RATES OF STAR FORMATION

Richard B. Larson
Yale University Observatory

I. INTRODUCTION

Perhaps the most fundamental fact of extragalactic astronomy is the existence of the Hubble sequence of galaxy types, embodying a correlation between the relative prominence of a spheroidal component, the form of the spiral pattern (if present), and the degree of resolution into young stars (Hubble 1936; de Vaucouleurs, this conference). The colors and spectral types of galaxies vary systematically along this sequence, indicating a variation in the relative content of old and young stars; the correlation between stellar content and the central concentration of the light distribution has been particularly emphasized by Morgan (1958). There is also a systematic variation of fractional gas content along the Hubble sequence, albeit with considerable scatter, such that the relative content of young stars is positively correlated with the gas content (Roberts 1976; de Vaucouleurs, this conference.) Because the correlations are by no means perfect, it has been suggested that the disc/bulge ratio is not a fundamental parameter of the Hubble sequence (Sandage, Freeman, and Stokes 1970, Freeman 1976), and that the disc/bulge ratio and the gas content should be treated independently in a two-parameter classification scheme (van den Bergh 1976a, 1977). It nevertheless remains true that even in van den Bergh's two-dimensional classification the majority of galaxies fall roughly along a diagonal band, indicating a positive correlation between the disc/bulge ratio and the relative content of gas and young stars.

The generally accepted explanation of the variation in gas and stellar content along the Hubble sequence is that galaxies of different types all have about the same age but have turned their gas into stars at different rates, the elliptical galaxies having consumed nearly all of their gas in a rapid early burst of star formation, while the irregular galaxies have experienced much slower star formation and therefore still contain large amounts of gas (Baade 1963). Synthetic stellar population models have provided support for this picture by demonstrating that the variation in color along the Hubble sequence can be explained as due entirely to different rates of star formation, without invoking either age differences or a variation in the stellar mass spectrum (Tinsley 1968; Searle,Sargent, and Bagnuolo 1973). In this view, the star formation rate is regarded as being determined uniquely by the Hubble type, presumably through more fundamental underlying parameters such as the distribution of mass or angular momentum in the system. The lack of an exact correlation between gas content and structure might then be explainable as a result of gas inflows or outflows between galaxies and their surroundings (Larson 1972, van den Bergh 1977).

The star formation rate plays a crucial role also in collapsing protogalaxies, and is the primary factor determining the relative prominence of the disc and bulge components of the resulting system (Larson 1976a,b, Gott 1977). The formation of a spheroidal component requires a period of rapid early star formation, while the formation of a disc component requires a longer period of slow star formation that allows most of the residual gas to settle into a disc before turning into stars. Thus, if the conventional collapse picture is correct, the disc/bulge ratio of a galaxy is determined by the history of star formation during the collapse. The Hubble sequence can then be understood as resulting from differences in the timescale for star formation during both the initial collapse and the later evolution of galaxies, such that elliptical galaxies are formed when star formation is very rapid, and late spiral and irregular galaxies are formed when it is very slow.

A theoretical understanding of the formation and evolution of galaxies therefore depends on an understanding of star formation, and especially of the factors influencing the rate of star formation. We shall review below some of the theoretical problems of star formation in galaxies, some approaches that have been considered in models of galaxy evolution, and some possible observational tests that may help to clarify which processes or models are most relevant.

2. POWER-LAW MODELS FOR STAR FORMATION

In the absence of a detailed understanding of star formation, it has often been supposed that the star formation rate depends primarily on the gas density, and many attempts to model the evolution of galaxies have been based on the assumption originally made by Schmidt (1959) that the star formation rate per unit volume is proportional to a power n of the local gas density. Regarded as a physical law, this assumption was later dropped by Schmidt (1963), who used a power law simply as a convenient way of parameterizing the relation between the star formation rate and the surface density of gas in the galactic disc; nevertheless, many subsequent authors have continued to assume a power-law function of gas density. Schmidt (1959) found that a value for n of about 2 was most consistent with various observations, and many (but not all) subsequent studies have found similar values (Madore 1977). If $n > 1$, the timescale for gas consumption is proportional to a negative power $-n+1$ of the gas density, and therefore the gas is turned into stars most rapidly in the regions of highest initial density. This prediction is qualitatively consistent with the fact that the denser inner parts of most spiral galaxies appear to have processed their gas more completely into stars than the outer parts. It might even be possible to explain the Hubble sequence in this way, if the initial gas density is the main parameter determining the Hubble type (Larson 1976b, Gott 1977).

However, a number of problems and ambiguities arise if it is attempted to apply the power-law assumption to the interpretation of the observations in all situations. A number of attempts to determine the exponent n empirically by correlating the distributions of young stars and gas in external galaxies have yielded different results for different regions; for example, different values of n are required to fit the observations in the inner and outer parts of galaxies (e.g. Madore et al. 1974). In addition, although there is a broad correlation, there is often no close correspondence between the total star formation rate and the gas content of galaxies, particularly among early-type galaxies (e.g. Balick et al. 1976); evidently other factors in addition to the gas content must play a role in determining the star formation rate. It is possible that on a small scale the gas is more condensed in some galaxies than in others, but then the "gas density" is no longer a very well-defined quantity, since its value depends on the scale on which it is measured; also, it is clear that other factors besides the mean gas density, such as those that determine the clumpiness of the gas, must play a role in determining the star formation rate.

Protogalactic collapse models in which the star formation rate is assumed to vary as a power of the gas density can provide satisfactory models for the formation of elliptical galaxies (Larson 1975), but cannot account for the formation of systems with both bulge and disc components (Larson 1976a, b). To obtain a significant disc component, it is necessary to assume a much greater drop in the star formation rate during the later stages of the collapse than is given by a power-law function of the gas density. Here again, it seems clear that other factors besides the gas density must be important; either star formation must be greatly enhanced during the earliest stages of the collapse, perhaps by strong collisions and shock compression, or it must be suppressed at later times, perhaps by tidal forces, relative to a simple power-law dependence on gas density.

