1980ApJ...237..692L

THE ASTROPHYSICAL JOURNAL, **237**:692–707, 1980 May 1

THE EVOLUTION OF DISK GALAXIES AND THE ORIGIN OF S0 GALAXIES

RICHARD B. LARSON, BEATRICE M. TINSLEY, AND C. NELSON CALDWELL
Yale University Observatory
Received 1979 July 25; accepted 1979 October 31

ABSTRACT

We reconsider the relation between spiral and S0 galaxies in the light of recent data on the colors and morphology of disk systems, and on the content of clusters at different redshifts. Star formation will strongly deplete the gas in most spirals in a fraction of the Hubble time, so we suggest that the gas in spirals has been replenished by infall from residual envelopes, probably including gas-rich companions and tidal debris. S0's may then be disk systems that lost their gas-rich envelopes at an early stage and consumed their remaining gas by star formation. This picture is consistent with the colors of S0's if most of their star formation stopped at least a few gigayears ago, and it is consistent with their small disk-to-bulge ratios relative to spirals, since this is a direct result of the early truncation of star formation. Numerical simulations show that the gas envelopes of disk galaxies in clusters are largely stripped away when the clusters collapse, but star formation can continue in the spirals for several gigayears while their remaining disk gas is consumed. These results can explain the blue galaxies observed by Butcher and Oemler in two condensed clusters at $z \sim 0.4$: these clusters are seen just before most of their galaxies run out of gas, so that star formation is still occurring in them but will soon die out, causing the spirals to evolve into S0's with normal present colors. A rapid evolution of the galaxy content of condensed clusters is predicted at moderate redshifts, ranging from a large fraction of blue galaxies at $z \sim 0.4$ to very few at $z \sim 0$.

Subject headings: galaxies: evolution — galaxies: formation

I. INTRODUCTION

Disk galaxies are conventionally characterized by two main parameters: the relative prominence of the disk and bulge components, and the relative content of gas and young stars in the disk. Of these, the gas content seems to be the one most closely related to the Hubble type, and it has often been suggested in particular that S0 galaxies differ from spirals basically just by their lack of gas and young stars (e.g., Baade 1963; Sandage, Freeman, and Stokes 1970; van den Bergh 1976*a*). A popular idea has been that S0's are "stripped spirals" from which the gas has been removed, either by collisions (Spitzer and Baade 1951), or by sweeping by intergalactic gas in clusters (Gunn and Gott 1972), or by galactic winds (Faber and Gallagher 1976). The importance of sweeping by cluster gas is suggested by the presence in many clusters of a relatively high proportion of S0 galaxies (Oemler 1974; Melnick and Sargent 1977) and "anemic" spirals (van den Bergh 1976*a*). However, sweeping cannot account for the many field S0's, for which an internal mechanism for removing the gas appears necessary; also, Dressler (1980) has shown that the ratio of S0's to spirals increases with the local density of galaxies, and exceeds the ratio in the field even in regions where the density is too low for gas to be stripped from the disks of spirals.

Several recent photometric studies place strong constraints on any evolutionary picture in which S0 galaxies are former spirals whose gas has been depleted. Sandage and Visvanathan (1978*b*), in an extensive study of the colors of early-type galaxies, find no evidence for recent star formation in S0's, and argue that they are not produced by the stripping of spirals but have a separate origin. Burstein (1979*a*, *b*) also argues that S0's cannot be stripped spirals because they have smaller disk-to-bulge ratios and possess a "thick disk" component not found in spirals. Dressler (1980) finds in addition that the absolute bulge sizes of S0's are systematically larger than those of spirals, in apparent contradiction to an origin of S0's as spirals whose gas has been depleted.

On the other hand, Butcher and Oemler (1978*a*, *b*) present evidence favoring the recent transformation of spirals into S0's in clusters: two centrally condensed clusters at redshifts ~0.4 are found to contain large numbers of blue, presumably spiral, galaxies, whereas nearby clusters that are structurally similar contain few spirals and consist mainly of E and S0 galaxies. While Butcher and Oemler's observations favor an evolutionary relation between spiral and S0 galaxies, they pose a problem for a simple stripping model, since it is difficult to understand why the spirals in the distant condensed clusters have not already been stripped.

In this paper, we consider again whether spiral galaxies can evolve into S0's as a result of gas loss, and whether this is consistent with all of the observations mentioned above. First we discuss the evolution of the gas content of disk galaxies (§ II), and show that in most spiral galaxies the time scale for gas consumption by star formation is considerably shorter than the Hubble time. Thus, within a Hubble time, most spirals may evolve into gas-poor systems resembling S0's, unless their gas is somehow replenished. We therefore suggest that the lifetimes of most spirals have been prolonged by accretion of gaseous matter from their surroundings, perhaps including gas-rich companion galaxies and tidal debris as well as primordial gas; and we propose that the S0 galaxies are former spirals that lost their external gas supply at a relatively early stage. In § III, we use new models to discuss the color evolution of disk galaxies whose star formation is truncated at various times, and we show that the observed colors of S0's can be explained if most of their star formation ceased at least a few gigayears ago. In § IV, we show that since the disk luminosity of a galaxy is significantly reduced by early truncation of star formation, the smaller disk-to-bulge ratios of S0's are consistent with their origin as former spirals. The "thick disks" of S0's may also be present in spirals, but normally masked by the much brighter thin disk.

In § V, we consider the implications of this picture for the evolution of disk galaxies in clusters, and we show by numerical simulations that any extended gas envelopes that may initially surround spiral galaxies will be mostly swept away by collisions at the time a cluster first collapses. Thereafter the spiral galaxies will begin to exhaust their remaining gas, and gradually evolve into S0's after a delay of typically several gigayears. We estimate that significant star formation in the blue galaxies in the distant clusters will die out soon after these clusters are observed, and that the galaxy content of these clusters will then evolve to resemble that of nearby dense clusters by the present time. Finally, we predict schematically how the relation between the structure and galaxy content of clusters should depend on redshift, and we suggest further observations to test this picture. The conclusions are summarized in § VI.

II. EVOLUTION OF THE GAS CONTENT OF DISK GALAXIES

a) Gas Consumption by Star Formation

If there is no transfer of matter between a galaxy and its surroundings, star formation will eventually exhaust the gas and cause a spiral galaxy to evolve into a gas-poor system resembling an S0. The crucial question is whether this can be expected to happen within the Hubble time. To try to answer this question, we have estimated the time scales for gas consumption in many spiral galaxies whose gas contents are known. Such estimates were made previously by Roberts (1963), who concluded that the time required to exhaust the gas in most spirals exceeds the Hubble time, and that therefore no significant evolutionary changes in morphological type can be expected. This result supported the view that the distinction between spiral and S0 galaxies is intrinsic and not due to evolution (e.g., Sandage, Freeman, and Stokes 1970). However, Roberts included a large correction factor increasing the time scale to allow for gas recycling, an effect which is now considered to be much less important; if this factor is revised to agree with current estimates, Roberts's gas-consumption time scales become shorter than the Hubble time.

The most direct estimate of a gas-consumption time scale can be made for the solar neighborhood, where the data are most complete. The surface mass density of gas in the solar neighborhood is about 8 $M_\odot$ pc^{-2} (Gordon and Burton 1976), while recent estimates of the rate at which this gas is being consumed by star formation range from $\sim$3–7 $M_\odot$ pc^{-2} Gyr^{-1} (Miller and Scalo 1979) to $\sim$10 $M_\odot$ pc^{-2} Gyr^{-1} (Tinsley 1980). The time scale for gas consumption, $\tau_{\rm gas} \equiv |M_{\rm gas}/\dot{M}_{\rm gas}|$, is then between $\sim$0.8 and 2.7 Gyr, which is much shorter than the age of our Galaxy. Short time scales for gas consumption had previously been estimated by van den Bergh (1957) and by Larson (1972), who obtained values of $\sim$0.7 and $\sim$3 Gyr, respectively; these authors suggested that the local gas supply might have been sustained by inflow of gas, either from elsewhere in the Galaxy or from outside. A similar problem exists for our Galaxy as a whole: the total amount of gas is from $\sim 4 \times 10^9$ $M_\odot$ (Gordon and Burton 1976) to 6×10^9 $M_\odot$ (Solomon, Sanders, and Scoville 1979), and the total rate of gas consumption is estimated to be $\sim$4 $M_\odot$ yr^{-1} (Smith, Biermann, and Mezger 1978) or $\sim$2.6 $M_\odot$ yr^{-1} (Mezger, private communication); the resulting time scale for our Galaxy is thus $\sim$2 Gyr.

Estimates of gas lifetimes for other galaxies can be made using measured total H I contents together with star formation rates (SFRs) obtained from *UBV* colors via the models of Larson and Tinsley (1978, hereafter LT). We have used data from the *Second Reference Catalogue of Bright Galaxies* (de Vaucouleurs, de Vaucouleurs, and Corwin 1976) to calculate $\tau_{\rm gas}$ for most of the "normal" galaxies (LT's *Hubble Atlas* sample) for which H I indices and *UBV* colors are available. SFRs per unit stellar mass, $\dot{M}_{\rm star}/M_{\rm star}$, have been derived from the colors using Table 2 of LT, and $\dot{M}_{\rm gas}$ has been taken as 0.8 times $\dot{M}_{\rm star}$ to allow for gas recycling (Tinsley 1980). To relate $M_{\rm star}$ and hence $M_{\rm gas}$ to the blue luminosity L_B, the model mass-to-light ratios have been multiplied by a factor of 2 to bring them into agreement with the M/L ratios observed in the luminous parts of spiral galaxies (Faber and Gallagher 1979); this can be regarded as a correction for unseen faint stars not included in the models (LT). The total gas mass per unit luminosity, $M_{\rm gas}/L_B$, has been taken as twice the ratio of H I mass to L_B given by de Vaucouleurs, de Vaucouleurs, and Corwin (1976) in order to include the mass of helium and an allowance for molecular gas that assumes that 36% of the hydrogen is molecular; this is intermediate between the molecular fraction of $\sim$50% estimated for our Galaxy (Gordon and Burton 1976) and the fraction $\sim$15% found for two late-type galaxies by Morris and Lo

(1978). Five galaxies whose *UBV* colors lie too far from the normal ($U - B$, $B - V$) relation for an unambiguous estimate of their SFR have been excluded; four of these have $B - V < 0.4$, and the fifth, NGC 253, has $U - B$ bluer than normal for its $B - V$ by more than 0.1 mag.

