THE ASTROPHYSICAL JOURNAL, Vol. 160, June 1970

# THE INTRINSIC FLATTENING OF E, S0, AND SPIRAL GALAXIES AS RELATED TO GALAXY FORMATION AND EVOLUTION

ALLAN SANDAGE,* KENNETH C. FREEMAN, AND N. R. STOKES
Mount Stromlo and Siding Spring Observatories, Research School of
Physical Sciences, Australian National University
*Received 1969 August 19*

## ABSTRACT

The distribution of apparent axial ratios for 168 E, 267 S0 + SB0, and 254 ordinary spiral galaxies confirms previous conclusions that ellipticals have only moderate intrinsic flattenings, $q$, which range from $q = 1$ to $q \simeq 0.3$. Ordinary spirals and S0's are intrinsically flatter, possessing thin disks of $\langle q\rangle \simeq 0.25$ with only small dispersion. Because of this difference, true spirals and S0 systems cannot have evolved either from, or into, E systems.

Evolutionary questions concern (*a*) why S0 → Im galaxies have flattened to a disk, while E galaxies have not; (*b*) why the spheroidal components of all galaxies contain only old stars, probably formed at a single epoch; and (*c*) why S0 and early Sa galaxies have lost their young spiral-arm population, while Sb → Im galaxies have not.

Analysis of the four properties, (1) the disk blue-light surface brightness at $r = 0$, (2) the absolute length scale of the disk, (3) the size ratio of the spheroidal component to the total radius of the galaxy and (4) the mean total mass density $\langle \rho \rangle$, shows that galaxy type does not depend uniquely on systematic variations of these parameters along the Hubble sequence. The only strongly systematic variable is $\langle \rho \rangle_{\rm H\,I}$.

Observations suggest that the spheroidal subsystem is relaxed. If relaxation occurs by Lynden-Bell's process of a rapidly changing ($\tau_R \simeq$ free-fall time) gravitational potential in the collapsing protogalaxy, then stars in the spheroidal component must be the same age to within the collapse time (several times $10^8$ years). Because the mean angular momentum per unit mass, $h$, of the spheroidal component is low, we suggest that the relative size of this component is determined by the relative amount of low-$h$ matter in the protogalaxy, and that it is this low-$h$ matter which fragments into stars during the rapid collapse phase; the high-$h$ matter subsequently settles in gaseous form to the disk.

The fundamental distinction between E and spiral systems probably lies in their primeval mass-angular momentum distribution.

The morphological type of a spiral or S0 system appears to be essentially defined at the time of formation of the old disk stars.

## I. INTRODUCTION

Galaxies of type S0 are of particular interest because of their importance for the yet unsolved problem of galactic evolution.

Hubble recognized S0 systems as a separate class only after 1936. They were not included in his early classification system (Hubble 1926). He had not isolated them observationally at the time of his writing *The Realm of The Nebulae* (Hubble 1936), although he had come to believe that such systems must exist. The "transition" from ellipticals to true spirals in the 1926 system appeared to be so abrupt as regards the presence or absence of spiral arms that he wrote: "The junction [between E and S] may be represented by the more or less hypothetical class S0—a very important stage in all theories of nebular evolution. Observations suggest a smooth transition between E7 and SBa, but indicate a discontinuity between E7 and Sa in the sense that Sa spirals are always found with arms fully developed [whereas no trace of arms exists in E systems]. Speculation concerning the discontinuity will be unprofitable until more detailed information has been assembled from large-scale photographs" (Hubble 1936, chap. 2).

Observational recognition of S0's came only after long-exposure plates, taken primarily with the 100-inch telescope, became available. The change in the classification

---

* On leave from Mount Wilson and Palomar Observatories.

system to include these galaxies has been described elsewhere (Sandage 1961, hereinafter called *Atlas*); S0 and SB0 galaxies (or SA0 and SB0 in de Vaucouleurs's notation 1956*b*, 1959*a*, 1963*a*) are defined by many type examples such as NGC 1201, NGC 7457 (SA0, *Atlas*, p. 4), NGC 1291 (SB0, de Vaucouleurs 1956*b*), NGC 2859, and NGC 2950 (SB0, *Atlas*, p. 42).[1]

The principal classification features of S0 systems are (1) elliptically shaped images whose apparent axial ratios range from $b/a = 1$ (NGC 524, *Atlas*, p. 6), through more flattened shapes (NGC 4215, *Atlas*, p. 6; $b/a \simeq 0.3$), to forms which are flatter than any known elliptical (e.g., NGC 4762, *Atlas*, p. 8; $b/a \simeq 0.1$); (2) no spiral arms; (3) a bright central region which, in general, resembles a small E galaxy whose diameter can range from about 0.3 to 0.8 of the total extent of the visible galaxy on well-exposed plates; and (4) an extensive outer envelope whose intensity gradient is less steep than that of normal E systems. On large-scale plates exposed to the sky limit, it is usually easy to distinguish S0 and E galaxies solely by the existence of this envelope.

S0 galaxies have the appearance of spirals without arms, and the type was described in this way by Spitzer and Baade (1951). Although these authors explained the class as true spirals stripped of dust, gas, and arms by collisions in clusters of galaxies, it now appears that the S0 phenomenon is more fundamental than this because (1) very many S0 systems exist outside large clusters and (2) the stripping efficiency in clusters is less than originally calculated due to the increased distance scale established after 1957 (cf. Sandage 1958, 1962). Nevertheless, the term "stripped" spiral does describe the form, as shown by measurements of the intensity distribution of the disks of true spirals and S0's.

De Vaucouleurs (1959*a*) has shown that $I(r)$ for the disks of spirals is quite different from that of either the central regions of spirals or of E galaxies. The intensity law for spiral disks is exponential, $I(r) = I_0 \exp(-\alpha r)$, in all cases studied, whereas a steeper law applies to E systems (cf. de Vaucouleurs 1959*a*; King 1966). These conclusions have been strengthened by modern studies of typical spirals such as M31 (Fricke 1954; de Vaucouleurs 1958), M33 (de Vaucouleurs 1959*b*), NGC 300, 6744, 4945 (de Vaucouleurs 1962, 1963*b*; 1964), and many galaxies observed by van Houten, Oort, and Hiltner (1954).

The first photometry of the envelope of an S0 galaxy was done by de Vaucouleurs (1956*b*) for the southern SB0 system NGC 1291, with the result that the disk *follows the exponential law for spirals*. Extensive work by Liller (1960), Johnson (1961), Hodge and Webb (1964), and Hodge and Merchant (1966) has confirmed the result.

## II. INTRINSIC FLATTENING OF S0 SYSTEMS

The similarity in intensity distributions between the "envelopes" of S0 galaxies and the disks of spiral galaxies does not in itself prove that the two types are similar in other properties. For this, it must be shown that the *flattening* of the S0 envelopes is similar to that of the spiral disks, and dissimilar to the intrinsic flattenings of ellipticals.

