such as nuclear regions, the consumption time is short and the gas density will no exceed ρ_{vir} for very long. Then SF ends quickly after it begins, and it must wait for the gas density to increase above the critical value before it can begin again. The starburst process evidently consists of three phases (1) the gas density first increases above ρ_{vir} , which is extremely large, (2) SF proceeds at high density, and

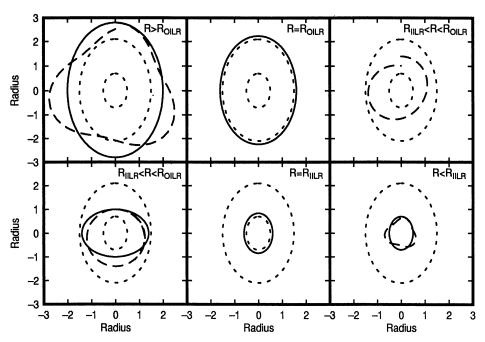


Fig. 4. Star orbits in a bar potential: dashed curves are the orbits without bar forcing at various radii. The bar locks the orbits into alignment parallel or perpendicular to the bar, depending on radius, as shown by the solid curves (see explanation in text). The elliptical dotted curves are the inner and outer inner Lindblad resonances.

(3) the gas density decreases below ρ_{vir} by conversion into stars and by pressurized removal from the region. Various processes can increase the gas density when radial accretion is involved. Galaxy interactions can lead to strong tidal arms which remove angular momentum from the gas and cause it to spiral inward. This increases the gas density slightly everywhere and by a large amount in the center. Bars can also remove gas angular momentum in the bar region, channeling it to the center along leading dust lanes in the bar. If there are two inner Lindblad resonances (ILR), then this accreting gas can accumulate in a ring between them. Here we discuss one process of such ring formation. First it is illustrative to consider the amount of gas available for accretion and the average accretion rate. A typical bar in an early type galaxy (which is the type that has two inner Lindblad resonances) will contain around 10⁹ M_☉ of stars. If the average rate of mass expulsion from each star is $\sim 10^{-10}~{\rm M}_{\odot}$ per year from planetary nebulae, supernovae and steady winds, then the rate at which new material becomes available for accretion is $\sim 0.1~M_{\odot}$ per year. Typical accretion speeds might be around 10 km s⁻¹ (Quillen et al. 1995), and typical bar radii ~ 3 kpc, so the accretion time is slightly less than 10⁹ years. This means that the bar may contain $\sim 10^8 \ \mathrm{M}_{\odot}$ of gas and accrete with a steady rate of $\sim 0.1 \ \mathrm{M}_{\odot}$ per year. We choose numbers here that give an average accretion rate comparable to the gas production rate in order to explain the common observation of bar dust lanes, which is where the accretion occurs. The SF rate in an inner ring is often comparable to that in the rest of the disk, which is several M_{\odot} per year. This means that $\sim 10\%$ of the time is spent in a burst, or that $\sim 10\%$ of early type barred galaxies have an inner Lindblad ring burst. These numbers are very approximate, but they are reasonable, suggesting that the gas in the ring burst may be replenished by stellar evolution in the bar, i.e., it need not come from the outer disk and the galaxy need not be interacting. Why does the gas accrete in a ring instead of going all the way to the nucleus? Ring accretion occurs when there are two inner Lindblad resonances, as mentioned above. (Ring accretion may also occur for a different reason if there is a nuclear minibar, as discussed by D. Friedli at this conference). Figure 3 shows the angular speed Ω and the angular precession rate of close elliptical orbits, $\Omega - \kappa/2$. It also shows the pattern speed of the bar, Ω_p . Where the line Ω_p crosses the curve $\Omega - \kappa/2$, there is an ILR. In this figure there are two ILR because the curve $\Omega - \kappa/2$ has a maximum above Ω_p . This tends to occur when there is a large core radius in a strong central bulge, because then the central density is about constant, Ω approaches a constant at small radii, and $\Omega \sim \kappa/2$ when Ω is constant (i.e., $\Omega - \kappa/2 \rightarrow 0$ at small r).