The resulting time scales for 36 galaxies (nearly all spirals) show no significant correlation with color or morphological type, except that the two largest values are both for late-type galaxies, so the results are presented in a single histogram in Figure 1. The values of τ_{gas} range from 0.9 to 15 Gyr, with a median of 3.9 Gyr that is considerably shorter than the Hubble time. It should be noted that these time scales are based, following Faber and Gallagher (1979), on an assumed Hubble constant H_0 of 50 km s^{-1} Mpc^{-1}, and they would decrease as $H_0{}^{-1}$ if larger values of H_0 were considered.

These results suggest that if no gas replenishment occurs, most spiral galaxies will consume most of their gas in much less than a Hubble time. Although we cannot predict their detailed subsequent evolution, it is plausible that many of them will become extremely gas-poor systems with negligible star formation, resembling S0's. We therefore postulate that S0 galaxies are simply former spirals that consumed their gas by star formation. The problem then becomes not to understand why the S0 galaxies are gas-poor, but why many spirals have remained gas-rich despite rapid consumption of their gas by star formation. Some replenishment of the gas in spiral galaxies seems required, and we discuss next the possibility that continuing accretion of gas plays an important role in their evolution.

b) *Gas Replenishment*

Accretion of diffuse peripheral gas left over from galaxy formation was proposed by Oort (1970) and Larson (1972) as a continuing source of gas in spiral galaxies. They concluded that a present infall rate of order 1–2 $M_\odot$ yr^{-1} is theoretically plausible and is consistent with observations of the "high velocity clouds" around our Galaxy. Oort and Hulsbosch (1979) and Verschuur (1978) have suggested that several complexes of these clouds may represent streams of matter falling into the Galaxy. Sciama (1972) and Cox and Smith (1976) have also shown that infall of hot intergalactic gas at about the above rate would be consistent with X-ray observations. At infall rate of, say, 2 $M_\odot$ yr^{-1} would be comparable to the total rate of gas consumption by star formation in the Galaxy (§ II*a*), and it could therefore significantly prolong the galactic gas content. Gas inflow at such a rate would also help to explain the slow evolution of metallicity in the solar neighborhood, and would generally be consistent with constraints from chemical evolution (Tinsley and Larson 1978).

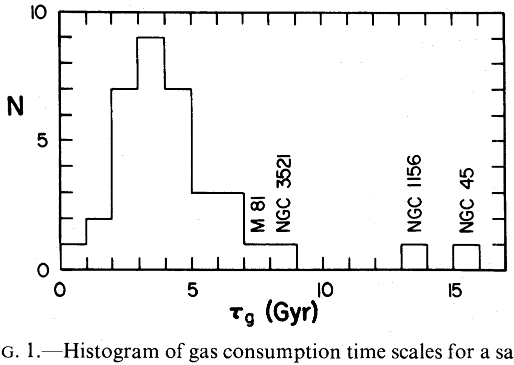

FIG. 1.—Histogram of gas consumption time scales for a sample of "normal" galaxies. A few galaxies with long time scales are labeled.

Recent 21 cm searches (Lo and Sargent 1979; Haynes and Roberts 1979; Haynes 1979; Davies 1979) have mostly found little or no evidence for isolated intergalactic hydrogen clouds in nearby groups of galaxies. Thus cool intergalactic clouds apparently do not provide a significant source of gas for spiral galaxies, at least at the present epoch. However, extended H I envelopes or streams that appear to be tidal debris are often found in groups of galaxies (Haynes 1979); the Magellanic stream is a local example of this phenomenon. In some cases a simple tidal model of material drawn out from a disk successfully explains the observations, while in other cases, such as M51 (Haynes, Giovanelli, and Burkhead 1978), a more complicated picture possibly involving some primordial gas seems required. If tidal interactions disperse the peripheral gas of a spiral galaxy into an extended envelope, this will have important implications for the evolution of the system, since the dispersed gas probably cannot form stars efficiently until it has fallen back into a disk. Thus envelopes or streams of tidal debris could serve as reservoirs of gas to fuel later star formation in spiral galaxies.

The time required for such tidal debris to settle back into a disk is at least ~1 Gyr, and is comparable to the orbital periods of the interacting galaxies. Thus repeated tidal encounters could keep the outermost gas of a spiral galaxy stirred up and relatively widely dispersed for many Gyr; moreover, any remaining primordial gas might also be kept stirred up and prevented from settling rapidly to a disk. This would prolong the time scale for gas consumption in the system; a possible example of this effect is the well-known tidally influenced galaxy M81, which has an unusually long gas consumption time of 8 Gyr (Fig. 1). Tidal interactions were probably much more frequent in the past (Toomre 1977), so such processes could plausibly have played a role in extending the lifetime of a large fraction of spiral galaxies.

A related phenomenon that may have sustained the gas supply of many spirals is the accretion of small gas-rich companion galaxies; for example, the Magellanic Clouds will probably eventually be accreted by our Galaxy (Tremaine 1976). Since small galaxies are generally more gas-rich than large ones, and some of them apparently consist mostly of gas (Lo and Sargent 1979), the accretion of smaller companion galaxies could replenish the gas supply of large spirals and significantly prolong their lifetimes.[1] For example, the total H I mass in the Magellanic Clouds and the Magellanic Stream is nearly $2 \times 10^9 M_\odot$ (Mathewson, Cleary, and Murray 1974), which is comparable to the mass of H I in the disk of our Galaxy.

We therefore suggest that most spirals may have possessed, for a significant fraction of their lives, extended gas-rich envelopes, perhaps containing gas-rich companion galaxies and tidal debris in addition to any remaining primordial gas, and that continuing accretion of this matter may have helped to sustain their gas contents up to the present time.

c) Gas in S0 *Galaxies*

The above picture of gas replenishment in spiral galaxies suggests that S0 galaxies may be disk systems that never possessed significant gas-rich envelopes or that lost their envelopes at a relatively early stage, and then exhausted their remaining gas by star formation. (Elliptical galaxies may then be systems whose gas supply was depleted or cut off even earlier, before the formation of any significant disk component [cf. Ostriker 1977].) Other proposed mechanisms for removing the gas from the disks of S0 galaxies may remain relatively ineffective until star formation has already reduced their gas content to a low level, since galactic winds do not affect the dense gas in the disk of a typical spiral galaxy (Bregman 1978), and sweeping by intergalactic gas is also relatively ineffective in the disks of gas-rich galaxies with a substantial rate of mass loss from young stars (Gisler 1979). Winds or sweeping become even less effective if much of the gas is in dense clouds, as is the case in our Galaxy. Thus the gas-poor nature of S0 galaxies is probably due mainly to gas consumption by star formation after replenishment by infall of peripheral gas has ceased to be significant.

The high ratio of S0's to spirals in dense clusters can then be understood as resulting from the sweeping away of residual gas-rich envelopes of disk galaxies when the clusters collapse (§ V*b*). In this sense, sweeping is still the process responsible for the formation of S0 galaxies in clusters, but the material swept is diffuse outlying gas rather than the dense gas in galactic disks, and the cessation of star formation is delayed until the remaining disk gas has been consumed. In § V*c* we discuss further how this delay may account for Butcher and Oemler's (1978*b*) observations of blue galaxies in some distant clusters that apparently have already collapsed.

Whatever initially depletes the gas in S0 galaxies, gas must be continually removed from them, or stellar mass loss will produce more than the observed amount of gas in $\lesssim 1$ Gyr (Faber and Gallagher 1976). Sweeping by intergalactic gas can plausibly keep the S0's in clusters gas-free once star formation has reduced their gas contents to a sufficiently low level, but field S0's require an internal removal mechanism. Faber and Gallagher (1976) favored galactic winds, but successful wind models require a supernova rate as high as that in elliptical galaxies (Bregman 1978), whereas the observed supernova rate in S0's is much lower (Oemler and Tinsley 1979). It is therefore doubtful that S0 galaxies can sustain winds.

Continuous star formation as a gas removal mechanism was tentatively rejected by Faber and Gallagher (1976), on the grounds that it would lead to colors bluer than are observed; we also conclude in Appendix A that ongoing star formation is probably rare in the sample of "normal" S0 galaxies discussed in § III and Appendix A. However, it seems possible that bursts of star formation occurring at intervals of $\lesssim 1$ Gyr could reduce the gas content of a galaxy below the observed limits, while leaving it looking like a normal S0 most of the time. We have used the models of LT to estimate the color changes that would occur during and after bursts if the average SFR equals the stellar mass-loss rate. Provided that star formation is in progress for $\lesssim 10\%$ of the time, the colors of the galaxy would resemble an early spiral for $\lesssim 25\%$ of its life, and it would be essentially as red as an undisturbed old population for most of the time. During star formation such a galaxy would probably have a smooth disk with patches of dust and young stars, and might resemble some of the smooth but noticeably dusty disk systems classified as $S0_3$ in the *Hubble Atlas* (Sandage 1961), some of which are also called anemic spirals by van den Bergh (1976*a*). Several of these galaxies are unusually blue for S0's, but they are not in the sample of "normal" S0's discussed below (§ III; Appendix A). Within this sample, NGC 5102 has unusually blue colors, and detailed studies have shown it to have patches of star formation (van den Bergh 1976*b*; Danks, Laustsen, and van Woerden 1979; Pritchet 1979). Thus it is possible that field S0's undergo occasional episodes of star formation during which they appear temporarily bluer than normal and are generally classified as slightly later morphological types, such as $S0_3$ or Sa, or as peculiar types. The large scatter in the gas contents of early-type disk galaxies predicted by this picture is also consistent with the observed scatter (Balick, Faber, and Gallagher 1976).

III. EVOLUTION OF COLORS

The colors of S0 galaxies give information on how long ago most of their star formation must have ceased, since relatively recent star formation would lead to bluer colors. It is particularly useful to compare the colors of S0 galaxies with those of ellipticals, whose stellar populations are believed to be almost entirely very old (Faber 1977). In this section, we first discuss recent data on E and S0 colors, then interpret the results in terms of constraints on past star formation in S0's.

a) Colors of Nearby E *and* S0 *Galaxies*

The main factor affecting the colors of E and S0 galaxies is believed to be metallicity, which increases

---

[1] To paraphrase Saslaw (1973), "galaxies are very nutritious."

systematically with galaxy luminosity, producing the observed correlation between color and magnitude (Faber 1973, 1977). Metallicity effects must therefore be taken into account before differences in star formation can be assessed.