Theoretically, it is clear that the occurrence and the rate of star formation cannot depend only on the gas density, even locally. For example, if the gas density in a galactic disc is too low, then tidal shearing forces dominate over the self-gravity of the gas, and gravitational collapse cannot occur. Hence there is a threshold gas density for star formation which depends on the mean density of matter interior to the region considered. Also, the self-gravity of the gas must overcome thermal pressure and any other effects such as magnetic forces tending to prevent collapse, so the thermodynamics and temperature of the gas, as well as possible magnetic effects, will also play a role in star formation.

A major part of the star formation problem therefore consists in understanding how the requisite conditions of density, temperature, etc. are produced; that is, we need to understand how and at what rate sufficiently dense, cool, gravitationally unstable clouds are formed. Since this must involve the larger-scale dynamics of the interstellar medium, it is appropriate to consider in more detail the conditions required for star formation and how they might be produced by various large-scale dynamical processes.

3. STAR FORMATION PROCESSES

(a) Conditions Required

The condition that the self-gravity of the gas must exceed external tidal forces is essentially the same as the classical Roche condition for the formation of gravitational condensations in a primordial solar nebula, and it requires that the gas density be higher than a critical value which is approximately the mean density of matter interior to the point considered. The growth of gravitational instabilities in the gas in a differentially rotating galactic disc was studied by Goldreich and Lynden-Bell (1965), and they derived a critical density depending on the Oort constants of galactic rotation that is about 2-3 cm^{-3} in the solar region of our galaxy. This is greater than the observed gas density, but not by a large factor, so that gravitational instability could be triggered by a modest compression of the gas; such compression might result, for example, from the dissipation of random motions and the settling of the gas into a thinner and denser layer, or from large-scale dynamical processes such as a compression wave or shock front in the gas layer. The most unstable wavelength is several times the thickness of the gas layer, or roughly one kiloparsec, and the amount of gas in a region of this size is several million solar masses. This is comparable to the masses of the large interstellar cloud complexes in which star formation is observed to be taking place in our galaxy, suggesting that gravitational instability has played an important role in the formation of these cloud complexes.

The second condition required for star formation is the well-known Jeans condition: a cloud can collapse only if its mass exceeds a critical value depending on its density and temperature. For the typical parameters of interstellar neutral hydrogen, the Jeans mass is of the order of a few thousand solar masses, becoming smaller in the denser clouds (and much smaller in the molecular clouds). The fact that the Jeans mass is much smaller than the mass contained in a typical unstable region of the galactic gas layer suggests that tidal forces may provide a more important limit to star formation; once tidal forces are overcome, thermal pressure is a relatively unimportant barrier to gravitational collapse. Again it is interesting to note that in large star-forming complexes containing millions of solar masses of gas, individual collapsing clouds like that associated with the Orion nebula have masses which are typically a few thousand solar masses, in rough agreement with the Jeans mass. However, there are also smaller molecular clouds with masses of only tens or hundreds of solar masses that appear to be gravitationally bound and perhaps collapsing, suggesting that mechanisms other than internal self-gravity have acted to compress the gas in these clouds.

A third condition that is required if massive clouds are to fragment into stellar objects is the presence of efficient cooling mechanisms that can radiate away the collapse energy within a free-fall time. Both theory and observations indicate that collapsing molecular clouds can cool efficiently to temperatures below 10 °K, so that the cooling requirement does not pose a problem for star formation in the galactic disc. However, the cooling condition may well be more important in protogalaxies where the gas density and/or the abundance of heavy elements is lower, and may limit the sizes and masses of galaxies (Rees and Ostriker 1977).

(b) Ways of Achieving These Conditions

If gravitational instability of the galactic gas layer is an important cause of star formation, then the rate of star formation depends on the rate at which the gas is compressed to the required threshold density. If there are no large-scale disturbances or non-circular motions, cooling and the dissipation of turbulent motions will cause the gas to settle into a thinner and denser layer and eventually become unstable, triggering star formation. The formation of hot stars, supernovae, etc. will feed energy back into the gas and tend to limit the rate at which it can settle into a thinner layer. The rate of star formation is then determined by a complicated feedback cycle in which the controlling factor is the rate of dissipation of turbulent motions. Galactic evolution models of this type were studied by Talbot and Arnett (1975), who assumed in effect that the dissipation timescale is comparable to the period of vertical oscillations through the galactic plane, and thus depends on the local surface density of matter in the disc; the timescale for star formation then depends on the disc surface density, decreasing with increasing surface density. A somewhat similar situation occurs in some of the collapse models of Larson (1976a), in which the rate at which gas settles into a thin disc is limited by the rate of dissipation of turbulent motions. Regardless of the details of specific models, if the rate of star formation is controlled by dissipation, it may be expected to depend on dynamical timescales which are shortest in the regions of highest total mass density.

Since the most conspicuous signs of star formation in galaxies (HII regions) are often concentrated in well-defined spiral arms, much theoretical attention has centered on the relation between star formation and the dynamics of spiral structure. There are several ways in which the large-scale dynamics of galaxies can interact with the small-scale dynamics to bring about the conditions required for star formation. A bar- or spiral-shaped density wave can be generated in a galactic disc by a global instability of the disc (Bardeen 1975) or by a tidal interaction with another galaxy (Toomre 1974). The effect of an underlying density wave in the stars, whether bar- or spiral-shaped, is to induce a trailing spiral compression wave in the gas, which becomes a shock front if the amplitude of the wave is large enough (Woodward 1975, Sanders and Huntley 1976). Even if the amplitude is modest and no shock is produced, such a compression wave may act to trigger or enhance star formation by raising the gas density to the threshold required for local gravitational instability; star formation might then occur in clumps which on a small scale have an irregular and patchy distribution, but on a large scale are organized into a broad spiral pattern by the underlying density wave.

If the gas motions are supersonic and a shock is formed, the detailed response of the gas depends on its physical state before entering the shock. If much of the gas is relatively cool ($\sim$ 100 K) and smoothly distributed, a sharply defined shock front can form and compress the gas by a large factor which is approximately the square of the Mach number of the shock; for example, if the gas velocity perpendicular to the shock is 10 km/s, the compression factor is of the order of 100, enough to reduce the Jeans mass by an order of magnitude to a few hundred solar masses. Enhanced cooling behind the shock front will probably cause a further decrease in the temperature and Jeans mass. Another possibility is that the interstellar medium consists of dense, cool clouds in pressure equilibrium with an intercloud medium at a temperature of $\sim 10^4$ K; then the primary shock front occurs in the intercloud gas, and clouds entering the shock are compressed by an increase in the ambient pressure which can cause them to collapse (Shu et al. 1972, Woodward 1976). In either case, one might expect a fairly well-defined region in which clouds of more than a few hundred solar masses collapse to form stars. On the other hand, if the interstellar medium contains no smoothly distributed component at all, the shock region may have a much more complex structure, possibly appearing as a broad zone within which cloud collisions are a dominant process.