Star orbits outside the outer ILR (=OILR) are aligned along the bar. This is essentially because the angular precession rate of closed orbits in the absence of the non-axisymmetric forcing from the bar, $\Omega - \kappa/2$, is smaller than the angular rate of the bar pattern at r > OILR. Then orbits starting at their apocenter at the bar end

would loop back before they reach the other bar end, as shown by the dashed line in the upper left of Figure 4, in the absence of bar forcing. Aligned orbits overcome this problem with bar forcing because they have the slowest parts of their orbits (at apogalacticon, where they spend the longest proportion of time) at the position where bar forcing is strongest, on the bar. Thus the time-average centripetal force along their orbits is larger than the time average for an orbit starting at the same position without bar forcing. In effect, they get an extra kick at apocenter which flings them around to the other end of the bar just when their epicycle returns to the initial phase at apocenter (solid curve in the upper left figure). Between the two inner resonances, the precession rate of closed elliptical orbits is larger than the angular rate of the bar pattern (i.e., $\Omega - \kappa/2 > \Omega_p$), so orbits starting with apocenter on the bar and no bar forcing will turn around in radius too late, after they pass the other end of the bar (dashed curve in the upper right of Fig. 4). Stable orbits are the ones that have less than this average forcing, and so they put the slowest part of their orbit (at apocenter) where the bar forcing is weakest, on the bar minor axis (solid curve in lower right; the dashed curve in the lower right would be the orbit in the absence of bar forcing for the same initial condition). Thus stellar orbits between the two ILRs tend to be elongated perpendicular to the bar. Inside the IILR, they can be parallel to the bar again, but need not be, depending on where the bar forcing is strongest, i.e., on the major or minor axis (if it is strongest on the minor axis inside the IILR, then the star orbits can remain somewhat perpendicular to the bar there).

Gas orbits differ from star orbits because the gas dissipates energy, which tends to decrease its radius. This means that outside the OILR, the precession rate of closed orbits increases as energy dissipates. As a result, the gas orbit ellipses are thrown forward relative to the bar for decreasing radius, and orbit crowding then gives a trailing spiral or gas/dust lane. This is shown in the middle of Figure 3. The slowest parts of the gas orbits (their apocenters) lead the bar with this configuration, so the gas is systematically pulled back by the bar force, causing the gas to lose angular momentum. The angular momentum loss and the energy loss both lead to gas accretion. Inside the IILR, the dissipation of gas energy still causes the gas to move inward, but here the decreasing radius leads to increasing lag between the closed orbits and the bar (because $\Omega - \kappa/2$ decreases with radius). Figure 3 (right), shows how this produces a leading spiral. Now the torques on the gas are in the forward direction of rotation, so the gas gains angular momentum. If this gain is more than enough to compensate for the loss from orbital viscosity, then the gas moves outward in radius at r < IILR. As a result of these gas motions, a ring forms between the two ILRs (see Buta & Combes 1996). An interesting variation occurs if there is a massive central object that dominates the rotation curve inside the ILR (Combes, private communication). Then Ω will continue to increase as $r \to 0$ and there will be only 1 ILR. The gas orbits for this case are shown in Figure 5. Because $\Omega - \kappa/2$ increases continuously as $r \to 0$, all orbits twist forward at smaller radii. Then orbit crowding produces a trailing spiral for all radii and all the gas accretes to the center without producing a ring. Whether or not a ring forms, the accretion of gas to the central region can increase the density to above the critical value, leading to SF by normal processes. Because of the high critical density at small r, the SF rate can be enormous and the gas consumption time very short. This is discussed in the next section.

3.2. Star Formation Rates in Nuclear Starburst Rings

We may use the observed scaling properties for SF in galaxies to infer that SF everywhere has a threshold density comparable to the local virial density. In section 2, this property was demonstrated only for average galaxy dimensions, but if the property is general, then it should apply to starburst regions too, as long as the local virial density, rather than the average, is used. The SF rates obtained in this way are reasonable. We take for a local SF rate the ratio of the local gas density, which is around ρ_{vir} at threshold, to the SF time, which is $t_{SF} \sim 0.2t_{orbit}$, all multiplied by some efficiency ϵ , which is essentially a measure of the fraction of the gas mass that goes into stars before the clouds are disrupted. Thus the SF rate is

$$SF\ Rate(\text{gm cm}^{-3} \text{ s}^{-1}) = \frac{\epsilon \rho_{gas}}{t_{SF}} \sim \frac{5\epsilon \rho_{vir}}{t_{orbit}} \sim \frac{70\epsilon}{Gt_{orbit}^3},$$
 (11)

for t_{orbit} in seconds, or

$$SF \ Rate(M_{\odot} \ pc^{-3} \ yr^{-1}) \sim \frac{0.015\epsilon}{(t_{orbit}/M_{Y})^{3}},$$
 (12)

for t_{orbit} in units of 10^6 yr. This expression should be valid everywhere above the threshold density. In the main disks of galaxies, $t_{orbit} \sim 250$ My and the volume is $\sim 200\pi \times 12000^2$ pc³, so the total SF rate is $\sim 10^2 \epsilon$ M_{\odot} yr⁻¹. This is the observed SF rate for a reasonable efficiency of $\epsilon \sim 0.05 - 0.1$. In nuclear starburst