We consider the large and homogeneous set of colors and magnitudes for E and S0 galaxies obtained by Sandage and Visvanathan (1978*b*, hereafter SV). Figure 2 shows SV's $u - V$ colors, corrected for galactic reddening, aperture, and redshift, plotted against absolute magnitude for six subsamples of their data. The figure includes only the galaxies they classified as E or S0, and omits intermediate and peculiar types and those with uncertain colors; galaxies fainter than $M_V = -20$ are also not shown, but they would not affect our discussion. The galaxies are subdivided in Figure 2 into three classes according to whether SV assigned them membership in clusters or dense groups (both called "clusters" here), looser groups, or the field. To aid in comparisons, a box including most of the cluster E's has been reproduced in each panel of Figure 2.

Inspection of this figure suggests some possible differences in the color-magnitude (C-M) distributions of the different subsamples. In particular, S0's in all environments seem to show more scatter about a mean C-M relation than do the cluster E's, and cluster S0's near $M_V = -21$ look redder than the corresponding E's. Also, group and field E's show more scatter than the cluster E's, and field E's brighter than $M_V = -22$ seem bluer than the corresponding cluster E's. Similar differences were noted previously by Faber (1977) and Burstein (1977) for different samples. In Appendix A we discuss the statistical significance of these differences, and suggest that the C-M relation for cluster E's reflects mainly a relation between metallicity and luminosity, with some intrinsic scatter, while the bluer colors of some field E's plausibly reflect residual star formation. However, since the S0's show no signs of being bluer, on average, than cluster E's of the same magnitude, it is important to discuss in more detail whether they could have younger stellar populations.

We must first know what colors S0 galaxies would have if their bulges and disks contained only very old stars. Stars in the disks of S0's are expected to be metal-rich, on both observational and theoretical grounds: (1) The old-disk stars in the solar neighborhood mostly have nearly solar metallicity, and the mean disk metallicity is probably even greater at smaller galactic radii. (2) Theoretical models for disk formation suggest a substantial early enrichment by massive stars in the young halo (Ostriker and Thuan 1975; Tinsley and Larson 1978). (3) If star formation in the disk is sustained by infall, or if the disk gas is completely consumed, as proposed in § II, models for chemical evolution predict that the mean stellar metallicity should approach the "yield," a quantity near the solar metallicity (e.g., Tinsley 1980). We therefore expect that the disk stars in S0's have a mean metallicity close to the yield. Possible evidence against such a high metallicity for all S0 disks is provided by data of Faber (1977, and private communication) indicating relatively weak metallic line indices for some S0 disks; however, line weakening would also occur if the disk population were relatively young, since the lines studied are weaker in bluer stars. The color of an old population with the yield metallicity can be estimated from the colors of cluster E's, since theories (e.g., Larson 1974; Tinsley and Larson 1979) predict a maximum metallicity for large ellipticals asymptotically equal to the yield; the appropriate color is therefore the maximum value for cluster E's, which is seen in Figure 2 to be $u - V \approx 2.5$.

A schematic model can now be constructed to predict the colors of purely old S0's as a function of luminosity, as illustrated in Figure 3. Each S0 is assigned a typical disk-to-bulge ratio of 1.2 in visual light (Burstein 1979*a*); the bulge component is assumed to have the color of an old E galaxy of its luminosity, according to the curve in Figure 3 labeled

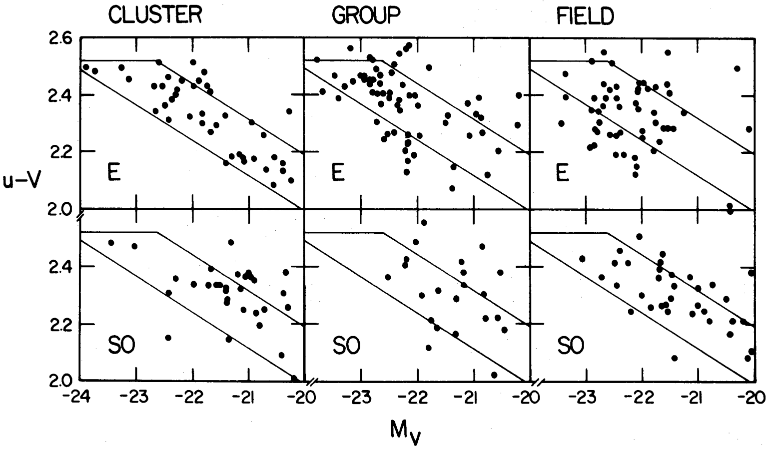

FIG. 2.—Color-magnitude plots from the data of Sandage and Visvanathan (1978*b*, Table 1). The ordinate is $u - V$ corrected by SV for reddening, aperture, and redshift effects. The boxes are drawn to aid comparison; they are the same in each panel, and chosen to include most cluster E's.

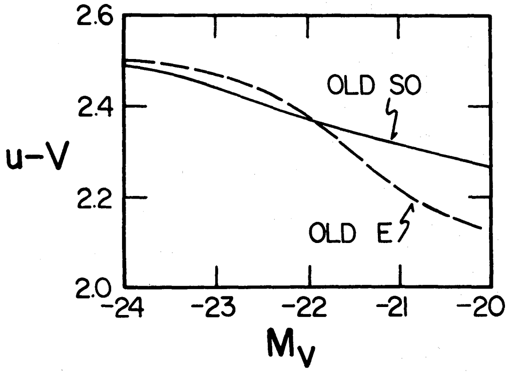

FIG. 3.—Color-magnitude relations for purely old stellar systems. The line "old E" is an eye-drawn curve through SV's cluster E's. The line "old S0" is a model for S0's with bulges following the "old E" relation and metal-rich disks with $u - V = 2.50$; S0 galaxies would have such a C-M relation if all their stars were old.

"old E," which is eye-drawn through SV's cluster E data; and the disk is assigned a color $u - V = 2.50$. This simple model predicts that S0's fainter than $M_V \sim -22$ would be redder than the corresponding cluster E's if their disk stars were all old.

We can therefore interpret the scatter in the colors of S0's in Figure 2 as follows. Those S0's that are redder than cluster E's of the same magnitude are redder because of the higher metallicity of their disks, which must be quite old. Most S0's are not so red, and this suggests that many of them contain some younger stars. The fact that the cluster S0's appear somewhat redder than the field S0's can then be explained if star formation is truncated sooner in the cluster galaxies (§ II*c*). We can estimate the maximum amount by which younger stars affect the colors of most S0 galaxies by comparing the data of Figure 2 with the curve labeled "old S0" in Figure 3; this gives a maximum color change of about 0.1 mag in $u - V$ for most S0's. The corresponding difference in $B - V$ is about 0.05 mag, according to the mean $(u - V, b - V)$ relation from SV's data and the transformation between $b - V$ and $B - V$ given by Sandage and Visvanathan (1978*a*). We use this result below to set limits on how recently star formation can have ceased in most S0's, although we note that the estimated maximum color change due to younger stars is rather uncertain and could, for example, be too small if internal reddening is significant.

b) *Constraints on the History of Star Formation in* S0 *Galaxies*

Theoretical models for the evolution of colors can now be used to translate the above limit on color changes into constraints on when star formation stopped. We have calculated an extensive set of models in which galaxies have a variety of star formation rates (SFRs), including cutoffs at various past times; Appendix B gives details. Similar models were considered by Biermann and Tinsley (1975), but the more accurate data now available call for updated models. To be definite, we adopt an age of 15 Gyr for all present-day galaxies, and we assign all models the same metallicity, so that color differences are due only to differences in the history of star formation. The predicted color for a purely old model is $B - V = 0.99$, which is consistent with observations of giant E galaxies. Acceptable models for typical S0 galaxies should then be bluer than this by no more than 0.05 mag. (Since we are concerned only with color differences, uncertainties in absolute colors due to galactic reddening, the adopted age and metallicity, etc. are unimportant.)

Figure 4 illustrates the effects on the present $B - V$ color if star formation is cut off at various times in plausible models for spiral galaxies. The abscissa $(B - V)_0$ is the present color attained by each model if star formation is not truncated, and the ordinate $\Delta(B - V)$ is the difference between the present color of a truncated model and that of a purely old system (0.99). Each curve is labeled by the number of gigayears ago that star formation stopped. We can use Figure 4 to say how long ago star formation must have stopped in galaxies that would have become spirals with various present colors $(B - V)_0$, if instead they were to become S0's with $\Delta(B - V)$ less negative than -0.05. For example, a galaxy with $(B - V)_0 = 0.75$ would have to have stopped forming stars at least 1.5 Gyr ago in order to become red enough; an early-type spiral with $(B - V)_0 > 0.9$ could have stopped forming stars as recently as 10^8 years ago; but a late-type spiral with $(B - V)_0 < 0.55$ must have been truncated more than 5 Gyr ago.

The models in Figure 4 assume an initial burst of star formation in the bulge, followed by star formation at a constant rate in the disk, as suggested by recent data on the ages of stars in the solar neighborhood (Twarog 1980). We have also studied models with a declining

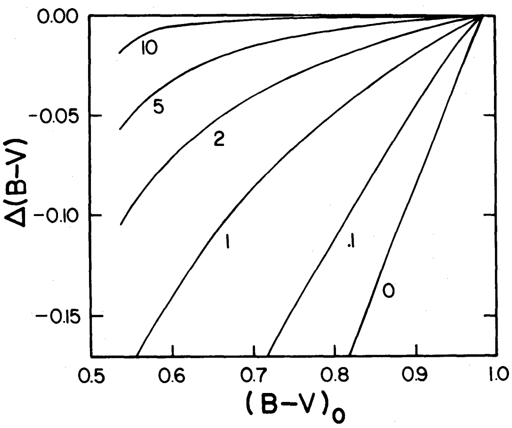

FIG. 4.—Effects on $B - V$, at age 15 Gyr, of truncating star formation at the number of Gyr ago shown on each line. The abscissa $(B - V)_0$ is the color a model would have at age 15 Gyr if star formation continued. The ordinate $\Delta(B - V)$ is the difference between the color of the truncated model and that of one cut 14 Gyr ago (0.99).

