SURFACE PHOTOMETRY AND THE STRUCTURE OF ELLIPTICAL GALAXIES

John Kormendy
Dominion Astrophysical Observatory, Herzberg Institute of Astrophysics, Victoria, British Columbia V8X 4M6, Canada

S. Djorgovski
Division of Physics, Mathematics, and Astronomy, California Institute of Technology, Pasadena, California 91125

1. INTRODUCTION

This paper reviews surface photometry of bulges and elliptical galaxies. Work prior to 1982 is discussed by Kormendy (1982a; hereafter K82). Since then, the subject has gone through a revolution. CCD detectors have come into common use, providing photometry accurate enough to measure new classes of subtle properties of ellipticals. Together with improvements in seeing, CCDs have allowed the resolution and study of galaxy cores and nuclei (Section 2). Newly discovered structural details, such as dust, shells, and dynamical subsystems, show the importance of accretion events in galaxy evolution (Sections 3–6). Better measurements of parameter scaling laws have led to improved constraints on galaxy formation (Section 8). Finally, CCDs provide accurate measurements of departures from elliptical isophotes (Section 9) and color gradients (Section 10). These

1 Visiting Astronomer, Canada–France–Hawaii Telescope, operated by the National Research Council of Canada, the Centre National de la Recherche Scientifique of France, and the University of Hawaii.
2 Alfred P. Sloan Fellow.
observations are currently producing a quantum jump in our understanding of elliptical galaxies.


Unless otherwise noted, we assume that the Hubble constant is \( H_0 = 50 \text{ km s}^{-1} \text{ Mpc}^{-1} \).

2. CORES AND NUCLEI

Two kinds of structure are commonly seen at the centers of early-type galaxies. When observed with sufficient resolution, the steep brightness profile of an elliptical usually flattens into a nearly constant surface brightness core. In addition, a nucleus is sometimes seen inside the core. By this, we mean a dynamically distinct cluster of stars that is much smaller and denser than the core. Cores are reviewed by Kormendy (K82, 1984, K87, 1987c) and Lauer (1988b); here we summarize their properties in Section 2.1. Nuclei are less well studied and understood; we review them in Section 2.2.

2.1 Summary of Core Properties

Reliable core photometry became available only when problems of seeing and photographic photometry were resolved. All photographic core photometry proves to be unreliable. CCDS do better: Profiles derived by different authors routinely agree to \( \lesssim 0.1 \text{ mag arcsec}^{-2} \) (Lauer 1985a, K87). Seeing affected early work on cores so strongly that it was not clear whether most galaxies have cores at all (Schweizer 1979, 1981b). Seeing corrections could be derived if one assumed that galaxies have cores, but the derived core parameters were model dependent (K82, Kormendy 1984). Then CCD observations by Lauer (1985a,b) found nearly isothermal cores in about a dozen ellipticals, increasing the sample of resolved cores by a factor of four. More recently, the Canada–France–Hawaii Telescope (CFHT) has provided seeing a factor of about two better than available previously; nearly isothermal cores are now resolved in almost all bright, nearby ellipticals (Kormendy 1985a, 1987c, K87).

CCD data are accurate enough for seeing corrections based on deconvolutions. Simple techniques are used by Djorgovski (1983) and Lauer (1985a,b), and more powerful techniques are discussed by Bendinelli et al.
(1982, 1984a, 1985, 1986, 1988, and references therein). The latter techniques have a mixed record, successfully revealing a stellar nucleus in M81 but finding a similar nucleus in M32 that is not confirmed by higher resolution observations (Section 2.2). Their weakness is that they magnify nonrandom errors, such as kinks in the observed profile. Lauer’s simpler technique mines less resolution from the data but finds fewer spurious features. Better still is resolution good enough to require little correction. By the time this article appears, the Space Telescope may provide a long-awaited factor-of-five additional improvement in resolution.