Whatever the details, if density waves are important in initiating or enhancing star formation, the rate of star formation will depend on the rate at which the gas is processed through the associated compression waves or shock fronts. In standard versions of density wave theory, a quasi-stationary spiral wave pattern is postulated (although at present this assumption is not strongly supported by either theory or observations). The governing timescale for star formation is then the period of passages through the spiral arms, and this is closely related to the orbital period and generally decreases toward the center of a galaxy. A secondary parameter that may influence the rate of star formation is the "shock strength", or the perpendicular velocity of gas entering the shock front.

If large-scale dynamical phenomena are important for star formation in disc systems, they are likely to play an even more important role in proto-galaxies or young galaxies where the gas has not yet settled into a disc and high-velocity non-circular motions and collisions between gas streams occur. For example, if two gas streams collide at a velocity of 100 km/s and the gas is compressed in a shock front, the density is increased by a factor of $\sim$ 100 if the gas cools to 10^4 K and by a factor of $\sim 10^4$ if it cools to 100 K. In the latter case the conditions are comparable to those existing in interstellar clouds in our galaxy and the Jeans mass is reduced below 10^3 solar masses, so that collapse and star formation seem very likely to occur.

In addition to whatever dynamical processes may act to compress the gas, the star formation rate may also be influenced by gas flows between the star-forming region and its surroundings. For example, the absence of detectable gas and young stars in most elliptical and S0 galaxies may indicate removal of the gas by a galactic wind (Faber and Gallagher 1976) or by motion of the galaxy through an intergalactic medium (Gisler 1976). The apparently low gas contents and star formation rates of van den Bergh's (1976a) anemic spirals may also reflect gas loss by a sweeping process. On the other hand, in gas-dynamical models for the formation of galaxies (eg. Larson 1974, 1976a), gas continues to flow <u>into</u> the star-forming regions for an appreciable period of time, and the rate of star formation is largely controlled by the inflow rate. Some galaxies may also continue to accrete gas from an extended circumgalactic envelope or intergalactic medium, or even from neighboring galaxies, and this may help to explain why some galaxies have relatively high gas contents and star formation rates for their types (Larson 1972). The infall rate estimated by Oort (1970) for our galaxy is in fact comparable to the total star formation rate, but the interpretation of the observations is very uncertain, and it remains an important question for galactic evolution to establish whether infall is occurring at this rate.

4. OBSERVATIONAL INDICATIONS AND TESTS

Now that we have reviewed some of the theoretical possibilities, it is of interest to consider whether observations can help to clarify the main processes or factors influencing the rates of star formation in galaxies. Most of the processes or models that have been discussed involve dynamical timescales that decrease with increasing total density of the region of interest, and they predict that star formation should proceed fastest in the densest regions. Thus, if gas flows are not important, the conversion of gas into stars should be most complete in the densest galaxies and in the densest parts of galaxies; for example, the fractional surface density of gas should decrease inwards in the discs of spiral galaxies. These expectations are in general agreement with the observed trends in the content and distribution of hydrogen in galaxies (Roberts 1976). This supports the notion that dynamical timescales are important for star formation; however, it does not provide a very specific test for theories of star formation, since most of the theories make qualitatively similar predictions in this respect.

An important constraint on models for our own galaxy is provided by the finding of Gordon and Burton (1976) that the fractional surface density of gas in all forms in our galaxy is approximately constant for radii between 5 and 15 kpc, although it is not clear at present how this result is to be interpreted. If the thickness of the gas layer is somehow maintained constant, there is a threshold surface density required for gravitational instability which varies with the total interior density of all matter; the observations may then reflect the importance of gravitational instability and of a threshold gas density (cf. Quirk and Tinsley 1973). However, it is also possible that the gas content in the inner part of the 5 - 15 kpc zone is maintained by a gas inflow from other regions, either inside or outside the disc.

A secondary general prediction of most models is that the heavy element abundance in the disc of a spiral galaxy should decrease outward because of the decreasing completeness of conversion of gas into stars. Again, this prediction appears to be in qualitative accord with the observations, as reviewed by Jensen, Strom, and Strom (1976). These authors noted correlations between their chemical composition indices and parameters representing the "processing speed" and the "shock strength" in density wave models fitted to many galaxies, finding consistency with the predictions of density wave theory. However, this is only an indirect test, and similar correlations would probably be predicted by any theory in which the star formation timescale decreases with decreasing rotation period.

The role of gravitational instabilities or density waves for star formation can be more directly tested by observing the spatial distribution of young stars in spiral galaxies. Local gravitational instabilities alone would lead to star formation in irregular, disconnected spiral arm fragments (Goldreich and Lynden-Bell 1965), while a density wave should introduce some large-scale organization into the spiral pattern. In support of the gravitational instability picture, it can be noted that in a significant fraction of spiral galaxies no large scale pattern is evident and only short, irregular, disconnected spiral features are seen. Most spiral galaxies show some mixture of regular and irregular features, but only about one-third of them have a large-scale symmetrical spiral pattern. Even in galaxies with a large-scale pattern there is often evidence for star formation in regions outside the main spiral arms; in some cases, patchy regions not connected with the main spiral arms are as blue as the arms themselves and are experiencing equally vigorous star formation (Talbot, this conference).