SFR (Appendix B), and find that for a given $(B - V)_0$ they must stop forming stars earlier than the models with constant SFR in order to attain the colors of S0's by the present time. For example, a spiral with $(B - V)_0 = 0.75$ and an exponentially declining SFR must stop forming stars at least 3 Gyr ago in order to become red enough. However, the limit of 5 Gyr for galaxies with $(B - V)_0 < 0.55$, which comprise as many as a fifth of all spirals, is relatively model-independent.

Star formation stops abruptly in these models, whereas it may be more realistic to consider a gradual fading. The constraints in this case would be very similar. For example, a galaxy with $(B - V)_0 = 0.55$ would have to have a substantial reduction in its SFR at least 5 Gyr ago, so that the colors would become almost as red as a galaxy with no star formation within ~ 1 Gyr of the present, and then the last residual star formation could stop very recently. (Biermann 1977 has presented some models of this type.)

The constraints on past star formation in S0's can therefore be summarized as follows: any S0 that was formerly a very blue spiral must have undergone a large reduction in its SFR at least 5 Gyr ago, while S0's that began as spirals with typical colors could have been making stars until only a few gigayears ago, and the last residual star formation in S0 disks (including sporadic consumption of recycled gas) could have continued until almost the present time. These results agree with SV's conclusion that a constant rate of turning off spirals to make S0's would yield too many blue S0's, but the colors of the S0 galaxies are still consistent with their being former spirals, provided that most of their star formation ceased at least a few gigayears ago.

IV. STRUCTURAL DIFFERENCES BETWEEN SPIRAL AND S0 GALAXIES

There are several apparent structural differences between spiral and S0 galaxies (§ I), and we discuss here whether these are consistent with the hypothesis that spirals and S0's have evolved from common ancestors.

a) Disk-to-Bulge Ratios

S0 galaxies have smaller average disk-to-bulge (D/B) ratios than spirals, and this has often been cited as evidence that stripping of spirals cannot produce most S0's (e.g., Faber and Gallagher 1976; SV; Burstein 1979*a*). However, since the early termination of star formation in S0's leads to smaller disks than those of spirals, this difference is not necessarily evidence against a stripping or gas-depletion origin for S0's.

The effect can be modeled in a simple way as follows. We assume that spiral galaxies form stars at a constant rate in their disks (§ III*b*). If a spiral galaxy forms disk stars up to the present time t_0, but an S0 does so only up to a time $t_s < t_0$ when star formation stops, the present ratio of spiral to S0 disk masses is t_0/t_s. If we assume from the discussion in § III*b* that $t_0 \approx 15$ Gyr and $t_s \lesssim 10$ Gyr, the ratio t_0/t_s is then $\gtrsim 1.5$. Thus the D/B ratios of spirals should ecceed those of S0's by a factor $\gtrsim 1.5$ in mass. When luminosities are considered, the relative prominence of the disks in spirals is further enhanced by the greater luminosity per unit mass of their young stars. From the models in Appendix B, we find that the factors by which the D/B luminosity ratios of spirals should exceed those of S0's are $\gtrsim 4$ in V light and $\gtrsim 6$ in B light.

In the most quantitative study to date, Burstein (1979*a*) has decomposed the blue-light profiles of 12 S0 galaxies into bulge and disk components, finding D/B values between 0.2 and 3, with a mean of 1.3. The above simple model then predicts that if these galaxies had continued forming stars, the resulting spirals would have attained D/B values of at least 1.2–18, with a mean of $\gtrsim 8$. We can compare this prediction with the data discussed by Burstein for three spiral galaxies; these data imply D/B ratios of ~ 3, 12, and 12 in blue light for M31, NGC 4321, and NGC 6744, respectively. (Burstein 1979*a* gives somewhat smaller D/B ratios because he subtracted an estimate of the disk light due to very young stars.) These data are too limited to be definitive, but they are consistent with the idea that spiral and S0 galaxies evolved from common ancestors ~ 5 Gyr ago.

A fourth spiral discussed by Burstein (1979*a*) is M101, for which the data correspond to a D/B of ~ 25 in blue light, indicating a negligible bulge component. This galaxy has a very blue integrated color of $B - V = 0.40$ (de Vaucouleurs, de Vaucouleurs, and Corwin 1976), so most of its stars must be less than 5 Gyr old according to the models of Appendix B. Thus if star formation had been cut off ~ 5 Gyr ago in such a galaxy, there would not have been much of a galaxy left! S0's formed from spirals with almost no bulge component are in general likely to be very hard to find, being almost pure disk systems of very low surface brightness. Sandage, Freeman, and Stokes (1970) note the absence of S0 galaxies with D/B ratios as large as those in some late-type spirals, and we suggest that this may be due to a selection effect: many late-type galaxies are very blue and so probably have rather young disks; therefore, if they had been truncated in the past they would have become very faint objects, and may well have escaped detection or inclusion in the well-studied samples of bright galaxies.

b) "Thick Disks"

Another possible structural difference between spirals and S0's is a "thick disk" component found in some S0's (Burstein 1979*b*). Burstein measured surface-brightness profiles perpendicular to the major axes of five S0's, and tried to fit them with models consisting of the known bulge component plus a projection of the disk profile measured along the major axis, assuming the disk to be very thin. An excess surface brightness over that predicted by these models appears at 24 mag arcsec^{-2} in B, and he called this a "thick disk." By contrast, a corresponding surface-brightness profile of M31 is matched perfectly by a projected thin disk, suggesting that spirals may not possess a "thick disk." However, these data do not rule out the possibility that M31 has a thick disk of old stars at the 24 mag arcsec^{-2} level which is hidden by a younger and much brighter thin disk. Let us imagine what would happen if the thin-disk components of Burstein's S0's were made brighter by 2 mag, which corresponds to the factor of 6 mentioned above as the ratio of blue luminosities of spiral and S0 disks. From his figures, one can see that such a bright projected thin disk would largely hide the thick component, down to the surface brightness limit of the observations.

Moreover, at least one well-studied edge-on spiral, NGC 4565, does appear to have a thick disk. This may be seen by comparing van der Kruit's (1979) isophotes of NGC 4565 with Burstein's (1979*b*) perpendicular profiles of NGC 4762, which is the most nearly edge-on galaxy in his sample. At comparable magnitude levels, the isophotes of NGC 4565 are comparable in thickness to those of NGC 4762; for example, along a perpendicular profile passing 7.5 kpc from the center ($H_0 = 50$), the isophotal thickness of NGC 4565 at a blue surface brightness of 24.5 mag arcsec^{-2} is about 8 kpc, while that of NGC 4762 is about 7 kpc. Jensen and Thuan (1980) also report photometry of NGC 4565 which they say shows evidence for a "fat disk" similar to that of NGC 4762. The deep infrared photometry of NGC 4565 by Spinrad *et al.* (1978) also clearly shows a thick disk or flattened halo component that becomes progressively rounder at fainter magnitudes.

In summary, the present very limited data do not appear to support the suggestion that thick disks are a feature unique to S0 galaxies.

c) *Bulge Sizes*

Dressler (1980) has emphasized that S0 galaxies not only have smaller D/B ratios than spirals but also have absolutely larger bulges. He notes that, for his magnitude-limited sample, this observed difference could be due to selection against S0's with small bulges if their disks are fainter than those of spirals by 2 mag. This is just the difference in disk luminosity that we expect typically to result from early truncation of star formation in S0's (§ IV*a*), suggesting that the bulge size distributions of spirals and S0's may in fact be the same.

One would then expect the total luminosities of S0's to be fainter than those of spirals; for example, with a D/B for S0's of 1.3, the expected difference in luminosity is 1.5 mag. No difference between the overall luminosity functions of spirals and S0's appears in Dressler's (1980) data, and he therefore rejects the possibility that the apparent difference in bulge sizes is due to selection. Tammann, Yahil, and Sandage (1979) find that the luminosity function of S0's is about 0.5 mag fainter than that of spirals (if magnitudes uncorrected for internal absorption are used, as was done by Dressler); this difference supports the suggestion that S0's have fainter disks than spirals, but it is still not large enough to allow the difference in bulge sizes to be ascribed entirely to selection.

In any case, it would not be surprising to find that S0's have somewhat larger bulges than spirals, because Dressler's (1980) statistics show that (1) the bulge sizes of both spirals and S0's increase with the density of their surroundings, and (2) the ratio of S0's to spirals increases with density. Therefore, even if the apparently larger bulges of S0's at a given density are due to the selection effect just mentioned, these two trends must combine to give S0's larger bulges than spirals when averaged over all regions. In other words, if S0's evolve from spirals, their ancestors are not a *random* sample of spirals but are preferentially those in the denser regions of space, which are observed to have larger bulges.

V. EVOLUTION OF THE SPIRAL AND S0 CONTENT OF CLUSTERS

Disk galaxies in clusters provide a strong constraint on any theory relating spirals and S0's because of the strong correlations between the ratio of spirals to S0's and the density and redshift of a cluster (§ I). We discuss in this section recent observations of cluster populations, and show that they can be understood if the spirals evolve into S0's after their peripheral gas is swept away during the collapse of the cluster, as proposed in § II.

a) *The Butcher-Oemler Effect*

Butcher and Oemler (1978*a*) found that two rich, centrally-concentrated clusters at redshifts ~ 0.4 contain large populations of blue, presumably spiral galaxies. This phenomenon is not observed in nearby clusters, since those with significant central concentration are all spiral-poor, and only those with little or no concentration are spiral-rich (Morgan 1962; Oemler 1974; Butcher and Oemler 1978*b*). Dressler (1980) has also demonstrated a close relation between spiral content and galaxy density in nearby clusters, the fraction of spirals decreasing and the fraction of S0's increasing monotonically with the density of the surroundings.

These observations support the idea that the spirals in the denser clusters were turned into S0's by depletion of their gas after the clusters collapsed, and it is natural to identify the distant blue clusters as direct ancestors of the nearby condensed clusters, seen while they still had spirals. However, before we can conclude that this is a plausible interpretation, we must answer the following questions raised by Butcher and Oemler's results: (1) Can clusters that have already collapsed still contain the proportion of star-forming galaxies that is observed in the distant blue clusters? (2) Can these blue galaxies then evolve rapidly enough into red galaxies to become typical present-day S0's? (3) Why are there no recently collapsed clusters at

1980ApJ...237..692L

small redshifts that still contain a large population of blue galaxies?