Core profile shapes have been examined systematically by Lauer (1985b). His CCD data show that virtually all cores have slightly non-isothermal brightness profiles. Kormendy’s (1985a, K87) high-resolution CCD photometry shows further that core profile shape correlates weakly with galaxy luminosity. A few galaxies, including some first-ranked galaxies in clusters, have isothermal profiles. Fainter galaxies have profiles that do not flatten completely into a core. The faintest galaxy with a well-resolved core is M31; it is even less isothermal than the ellipticals (Kent 1983). Such profiles have been interpreted as evidence for central black holes (Young et al. 1978, 1979) or anisotropy (Kormendy 1985a).

Some cores also show kinematic evidence for anisotropy (K87). For example, the core of NGC 1600 falls far below the “oblate line” describing isotropic oblate rotators in the $V_{\text{max}}/\sigma - \epsilon$ diagram (Illingworth 1977). Like the overall shapes of ellipticals, the E3 shape of the core of NGC 1600 must be maintained by anisotropy. Dynamical modeling is required to explore the orbital distribution functions implied by these observations.

Considerable effort has also gone into the measurement of characteristic parameters of cores, i.e. central surface brightnesses $\mu_0$, core radii $r_c$ at which the surface brightness has fallen by a factor of two, and central velocity dispersions $\sigma$. Structural scaling laws revealed by these data are discussed in Section 8.

2.2 Nuclei

Dense nuclear star clusters superposed on much larger cores have been recognized in at least six bright galaxies. The best example is in M31 (Light et al. 1974); this has $r_c \approx 0.4$ and $\mu_{0V} \approx 12.4$ mag arcsec$^{-2}$, compared with $r_c = 17''$ and $\mu_{0V} = 15.7$ mag arcsec$^{-2}$ for the bulge. Tremaine & Ostriker (1982) have shown that the nucleus and bulge of M31 are dynamically independent. We refer to these central star clusters as nuclei and distinguish them from bulges with cuspy brightness profiles (e.g. a pure power law like that in M32; Tonry 1984b). We also distinguish nuclei from nonthermal point sources such as those in Seyfert galaxies, quasars, and M87.
Besides M31, nuclei are found in M81 (Kormendy 1985a, Bendinelli et al. 1986) and in other nearby bulges (Kormendy 1985a). Their detection is limited by poor resolution (Lauer 1988b). Nuclei are also seen in many dwarf spheroidal galaxies (Reaves 1977, 1983, Romanishin et al. 1977, Caldwell 1983, 1987, Binggeli et al. 1984, 1985, Ichikawa et al. 1986, van den Bergh 1986, Caldwell & Bothun 1987) and in many disk galaxies that are late enough in type so that they do not contain bulges (e.g. M33; Gallagher et al. 1982, and references therein).

Nuclei are rarely seen in ellipticals. For example, in M87, the nuclear spectrum shows no stellar absorption lines when the spectrum of the underlying core is subtracted (Dressler 1988, Kormendy 1989). A small central excess of brightness above an isothermal core in NGC 3379 (de Vaucouleurs & Capaccioli 1979, K82, Nieto & Vidal 1984) is not due to a stellar nucleus either, but only to a nonisothermal core exactly like those in other ellipticals (Bedinelli et al. 1984b, Kormendy 1985a). Nuclei suspected to exist in M32 and NGC 4649 (Bedinelli et al. 1982, Lauer 1988b) are not confirmed at better resolution (J. Kormendy, in preparation). Nuclei should be relatively easy to see in bright ellipticals because they have large, well-resolved cores. Their apparent scarcity could be due to the existence of a maximum luminosity for nuclei. Also, few ellipticals are as close to us as the bulges that are known to contain nuclei.