In the standard density wave theory, gas flows through a quasi-stationary spiral wave pattern and star formation is initiated by a shock front at the inner edge of a spiral arm; thus there is predicted to be a gradient of stellar content across the arm, with the youngest and bluest stars near the inner edge and an older, fainter, redder population near the outer edge. However, attempts to observe this effect have so far not shown very consistent or convincing agreement with the predictions. The most detailed data are provided by Schweizer's (1976) multicolor surface photometry of several bright spiral galaxies. Schweizer's data do not show clear evidence for the predicted color gradients across spiral arms, and Schweizer therefore attaches more weight to the surface brightness profiles, which are predicted to be peaked at the inner edge of the arm. In some cases this effect is observed, but in most cases the surface brightness profile is rather broad and symmetrical, showing no evidence for a systematic gradient across the arm. The distribution of HII regions across spiral arms also does not generally show the predicted concentration near the inner edge of the arm (e.g. Kennicutt and Hodge 1976). Thus there is at present little clear evidence that star formation in spiral arms takes place in a progressive wave with a sharply defined front at the inner edge of the arm.

What then is the relation between spiral structure and star formation? It is possible that density waves are often not quasi-stationary in form but are relatively transient and changing features that dissolve and re-form repeatedly; it is also possible that the primary role of a density wave, whether transient or long-lived, is not to create a sharp shock front but to enhance the gas density in a broad spiral-shaped compression zone and thus trigger gravitational instabilities, which are the basic mechanism for initiating star formation. Star formation might then occur more or less concurrently over a broad spiral-shaped region with a patchy small scale structure, a picture which seems to correspond more closely to the observations. A spiral shock front may also be produced if there is a smoothly distributed component of the gas, but may play a less dominant role in star formation, being responsible for only part of the star formation.

It is worth recalling that the main reason for the conventional view that most star formation occurs in spiral arms is that the most conspicuous young objects, namely HII regions, tend to line up along narrow arms. It may be, however, that HII regions are not typical of star formation in general but only appear in special circumstances associated, perhaps, with strong shock compression of the gas. Some 3-dimensional collapse calculations to be published later, as well as the observed locations of young stars and protostars of different masses, suggest that the most massive stars tend to form in the densest clouds. If the densest clouds are formed preferentially in narrow spiral shock regions, the mass spectrum of star formation in such regions may not be typical but may favor the most massive stars. The narrow arms defined by the HII regions may then not represent the loci of star formation in general, but only the loci of formation of the most massive stars. Galaxies without strong, well-defined spiral patterns may not even form stars massive enough to make HII regions; this would be consistent with the observation that some galaxies like M104 and NGC 2841, which have no strong spiral patterns, contain knots of young stars but no HII regions and hence presumably no O stars (van den Bergh 1976b; Kormendy, this conference). It would also be consistent with the fact that type II supernovae, which are believed to have massive progenitors, appear to occur only in the spiral arms of Sb and Sc galaxies, which have prominent spiral structure (Tinsley 1977).

5. MEASURES OF STAR FORMATION RATES IN GALAXIES

Another possible way to gain information about the factors or processes that control star formation is to obtain quantitative measures of the total star formation rates in galaxies and see if they are correlated with other properties such as morphology which may be indicative of the dynamics of the system. Although it would be even more useful to have detailed measures of star formation rates as a function of spatial position in galaxies, such data do not yet exist in quantity, so we shall here consider only integrated quantities. For statistical purposes it is desirable to have data for a large number of galaxies, and for comparison with synthetic models it is most useful to have colors on a standard photometric system; consequently, UBV colors are generally the most useful indicators of star formation rates in galaxies (Searle, Sargent, and Bagnuolo 1973; Larson and Tinsley 1977, hereafter LT). Other measures of star formation rates, such as emission line strengths, yield results that are generally consistent with those obtained from the UBV colors (LT).

If star formation is initiated or enhanced by large-scale dynamical disturbances or shock fronts, then the star formation rate should be particularly high in galaxies that have been violently disturbed by encounters with other galaxies, or in galaxies that are very young and still collapsing, if such systems can be found. This suggests that a study of the relation between color and morphology for various types of peculiar or interacting galaxies may be instructive. For this purpose LT compiled data on the UBV colors of normal galaxies in the Hubble Atlas (Sandage 1961) and peculiar galaxies in the Atlas of Peculiar Galaxies (Arp 1966), using colors primarily from the Second Reference Catalog of Galaxies (de Vaucouleurs, de Vaucouleurs, and Corwin 1976). They found that the Hubble galaxies nearly all fall along a narrow sequence in the two-color (U-B, B-V) diagram, and can be interpreted as systems with various monotonically declining rates of star formation, ranging from constant in time for the irregular galaxies to a single initial burst for the elliptical galaxies (cf. also Tinsley 1968, Searle, Sargent, and Bagnuolo 1973). In contrast, the peculiar galaxies show a much greater scatter in the 2-color diagram, both above and below the sequence of normal galaxies. Among the various effects possibly responsible for this large scatter in colors, the most important is probably anomalous star formation histories characterized by recent bursts, as suggested by Searle, Sargent, and Bagnuolo (1973) to explain the extremely blue colors of certain peculiar compact galaxies.

With the aid of a new grid of models for galaxies with bursts of star formation, LT were able to show that the distribution of Arp galaxies in the 2-color diagram can be explained if many of them experience bursts of duration $\gtrsim 2 \times 10^7$ yr, in which the mass of stars formed may be as high as a few percent of the mass of the sytem. A search for possible correlations between anomalous colors and morphological peculiarities yielded the striking result that nearly all of the large scatter of the Arp galaxies in the 2-color diagram is caused by interacting or possibly interacting systems; other morphological peculiarities apparently have relatively little effect on the colors, and the non-interacting Arp galaxies have a color distribution similar to that of the Hubble galaxies. This result provides very suggestive evidence that violent dynamical disturbances can produce bursts of rapid star formation. Further support for this inference is that interacting systems whose tidal distortions are less than $\sim 5 \times 10^7$ years old, as inferred from the absence of prominent tidal tails on the side of each galaxy away from the companion, have colors that correspond to burst ages smaller than $\sim 5 \times 10^7$ years, while systems with well-developed tidal tails have colors corresponding to burst ages up to $\sim 2 \times 10^8$ years, consistent with the expected lifetimes of strong tidal distortions (e.g. Toomre and Toomre 1972, Lynds and Toomre 1976).