To answer these questions, we must first know when most of the diffuse envelope gas is swept away from the spiral galaxies in a collapsing or collapsed cluster, and for how long after this the galaxies can continue to form stars before their remaining gas is exhausted. We also need to know how fast the galaxies will then evolve from blue to red colors after star formation stops. The time scales for gas consumption and for color evolution have already been considered in §§ II*a* and III*b*, and the question of when most of the loss of envelope gas occurs is considered next in § V*b*.

A cosmological model is needed to relate redshifts to times, and we adopt the nominal parameters $\Omega_0 = 2q_0 = 0.04$ and $H_0 = 60$, corresponding to a present age of 15.4 Gyr since the big bang. The lookback times for Butcher and Oemler's (1978*a*) clusters at $z = 0.39$ and $z = 0.46$ are then 4.6 Gyr and 5.1 Gyr, respectively, so for discussion we shall consider a lookback time of 5 Gyr. We also assume that all galaxies have a present age of 15 Gyr, so that the galaxies in the distant clusters are observed at an age of 10 Gyr. The conclusions would not be affected if other cosmological parameters consistent with constraints on the density and age of the universe were used instead.

b) Stripping of Gas Envelopes in a Collapsing Cluster

In this section, we present some simple numerical simulations of collisional stripping of gas envelopes to address the question of when most of the removal of peripheral gas takes place in a collapsing cluster. The simulations are based on an *N*-body model of a collapsing cluster calculated by S. Aarseth and described by Oemler (1973). One hundred particles were assigned relative masses proportional to the luminosities of the 100 brightest galaxies in the Coma cluster. and the calculations were started at the point of maximum expansion with the particles randomly distributed in a spherical volume with small velocities. For the present simulations we have used the positions of the particles at 100 equal time steps distributed over 4.2 free-fall times.

Our procedure is to assign a gaseous envelope to each particle, and then to note when the envelopes of the particles collide with each other. The envelope masses are assumed to be proportional to the particle masses, but the ratio need not be specified since we are concerned only with the fraction of envelope material that is lost, and since we do not follow any dynamical effects of mass loss. The envelopes are assumed to be spherical with density distributions $\rho \propto r^{-2}$, and their radii are determined from their masses according to either $R \propto M^{1/3}$ or $R \propto M$. Collisions are assumed to occur when the envelopes of two particles intersect, and the amount of mass lost is taken to be either the entire overlapping region or that part which is less dense than the overlapping part of the other envelope (cf. De Young 1978, Fig. 5). The radius of the envelope is assumed to shrink instantaneously to a size corresponding to its new mass, an assumption that underestimates the number and severity of collisions during the initial collapse, when most particles suffer several collisions within one time step.

Several models have been calculated using different assumptions for the maximum envelope radius, the relation between radius and mass, and the criterion for the amount of mass stripped. Table 1 gives for four models the average number of collisions per particle,

TABLE 1

MODELS FOR STRIPPING DURING CLUSTER COLLAPSE

	1	2	3	4
(a) Model Specifications				
Radius-mass relation	$R \propto M^{1/3}$	$R \propto M^{1/3}$	$R \propto M^{1/3}$	$R \propto M$
Maximum envelope radius (kpc)[a]	200	100	100	100
Condition for loss[b]	less dense	less dense	all	all
(b) Results[c]				
Average number of collisions				
$t = 1$	16	4.9	3.2	0.37
$t = 4.2$	26	10.2	4.9	0.66
Fraction of envelope mass lost				
$t = 1$	0.54	0.20	0.48	0.10
$t = 4.2$	0.74	0.46	0.68	0.30
Reduction in mean radius[d]				
$t = 1$	0.73	0.92	0.73	0.90
$t = 4.2$	0.62	0.84	0.63	0.78

[a] Radius of envelope of the most massive particle. The range of masses is a factor of 14.4, and the average particle is less massive than the largest by a factor of 3.75.

[b] Criterion for determining what parts of envelopes are lost (see text).

[c] t = time from maximum expansion, in units of the free-fall time.

[d] (Mean radius at time t)/(initial mean radius).

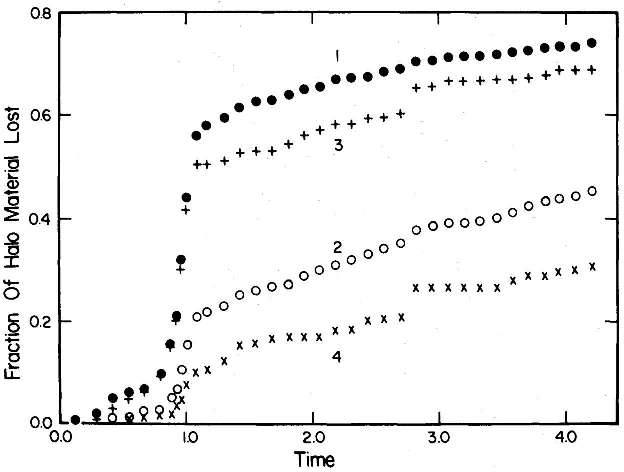

FIG. 5.—Cumulative fraction of gaseous envelope material that has been lost from galaxies in clusters, as a function of the time since maximum cluster expansion, in units of the free-fall time. Numbers refer to the models in Table 1. The bump between times 2.69 and 2.81 in models 3 and 4 is caused by a single severe collision between two massive envelopes.

the fractional mass loss, and the reduction in mean envelope radius at the time of the initial collapse and at the end of the simulation. The cumulative mass loss as a function of time is shown in more detail in Figure 5. The main results can be summarized as follows.

(1) Most of the mass loss occurs close to the time of collapse, one free-fall time after the start of the calculations. This is especially true of model 1, in which the average galaxy has experienced 16 collisions already by the collapse time.

(2) Collisions are much more effective if all the overlapping material is stripped than if only the less dense part is lost, as illustrated by models 2 and 3. The latter more conservative assumption is used in model 1, which would have even faster stripping at the collapse time if complete loss were to occur.

(3) The effectiveness of collisions increases rapidly with envelope radius. This can be seen by comparing models 1 and 2, which have different maximum envelope radii, and models 3 and 4, which have the same maximum radius but different average radii because of the different R–M relations. Results with many models suggest that there is a threshold size $R_T \sim 60$ kpc, such that for larger envelopes collisions are frequent during the collapse and tend to reduce the envelope radius to $\sim R_T$, while for smaller envelopes there is little stripping during the collapse. A similar phenomenon was found by Richstone (1974) in models for collisions between stellar systems, where the galaxies were stripped to a size ~ 50 kpc during the initial collapse of the cluster.

The conclusion of most importance for this paper is the first point above: most of the stripping that will ever occur takes place when the cluster first collapses. We shall therefore assume that most of the loss of gas-rich envelopes that causes spiral galaxies to evolve into S0's in clusters occurs at the time of cluster collapse.

c) *The Fate of Disk Galaxies in the Distant Blue Clusters*

The form of Butcher and Oemler's (1978*a*) distant blue clusters indicates that they have collapsed already, so the envelope gas has been stripped from their disk galaxies. Since the distant clusters resemble the Coma cluster in structure, it is possible that they collapsed at about the same time, estimated by Gunn and Gott (1972) to be about 6 Gyr after the big bang (when scaled to $H_0 = 60$). The unit of time in Figure 5 in this case is about 3 Gyr, and the zero point is 3 Gyr after the big bang; the Butcher-Oemler clusters are seen when the abscissa value is 2.3 units, and the present time is at about 4 units. Most of the stripping of the envelopes of spirals in the distant clusters would then have occurred about 4 Gyr before these clusters were observed.

This time interval is the same as the median time estimated in § II*a* (Fig. 1) for the gas in present-day spirals to be strongly depleted by star formation. If the same distribution of values of τ_{gas} applies to the disk galaxies in the distant clusters, then only about half of the spirals in these clusters would have run out of gas by the time the clusters are seen, and the presence of many remaining blue galaxies can therefore be understood. However, about three-quarters of the disk galaxies would have depleted their gas within ~ 1 Gyr after the clusters are seen, and it is likely that star formation will be slowed in all galaxies because of the loss of a continuing gas supply. To see whether the colors will become as red as those of typical S0's by the present time, we consider models in which star formation stops entirely 1 Gyr after the clusters are seen. The conclusions would not be altered if instead we allowed star formation to die out over a range of times between 1 and 7 Gyr after the collapse, as suggested by Figure 1.

The observed colors of galaxies in near and distant clusters are represented schematically in Figure 6, which is a C-M diagram analogous to Butcher and Oemler's (1978*a*) plots for the distant clusters and Coma. In both distant clusters, the brightest galaxies are red, and the fainter galaxies spread toward increasingly blue colors, occupying the region to the right of the diagonal line in Figure 6. The Coma cluster, by contrast, has only red galaxies (with a handful of exceptions), and is represented by a horizontal shaded strip with $0.94 \leq B - V \leq 0.99$; the red limit is the color of a model with only old stars at age 15 Gyr, and the width of the strip corresponds to the range of colors of most S0's, as discussed in § III. In this figure we have omitted the slope due to the metallicity-magnitude relation, since comparisons will be made below with models that all have the same metallicity.

A few representative models for galaxies in the distant clusters are shown as filled triangles in Figure 6. We assume that the observed range of colors is due to different histories of star formation, which we represent by models with an initial burst followed by a constant SFR for 10 Gyr (Appendix B), the reddest

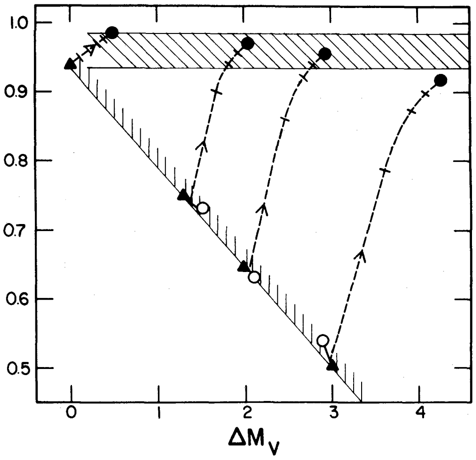

FIG. 6.—A color-magnitude diagram analogous to Butcher and Oemler's (1978*a*) Figures 7 and 9 for the distant clusters, whose C-M distributions are represented schematically by the area to the right of the diagonal line, and their Figure 15 for Coma, represented by the shaded strip at the top. The abscissa ΔM_V is the magnitude relative to the brightest cluster member. The triangles represent several models of age 10 Gyr, the open circles their positions at age 15 Gyr if star formation continues, and the filled circles their positions at age 15 Gyr if star formation stops at 11 Gyr. Evolution after star formation stops is shown by dashed lines, with tick marks at successive intervals of 1 Gyr.

model having only an initial burst and the bluest having no burst. The models shown are on the boundary of the region occupied by the distant clusters, so others with the same colors would lie to their right. If star formation were to continue at the same rate up to the present age, 15 Gyr, the models would evolve along the full lines to the positions shown by open circles. The dashed lines show the evolution if star formation stops at an age of 11 Gyr. The models then become rapidly redder and fainter, and reach the positions shown by filled circles at age 15 Gyr.