Since nuclei are poorly resolved, little further is known about them. Stellar kinematic data are available in M31 (K87, Dressler & Richstone 1988, Kormendy 1988b), NGC 3115 (J. Kormendy & D. O. Richstone, in preparation), and NGC 4594 (Kormendy 1988c). Rapid rotation and velocity dispersions of $\sim 100$ km s$^{-1}$ (after bulge subtraction) indicate that all three nuclei are disks. (Dressler & Richstone did not come to this conclusion, but they did not subtract the bulge spectrum; then detection of the cold component is difficult.) In NGC 3115 and NGC 4594, the disk structure is also seen in the isophotes.

The available data suggest that disklike nuclei are built out of gas that has fallen into the center (van den Bergh 1976, K82, Kormendy 1982b, 1988b,c, Gallagher et al. 1982, Kormendy & Illingworth 1983). This idea is a natural consequence of the hypothesis that black holes are fueled by infalling gas. If gas can reach the black hole, it may form stars along the way when the density gets high enough in the gravitational funnel. This may even be a necessary step in the formation of nuclear black holes, since collapse times of cores in giant ellipticals are long, whereas nuclei can evolve more rapidly (Kormendy 1988a,b,c). Further discussion is given in Shlosman & Begelman (1987) and in Duschl (1988a,b).

Nuclei may originate in other ways, too. Globular clusters sink toward the center by dynamical friction and may form nuclei (Tremaine et al.
1975). A large galaxy can accrete a small one with a compact core (Section 3). And black holes may produce central density cusps. Accreted nuclei should be distinguishable from black hole cusps: In general they should have smaller \( \sigma \) and a different rotation axis than the rest of the galaxy. However, accreted nuclei and ones grown by gas infall and star formation may be difficult to distinguish.

There are indications that nuclei in dwarf spheroidal and disk galaxies are similar to those seen in bulges. The nucleus of M33 is interpreted by Gallagher et al. (1982) as a composite-age stellar population, consistent with late infall of gas and subsequent star formation. Spectra of nuclei in dwarf spheroidal galaxies suggest that they are \( \gtrsim 5 \) Gyr old but sometimes contain a contribution from younger (A–F) stars; this is also consistent with secondary formation (Bothun et al. 1985, Caldwell & Bothun 1987, Bothun & Mould 1988). It is also possible that some “nuclei” in dwarf galaxies are really very low-luminosity bulges, since small bulges have small \( r_c \) and high \( \mu_0 \) (see K87 for a review).

3. DYNAMICAL SUBSYSTEMS IN GALAXY CORES: EVIDENCE FOR MERGERS

A major development in recent years has been the realization that galaxies accrete significant amounts of material in the form of gas and small companions. The next three sections discuss some of the evidence. We begin with observations of distinct dynamical subsystems in galaxy cores.

The first clear example was NGC 5813. Efstathiou et al. (1982) found a core-within-a-core brightness profile in this otherwise normal elliptical (i.e. its core contains a second, smaller core of higher surface brightness). The inner core rotates more rapidly than the outer, and, except for the central measurement, has a smaller velocity dispersion. Kormendy (1984) suggested that these observations are the signature of a merger between a low- and a high-luminosity elliptical. Low-luminosity ellipticals have smaller core radii and higher central surface brightnesses than giant ellipticals. Kormendy showed that the robust core of a small elliptical can survive a merger with a giant elliptical and form a distinct subsystem at the center. He predicted that the rotation axis of the subsystem should be oriented randomly with respect to the main galaxy. In practice, observed orientations may be somewhat nonrandom because merger cross sections depend on encounter geometry. Nevertheless, this provides a test of the merger hypothesis. Also, the nucleus should in general rotate more rapidly than the rest of the galaxy because low-luminosity ellipticals are rapid rotators (Davies et al. 1983). Finally, the Faber-Jackson (1976) relation predicts that the velocity dispersion should in some cases decrease toward
the center. These effects are also seen in $N$-body simulations (e.g. Balcells & Quinn 1988).