The problem is statistical and is similar to that already solved for the E galaxies by Hubble (1926). In principle, the frequency function of intrinsic axial ratios for a randomly oriented sample of galaxies can be obtained from the observed distribution of apparent flattening by solving an integral equation.

A preliminary discussion of the problem, applied to field S0 galaxies, and based on a sample of thirty-three S0 systems south of $\delta = -35°$, has been given by de Vaucouleurs (1956*b*, § 13; 1959*a*, § 5). He concluded that S0's are indeed highly flattened systems similar to spirals. We confirm these findings in the present study which is based on a

---

[1] Halftone reproductions are seldom satisfactory in showing the intensity distribution of E and S0 systems. However, the *Palomar Sky Survey* prints show well the distinctive differences between S0 and E galaxies. Besides the galaxies noted, a particularly good example of an S0 is NGC 5102 on *Sky Survey* print $13^h04^m$, $-36°$ in the Whiteoak Extension.

larger sample of E, S0, and SA galaxies listed in the *Reference Catalogue of Bright Galaxies* (de Vaucouleurs and de Vaucouleurs 1964).

In the following, we consider that the forms of S0 and spiral galaxies can be approximated by single ellipsoids of revolution, although a better approximation would be two concentric ellipsoids of different flattening, one representing the nuclear bulge, and the other the flattened disk. We only use the results differentially between spirals and S0 galaxies, so the distinction is unimportant here.

Consider an ellipsoid whose *intrinsic* axial ratio is $q = b_I/a_I$. Hubble (1926) showed that the apparent axial ratio $b/a$ is related to $q$ and to the inclination angle $i$ between the equatorial plane and the observer's line of sight by

$$\cos^2 i = \frac{1 - b^2}{1 - q^2}, \tag{1}$$

where $a_I = a = 1$, without loss of generality. He further showed that the relative frequency of galaxies of given $q$ having inclinations between $i_1$ and $i_2$ is

$$\sin i_2 - \sin i_1 .$$

For a given inclination, the apparent flattening is independent of the azimuth of the observer. Therefore, for a given $q$, and for a random orientation of equatorial planes, the probability of finding apparent flattening between $b$ and $b + db$ is

$$dN(q, b) = \sin (i + di) - \sin i \equiv d (\sin i) . \tag{2}$$

From equations (1) and (2) it follows that

$$dN(q, b) = \frac{bdb}{(1 - q^2)^{1/2}(b^2 - q^2)^{1/2}} . \tag{3}$$

The relative frequency in the discrete interval from $b_1$ to $b_2$ is given by

$$N(q, b_1, b_2) = \frac{1}{(1 - q^2)^{1/2}} \int_{b_1}^{b_2} \frac{bdb}{(b^2 - q^2)^{1/2}} \quad \text{for given } q ,$$

or

$$N(q, b_1, b_2) = \frac{1}{(1 - q^2)^{1/2}} [(b_2^2 - q^2)^{1/2} - (b_1^2 - q^2)^{1/2}] . \tag{4}$$

In an assemblage of galaxies with different intrinsic axial ratios, distributed according to a function $f(q)$, the relative frequency of galaxies between observed flattening $b_1$ and $b_2$ is given by

$$N(b_1, b_2) = \int_0^1 f(q)(1 - q^2)^{-1/2}[(b_2^2 - q^2)^{1/2} - (b_1^2 - q^2)^{1/2}]dq . \tag{5}$$

This is the basic equation of the problem.

We have solved equation (5) numerically for different assumptions of $f(q)$, and with assigned limits of $b_1$ and $b_2$, to obtain predicted distributions for comparison with the following observations. We chose to consider only those systems which appear in the Shapley-Ames catalog (1932) (but as classified in the *Reference Catalogue*), because such galaxies, being brighter than the *Reference Catalogue* limit, may have smaller classification errors than fainter systems.

Our sample of ordinary spirals includes both those classified SA and SAB in de Vaucouleurs's system. We have further restricted the sample to those spirals brighter than

$m_H = 12.5$ so as to (1) obtain comparable numbers of S0 and S galaxies and (2) to avoid to some extent the classification problem of spindle galaxies ($b/a \lesssim 0.3$) where confusion between S0 and spirals is likely for faint galaxies on plates of low resolution.

De Vaucouleurs has corrected the axial ratios for the subjective systematic errors first discussed by Holmberg (1946) and used by de Vaucouleurs (1959*c*). The tabulated axial ratios (which are expressed as log $a/b$ in the *Reference Catalogue*) were sorted into ten discrete intervals whose boundaries are given in Table 1. The intervals in $b/a$ were chosen to be nominally $\Delta(b/a) = 0.09$, but in some cases the actual interval varies by $\pm 0.01$ due to the requirement that log $a/b$ must be discretized to conform with the tabulation in the *Reference Catalogue.*

Table 1 lists the absolute frequencies of E, S0, SB0, Sa, Sb, and Sc galaxies sorted in this way. The histograms of Figure 1 show the resulting distribution for 168 E, 267 S0 and SB0, and 254 Sa, Sb, and Sc galaxies. We have formed these groupings because the individual histograms are not sensibly different for S0 and SB0 or for Sa, Sb, and Sc.

TABLE 1

OBSERVED FREQUENCY DISTRIBUTION OF FLATTENING

| | | | TYPE OF GALAXY | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | S0 | | | SA | | | |
| log $a/b$ | log $b/a$ | BIN No. | E | S0 | SB0 | Total S0 | Sa | Sb | Sc. | Total SA |
| 0.00–0.04........ | 1.00–0.91 | i | 40 | 15 | 6 | 21 | 2 | 11 | 16 | 29 |
| 0.05–0.08........ | 0.90–0.83 | ii | 32 | 17 | 8 | 25 | 3 | 9 | 15 | 27 |
| 0.09–0.13........ | 0.82–0.74 | iii | 29 | 24 | 13 | 37 | 5 | 9 | 10 | 24 |
| 0.14–0.19........ | 0.73–0.65 | iv | 24 | 22 | 11 | 33 | 8 | 15 | 14 | 37 |
| 0.20–0.25........ | 0.64–0.56 | v | 24 | 13 | 13 | 26 | 3 | 11 | 13 | 27 |
| 0.26-0.33........ | 0.55–0.47 | vi | 14 | 24 | 7 | 31 | 5 | 22 | 11 | 38 |
| 0.34–0.43........ | 0.46–0.37 | vii | 4 | 19 | 9 | 28 | 3 | 11 | 10 | 24 |
| 0.44–0.55........ | 0.36–0.28 | viii | 1 | 27 | 9 | 36 | 8 | 10 | 4 | 22 |
| 0.56–0.72........ | 0.27–0.19 | ix | 0 | 23 | 4 | 27 | 1 | 11 | 8 | 20 |
| 0.73–1.00........ | 0.18–0.10 | x | 0 | 2 | 1 | 3 | 0 | 2 | 4 | 6 |
| Total.......................... | | | 168 | 186 | 81 | 267 | 38 | 111 | 105 | 254 |

The principal result of Figure 1 is that the distribution of S0 galaxies is much different from that of ellipticals and more closely resembles that for ordinary spirals. The result, of course, is not new, but it strengthens previous conclusions because the sample is the largest yet available for accurately classified galaxies.