The models with truncated star formation evolve to colors within the range adopted for present S0's, except for the bluest, which has $B - V = 0.5$ in the distant clusters and is too blue by ~ 0.02 mag now. Only a small fraction of Butcher and Oemler's blue galaxies have colors corresponding to $B - V < 0.6$, and Figure 6 shows that these are predicted to become S0's that are several magnitudes fainter than the brightest cluster members. Few colors are available for such faint S0's in nearby clusters (Visvanathan and Sandage 1977), but among the brighter S0's in Figure 2 there are a few that are bluer than a purely old system by more than 0.05 mag in $B - V$, so the "excess" blueness of 0.02 mag in Figure 6 is not a significant discrepancy.

We conclude that the distant blue clusters can evolve into clusters as red as Coma if most of their star formation stops, as is plausible, soon after the time that they are seen.

d) *Dependence of Cluster Properties on Redshift*

Having shown that the distant clusters at $z \sim 0.4$ can evolve into present-day clusters like Coma, we can now indicate schematically how the properties of clusters should vary with redshift. It will be convenient to discuss the evolution of clusters in Butcher and Oemler's (1978*b*) plot of spiral content versus central concentration, reproduced here as Figure 7 with the distant blue clusters added (crosses).

We assume that at the time of maximum expansion, clusters begin in the upper-left corner of this diagram, with no central concentration and a high proportion of spirals. The subsequent evolution of a cluster then depends on the ratio of its free-fall time t_f to the characteristic time $\tau_{\rm gas}$ required for spirals to die out after the cluster collapses. A cluster that collapses at large redshift and has $t_f < \tau_{\rm gas}$ will first evolve to the right in Figure 7, increasing in central concentration with little change in galaxy content. After a time $\tau_{\rm gas}$ has elapsed, the cluster will then evolve rapidly downward, with decreasing spiral content but little further change in central concentration. This is the situation discussed in § V*c*, and the evolutionary path is shown schematically by the short-dashed curve in Figure 7.

On the other hand, a cluster that collapses at small redshift and has $t_f > \tau_{\rm gas}$ will suffer significant loss of spirals while the collapse is still in progress, and will evolve downward and to the right in Figure 7 along a sloping trajectory qualitatively like that indicated by the long-dashed curve. In such a cluster, the stripping and the fading of spirals could be well advanced in the core region even while the outer regions are still

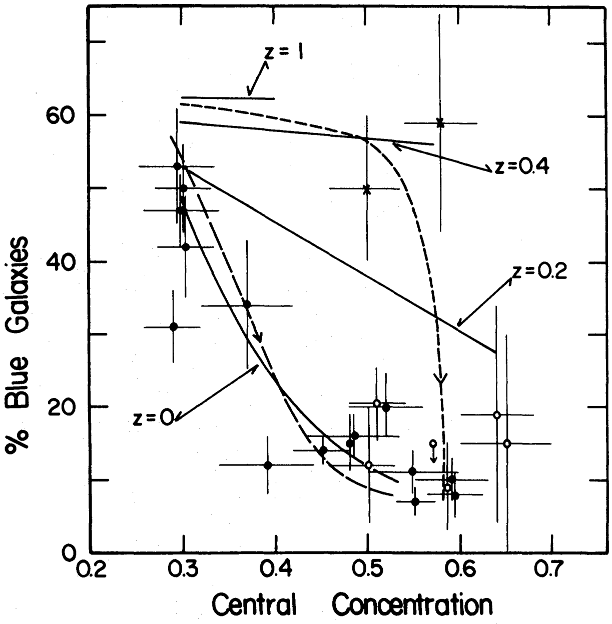

FIG. 7.—Spiral content versus central concentration of clusters. The ordinate is the percentage of galaxies with the colors of spirals, and the abscissa is a central concentration index which has a value of 0.3 for no concentration. Points with error bars are Butcher and Oemler's (1978*b*) data for nearby clusters, and crosses are their distant clusters; the Virgo cluster is at (0.37, 34) and Coma is at (0.55, 7). Broken lines are schematic evolutionary tracks, and solid lines are schematic isochrones labeled by redshift.

collapsing. This is apparently the situation in the Virgo cluster, which seems to be undergoing its first collapse now. There are relatively few spirals in the core, and many of its spirals are of the "anemic" type which van den Bergh (1976*a*) interprets as being in transition toward S0's.

These examples show why no collapsed clusters have large blue populations at the present epoch: the clusters that collapsed early have had time for the stripped disk galaxies to use up their internal gas, and those that collapsed recently were steadily depleted of spirals during the collapse.

From these arguments, we can sketch schematic isochrones (iso-redshift curves) in Figure 7, as indicated for a few redshifts by the solid lines. Each curve is a locus of clusters observed at the same redshift but having different initial densities and therefore different rates of evolution. Only the qualitative character of these curves has any meaning, since the horizontal coordinate in Figure 7 is not related unambiguously to time.

At $z = 1$, the time is 7 Gyr in the cosmological model adopted in § V*a*. The densest clusters have only recently collapsed but may not yet have reached their maximum central concentration, and no stripped spirals have yet run out of gas, so the isochrone is a short horizontal line.

By $z = 0.4$, the time is 11 Gyr and the densest clusters have collapsed to a greater degree of central concentration; the spirals are beginning to fade but most have not yet become S0's. Thus the isochrone is again almost horizontal, a little below the previous line to allow for the loss of some of the more gas-poor spirals. Note that all clusters at $z \gtrsim 0.4$ are predicted to have substantial blue populations.

At smaller redshifts, the spirals in the more-condensed clusters rapidly die out, while the less-condensed clusters still contain a substantial proportion of spirals. Thus the isochrone for $z = 0.2$ is shown schematically as sloping down to the right, indicating that the more dynamically evolved clusters have fewer spirals.

Finally, by the present time, even partially collapsed clusters must have suffered a significant loss of spirals, because the internal gas time scale of the galaxies is less than the time scale of cluster collapse, as discussed above. Only clusters that have not yet begun to collapse, and therefore have negligible central concentration, can remain spiral-rich. We have thus accounted qualitatively for the distribution of the nearby clusters in Figure 7, and the isochrone for $z = 0$ is drawn accordingly.

The most significant feature of the schematic isochrones in Figure 7 is that they predict rapid evolution of the spiral content of condensed clusters between redshifts of 0.4 and 0, and an increasing correlation between galaxy content and central concentration. These predictions are accessible to direct observational test, by means of further studies of galaxy populations in clusters with redshifts of a few tenths.

VI. CONCLUSIONS

In this paper we have reconsidered the question of whether there is any evolutionary relation between spiral and S0 galaxies. According to our estimates (§ II*a*), the median time required for star formation to exhaust the gas in spiral galaxies is $\sim 4(50/H_0)$ Gyr, considerably shorter than the Hubble time; this suggests that if their gas is not replenished, many spirals may within a Hubble time evolve into gas-poor systems resembling S0's. It is plausible that the gas in many spiral galaxies has been replenished by accretion from extended envelopes of debris left over from galaxy formation, probably including gas-rich companion galaxies and tidal debris as well as primordial gas (§ II*b*). The S0 galaxies can then be interpreted as disk systems that lost their gas-rich envelopes at an early stage, and then exhausted their remaining gas by star formation. Once their initial gas content has been depleted, the S0 galaxies can be kept swept clean by cluster gas, while those in the field might continue to consume their recycled gas by sporadic low-level star formation (§ II*c*).

We therefore propose a modified version of the "stripped spiral" hypothesis for S0 galaxies, in which the material stripped is diffuse outlying gas rather than dense gas in the disks of spirals, and star formation does not stop immediately but continues until the remaining disk gas has been consumed. The removal of outlying gas requires much less violent interactions than does the sweeping of disk gas, and this could explain Dressler's (1980) observation that the ratio of S0's to spirals increases with the local space density of galaxies even in regions where the density is too low for gas to be stripped from the disks of spirals; Dressler also notes that this result could be understood if disks form over a long period of time from diffuse gas that is easily removed by interactions with neighboring galaxies. Similarly, Strom and Strom (1978) quote evidence that the sizes of S0 disks are smallest in the densest regions, and suggest as an explanation that the growth of disks is truncated early in dense regions.

Although the colors of "normal" S0 galaxies are not bluer, on the average, than ellipticals of the same magnitude, this constraint on recent star formation in S0's is alleviated by the expected high metallicities and red colors of old disk stars (§ III*a*); thus the colors of S0's are consistent with their origin as spirals with truncated star formation, provided that most of the star formation stopped at least a few gigayears ago (§ III*b*). The fact that S0's have smaller disk-to-bulge ratios than spirals is also consistent with such an origin, since the early truncation of star formation reduces the disk masses and greatly reduces the present disk luminosities of S0 galaxies compared with spirals (§ IV*a*). The "thick disks" of S0's may also be present in spirals, but generally hidden by the much greater luminosity of the thin disk. The larger observed bulge sizes of S0's relative to spirals could be due partly to selection against S0's with small bulges, and partly to the tendency for the spirals with the largest bulges to be

1980ApJ...237..692L

in the densest regions of space, where they are most likely to become S0's (§ IV*c*). In any case, all of the available data appear to be consistent with the evolution of spiral and S0 galaxies with similar bulge sizes from common ancestors perhaps ~5 Gyr ago.