A number of galaxies that dramatically show this behavior have now been found. Franx & Illingworth (1988), Jedrzejewski & Schechter (1988a), and Bender (1988b) have found seven elliptical galaxies whose cores are kinematically distinct from the rest of the galaxy. In four cases, the inner and outer parts rotate in opposite directions. This is strong evidence for accretion. Further support is provided by the observation of isophote twists between the two subsystems (Efstathiou et al. 1982, Bender 1988b).

In the new cases, no core-within-a-core structure is seen. This is not surprising. Efstathiou et al. (1982) warned us that their NGC 5813 photometry is not of high quality. It should be checked, especially since dust can counterfeit a core-within-a-core structure. Also, distinct cores are only expected in extreme cases, e.g. when a galaxy like M87 eats one like M32 (Figure 3 in Kormendy 1984). Such events should be rare, because faint ellipticals are rare (Binggeli et al. 1985, Sandage et al. 1985b). Mergers between nearly equal galaxies are not likely to leave a signature in the brightness profile.

The merger interpretation is attractive, but alternatives are possible. For example, if the figure rotation velocity in an elliptical is backward with respect to the streaming velocity of the stars, the sum (which is what we observe) can change sign. However, counterstreaming is difficult to achieve (Vietri 1986, 1988) and does not explain subsystems that rotate at right angles to their galaxies (NGC 4406; Bender 1988b, Franx 1988). This interpretation seems improbable. Another possibility is that the inner and outer parts of a galaxy acquire different angular momenta through tidal torques (Binney 1987, Barnes & Efstathiou 1987). This cannot be excluded, although it is least likely near the center. Like the above authors, we conclude that the observed dynamical subsystems result from mergers. These could be mergers of bulges or ellipticals, or ones involving gas infall and star formation (Section 4). IC 1459 and NGC 5322 may be examples of the latter: Their nuclear subsystems appear to be counterrotating stellar disks (Franx & Illingworth 1988, Bender 1988b).

Only a fraction of all merger remnants can be recognized from observations like the above. The fact that about one third of the ellipticals examined so far show nuclear subsystems (Bender 1988b, Jedrzejewski & Schechter 1988a,b) suggests that mergers affect a significant fraction of galaxies.

4. DUST IN ELLIPTICAL GALAXIES

According to classical definition (Hubble 1926, de Vaucouleurs 1959, Sandage 1961), elliptical galaxies contain no dust. Galaxies with dust have
usually been given S0 or later-type classifications. Now sensitive searches are finding that even the remaining, classical ellipticals often contain dust. This section summarizes its properties. Other recent reviews have been given by Schweizer (1987), Bertola (1987), and Nieto (1988; hereafter N88). A catalog of dusty ellipticals has been published by Ebneter & Balick (1985).

Progress in this subject has depended critically on the ability to detect subtle absorption features superposed on steep brightness gradients. CCD surveys are especially powerful: Their dynamic range is large, and the data can easily be subjected to digital “unsharp masking” (Sandage & Miller 1964, Malin 1977). This is done by dividing the image by a model of the overall brightness distribution without the fine structure. The model can be a smoothed version of the original image, or a synthetic galaxy image with the best-fitting elliptical isophotes, or an image taken in a redder bandpass. (In the last case, the ratio is a color image.) These techniques show that $\gtrsim 50\%$ of bulges and elliptical galaxies contain dust.

4.1 Frequency of Occurrence of Dust

A few dusty ellipticals have been known for years. They received little systematic attention until Bertola & Galletta (1978) pointed out that several ellipticals have dust lanes along their minor axes and therefore may be prolate. This had immediate impact because of the recent discovery (Bertola & Capaccioli 1975, Illingworth 1977) that most bright ellipticals are dynamically supported not by rotation but by velocity dispersion anisotropy, which suggests that they are triaxial (Binney 1976, 1978a,b, 1982a,b).