Thus there is evidence that large enhancements in the star formation rate can be triggered by violent dynamical disturbances, although the detailed mechanisms of star formation involved are not clarified by these results, which refer only to the integrated colors of galaxies. It is possible, for example, that rapid star formation could be caused by a transient tidally-induced density wave of large amplitude, by a direct high-velocity collision of the gas in the two galaxies, or by the rapid infall of gas into the dense nuclear region of one or both galaxies. Further observations to determine the location of the regions of most active star formation will be important to distinguish between these possibilities. In at least one case, detailed observations of a highly distorted and probably colliding system of galaxies show a sharply defined, narrow zone of intense star formation at the edge of one of the galaxies, probably caused by a shock front (C. R. Lynds, private communication). It is clear that detailed color mapping of galaxies will play an important role in further attempts to clarify the mechanisms of star formation in different types of galaxies.

6. CONCLUSIONS

The basic process of star formation is the occurrence of gravitational instability and collapse in the densest parts of the gas in galactic discs and protogalaxies, and therefore the rate of star formation depends on the rate at which the requisite conditions of high density and low temperature are produced. In the absence of large-scale dynamical effects, gravitational instabilities can be triggered by the dissipation of turbulence and the settling of the gas in a galactic disc into a thinner and denser layer; the controlling timescale for star formation is then the dissipation timescale, which depends on the period of motions perpendicular to the galactic plane and hence on the density of matter in the disc. Gravitational instabilities can also be triggered by shock fronts or compression waves resulting from the collapse of protogalaxies, collisions between galaxies, or density waves generated in galactic discs by large-scale instabilities or tidal interactions. If density waves (whether of bar or spiral form) are important in mediating star formation, the governing timescale is basically the orbital period, which depends on the mean density of matter interior to the point considered. Likewise, in a collapsing protogalaxy the dominant timescale for collisions, etc., is the free-fall time, which depends on the mean interior density. If interactions between galaxies are important, the relevant timescale depends on the frequency of interactions, which depends on the density of surrounding galaxies.

A common feature in all cases is the importance of dynamical processes and dynamical timescales that depend on the total density of matter in the region considered; depending on the scale of the dominant processes, the relevant density may be the local density of a galactic disc, the mean interior density of the galaxy, or the density of galaxies (or pre-galactic gas clouds) in a region of strong interactions or collapse. The resulting prediction that the timescales for star formation should be shortest in the densest regions is consistent with many observations: (1) The fractional gas content is generally smallest in the most condensed stellar systems, i.e. the elliptical galaxies and the bulge components of early-type spirals, and highest in the loosest systems, such as irregular galaxies and the outermost parts of late-type spirals. (2) The fractional surface density of hydrogen generally increases outward in the discs of spiral galaxies. A simple quantitative explanation of this based on density wave theory was suggested by Oort (1974). (3) The observed outward decrease in the abundance of heavy elements in spiral discs is qualitatively explainable by the decreasing completeness of conversion of gas into stars. (4) The gas content of galaxies is correlated with their surroundings in that galaxies in dense surroundings tend to be gas-poor (e.g. Oemler 1974), while the most gas-rich galaxies are found in low-density surroundings. Also, the ratio of elliptical galaxies, which result from rapid star formation, to disc galaxies, which result from slower star formation, is highest in the densest regions.

Thus many of the most basic properties of galaxies can be qualitatively understood if the timescale of star formation depends, in zero-order approximation, on the total mass density in the region considered. Although this differs from the original assumption of Schmidt (1959) that the star formation rate depends only on the gas density, the two hypotheses lead to similar conclusions for many purposes, since most star formation occurs in circumstances where the gas density is at least locally comparable to the total density. Thus, for example, both power-law functions of gas density and dynamical-timescale assumptions for the star formation rate in a collapsing protogalaxy can be made to yield acceptable models for the formation of elliptical galaxies (Larson 1974, 1975).

Of course, it is still greatly oversimplified to regard the timescale for star formation as a function of only a single parameter, especially if small scale features or rapid time variations are of interest. For example, star formation will be inhibited in local regions where the gas density is below the threshold density required for self-gravity to exceed tidal forces. Gas flows can also play a crucial role in determining the rate of star formation, especially if the gas density is close to a threshold value, since the star formation rate may then be very sensitive to the amount of gas present. Perhaps a good example of all of these effects at work occurs in active galactic nuclei showing evidence for vigorous star formation or more exotic but related forms of activity. Since the density of galactic nuclei is very high, all of the timescale arguments would suggest a very short timescale for the conversion of gas into stars or collapsed objects; hence the observed presence of gas in these nuclei is probably an indication of recent gas inflow. Because the gas must accumulate to a high threshold density before gravitational instability and star formation can occur (cf. also Mathews 1972), the formation of particularly massive stars may be favored. The rapid collapse of a dense gas cloud in a galactic nucleus may then lead to an intense burst of formation of massive stars, supernovae, pulsars, etc., possibly accounting for many of the phenomena seen in active nuclei.

As usual, the strongest conclusion that can be drawn is the necessity for further observations to clarify how and where star formation takes place in galaxies, and what factors determine its rate. Peculiar as well as normal galaxies will be important to study, since they may provide ready-made "experiments" in which galaxies have been subjected to various extreme conditions, and the effect of these on star formation can be observed. The multicolor detailed photometric mapping of galaxies which is now possible holds much promise in this regard, but it is also important to have more detailed information on the distribution and physical properties of the gas in galaxies, since an understanding of star formation must ultimately depend on understanding the dynamics of the gas. Thus observations at many wavelengths, including radio, will contribute indispensably to a better understanding of star formation.

REFERENCES

Arp, H. C., 1966. Atlas of Peculiar Galaxies. California Institute of Technology. Also in Ap.J. Suppl., 14, 1.

Baade, W., 1963. The Evolution of Stars and Galaxies. Harvard University Press. Reprinted by MIT Press, 1975.

Balick, B., Faber, S. M., and Gallagher, J. S., 1976. Ap.J., 209, 710.

Bardeen, J. M., 1975. IAU Symposium No. 69, Dynamics of Stellar Systems, ed. A. Hayli, p. 297. D. Reidel Publishing Co.

de Vaucouleurs, G., and de Vaucouleurs, A., and Corwin, H. G., 1976. Second Reference Catalog of Bright Galaxies. University of Texas Press.

Faber, S. M., and Gallagher, J. S., 1976. Ap.J., 204, 365.