Numerical simulations of the stripping of gas envelopes of galaxies in clusters indicate that such envelopes are rapidly stripped down by collisions to radii of $\lesssim 60$ kpc, and that most of the loss of envelope matter occurs at the time the cluster collapses, one free-fall time after maximum expansion (§ V*b*). Thus in a collapsed cluster most of the spirals will have lost their gas-rich envelopes and ceased to accrete gas, but there will be a period of several gigayears while their remaining disk gas is consumed by continuing star formation. This delay allows us to explain the blue galaxies observed by Butcher and Oemler (1978*a*) in two distant ($z \sim 0.4$) condensed clusters. If these clusters are like the Coma cluster, they could have collapsed about 6 Gyr after the big bang, which is about 4 Gyr before they are observed. Thus when the clusters are seen, many of their disk galaxies will be running out of gas, but there will still be a substantial number of galaxies with gas and young stars. Soon afterward the remaining gas will be depleted, and by the present time nearly all of the disk galaxies can evolve into S0's with normal colors (§ V*c*). Thus our picture predicts rapid evolution of the galaxy content of clusters in the redshift range ~0.4–0, such that at $z \gtrsim 0.4$ all clusters should still contain many blue galaxies, while by the present time only diffuse, uncollapsed clusters should still contain a significant fraction of blue galaxies (§ V*d*).

Several observational tests of this picture can be made. Further observations of the galaxy content of clusters with redshifts between 0 and 0.4 will be particularly important to test the prediction that the galaxy content of clusters varies rapidly with redshift in this range, and correlates increasingly with cluster concentration, as indicated by the schematic isochrones in Figure 7. When structural studies of galaxies at redshifts of several tenths become possible—for example, with the Space Telescope—the predicted evolutionary changes in the disk-to-bulge ratio of disk galaxies can be looked for; in general, the differences in structure between spiral and S0 galaxies should have been smaller in the past. Detailed photometric studies and population syntheses of S0 galaxies should help to test our expectation that their disks are both younger and more metal-rich than their bulges or elliptical galaxies of the same luminosity.

Finally, we note that the proposed gradual buildup of spiral galaxies by accretion of outlying matter, including small gas-rich companion galaxies and tidal debris, parallels the suggestion that elliptical galaxies are built up by mergers of smaller systems (Toomre 1977; Tinsley and Larson 1979). In any process of galaxy formation by aggregation of smaller units, stars in the interacting subsystems will be dispersed to form a "spheroidal" component, while the gas will tend to settle to a disk, unless violent tidal effects disrupt any forming disk. Thus elliptical galaxies and large bulges might preferentially be formed in dense regions where violent interactions between systems of comparable size are common (cf. Ostriker 1977), whereas disks might generally be formed when galaxies are more isolated and interact with and accrete only diffuse matter or small companions. Relatively weak interactions would suffice to remove the diffuse matter around a forming spiral galaxy and thus cause it to evolve into an S0. We can then understand why the elliptical galaxies are strongly concentrated into the densest regions, while S0's are more widely distributed in less dense environments, and spirals populate the least dense regions, with a tendency for bulge size to decrease with decreasing density of the surroundings (Oemler 1974; Dressler 1980).

It is a pleasure to thank Gus Oemler for many valuable discussions, and to acknowledge useful comments from D. Burstein, A. Dressler, S. Faber, S. Strom, and S. van den Bergh. This work was supported in part by the National Science Foundation (grant AST 77-23566) and the Alfred P. Sloan Foundation.

APPENDIX A

FURTHER COMMENTS ON THE COLORS OF E AND S0 GALAXIES

In § III*a* several apparent differences were noted among the C-M diagrams of E and S0 galaxies in various environments (Fig. 2). This Appendix discusses the statistical significance of the differences, and some points of interpretation.

a) Differences among Subsamples

The following differences may be seen in Figure 2.

1. Cluster E's show a smaller scatter in the C-M diagram than do the other subsamples. This visual impression is confirmed by the correlation coefficients between $u - V$ and $|M_V|$, which are 0.87, 0.65, and 0.43 for E's in clusters, groups, and the field, and 0.69, 0.59, and 0.66 for the respective S0's (including galaxies fainter than $M_V = -20$). It can be shown that the value for the cluster E's exceeds the others at the 2 σ level, while the values for the other subsamples are not significantly different. SV note that their observational errors in $u - V$ are $\lesssim 0.02$ mag, so the scatter in all subsamples is real. These differences in scatter show both that E's and S0's have different color distributions (at least in clusters), and that the environment affects the color distributions of galaxies (at least of E's).

2. Field E's are systematically bluer than cluster E's, at least in the magnitude range of most of the field E's

in this sample. In particular, the mean value of $u - V$ for field E's with $-21.5 \geq M_V \geq -23.5$ is 0.063 mag bluer than that for cluster E's (although their mean magnitudes differ by only 0.008 mag). A *t*-test gives a 99% confidence level for the reality of this difference, which corresponds to 0.04 mag in $B - V$. Again, we find that colors depend on environment.

3. The fainter S0's in the magnitude range of Figure 2 are redder than cluster E's of the same magnitude. When all S0's and E's are combined, this effect is not statistically significant; but for the cluster members alone, a *t*-test gives a 90% probability that S0's are redder than E's in the magnitude interval -20.5 ± 0.5. This result gives further evidence that S0's and E's are not identical in their C-M relations.

The above three differences were noted by Faber (1977) and Burstein (1977), in their study of a different sample of galaxies.

4. A final obvious difference is that there are many fewer bright S0's than bright E's, as pointed out by van den Bergh and McClure (1979) for SV's Shapley-Ames sample (most of their galaxies), and by Burstein (1978) for a different sample. It is therefore hard to tell whether the brightest E's and S0's have the same colors.

The overall correlation between color and magnitude is interpreted as a correlation between metallicity and luminosity (§ III*a*). Departures from a mean C-M relation could be due to scatter in this correlation, or to variations in the history of star formation.

The galaxies least likely to contain young stars are cluster E's, since sweeping of the gas shed by old stars is most likely in a cluster, and since cluster members have emission lines less often than other E's (Gisler 1978). Their relatively well-defined C-M diagram can therefore be interpreted as the locus of $u - V$ due to metallicity variations in very old populations, with some intrinsic scatter in the correlation with magnitude.

Star formation could be responsible for the blueward scatter of field E's and some group E's in Figure 2. Oemler and Tinsley (1979) have suggested that the supernovae in E galaxies are due to a small finite SFR, which would make their colors bluer, on average, by a few hundredths of a magnitude; the SFR required to account for the supernovae would use up gas as fast as it is shed by dying old stars, and any associated morphological perculiarities could plausibly have escaped attention.

b) Ongoing Star Formation in S0*'s?*

On the other hand, S0 galaxies have a much lower supernova rate than E's, a result which Oemler and Tinsley (1979) interpret as indicating much less current star formation in S0's. Perhaps this is because disk galaxies with even a little ongoing star formation tend to be classified as early-type spirals rather than S0's, as suggested in § II*c*; SV note that their galaxies with classifications slightly later than S0 are substantially bluer. An exception is NGC 5102, a well-known S0 with star formation (van den Bergh 1976*b*; Danks, Laustsen, and van Woerden 1979). This is the bluest galaxy in SV's sample classified as a normal S0, and it is a group member at $M_V = -18.98$, $u - V = 1.83$. Only one of their S0's is further below their mean C-M relation, viz., NGC 4382. Remarkably, NGC 4382 is the only bona fide S0 with a supernova in Oemler and Tinsley's (1979) list of 24 supernova producers that have been classified as S0 or E. The fact that these two S0's stand out so anomalously in the C-M diagram suggests that star formation is indeed very rare in galaxies classified as morphologically normal S0's.

Ultraviolet photometry (at wavelengths below 3300 Å) would be helpful in determining whether galaxies with slightly blue optical colors have a little ongoing star formation or relatively recent truncations. However, even old stellar populations could have substantial ultraviolet fluxes from hot stars such as blue stragglers and horizontal branch stars (Wu *et al.* 1980), so the answer might not be clear in individual cases. The strongest tests could be made if ultraviolet colors were available for many E and S0 galaxies, including those with minor morphological peculiarities, in various environments.

APPENDIX B

EVOLUTIONARY MODELS FOR SPIRAL AND S0 GALAXIES

This appendix gives some details of the models used for evolution of colors and magnitudes.

It has previosuly been shown that models with continuously decreasing star formation rates (SFRs) have colors at ages ~5–20 Gyr that fall in the range observed for spiral galaxies (e.g., Tinsley 1968; Searle, Sargent, and Bagnuolo 1973). If S0's differ from spirals only in that star formation has not continued up to the present, then we should be able to reproduce the colors of S0's using the same models with star formation stopped at some time in the past. Results for such models are given here, using the method of calculation and input quantities adopted by LT. Only solar (or old-disk) metallicities are used for the evolutionary tracks and colors of stars.

Two series of models have been studied: (1) models having exponentially declining SFRs with different time constants; and (2) models with a constant SFR plus an initial burst of 1 Gyr that makes a fraction f_b of the stars present at age 15 Gyr. In each series, star formation is truncated at different times. The only color discussed here is $B - V$, since this is less sensitive

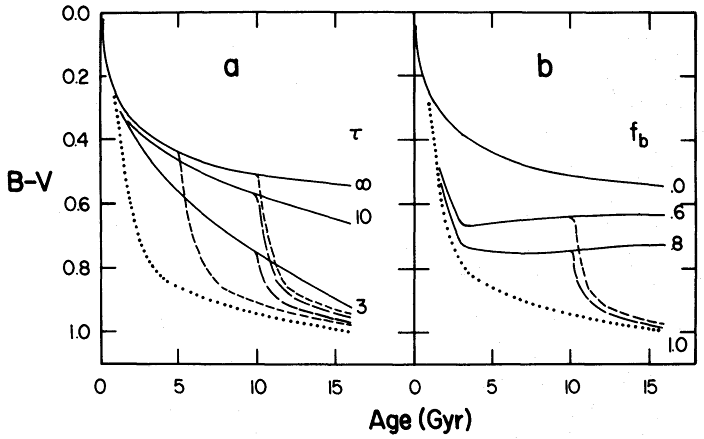

FIG. 8.—Evolution of $B - V$ in models with continuing (*solid lines*) and truncated (*broken lines*) star formation. (*a*) Models with exponential SFRs; label τ is the time constant in Gyr. (*b*) Models with a constant SFR after an initial burst lasting 1 Gyr; label f_b is the fraction of stars formed in the burst, measured at 15 Gyr in the case of continuing star formation.