Systematic surveys for dust followed, and detection rates increased as search techniques improved. Hawarden et al. (1981) examined carefully chosen diskless galaxies on the ESO/SRC IIIa-J and Palomar sky surveys and found a substantial number (40) with dust. Sadler & Gerhard (1985a,b) found dust in $23 \pm 7\%$ of ellipticals with mean diameters of at least 2′ on the ESO B survey. Like all such estimates, this is a lower limit. The dust is usually in well-defined, nearly edge-on disks; this implies that many face-on dust distributions are going undetected. Sadler & Gerhard estimated that the true fraction of ellipticals with dust is at least 40%. A CCD survey by Sparks et al. (1985) led to similar conclusions. More recently, Djorgovski & Ebneter (1986) and Ebneter et al. (1988) have detected dust in 36% of the 116 ellipticals they studied. Finally, CCD photometry with the CFHT (Kormendy & Stauffer 1987; J. Kormendy, to be published) shows a still higher detection frequency, because of the excellent seeing on Mauna Kea. Dust distributions are often so small that they are barely detected even with the CFHT. Many more may await discovery with the Space Telescope.
This dust was also found by the Infrared Astronomical Satellite (IRAS). Detection frequencies in co-added IRAS survey data on bright, nearby ellipticals are comparable to or larger than those seen optically (Jura et al. 1987). Optical and IRAS photometry both imply that typical dust masses are $\sim 10^3$–$10^6 \ M_\odot$; for canonical gas-to-dust ratios, this corresponds to $\sim 10^7$–$10^8 \ M_\odot$ of cold gas (e.g. Sadler & Gerhard 1985b, Sparks et al. 1985, Jura 1986, Jura et al. 1987, Véron-Cetty & Véron 1988).

It is now clear that dust in elliptical galaxies is not rare. This is one more piece of evidence that ellipticals contain substantial amounts of interstellar matter [see Schweizer (1987) for a review]. Some gas is acquired by accretion (see the next section), and some is expected from mass loss during stellar evolution (Sandage 1957, Faber & Gallagher 1976). With the discovery that ellipticals generally contain $10^9$–$10^{10} \ M_\odot$ of X-ray–emitting gas (e.g. Forman et al. 1985), the idea that they are surprisingly free of interstellar matter has disappeared.

Although the precise frequency is uncertain because of classification bias, the above surveys show that bulges contain dust still more often than ellipticals. Even prototypical bulges can be riddled with dust (e.g. M31; Johnson & Hanna 1972, Kent 1983, McElroy 1983), as well as ionized (Ciardullo et al. 1988) and other gas.

### 4.2 Origin of Dust: Further Evidence for Galaxy Mergers

There is strong evidence that many large-scale dust and H I gas distributions are accreted. The most convincing evidence is kinematic: The gas and dust are usually in disks rotating at random orientations with respect to the optical major axis [H I (e.g. Gallagher et al. 1977); H II (Schweizer 1980, 1981a, 1982, Davies & Illingworth 1986, Caldwell et al. 1986); H II associated with dust (Burbidge & Burbidge 1959, Graham 1979, Marcelin et al. 1982, Möllenhoff 1982, Sharples et al. 1983, Caldwell 1984, Bertola et al. 1985, Möllenhoff & Marenbach 1986, Wilkinson et al. 1986, Bland et al. 1987, Varnas et al. 1987, Galletta 1987, Bertola & Bettoni 1988, Bertola et al. 1988a,b, Möllenhoff & Bender 1988)]. Minor-axis dust lanes rotate at right angles to the stars. Sometimes dust lanes and stars even rotate in opposite directions. This gas cannot come from internal mass loss. Accretion is also suggested by the morphology (although dust is not correlated with the presence of ripples and shells; Schweizer & Ford 1985). At large radii, dust lanes often show S-shaped warps or transitions from regular disks to irregular distributions. Such behavior is expected for material just settling into equilibrium, since orbital clocks run slower at larger radii. Note that accretion does not require cannibalism; gas can be donated by a galaxy that gets away (Schweizer 1987).
Small dust lanes are more common than large-scale dust distributions. Whether or not they have the same origin is not clear. They are usually well-defined rings or disks near the center, often oriented parallel to the major axis. Many resemble the inner dust lanes commonly seen in S0 and spiral galaxies (Sandage 1961). Some or even most may have an internal origin. However, a folklore is developing, perhaps prematurely, that dust in ellipticals is always accreted. Kinematic constraints are badly needed on the fraction of inner dust disks that are accreted. Is the fraction of counterrotating cases near 50% (as it is for large-scale dust lanes; Bertola et al. 1988a,b), or is it much smaller? At stake is a better understanding of how much secular evolution results from mergers and how much from internal processes.