Superposed on the histograms in Figure 1 are the frequency polygons showing the theoretical predictions obtained from equation (5). The weighting functions $f(q)$ are plotted to the right in Figure 1. The integrations were performed numerically on the computer of the Mount Stromlo Observatory using steps of 0.01 in $db$ and $dq$.

Three different forms of the assumed intrinsic $f(q)$ distributions are presented for E galaxies:

a) $f(q) = 1\,, \quad 0.3 \leq q \leq 1$ (a uniform distribution) ,

b) $f(q) \propto \exp\,[-(q - q_0)^2/2\sigma^2]$ (a Gaussian distribution) ,

c) $f(q) \propto \left(1 + \frac{q - q_0}{a}\right)^p \exp\,[-p(q - q_0) \quad ]$(skewed binomial distribution) .

The uniform case does not agree particularly well with the histogram. It predicts more E0, E6, and E7 galaxies and fewer E2, E3, and E4 galaxies than actually observed. Comparison of the observed and predicted frequencies given in Table 2 shows that the uniform case is outside the $\sqrt{N}$ standard deviation for many of these subclasses.

A range of parameters for predicted curves *b* and *c* has been considered. The curves in Figure 1 are representative of those which give improved fits to the observations. All but one point in case *b* and all points in case *c* are within the $\sqrt{N}$ tolerance. Case *c* is particularly attractive because it predicts *no* elliptical galaxies with an intrinsic axial ratio less than 0.3 but still allows some truly spherical systems.

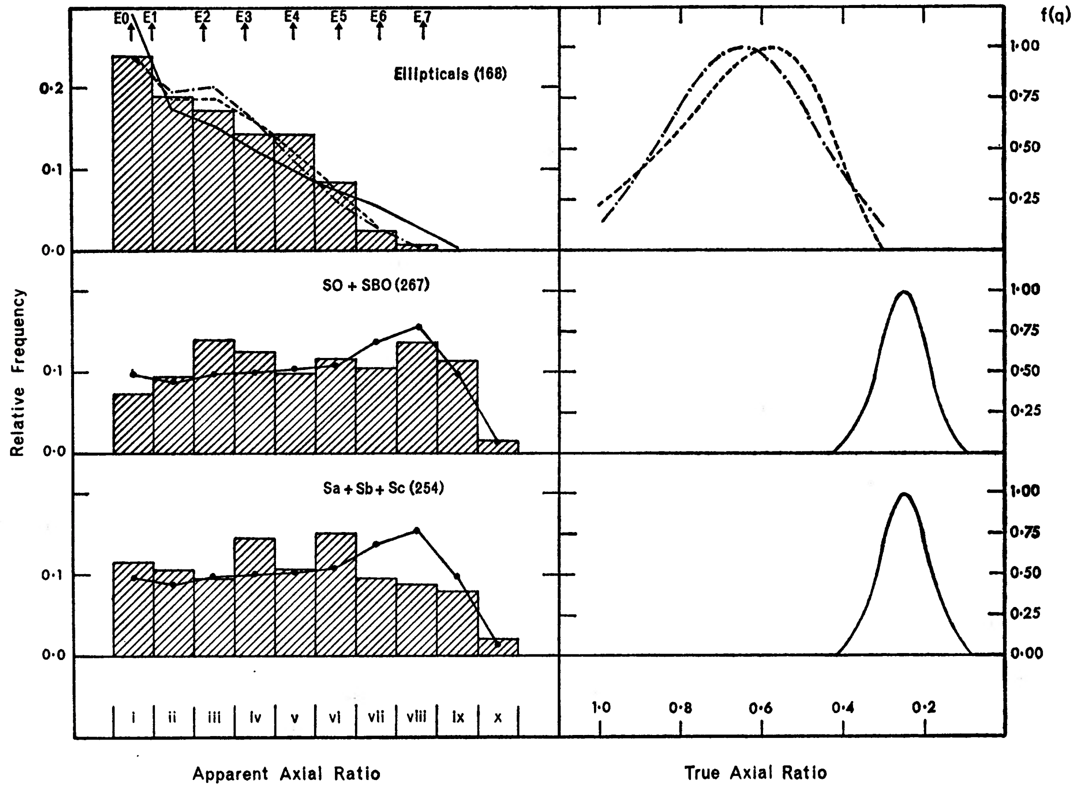

FIG. 1.—*Left*, histograms of the distribution of apparent axial ratios for E, S0, and spiral galaxies, sorted into class intervals defined in Tables 1 and 2. The curves are predicted ratios for various assumptions of the distribution of intrinsic flattening. *Right*, assumed intrinsic distributions corresponding to the curves on the left. The case of uniform flattening for E galaxies is a straight horizontal line truncated at $b/a = 0.3$ and is not shown. Dot-dash curve is a Gaussian with $q_0 = 0.65$ and $\sigma = 0.18$. Dashed curve is a skewed binomial distribution defined in the text, with $q_0 = 0.58$, $a = 0.31$, and $p = 3.0$. The two solid curves are Gaussians with $q_0 = 0.25$ and $\sigma = 0.06$.

The representation of the elliptical galaxies by simple ellipsoids must be a very good approximation, and one would expect that predicted curves of frequency distribution of intrinsic axial ratios could be found which fit the observations well. But the representation of lenticulars and spirals by ellipsoids of revolution is an approximation of the first order only. We found that frequency distributions centered at about $q = 0.25$, and having small dispersion about that $q$, came closest to describing the S0 and S observations. A single Gaussian with $q_0 = 0.25$ and $\sigma = 0.06$ is shown for both the S0 and SB0 and the Sa, Sb, and Sc groups in Figure 1.

The result that E galaxies may not be uniformly distributed in $q$ but may have a peak near $q \simeq 0.6$ is of some cosmogonic interest, but we are not convinced that either this result or the difference between the observed and predicted frequencies for spirals is necessarily real because of (1) the smallness of the sample and (2) possible remaining systematic errors in the corrected estimates of $b/a$. However, our major conclusion is the same as that of Hubble (1926) and of most others who have considered the problem, (cf. Bruggencate 1930; Machiels 1930; de Vaucouleurs 1959*a*). E galaxies exist for all intrinsic flattening values from $q = 1.0$ to $q = 0.3$, whereas the observed flattenings of spirals (and now S0's) can be explained only if they are all intrinsically flat disks with nearly the same value of $q$. Similar conclusions were reached by Rood and Baum (1967), who used a smaller sample of galaxies which they classified in the Coma cluster.

We are forced to conclude from these results that S0's cannot have evolved either to or from ellipticals because the flattenings are so fundamentally different. The flattening of systems of stars alone must be an endowed property. It cannot temporally change if galaxy ages are less than several relaxation times ($\sim 10^{12}$ years in the equilibrium state).