Freeman, K. C., 1976. Galaxies, Sixth Advanced Course of the Swiss Society of Astronomy and Astrophysics, ed. L. Martinet and M. Mayor, p. 1. Geneva Observatory.

Gisler, G. R., 1976. Astr. Ap., 51, 137.

Goldreich, P., and Lynden-Bell, D., 1965. M.N.R.A.S., 130, 125.

Gordon, M. A., and Burton, W. B., 1976. Ap.J., 208, 346.

Gott, J. R., 1977. Ann Rev. Astr. Ap., 15, in press.

Hubble, E., 1936. The Realm of the Nebulae. Yale University Press. Reprinted by Dover Publications, 1958.

Jensen, E. B., Strom, K. M., and Strom, S. E., 1976. Ap.J., 209, 748.

Kennicutt, R. C., and Hodge, P. W., 1976. Ap.J., 207, 36.

Larson, R. B., 1972. Nature, 236, 21.

Larson, R. B., 1974. M.N.R.A.S., 166, 585.

Larson, R. B., 1975. M.N.R.A.S., 173, 671.

Larson, R. B., 1976a. M.N.R.A.S., 176, 31.

Larson, R. B., 1976b. Galaxies, Sixth Advanced Course of the Swiss Society of Astronomy and Astrophysics, ed. L. Martinet and M. Mayor, p. 67. Geneva Observatory.

Larson, R. B., and Tinsley, B. M., 1977. Ap.J., under deliberation.

Lynds, C. R., and Toomre, A., 1976. Ap.J., 209, 382.

Madore, B. F., 1977. M.N.R.A.S., 178, 1.

Madore, B. F., van den Bergh, S., and Rogstad, D. H., 1974. Ap.J., 191, 317.

Mathews, W. G., 1972. Ap.J., 174, 101.

Morgan, W. W., 1958. Publ. A.S.P., 70, 364.

Oemler, A., 1974. Ap.J., 194, 1.

Oort, J. H., 1970. Astr. Ap., 7, 381.

Oort, J. H., 1974. IAU Symposium No. 58, The formation and Dynamics of Galaxies, ed. J. R. Shakeshaft, p. 375. D. Reidel Publishing Co.

Quirk, W. J., and Tinsley, B. M., 1973. Ap.J., 179, 69.

Rees, M. J., and Ostriker, J. P., 1977. M.N.R.A.S. 179, 541.

Roberts, M. S., 1976. Galaxies and the Universe, ed. A. Sandage, M. Sandage, and J. Kristian, p. 309. University of Chicago Press.

Sandage, A., 1961. The Hubble Atlas of Galaxies. Carnegie Institution of Washington.

Sandage, A., Freeman, K. C., and STokes, N. R., 1970. Ap.J., 160, 831.

Sanders, R. H., and Huntley, J. M., 1976. Ap.J., 209, 53.

Schmidt, M., 1959. Ap.J., 129, 243.

Schmidt, M., 1963. Ap.J., 137, 758.

Schweizer, F., 1976. Ap.J. Suppl., 31, 313.

Searle, L., Sargent, W. L. W., and Bagnuolo, W. G., 1973. Ap.J., 179, 427.

Shu, F. H., Milione, V., Gebel, W., Yuan, C., Goldsmith, D. W., and Roberts, W. W., 1972. Ap.J., 173, 557.

Talbot, R. J., and Arnett, W. D., 1975. Ap.J., 197, 551.

Tinsley, B. M., 1968. Ap.J., 151, 547.

Tinsley, B. M., 1977. Supernovae, ed. D. N. Schramm, p. 117. D. Reidel Publishing Co.

Toomre, A., and Toomre, J., 1972. Ap.J., 178, 623.

Toomre, A., 1974. IAU Symposium No. 58, The Formation and Dynamics of Galaxies, ed. J. R. Shakeshaft, P. 347. D. Reidel Publishing Co.

van den Bergh, S., 1976a. Ap.J., 206, 883.

van den Bergh, S., 1976b. A.J., 81, 797.

van den Bergh, S., 1977. The Galaxy and the Local Group, ed. R. J. Dickens and J. E. Perry, Royal Obs. Bull., 182, in press.

Woodward, P. R., 1975. Ap.J., 195, 61.

Woodward, P. R., 1976. Ap.J., 207, 484.

DISCUSSION

FABER: You showed UBV colors of models with bursts of star formation in initially very red systems. What happens to your conclusions if the burst occurs in a galaxy somewhere else on the Hubble sequence?

LARSON: We studied a whole series of models with bursts in galaxies of different initial colors. If the initial galaxy is bluer, the evolution in the two-color diagram is rather similar except that the starting point is displaced toward bluer colors. The case I showed, i.e. bursts in a red galaxy, leads to the least extreme burst parameters needed to account for any given UBV colors.

SEGALOVITZ: How do you explain the colors of galaxies at the lower right hand side of your U-B vs. B-V diagram of peculiar galaxies? These galaxies are outside the bounds of your theoretical curves.

LARSON: I would guess that these very red colors are probably due mainly to internal reddening in these galaxies, which would displace their colors in this direction. Although internal reddening has in principle been corrected for, some galaxies (especially peculiar ones) may have more internal reddening than was assumed for them on the basis of morphological type.

HUCHRA: How can you get corrected colors for interacting galaxies when the de Vaucouleurs corrections depend on morphological type and the interacting systems are not easily typable?

LARSON: We did not attempt to rediscuss the corrections but assumed that all of the required corrections had been properly made by de Vaucouleurs. I cannot say how well it is possible to determine these corrections for peculiar galaxies, but they might well be very uncertain. Because of this, we replotted all of the samples using colors corrected only for local galactic reddening. Although some individual colors were significantly affected, the overall distributions were not altered and the same correlations were seen.

G. DE VAUCOULEURS: The calculated colors for galaxies with bursts are for stellar populations only. The observed colors include the contribution of emission lines, and I suspect that this contribution is not negligible. It is, therefore, dangerous to draw too precise conclusions from a comparison of models and observed colors.

LARSON: The contribution of emission lines was estimated by Huchra to be unimportant in most cases. Perhaps he would like to comment on this point.