4.3 Three-Dimensional Shapes of Ellipticals Containing Dust

Bertola & Galletta’s (1978) pioneering paper raised the hope that dust-lane geometry could be used to measure galaxy shapes. However, the large number of free parameters make this complicated. The results provide further evidence that ellipticals are triaxial, and they sometimes favor an oblate or a prolate configuration, but they have not securely told us the shape of any individual galaxy.

This subject is reviewed in detail by Merritt & de Zeeuw (1983) and will also be reviewed in the next volume of this series by de Zeeuw (1990). Thus our summary of the predictions is brief. Gas in a spheroidal or triaxial potential settles into certain preferred planes through differential precession and dissipation (Kahn & Woltjer 1959, Gunn 1979, Lake & Norman 1983). Consider first the simplest case, in which the shape of the potential does not rotate. Then the gas settles into one of two planes, i.e. perpendicular to the shortest or to the longest axis (e.g. Heiligman & Schwarzschild 1979, Tohline et al. 1982, Steiman-Cameron & Durisen 1982). In a spheroidal galaxy, only the equatorial orbits are stable; polar orbits gradually tip over into the equatorial plane. If we knew that ellipticals are spheroids, then those with minor-axis dust lanes would be prolate and those with major-axis dust lanes would be oblate. But ellipticals can be triaxial. Then, for some infall angles and galaxy shapes, gas is captured into polar orbits. Therefore, a dust lane along a particular axis is consistent with either oblate or prolate structure. Already there is no unique relationship between galaxy shape and dust-lane geometry.

The next complication is that the figure can tumble (angular velocity $\Omega_p \neq 0$). However, the angular velocity we measure is that of the stars, and they stream through the figure (as they do through bars and spiral
arms). Therefore, we do not know $\Omega_p$; in fact, we are as interested in estimating $\Omega_p$ as in finding the shape of the galaxy. Tumbling elliptical galaxies allow additional equilibrium orbits, as summarized in Figure 1.

Half of the configurations shown in Figure 1 may be uncommon. If a galaxy tumbles about its long axis, stellar rotation velocities will be large along the minor axis and zero along the major axis. This has been observed in only a few galaxies (e.g. NGC 4261; Davies & Birkinshaw 1986, Wagner et al. 1988). We assume that ellipticals usually tumble about their short axes. Then stable major-axis dust lanes should be prograde. Minor-axis dust lanes should be perpendicular at small radii and should twist at large radii and show retrograde rotation.

What do we observe? Major-axis dust lanes counterrotate in two of the four ellipticals studied (Bertola et al. 1988a); retrograde gas velocities are also seen in the SB0 galaxy NGC 4546 (Galletta 1987). Of seven minor-axis dust lanes measured so far, three show retrograde-rotating twists [NGC 1316 (Schweizer 1980), NGC 5363 (Sharples et al. 1983, Bertola et al. 1985), and A0609 – 33 (Möllenhoff & Marenbach 1986)] and four show prograde twists [NGC 4589 (Möllenhoff & Bender 1988), NGC 5128 (e.g.