TABLE 2

COMPARISON OF OBSERVED AND PREDICTED FREQUENCY DISTRIBUTION OF FLATTENING FOR VARIOUS MODELS

| | | ELLIPTICALS | | | | S0+SB0 | | Sa+Sb+SC | |
|---|---|---|---|---|---|---|---|---|---|
| | | | Predicted | | | | | | |
| log $a/b$ | log $b/a$ | Observed | Model a | Model b | Model c | Observed | Predicted Model | Observed | Predicted Model |
| 0.00–0.04.. | 1.00–0.91 | 40 | 49.5 | 39.8 | 40.7 | 21 | 25.8 | 29 | 24.5 |
| 0.05–0.08.. | 0.90–0.83 | 32 | 29.0 | 33.0 | 31.2 | 25 | 23.2 | 27 | 22.1 |
| 0.09–0.13.. | 0.82–0.74 | 29 | 25.8 | 33.7 | 31.8 | 37 | 26.4 | 24 | 25.1 |
| 0.14–0.19.. | 0.73–0.65 | 24 | 20.5 | 27.1 | 26.9 | 33 | 26.8 | 37 | 25.5 |
| 0.20–0.25.. | 0.64–0.56 | 24 | 16.1 | 18.5 | 20.2 | 26 | 27.6 | 27 | 26.3 |
| 0.25–0.33.. | 0.55–0.47 | 14 | 12.2 | 10.3 | 12.1 | 31 | 29.0 | 38 | 27.6 |
| 0.34–0.43.. | 0.46–0.37 | 4 | 9.4 | 4.7 | 4.7 | 28 | 37.0 | 24 | 35.2 |
| 0.44–0.55.. | 0.36–0.28 | 1 | 4.8 | 0.9 | 0.4 | 36 | 41.5 | 22 | 39.5 |
| 0.56–0.72.. | 0.27–0.19 | 0 | 0.7 | 0 | 0 | 27 | 26.1 | 20 | 24.8 |
| 0.73–1.00.. | 0.18–0.10 | 0 | 0 | 0 | 0 | 3 | 3.6 | 6 | 3.4 |
| Total.................. | | 168 | 168 | 168 | 168 | 267 | 267 | 254 | 254 |

Therefore, if galaxy evolution is taking place along the sequence of forms, there has been no exchange between the E and S branches of the Hubble sequence. If an evolutionary origin of S0 galaxies is sought, it seems clear that they must have originated from previous disklike structures, whose spatial distribution resembles the old disk population of present-day spirals. But how?

## III. DOES EVOLUTION OCCUR ALONG THE SPIRAL SEQUENCE?

Perhaps the most remarkable property of the Hubble classification is that galaxies classified on the basis of form alone (the criteria are [1] openness of the spiral pattern, [2] apparent ratio of the central bulge to the exponential disk, and [3] the degree of resolution of the arms into luminous stars) show a systematic change in stellar content *which follows the same order as the form sequence.* In the order Sd to Sa (including both ordinary and barred), these progressive changes are (1) decreasing absolute luminosity for the brightest stars in regions of the spiral arms, (2) decreasing percentage of mass in the form of gas and dust, (3) decreasing numbers and sizes of spiral-arm H II regions (undoubtedly related to point 1), and (4) increasingly red mean $(B - V)$ and $(U - B)$

colors, despite an approximately constant mean mass-to-blue-light ratio (Roberts 1970).

The differences could possibly be explained in terms of the progressive evolution of the stellar content by well-understood processes. Consider, for example, what must eventually happen to the present stellar content of a contemporary Sc galaxy as the stars evolve. Luminous stars are obviously still forming, as shown by the presence of H II regions and O, B, and F supergiants. New generations of stars are made from the abundant gas and dust still present. But as time proceeds, each succeeding generation will progressively deplete the available pre-stellar material, locking more and more mass into dwarfs fainter than $M_V \simeq +5$—stars which cannot replenish the gas supply by mass loss in times less than $\sim 10^{10}$ years.

After star formation either has ceased or has been reduced to an unimportant level, the H-R diagram of the stellar content will evolve into an NGC 188 configuration and beyond. The final stage of the stellar content of the depleted galaxy, with no new stars or spiral arms, will resemble that of present day S0 systems. However, it is not clear that this depletion process has had enough time to turn an Sc system into an S0. For example, Roberts (1963) shows that spiral systems probably have sufficient gas to sustain their present rates of star formation for more than $10^{10}$ years.

We are now faced with the following problem. All galaxies appear to be nearly the same age. The evidence comes from the presence of an underlying old disk population and an extreme Population II (globular clusters with low metal abundances) which seem to be similar in stellar content in such young-appearing galaxies as the Large Magellanic Cloud and M33, as well as in the central regions of M31 and all elliptical systems. Why, then, are some galaxies already in the final S0 stage while others, such as Sc, Sd, and Sm galaxies, have not progressed far along the Hubble sequence? No one doubts that present S0 galaxies once had bright OB stars, gas, and possibly spiral arms. Why have these particular systems lost their young population while others, of the same age, have not?

### *a) Some Photometric and Integrated Properties of Spiral Systems*

The stellar content is not the only quantity that varies systematically along the spiral sequence. We list here some others that may be relevant to the evolution problem.

1. Roberts (1970) shows that the mean mass-to-blue-light ratio $\mathfrak{M}/L$ is approximately constant along the spiral sequence, at $\mathfrak{M}/L \simeq 12$ ($L$ is on the $B(0)$ system of the *Reference Catalogue*).

2. As mentioned earlier, the radial luminosity distribution in the disks of spiral and S0 systems follows closely the law $I(r) = I_0 \exp(-\alpha r)$, whereas the nuclear-bulge component follows a steeper law. Freeman (1970) has derived photometric quantities for the *disks* of thirty-six galaxies by extrapolating the observed exponential function to $r = 0$. These include (*a*) the *central* surface brightness (corrected to the face-on value) $B_0 = -2.5 \log I_0$ of the disk component alone, (*b*) the absolute length scale $\alpha^{-1}$ (in kpc) for the disks, and (*c*) the apparent size ratio $R_{\rm sph}/R_{26.5}$ of the spheroidal-bulge component to the disk radius at a (face-on) surface brightness $B_0 = 26.5$ mag per square second of arc. He finds:

A. For twenty-eight of the thirty-six galaxies, the central-disk surface brightness corrected for galactic absorption is nearly *constant* at $B_0 = 21.6$ mag $\pm\ 0.3(\sigma)$ along the entire form sequence from S0 to Im, despite the range in absolute luminosity from $-16.0 > M_B > -20.6$. This remarkable result is violated by only seven of the thirty-six galaxies, which average 2 mag brighter in $B_0$ (the violators themselves are of all types from S0 to Sdm), and by IC 1613 (Im) with $B_0 = 23.7$ mag per square second of arc (Freeman 1970, Fig. 5).