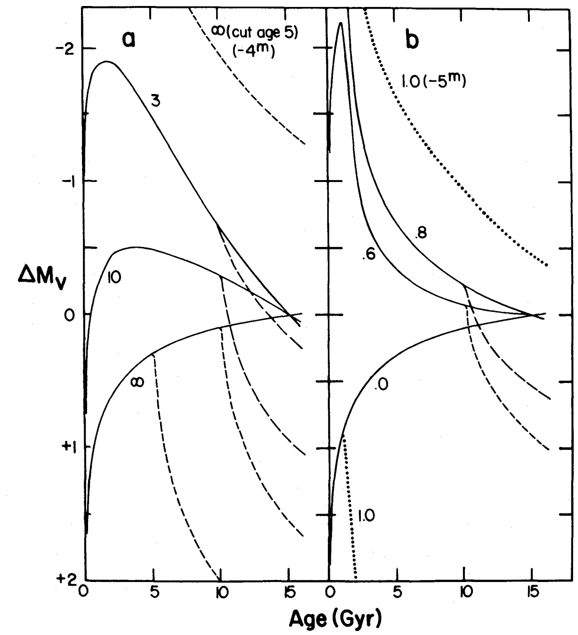

FIG. 9.—Evolution of M_V in models labeled as in Fig. 8. Each model is normalized so that $\Delta M_V = 0$ at age 15 Gyr in the case of continuing star formation. Note that two cases with large magnitude changes have their plots continued (with a revised zero-point) at the top.

to metallicity than, for example, $U - V$ or $V - K$, and since the stars contributing at B and V are better understood in an evolutionary sense than those dominating at shorter or longer wavelengths.

Figures 8 and 9 show the evolution of $B - V$ and M_V for some of the models. It is clear that they become rapidly redder and fainter after star formation stops, and immediately after the initial burst in the second series.

The exponential models get monotonically redder with age (Fig. 8*a*), but those with an initial burst can evolve toward bluer colors (Fig. 8*b*, models with $f_b = 0.6$ and 0.8). This is because, immediately after the burst ends, the colors become almost as red as if there were no star formation; but later the old population becomes fainter, while the integrated luminosity of stars formed at the continuing constant rate increases, leading eventually to bluer total colors.

Exponential models have often been considered as the most convenient way of parametrizing the range of present SFRs in spiral galaxies. However, models in which the SFR in the disk is constant after rapid formation of the bulge may be more realistic. The solar neighborhood is the only region where stellar ages give direct information on the past SFR, and a recent study by Twarog (1980) indicates that the rate has not decreased significantly during the life of the disk.

REFERENCES

Baade, W. 1963, *Evolution of Stars and Galaxies* (Cambridge: Harvard University Press).

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

Biermann, P. 1977, *Astr. Ap.*, **54**, 619.

Biermann, P., and Tinsley, B. M. 1975, *Astr. Ap.*, **41**, 441.

Bregman, J. N. 1978, *Ap. J.*, **224**, 768.

Burstein, D. 1977, in *The Evolution of Galaxies and Stellar Populations*, ed. B. M. Tinsley and R. B. Larson (New Haven: Yale University Observatory), p. 191.

———. 1979*a*, *Ap. J.*, **234**, 435.

———. 1979*b*, *Ap. J.*, **234**, 829.

Butcher, H., and Oemler, A. 1978*a*, *Ap. J.*, **219**, 18.

———. 1978*b*, *Ap. J.*, **226**, 559.

Cox, D. P., and Smith, B. W. 1976, *Ap. J.*, **203**, 361.

Danks, A. C., Laustsen, S., and van Woerden, H. 1979, *Astr. Ap.*, **73**, 247.

Davies, R. D. 1979, in *The Origin and Early Evolution of Galaxies*, meeting at the Royal Society, London, *Phil. Trans. Roy. Soc. London A*, in press.

de Vaucouleurs, G., de Vaucouleurs, A., and Corwin, H. G. 1976, *Second Reference Catalogue of Bright Galaxies* (Austin: University of Texas Press).

De Young, D. S. 1978, *Ap. J.*, **223**, 47.
Dressler, A. 1980, *Ap. J.*, **236**, 000.
Faber, S. M. 1973, *Ap. J.*, **179**, 731.
———. 1977, in *The Evolution of Galaxies and Stellar Populations*, ed. B. M. Tinsley and R. B. Larson (New Haven: Yale University Observatory), p. 157.
Faber, S. M., and Gallagher, J. S. 1976, *Ap. J.*, **204**, 365.
———. 1979, *Ann. Rev. Astr. Ap.*, **17**, 135.
Gisler, G. R. 1978, *M.N.R.A.S.*, **183**, 633.
———. 1979, *Ap. J.*, **228**, 385.
Gordon, M. A., and Burton, W. B. 1976, *Ap. J.*, **208**, 346.
Gunn, J. E., and Gott, J. R. 1972, *Ap. J.*, **176**, 1.
Haynes, M. P. 1979, in *IAU Symposium No. 84, The Large-Scale Characteristics of the Galaxy*, ed. W. B. Burton (Dordrecht: Reidel), p. 567.
Haynes, M. P., and Roberts, M. S. 1979, *Ap. J.*, **227**, 767.
Haynes, M. P., Giovanelli, R., and Burkhead, M. S. 1978, *A.J.*, **83**, 938.
Jensen, E. B., and Thuan, T. X. 1980, preprint.
Larson, R. B. 1972, *Nature*, **236**, 21.
———. 1974, *M.N.R.A.S.*, **169**, 229.
Larson, R. B., and Tinsley, B. M. 1978, *Ap. J.*, **219**, 46 (LT).
Lo, K. Y., and Sargent, W. L. W. 1979, *Ap. J.*, **227**, 756.
Mathewson, D. S., Cleary, M. N., and Murray, J. D. 1974, *Ap. J.*, **190**, 291.
Melnick, J., and Sargent, W. L. W. 1977, *Ap. J.*, **215**, 401.
Miller, G. E., and Scalo, J. M. 1979, *Ap. J. Suppl.*, **41**, 513.
Morgan, W. W. 1962, *Ap. J.*, **135**, 1.
Morris, M., and Lo, K. Y. 1978, *Ap. J.*, **223**, 803.
Oemler, A. 1973, Ph.D. thesis, California Institute of Technology.
———. 1974, *Ap. J.*, **194**, 1.
Oemler, A., and Tinsley, B. M. 1979, *A.J.*, **84**, 985.
Oort, J. H. 1970, *Astr. Ap.*, **7**, 381.
Oort, J. H., and Hulsbosch, A. N. M. 1979, in *Astronomical Papers Dedicated to B. Strömgren*, ed. A. Reiz, in press.
Ostriker, J. P. 1977, in *The Evolution of Galaxies and Stellar Populations*, ed. B. M. Tinsley and R. B. Larson (New Haven: Yale University Observatory), p. 369.
Ostriker, J. P., and Thuan, T. X. 1975, *Ap. J.*, **202**, 353.
Pritchet, C. 1979, *Ap. J.*, **231**, 354.
Richstone, D. O. 1974, *Ap. J.*, **204**, 642.
Roberts, M. S. 1963, *Ann. Rev. Astr. Ap.*, **1**, 149.
Sandage, A. 1961, *The Hubble Atlas of Galaxies* (Washington: Carnegie Institution of Washington).
Sandage, A., Freeman, K. C., and Stokes, N. R. 1970, *Ap. J.*, **160**, 831.
Sandage, A., and Visvanathan, N. 1978*a*, *Ap. J.*, **223**, 707.
———. 1978*b*, *Ap. J.*, **225**, 742 (SV).
Saslaw, W. C. 1973, *Pub. A.S.P.*, **85**, 5.
Sciama, D. W. 1972, *Nature*, **240**, 456.
Searle, L., Sargent, W. L. W., and Bagnuolo, W. G. 1973, *Ap. J.*, **179**, 427.
Smith, L. F., Biermann, P., and Mezger, P. G. 1978, *Astr. Ap.*, **66**, 65.
Solomon, P. M., Sanders, D. B., and Scoville, N. Z. 1979, in *IAU Symposium No. 84, The Large-Scale Characteristics of the Galaxy*, ed. W. B. Burton (Dordrecht: Reidel), p. 35.
Spinrad, H., Ostriker, J. P., Stone, R. P. S., Chiu, L.-T. G., and Bruzual, A. G. 1978, *Ap. J.*, **225**, 56.
Spitzer, L., and Baade, W. 1951, *Ap. J.*, **113**, 413.
Strom, S. E., and Strom, K. M. 1978, in *IAU Symposium No. 77, Structure and Properties of Nearby Galaxies*, ed. E. M. Berkhuijsen and R. Wielebinski (Dordrecht: Reidel), p. 69.
Tammann, G. A., Yahil, A., and Sandage, A. 1979, *Ap. J.*, **234**, 775.
Tinsley, B. M. 1968, *Ap. J.*, **151**, 547.
———. 1980, *Fundamentals of Cosmic Physics*, in press.
Tinsley, B. M., and Larson, R. B. 1978, *Ap. J.*, **221**, 554.
———. 1979, *M.N.R.A.S.*, **186**, 503.
Toomre, A. 1977, in *The Evolution of Galaxies and Stellar Populations*, ed. B. M. Tinsley and R. B. Larson (New Haven: Yale University Observatory), p. 401.
Tremaine, S. D. 1976, *Ap. J.*, **203**, 72.
Twarog, B. A. 1980, Ph.D. thesis, Yale University.
van den Bergh, S. 1957, *Z. Ap.*, **43**, 236.
———. 1976*a*, *Ap. J.*, **206**, 883.
———. 1976*b*, *A.J.*, **81**, 795.
van den Bergh, S., and McClure, R. D. 1979, *Ap. J.*, **231**, 761.
van der Kruit, P. 1979, *Astr. Ap. Suppl.*, **38**, 15.
Verschuur, G. L. 1978, *Bull. AAS*, **10**, 618.
Visvanathan, N., and Sandage, A. 1977, *Ap. J.*, **216**, 214.
Wu, C.-C., Faber, S. M., Gallagher, J. S., Peck, M., and Tinsley, B. M. 1980, *Ap. J.*, **237**, in press.

C. N. Caldwell, R. B. Larson, and B. M. Tinsley: Yale University Observatory, Box 6666, New Haven, CT 06511