HUCHRA: For the bluest galaxies, the colors of the gaseous emission lines and continuum resemble the colors of the exciting stars (at low redshifts). For systems much redder than U-B $\sim$ -0.3, the contribution of emission to the integrated colors is less than a few percent.

LASKER: With regard to the UBV studies of the Arp interacting galaxies without tails, which form a highly selected group, is there any further information to be found in the kinds of galaxies that are interacting?

LARSON: This is an interesting question, and I have the impression that several of the interacting systems without tails have a distinctive appearance that suggests that the galaxies are falling together (although this is admittedly a subjective impression). The galaxies in our sample which seem to define a distinct morphological group with common properties are Arp 118, 125, 140, 141, 142 and 143. C.R. Lynds (private communication) has also remarked upon the distinctive appearance of these systems and has noted that they appear to contain both a smooth old component and a very irregular, distorted young component (see also Freeman, this conference). It would evidently be very interesting to try to discover by more detailed studies what is happening in these systems. Perhaps Alar Toomre would like to comment on the possibility that they are falling together?

TOOMRE: I have no eye for telling whether galaxies are approaching each other!

RUBIN: NGC 5426/5427 is an interacting double galaxy in the Arp Atlas, and velocity measures show the bridges between the two to be stretching. That is, the sense of rotation in each galaxy is to pull the bridge apart. I think it is difficult to determine the sign of ΔV from the appearance of the galaxy.

KING (to Larson): How does your picture cope with the fact that the ellipticals apparently had the most vigorous rate of star formation at a time when their density was very low? Even worse, what about the Galactic halo?

LARSON: Although the density in the outer envelope of an elliptical galaxy may be low, it is located in a region of space with a higher than average density, both because elliptical galaxies are relatively concentrated and because they are preferentially located in dense clusters. This implies that dynamical timescales are relatively short, so processes whose rate depends on dynamical timescales could be relatively fast. An additional factor that may help star formation in elliptical galaxies is that the motions are highly non-circular, so that the relative velocities between colliding gas elements are very large and the potential amount of shock compression is also large. I would agree that it seems difficult to form stars in gas of very low density, but I have always imagined that even in a region of low average density the stars form in small regions of locally greatly enhanced density, produced either by shock compression or by gravitational collapse of parts of the protocloud to form at least temporarily bound systems such as star clusters or small galaxies which are later dispersed.

ICKE: Don't you think that a Parker-type magnetohydrodynamic instability could equally well initiate star formation?

LARSON: Maybe, but I have seen little evidence for this. The only evidence that I know of comes from polarization studies in the vicinity of young clusters which seem to show that the magnetic field lines are displaced toward the galactic plane in the way predicted by the Parker instability. Whether this means that star formation has occurred as a result of a Parker instability or simply that the magnetic field has been entrapped in a collapsing region without much affecting its dynamics is not clear. I don't know of any definitive tests that could distinguish whether the Parker mechanism is of general importance for star formation.

BAUM: You mentioned some data concerning the gradient of stellar type across a spiral arm, and you sounded skeptical that the data show such a gradient. The presence of a detectable gradient may depend on which galaxies we look at. Using an 8-color system, I made some photometric scans across arms in M74 and found evidence there for a stellar type gradient, roughly the amount one might expect if material flows through a nearly stationary arm. The results were published in IAU Symposium No. 24, *Spectral Classification and Multicolor Photometry*, p. 288, (1966).

LARSON: It is interesting that these results show evidence for star formation near the *outer* edge of the arm; current density wave theory would predict star formation near the *inner* edge.

YUAN: I would like to make some comments on the results of Schweizer which you quoted. The statement that the color variation across the spiral arm seems much broader than the theory predicts is based on the assumption that newly formed stars will follow the streamlines in the post-shock region. This is not true. The clouds in the post-shock region usually have a circular velocity lower than the local mean circular velocity. Therefore, after stars are formed, they will deviate from the streamlines and move radially inwards for a few tens of millions of years, sufficiently broadening the spiral arms as outlined by galactic shocks. This is also confirmed by Dr. R. Wielen's recent calculations. In addition, according to the results of shock-driven implosion calculations by Dr. P. Woodward, the star formation process will take place for about 10 million or even 20 million years. This also broadens the post-shock regions. So the dust lanes as well as the giant HII regions may define sharp spiral arms, but the color variation across the spiral arms should not be sharp. In fact, it is expected to be relatively broad, according to density wave theory.

LARSON: It is true that the effects you mention could smear out the region of star formation and reduce or eliminate any color gradients across spiral arms. However, one is then left without any unambiguous test of the idea that star formation takes place in a progressive wave propagating with respect to the gas. Although there is evidence for density-wave-driven shocks in the narrow dark lanes seen at the inner edges of spiral arms, there is little direct evidence that star formation is initiated by these dark-lane shock fronts. The dark lanes might only be incidental, or they might be responsible for only a part of the star formation.

TOOMRE: Despite your good cautionary remarks, we ought not to doubt that stars (or at least some stars) are nudged into existence by density and/or shock waves. After all, narrow strings of HII regions virtually guarantee that such mechanical persuasion must be considerable in settings like M51. And it is also reassuring that Nature seems to have carried out some analogous experiments with more circular waves of compression in the so-called ring galaxies - like Zwicky's old "Cartwheel" studied recently by R.A.E. Fosbury and T.G. Hawarden (1977, M.N.R.A.S. 178, 473). There again the rule seems to be: You crunch gas, and you get plenty of light!

BASH: I would like to report some work which I, Elizabeth Green and William Peters III have done which tends to support the density wave as the origin of star formation, at least in our Galaxy. We have devised a model which describes the Galactic orbits of CO-emitting molecular clouds. The model assumes that the molecular clouds are launched in the Galactic density wave and it predicts that the clouds will produce detectable CO emission for 30×10^6 years from birth, after which time there is an abrupt cut-off of the CO emission. A separate set of observations of the CO emission associated with young open star clusters reveals associated CO emission if the cluster contains O stars, but if the main-sequence turn-off is B0 or later, no CO is present. These later observations confirm the cut-off predicted by the model and, in addition, may allow us to use CO emission to trace the distribution of young open star clusters and to deduce their birth rates.