B. The length scale $\alpha^{-1}$ of the disk depends on morphological type (Freeman 1970, Fig. 6). *Only* galaxies of earlier types from S0 to about Sbc have large disks ($\alpha^{-1} \simeq 5$ kpc), whereas galaxies of *all* types from S0 to Sm have small disks ($\alpha^{-1} \simeq 1$ kpc).

The conclusion from facts 1, 2A, and 2B is that *all* late-type spirals (Sd–Sm) have *low-mass disks*, while S0 and early-type spirals include systems with both high- and low-mass disks. This follows from the constancy of both $B_0$ and $\mathfrak{M}/L$ with type, and the observed restriction of large $\alpha^{-1}$ values to early types.

C. Among the spirals (Sa–Sm) there is a very loose correlation of $R_{\mathrm{sph}}/R_{26.5}$ with type, such that the ratio is largest at Sa. This agrees with the visual impression of decreasing relative size of the bulge component along the sequence (one of the classification criteria), but the correlation has large scatter (Freeman 1970, Fig. 10). S0 galaxies cover the *entire* range of $R_{\mathrm{sph}}/R_{26.5}$ from 0.8 to 0.25, avoiding only the smallest values from 0.25 to ~0.1 given for some late-type Sc–Sm systems. The large range for S0's confirms the visual impression that small central bulges exist for some such galaxies (cf. NGC 4215, *Atlas*, p. 6; NGC 4762, *Atlas*, p. 8). Small bulge components exist for some Sa and Sb galaxies as well (cf. NGC 4941, *Atlas*, p. 10; NGC 4866, *Atlas*, p. 11; NGC 4293, *Atlas*, p. 11), a fact that violates one of the Hubble criteria for these particular systems.

TABLE 3

TWO PAIRS OF GALAXIES WITH SIMILAR PHOTOMETRIC PARAMETERS

| Parameter | NGC 4503 | NGC 1313 | NGC 1332 | NGC 5005 |
|---|---|---|---|---|
| Class | SB0⁻ | SB(s)d | S(s)0⁻ | SAB(rs)bc |
| $M_B$ | −18.4 | −18.7 | −19.9 | −20.1 |
| $R_{\mathrm{sph}}/R_{26.5}$ | 0.33 | 0.40 | 0.64 | 0.65 |
| $\alpha^{-1}$ (kpc) | 1.5 | 1.6 | 4.0 | 3.5 |
| $B_0$ (mag per sq. sec of arc) | 21.7 | 21.5 | 22.0 | 21.8 |

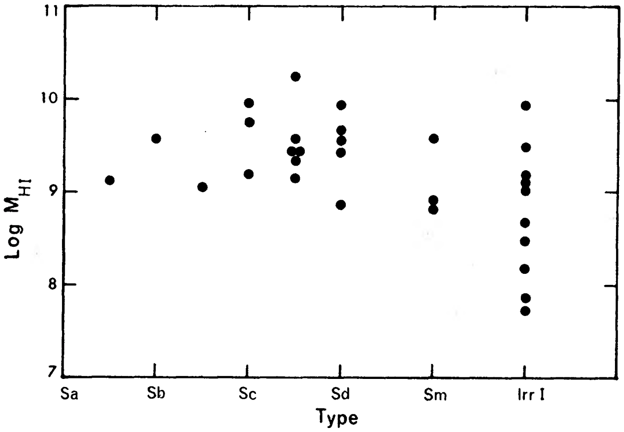

FIG. 2.—Total hydrogen mass in solar units from Roberts (1970) vs. morphological type. Except for the larger spread at Irr I, the correlation is weak or absent.

Table 3 gives data for two pairs of galaxies; each pair has very similar photometric parameters, but considerably different morphological types. The similarity of (*a*) total absolute luminosity (disk plus spheroidal component), (*b*) the length scale $\alpha^{-1}$, (*c*) the $R_{\mathrm{sph}}/R_{26.5}$ ratio, and (*d*) the disk $B_0$ value shows that the mass ratio of the spheroidal component to the disk is nearly the same for each pair of galaxies, despite their different morphological types.

3. Roberts (1970) gives the neutral-hydrogen mass $\mathfrak{M}_{\mathrm{H\,I}}$ for thirty spirals. Figure 2

shows that $\mathfrak{M}_{\mathrm{H\,I}}$ is not strongly correlated with type. The only trend is the large spread among the low mass Ir I systems. It is important to know whether the mean H I density $\langle\rho\rangle_{\mathrm{H\,I}}$ is a function of type. Although accurate H I dimensions are available for only a few of the thirty systems and Roberts points out that, since the H I dimensions may be 1.5 to 2 times the optical dimensions, estimates of the mean $\rho_{\mathrm{H\,I}}$ from optical diameters are uncertain, we have made a preliminary calculation.

Figure 3 shows $\log \mathfrak{M}_{\mathrm{H\,I}}$–$3 \log D(0)$, where $D(0)$ is the optical face-on diameter from the *Reference Catalogue* in arbitrary units (distances used are those given by Roberts 1970). This is a measure of $\log \rho_{\mathrm{H\,I}}$ if we assume that the H I layer in all systems has the same ratio of thickness to diameter and that the optical and H I dimensions are similar.

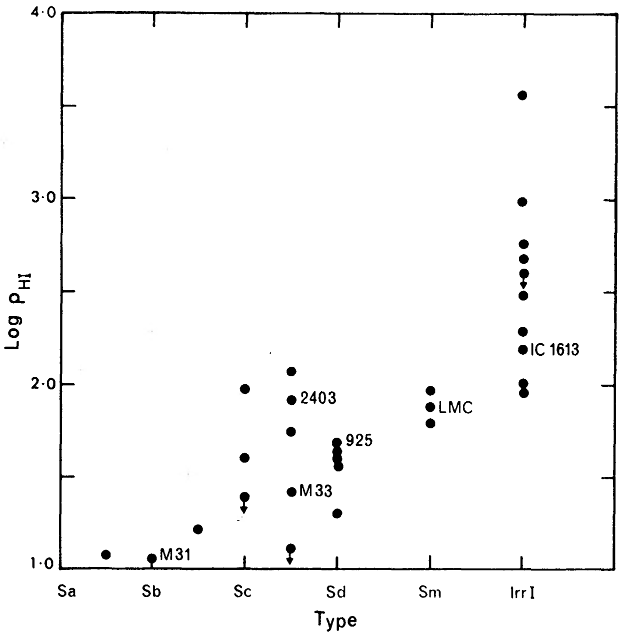

FIG. 3.—Variation of hydrogen mass density (in arbitrary units) with morphological type. Galaxy diameters are optical values. For LMC, M31, IC 1613, NGC 925, and M33, the optical diameter convolved with the 21-cm beam diameter is nearly equal to the observed H I diameter. Arrows denote galaxies whose optical diameters are known to be smaller than H I diameters, causing the true $\rho_{\mathrm{H\,I}}$ to be smaller than the plotted value.