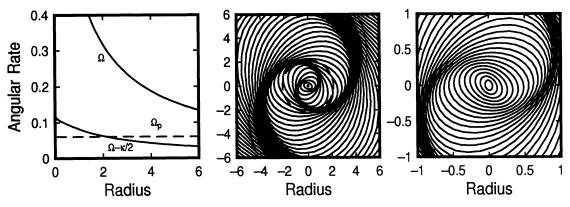


Fig. 5. Rotation properties and gas orbits when there is one inner Lindblad resonances.

rings, $t_{orbit} \sim 30$ My for a radius of R = 500 pc and an orbit speed of V = 100 km s⁻¹, and the volume is $\sim 200^2 \times 2\pi 500$ pc³, so the total SF rate is $\sim 10^2 \epsilon$ M_{\odot} yr⁻¹ again. Thus nuclear starburst rings should contain as much SF as a whole galaxy disk. This is observed to be the case, so expression (12) appears to be valid for both normal and nuclear ring starburst regions. We conclude as we did for dwarf galaxies above, that a starburst is normal SF in a region with a high background galaxy density. Normal SF presumably proceeds in four steps: (1) self-gravitational restructuring of the interstellar medium into large clouds, (2) compression from turbulence and existing generations of SF on smaller scales, (3) magnetic diffusion in the dense cores that result, and (4) a final collapse to stars. The main point of the present paper is that the first step in this process must wait until $\rho_{gas} > \alpha \rho_{vir}$ for alpha of order unity, depending on the distribution of mass in all forms (gas+stars+dark matter). In a nuclear ring, the first step makes the hotspots in the ring (Elmegreen 1993). What is different about SF in a starburst? High gas densities imply fast evolution for SF. Fast evolution implies that all stars can be young in an active region (i.e., many OB stars can be present simultaneously, not just a mixture of OB stars and red supergiants). This means that we will see unusual concentrations of OB stars (plus lower mass stars), and there will be high luminosity densities. High luminosity densities imply high turbulent speeds in the SF gas (scaling with size) and high thermal temperatures. If the turbulent speeds in the ambient medium are also high, then the clouds that form by gravitational instabilities will be more tightly bound than in normal regions. In that case, stars can form with less cloud disruption at a given efficiency, and end up with a higher total efficiency. This means more bound clusters (Elmegreen 1994), and because the gas density is high, more massive bound clusters, perhaps like those discussed by Dr. Ho at this conference. High thermal temperatures might also lead to larger stellar masses or a higher low-mass limit to the IMF, but this is uncertain. The thermal Jeans mass is proportional to the ratio $T^2/P^{1/2}$ for temperature T and pressure P. Thus T would have to increase by more than the fourth root of the pressure compared to normal regions before the characteristic stellar mass will increase. Since starburst regions are likely to have both high temperatures and high pressures, the change in $T^2/P^{1/2}$ could go either way.

REFERENCES

Blitz, L. 1993, in Protostars and Planets III, ed. E. H. Levy & J. I. Lunine (Tucson: Univ. Arizona Press), 125 Boulanger, F., & Viallefond, F. 1992, A&A, 266, 37

Buta, R. C. & Combes, F. 1996, Fundam. Cos. Phys., in press

Efremov, Yu. N. 1996, AJ, 110, 2757

Elmegreen, B. G. 1993, ApJ, 425, L73

_____. 1994, in Violent Star Formation from 30 Dor to Ultraluminous Galaxies, ed. G. Tenorio-Tagle (Cambridge: Univ. of Cambridge Press), 220

Elmegreen, B. G., Kaufman, M., & Thomasson, M. 1993, ApJ, 412, 90

Elmegreen, B. G., & Efremov, Yu. N. 1996, ApJ, in press

Elmegreen, B. G., Elmegreen, D. M., Salzer, J. J., & Mann, H. 1996, ApJ, in press

Kennicutt, R. C. 1989, ApJ, 344, 685

Quillen, A. C., Frogel, J. A., Kenney, J. D. P., Pogge, R. W., & DePoy, D. L. 1995, ApJ, 441, 549

Tully, R. B. & Fisher, J. R. 1977, A&A, 54, 661