TALBOT: I would like to show some slides produced by myself, R. Dufour, and E. Jensen which show the distribution of star formation in M83 (NGC 5236). The first slide shows a composite B, V and R photograph of this intermediate barred spiral galaxy. There are strong dust lanes coming out of the ends of the bar. They and the bright stars in their vicinity outline a classical two-armed spiral structure. U, B and V photographic photometry has been used to determine surface brightnesses for 500 x 500 pixels of 1".5 square. We assume that there is an underlying disk population which varies in intensity and color only over distances much greater than a pixel. We search for sets of adjacent pixels with uniform color and smoothly varying intensity; we identify them as the local underlying disk and subtract that intensity from the other pixels. The remaining light is assumed to be produced by groups of stars of a single age. With the assumption that these groups may be identified with clusters of varying age and amount of extinction, two-color-plane analysis gives the age of each group. A color-coded map is shown which gives the ages of the stars in each pixel.

The map of stellar ages shows what appears to be two separate patterns of star formation. In the region just outside the ends of the bar there are strong dust lanes, immediately outside of which there are well defined arms of HII regions and young blue stars. This pattern is a clear case of a two-armed spiral. Further out in the disk, however, the young stars are in an irregular patchy pattern. The pattern appears to be fragments of spiral form, but that is largely due to shearing of an irregular pattern.

It appears that there are two dynamical processes which produce star formation in the galaxy. The inner pattern appears to be a large-scale spiral shock wave (perhaps driven by the bar). I believe that the outer region may be attributable to the disk instability of Goldreich and Lynden-Bell. If that instability leads to chaotic motions of gas fragments and those fragments collide, there should be shocks and star formation in a fashion analogous to, but less well-organized than, the global spiral shock pattern.

HODGE (to Larson): In galaxies with a large number of HII regions, the half-width of the arm they define is the same, within the uncertainties, as the half-width of the arm in blue light. So I think you have a beautiful explanation of something that isn't generally the case.

LARSON: Isn't it true that sharply defined arms are rare in any case?

HODGE: Yes, they are rare, but even where they occur the HII regions have the same half-width as the intensity.

LARSON: So not even in that case do you find a separation between where the HII regions form and where the other stars form.

HODGE: Right. I think the Percival Lowell effect comes in here - you tend to see these HII regions in a nice line.

KING (to Larson): Regarding your suggestion that there might be two modes of star formation, with and without massive stars, you could as well have been talking about OB associations versus T associations. There ought to be an opportunity to follow this topic further by studying the distribution of these two types of associations in our own Milky Way.

LARSON: Indeed I also had in mind the distinction that you mention between OB and T associations; there seems to be considerable evidence that low-mass T Tauri stars can form widely throughout dark cloud regions, while the most massive stars form only near the centers of particularly dense and massive collapsing molecular clouds. In other words, the formation of massive stars requires higher densities than the formation of less massive stars, so that massive stars might only form under more restrictive conditions associated, perhaps, with strong shock compression or with the pile-up of dense gas in galactic nuclei. There are also some theoretical indications that the formation of massive stars may be favored in very massive and condensed clouds. One of these is that some three-dimensional collapse calculations on which I have been working recently indicate that the mass of the largest condensation that forms in a collapsing cloud is independent of the Jeans mass but increases with increasing cloud mass and decreasing angular momentum, so that the most massive objects form in the clouds that become most condensed. Another is that the upper mass limit for star formation set by effects such as radiation pressure in collapsing protostars increases with increasing ambient density, and may well not allow the formation of O stars except in the densest clouds.

ICKE: Currently, Frank Israel and I are doing a pilot study of the Fourier transforms of galaxies in Hα to look for exactly the thing Ivan King mentioned: two characteristic scales of star formation.

KORMENDY: Half of this comment consists of a complication of the galaxy morphology discussed earlier at this meeting, and half is directed at the question of how a density wave affects star formation.

When we speak of "spiral galaxies", we always think of beautiful two-armed spirals such as M83 or M51. But Sandage, in the Hubble Atlas, makes a strong distinction between these and a group of galaxies with very different behavior. These objects have many short spiral filaments, but no global spiral structure. Prototypes shown in the Hubble Atlas include NGC 2841, NGC 7217 and NGC 5055. Work is in progress to investigate the interpretation of these objects. However, the impression to date is that the main difference between NGC 2841 and two-armed spirals is that the former have failed to make a density wave.

In partial support of this, I can cite statistics of spiral structure types in a morphological survey of the 121 brightest northern barred spirals. Of galaxies late enough to show spiral structure, 55 had two-armed or multi-armed global patterns, commonly interpreted as density waves. None had pure NGC 2841 structure. Six objects were intermediate; the bars in these objects were relatively weak. On the other hand, even a cursory examination of the Hubble Atlas reveals many pure NGC 2841 spirals. This supports the widely-held feeling that a bar helps to drive a density wave: barred galaxies essentially never fail to produce one, non-barred spirals fail often.

The star formation evidence in NGC 2841-type spirals favors the existence of two types of star formation. Sandage has stated, and my Hα plates confirm, that NGC 2841 has essentially no HII regions. Similarly, the HII regions in NGC 5055 are unusually small for an Sc galaxy of its luminosity. But both galaxies are clearly still making stars. Evidently (1) NGC 2841 spirals succeed in making stars without any help from a density wave, but (2) two-armed spirals make stars more vigorously. In particular, the mass function seems different, in that NGC 2841 galaxies cannot make massive stars. Van den Bergh (1976, Astron. J. 81, 797) has also called attention to an abnormal mass function in spirals like M104, which are likely, on morphological grounds, to have NGC 2841-type structure. Finally, note that the difference between density wave and non-density wave star formation may only be one of degree, not a qualitative difference.

PUGET: Some recent data on the integrated infrared emission from the galactic plane indicate a change in the slope of the initial mass function with galactic radius. The number of less-massive stars formed in the 5 kpc ring of molecular clouds is relatively larger than the number of massive stars (as traced by HII regions), by a factor of 5 or so, when compared to the same ratio for the solar neighborhood (Serra and Puget, to be published).

LARSON: There seems to be increasing evidence that the customary assumption of a uniform IMF in galaxies may not be correct, and this would clearly have many fundamental implications for galactic evolution.