The labeled points are systems whose optical diameter convolved with the 21-cm beam diameter is nearly equal to the observed H I diameter. For these systems $\rho_{\mathrm{H\,I}}$ should be relatively accurate. Systems with arrows have convolved optical diameters significantly less than the observed H I diameter, so the calculated $\rho_{\mathrm{H\,I}}$ is too high. Figure 3 indicates that the mean $\rho_{\mathrm{H\,I}}$ is larger for the later-type systems by factors ranging up to $10^3$.

4. Holmberg (1964) finds that the mean total mass density $\langle\rho\rangle$ in spiral galaxies is a function of type: early-type systems have the greatest $\langle\rho\rangle$. He infers that evolution of an individual system along the sequence Sd→S0 is therefore unlikely because the necessary increase in $\langle\rho\rangle$ would be unacceptably high. But to calculate $\langle\rho\rangle$ Holmberg uses masses $\mathfrak{M}$ which have been derived statistically from the luminosity and color of the galaxy through the relation $\log \mathfrak{M}/L = 2.23C - 0.36$ ($L$ and $C$ are absorption-free integrated absolute photographic luminosity and color). For the seventeen systems in common to Holmberg's and Roberts's (1970) lists, we can compare the statistical masses with the dynamical masses given by Roberts as in Table 4 and Figure 4. There are no Sa galaxies in common. The mean Sb masses from the two lists appear to be in fair agree-

ment, but the difference $\log \mathfrak{M}$ (Roberts) − $\log \mathfrak{M}$ (Holmberg) $\simeq 0.6$ for Sc and Ir I systems. This difference is large enough to account entirely for the apparent difference $\langle\rho\rangle_{\mathrm{Sb}} - \langle\rho\rangle_{\mathrm{Sc}}$ found by Holmberg. Therefore, if the dynamical masses are correct, then $\langle\rho\rangle_{\mathrm{Sb}} \simeq \langle\rho\rangle_{\mathrm{Sc}}$ rather than varying with galaxy type. Until this difficulty can be resolved, we feel that the implications of Holmberg's $\langle\rho\rangle$-type relation for evolution along the spiral sequence must remain uncertain.

To summarize: (*a*) Table 3 indicates that the galaxy type does not depend on the background stellar distribution alone. (*b*) Late-type systems appear to have high H I densities and small values for the disk length scale $\alpha^{-1}$, while the S0 and early-type spirals have low $\rho_{\mathrm{H\,I}}$ and both large and small $\alpha^{-1}$. (*c*) The ratio $R_{\mathrm{sph}}/R_{26.5}$ is, in the mean,

TABLE 4

GALAXIES IN COMMON TO ROBERTS AND HOLMBERG

| Galaxy | Holmberg Type | $\log \mathfrak{M}_{\mathrm{Roberts}}$ | $\log \mathfrak{M}_{\mathrm{Holmberg}}$ |
|---|---|---|---|
| NGC 224 | Sb⁻ | 11.49* | 11.5* |
| NGC 598 | Sc⁺ | 10.59* | 9.9* |
| NGC 628 | Sc⁻ | 10.81‡ | 10.6† |
| NGC 925 | Sc⁺ | 10.82* | 10.1† |
| NGC 2403 | Sc⁺ | 10.96‡ | 10.0† |
| NGC 3031 | Sb⁻ | 11.15* | 11.4* |
| NGC 3034 | Ir II | 10.28‡ | 10.7† |
| NGC 4214 | Ir I | 9.58‡ | 9.1† |
| NGC 4244 | Sc⁺ | 10.79‡ | 9.3† |
| NGC 4472 | E | 12.11* | 12.0† |
| NGC 4631 | Sc⁺ | 10.63* | 10.4* |
| NGC 4656 | Ir I | 10.30‡ | 9.8† |
| NGC 5194 | Sc⁻ | 10.90* | 11.2† |
| NGC 5457 | Sc⁻ | 11.38‡ | 10.2† |
| IC 1613 | Ir I | 8.40‡ | 8.3* |
| IC 2574 | Ir I | 9.95‡ | 8.9† |
| Ho II | Ir I | 9.86‡ | 8.9† |

* Dynamical mass. † Statistical mass. ‡ H I profile mass.

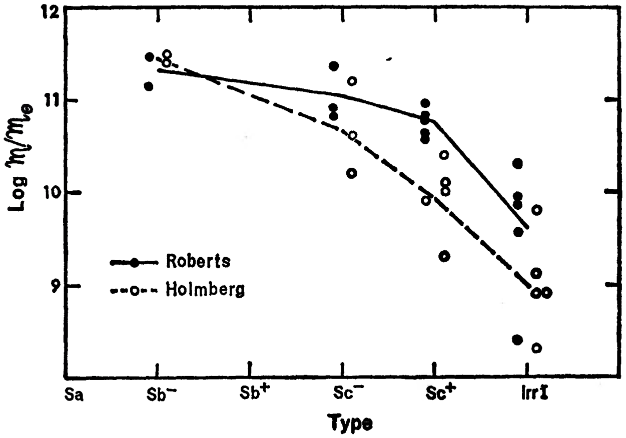

FIG. 4.—Comparison of Holmberg's masses (*open circles*) with Roberts's values (*filled circles*). Systematic difference at Sc and Irr is in the sense to reduce or remove Holmberg's correlation of $\langle\rho_{\mathrm{Total}}\rangle = f(\mathrm{type})$.

smaller for the later-type spirals, but the S0 systems can have all except the lowest values of this ratio in the range defined by the Sa–Sd galaxies.

### *b) The Formation Picture*

Eggen, Lynden-Bell, and Sandage (1962) infer from the motions of very old stars in the solar neighborhood that the Galaxy collapsed from its protocloud on a time scale of the order of $2 \times 10^8$ years. During this rapid collapse, some star formation occurred and the system became progressively enriched in metals. However, much of the protocloud remained in gaseous form throughout the collapse. This matter lost energy through inelastic cloud-cloud encounters, each element conserving its angular momentum, and settled into a disk in approximate centrifugal equilibrium. This disk is identified with the disk of the Galaxy, and contains stars at least as old as NGC 188 ($\sim 8 \times 10^9$ years). We further identify it with the exponential disks of spiral and S0 systems.

The spheroidal [bulge] components of spiral galaxies appear to follow de Vaucouleurs's (1959*a*) $r^{1/4}$ law for the radial distribution of surface brightness. King (1966) has shown that elliptical galaxies, whose surface brightnesses also follow the $r^{1/4}$ law fairly closely, are still better represented by the $I(r)$ distribution which corresponds to the quasi-isothermal stellar energy distribution function

$$f(E) = A(e^{-\beta E} - e^{-\beta E_{\text{escape}}}) , \tag{6}$$

where $E$ is the energy/mass and $A$, $\beta$, and $E_{\text{escape}}$ are parameters. This implies that E galaxies are almost relaxed, even though the *present* relaxation time for the equilibrium galaxy is so long (more than $10^{12}$ years). However, for a galaxy collapsing rapidly from the intergalactic medium, Lynden-Bell (1967) has shown that the distribution function of equation (6) is the most probable result of the relaxation that occurs because the gravitational potential is changing rapidly on the time scale of a stellar orbit.

It seems reasonable to infer from the $r^{1/4}$ law that the spheroidal components of spiral systems are nearly relaxed, and that this relaxation occurred through the rapid collapse via Lynden-Bell's violent relaxation process. Because of the long classical relaxation time, this means that the spheroidal component has not changed *dynamically* since its formation. Has the exponential disk changed dynamically since its formation? This disk seems to be a very general property of spirals and S0 systems. Freeman (1970) shows that there exists only one exponential disk of given total mass and angular momentum, and this strongly suggests that the disk also has not altered in its large-scale dynamical properties since it formed. Apart from the effects of stellar evolution, this means that the old stellar disk and spheroidal components, which contain most of the mass of spiral systems, have not changed since their formation.

What determined the relative sizes and masses of the spheroidal and exponential disk components? The exponential disk in centrifugal equilibrium has a fixed mass–angular momentum distribution $\mathfrak{M}(h)$, where $\mathfrak{M}(h)$ is the total mass with angular momentum per unit mass less than $h$. Because the spheroidal component is much more centrally concentrated than the disk, its mean angular momentum per unit mass is certainly low; for example, in M31, less than 5 percent of the total angular momentum lies in the spheroidal component (Takase 1967), although about 25 percent of the blue light comes from this component. It is also known from the kinematics of globular clusters and RR Lyrae stars that the angular momentum per unit mass of the halo population of our Galaxy is smaller than that of the disk. Kinman's results (1959) do not contradict this, as he included part of the spheroidal component in his calculation of $h$ for the disk. If the collapse picture described above is correct, then $\mathfrak{M}(h)$ is invariant for the system, and must be regarded as a property of the protocloud before its collapse. In particular, the exponential nature of the disk means that $\mathfrak{M}(h)$ is functionally similar for all protoclouds destined to be spiral or S0 systems, at least for the higher values of $h$ corresponding to the *disk* material. It then follows that the size of the *spheroid* depends on the amount of low-$h$ material originally present in the protocloud; a system like M33, with a weak spheroidal component, had in its protocloud only a small excess of low-$h$ matter above the requirements of its exponential disk. Further, *if the $r^{1/4}$ law implies relaxation, then it is the material with low angular momentum per unit mass that fragmented into stars during the rapid-collapse phase.* Rapid star formation in the spheroidal subsystem (on the time scale of the collapse) follows from the remarkable agreement between the observed $I(r)$ intensity distribution and that predicted by Lynden-Bell's violent relaxation mechanism.

Rapid star formation can also be deduced by an independent argument. Because E galaxies and the spheroidal subsystem of spirals are not flat, stars in this component must have large $Z$ energies. There are only two possibilities: either (1) the energy is that of the original protocloud before it was destroyed by gas-gas inelastic interactions or (2) the energy was put into the system at some later time. If possibility 1 applies, stars must have formed from the protocloud before the $Z$ energies decayed, otherwise all matter would be in a disk and no spheroidal component would exist. The decay time of the initial turbulence is generally estimated to be $\sim 5 \times 10^8$ years (cf. Schwarzschild 1964). Although possibility 2 cannot be excluded on energetic grounds (some sort of subsequent Seyfert galaxy or QSO-like input could be imagined), we would be left with the unexplained problem of why the spheroidal component appears to be relaxed. Since this is naturally explained by Lynden-Bell's violent relaxation process, which requires collapse, possibility 2 has no observational support.

We therefore believe that the low-angular-momentum component of the protocloud fragmented into stars almost completely to form the spheroidal subsystem, and did so in the short time it took the gravitational potential to change appreciably, namely, a few times $10^8$ years. This picture is consistent with the well-known difficulty (Mestel 1965*a*, *b*) of forming stars from high-$h$ matter. It is also consistent with (*a*) the absence of star formation at the present time in the central regions of many spiral systems and (*b*) the apparently old stellar evolutionary state (i.e., NGC 188 + some globular-cluster-like component) of all stars in the spheroidal regions of such galaxies. Observational evidence that globular clusters in the galactic halo are the same age to within less than $10^9$ years (Sandage 1969) is also consistent with this view.

The picture suggests that the fundamental distinction between spiral and elliptical systems lies in their primeval distribution of angular momentum $\mathfrak{M}(h)$. Elliptical systems appear almost relaxed, so their absence of a disk and their only moderate flattening (§ II) betray their low mean $h$. It is then not surprising that the whole mass of E systems was able to fragment *during* the rapid-collapse phase, and that star formation was essentially complete in the first $10^9$ years or less.

### *c) The Evolution Picture*

We adopt the following premises: (i) The old background stellar disk and bulge components in a spiral system have not changed (except by stellar evolution) since they formed. (ii) Galaxies later than about Scd have only small ($\simeq 2$ kpc) disk length scales, while systems earlier than about Sbc have any value of $\alpha^{-1}$ between 1 and 5 kpc. (iii) The present value of the mean $\rho_{\text{H I}}$ is greater for the later-type systems.

Consider again the pairs of systems described in Table 3. Each pair has members of widely different types but with similar values of $M_{\text{B}}$, $B_0$, $R_{\text{sph}}/R_{26.5}$, and the length scale $\alpha^{-1}$. This means that they have similar total masses and angular momenta (Freeman 1970). The probable distinction between these systems is then *the different values for their mean gas density $\rho_{\text{H I}}$*; this is strongly supported by Sanduleak's (1969) discovery that the density of very young stars ($M_{\text{pg}} \lesssim -6$) in the Small Magellanic Cloud is proportional to $\rho_g^{1.84\pm0.14}$, where $\rho_g$ is the local H I density (cf. Schmidt's [1959] result for the Galaxy).

If the old disk stars in these galaxies formed at about the same time $\tau_D$, then there are two basic ways to account for this distinction. (*a*) Immediately after $\tau_D$, each pair of systems had similar values of $\rho_{\text{H I}}$ and had similar appearances as Sd or later; the depletion of the interstellar medium then proceeded at different rates in these systems, despite their apparent structural and compositional similarity. (*b*) After the formation of the old disk stars at $\tau_D$, there was relatively little gas left, for example, in NGC 4503 (S0) and rather more ($\gtrsim$10 percent by mass) in NGC 1313 (Sd), and subsequent stellar evolution via new star formation has been so slow as not to interchange the initial gas ratios. We are distinguishing the two basic possibilities that the present morphological type of a spiral system either was set up at the time of formation of the old disk stars or has been determined since that time by the evolutionary rate of the individual system.

No one understands in any complete detail how spiral systems evolve. Nevertheless, it seems very likely that the evolutionary rate, i.e., the rate of depletion of the extreme Population I component, depends on some large-scale physical parameter. It is then difficult to see how two systems as similar in their background properties as NGC 4503 and 1313, or as NGC 1332 and 5005, could have evolved at very ***different*** rates if their $\langle\rho\rangle_{\text{H I}}$ after $\tau_D$ were ***similar*** as in proposition *a*. We therefore prefer to accept *b* above, that the present morphological type of spiral systems was determined by the relative amount of gas left over after the stars of the old disk population formed. Since this time, if galaxies are coeval, all spiral systems will have evolved to earlier types by depletion of their interstellar matter, but the extent of this evolution is severely limited by the present existence of Sd and Sm systems whose stellar content has not evolved far toward evolutionary "completion" (NGC 188 stage). This preference is consistent with Roberts's (1963) estimate that Irr, Sc, and Sb systems contain enough interstellar gas to continue their present rate of star formation for at least $10^{10}$ years, and with his conclusion that the identifying stellar feature of one morphological type will not change to those of another type during a period of order $10^{10}$ years. The preference is also consistent with current but incomplete observational data that broad-band colors (and the more detailed $I(\lambda)$ functions) of E, S0's, and the disks of early-type spirals are so similar. They would have been expected to differ if stars in the disks of S0's and early-type spirals had formed over all times, as evolution *along* the Hubble sequence would imply, compared with the *single* epoch of star formation in E galaxies.

### *d*) *Discussion*

If the conclusions of § III*c* are correct, then all we have done is to deduce that the morphological type of a spiral or S0 galaxy is essentially defined at the time of formation of the old disk stars, and that any subsequent change in type is small. To be sure, galaxies must have looked different in ancient times (the first $10^9$ years) when the old disk stars were of appreciably earlier mean spectral type. We mean only that the present forms of galaxies along the Hubble sequence are controlled more by the initial conditions of formation than by subsequent evolution along the sequence.

There are two immediate and probably related problems:

i) Why is the amount of gas left over after the formation of the old disk stars so different in systems of different type which are otherwise apparently fairly similar (in length scale, in $R_{\text{sph}}/R_{26.5}$, and therefore in total mass and angular momentum)?

ii) Why do most systems of type later than about Scd have small $\alpha^{-1}$ ($\simeq$2 kpc) and therefore low total mass and angular momentum? It is these systems that have the largest fraction of gas left over after the epoch of formation of old disk stars. (Earlier-type systems have any value of $\alpha^{-1}$ between about 1 and 5 kpc).

The following observational data would be most useful in resolving these difficulties: (*a*) Information about the stellar content of the old disks of spiral and S0 systems, from regions in these systems where the young Population I is obviously weak or absent. For example, does this content change with $\alpha^{-1}$ (or absolute magnitude) or with type? In particular, does this content change markedly near type Scd? Extensive multicolor observations of the underlying disks of M33, NGC 300, and the LMC compared with M31 and earlier galaxies would be a first step. (*b*) Surface photometry of more late type systems, to define problems (ii) more clearly.

Much of this section is necessarily speculative. However, from the observed regular variation of the young Population I stellar content and total gas content along the Hubble sequence, and from the apparent similarity in large-scale properties of the old disk and spheroid population in systems of very different morphological type, we are led to believe that the S0 galaxies are not produced by the evolution of late-type systems along the Hubble sequence. Rather, we believe that they and the other morphological types were essentially defined by the relative amounts of gas left over when the old disk stars formed, soon after the collapse from the intergalactic medium.

## REFERENCES

Bruggencate, P. ten. 1930, *Zs. f. Ap.*, **1**, 275.
Eggen, O. J., Lynden-Bell, D., and Sandage, A. 1962, *Ap. J.*, **136**, 748.
Freeman, K. C. 1970, *Ap. J.*, **160**, 811.
Fricke, W., 1954, *Zs. f. Ap.* **34**, 137.
Hodge, P. W., and Merchant, A. N. 1966, *Ap. J.*, **144**, 875.
Hodge, P. W., and Webb, C. J. 1964, *Ap. J.*, **140**, 681.
Holmberg, E. 1946, *Medd. Lund Obs.* (II), No. 117.
———. 1964, *Ark. f. Astr.*, **3**, 387.
Houten, C. J. van, Oort, J. H., and Hiltner, W. A., 1954, *Ap. J.*, **120**, 439.
Hubble, E. P. 1926, *Ap. J.*, **64**, 321.
———. 1936, *Realm of the Nebulae* (New Haven: Yale University Press).
Johnson, H. M. 1961, *Ap. J.*, **133**, 314.
King, I. R. 1966, *A.J.*, **71**, 64.
Kinman, T. D. 1959, *M.N.R.A.S.*, **119**, 559.
Liller, M. H. 1960, *Ap. J.*, **132**, 306.
Lynden-Bell, D. 1967, *M.N.R.A.S.*, **136**, 101.
Machiels, A. 1930, *Bull. Astr. Paris*, **6**, 317.
Mestel, L. 1965*a*, *Quart. J.R.A.S.*, **6**, 161.
———. 1965*b*, *ibid.*, **6**, 265.
Roberts, M. S. 1963, *Ann. Rev. Astr. and Ap.*, **1**, 149.
———. 1970, in *Stars and Stellar Systems*, Vol **9**, ed. A. and M. Sandage (Chicago: University of Chicago Press), chap. 11 (preprint).
Rood, H. J., and Baum, W. A. 1967, *A.J.*, **72**, 398.
Sandage, A., 1958, *Ap. J.*, **127**, 513.
———. 1961. *Hubble Atlas of Galaxies* (Carnegie Inst. of Washington, Pub. 618).
———. 1962, *Problems of Extragalactic Research*, ed. G. McVittie (New York: Macmillan Co.), p. 359.
———. 1969, *Ap. J.*, **157**, 515.
Sanduleak, N. 1969, *A.J.*, **74**, 47.
Schmidt, M., 1959, *Ap. J.*, **129**, 243.
Schwarzschild, M. 1964, in *Clusters and Stellar Evolution, R.O.B.*, **82**, 55–59.
Shapley, H., and Ames, A. 1932, *Harvard Ann.*, Vol. **88**, No. 2.
Spitzer, L., and Baade, W. 1951, *Ap. J.*, **113**, 413.
Takase, B. 1967, *Pub. Astr. Soc. Japan*, **19**, 427.
Unsöld, A. 1967, *Zs. f. Phys.*, **202**, 13.
Vaucouleurs, G. de. 1956*a*, *Mem. Commonwealth* (*Stromlo*) *Obs.*, Vol. **3**, No. 13.
———. 1956*b*, *Occas. Notes, R.A.S.*, **3**, 118 (No. 18).
———. 1958, *Ap. J.*, **128**, 465.
———. 1959*a*, *Hdb. d. Phys.*, **53**, 275.
———. 1959*b*, *Ap. J.*, **130**, 728.
———. 1959*c*, *A.J.*, **64**, 397.
———. 1962, *Ap. J.*, **136**, 107.
———. 1963*a*, *Ap. J. Suppl.*, **8**, 31.
———. 1963*b*, *Ap. J.*, **138**, 934.
———. 1964, *ibid.*, **139**, 899.
Vaucouleurs, G. de, and Vaucouleurs, A. de. 1964, *Reference Catalogue of Bright Galaxies* (Austin: University of Texas Press).