<table>
<thead>
<tr>
<th>FIGURE ROTATION AXI</th>
<th>TYPE OF ORBIT</th>
<th>DUST LANE APPEARANCE</th>
<th>DUST LANE KINEMATIC SIGNATURE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Short</td>
<td>Equatorial</td>
<td></td>
<td>Prograde</td>
</tr>
<tr>
<td></td>
<td>Anomalous</td>
<td></td>
<td>Perpendicular, then retrograde</td>
</tr>
<tr>
<td>Long</td>
<td>Equatorial</td>
<td></td>
<td>Retrograde</td>
</tr>
<tr>
<td></td>
<td>Anomalous</td>
<td></td>
<td>Perpendicular, then prograde</td>
</tr>
</tbody>
</table>

Figure 1  Stable orbits of gas in a rotating triaxial galaxy (adapted from Merritt & de Zeeuw 1983). As illustrated, the figure tumbles in the direction of stellar rotation ($\Omega_p > 0$); if $\Omega_p < 0$, the sense of gas rotation is reversed. Assume that the figure rotates about its shortest or longest axis (left). The second column gives the kind of orbit, and the third sketches resulting dust lanes seen edge-on. Anomalous orbits have different orientations at different radii (van Albada et al. 1982). They are the analogues of polar orbits in a stationary potential; at small radii, where $\Omega_p$ is unimportant, they are polar. At large radii, the figure rotates several times during an orbit and so is effectively oblate-spheroidal; then the orbit is equatorial (Simonson 1982). In between, the orbits have skew orientations determined by the Coriolis force. The schematic illustrations of dust lanes show the directions of stellar and gas motion; $\circ$ indicates approach, and $+$ indicates recession. The right column states the kinematic signature, i.e. the sense of rotation of the dust lane with respect to the stars.

The hypothesis that these dust lanes are in equilibrium can be saved if \( \Omega_p < 0 \) (e.g. Varnas et al. 1987). Although it is difficult (Vietri 1986), Vietri (1988) has succeeded in constructing at least one realistic dynamical model in which retrograde figure rotation is slow enough, and prograde stellar streaming large enough, so that the sum (i.e. the observed galaxy rotation velocity) is opposite to the tumbling direction at some radii (see also Freeman 1966). On the other hand, \( N \)-body models that collapse and become bar-unstable have always resulted in \( \Omega_p > 0 \) [see van Albada (1987) for a review]. It is not clear whether retrograde tumbling is a viable interpretation.

Therefore the observations suggest that many dust-lane warps are transient—that gas has settled to a preferred plane at small radii but still remembers the merger geometry in the warp (Tubbs 1980, Simonson 1982, Bertola et al. 1985, Wilkinson et al. 1986, Schweizer 1987, Schwarzschild 1987, Möllenhoff & Bender 1988). This possibility has existed from the beginning; it was resisted mainly because warps then tell us less about galaxy shapes. But the fact that dust lanes are often regular at small radii and irregular farther out (e.g. NGC 1316; Schweizer 1980) should already have convinced us that settling into principal planes is not always complete.

We dwell on this subject because it has seemed to be the most rigorous new method to measure the shapes of individual ellipticals. It remains promising. But even with photometry and kinematic data, it is difficult to unravel the many unknowns: the amount of triaxiality, the orientation of the galaxy, the pattern rotation speed, and the question of whether dust has settled into equilibrium. There are other uncertainties that we have not discussed. For example, a slowly rotating elliptical may not be exclusively oblate or prolate; it may at some radius change from one to the other. And some conclusions summarized in Figure 1 may be violated in special potentials. Simple deductions seem reasonably secure: (a) Ellipticals are generally triaxial; (b) some are prolate and others are oblate; and (c) some warps imply that \( \Omega_p \neq 0 \) and others imply nonstationary structure. But more detailed progress has been elusive.

We know of no simple remedies. As Merritt & de Zeeuw (1983) point out, better statistics would help. We may be basing far-reaching conclusions on configurations that turn out to be rare. New discoveries of systematic behavior may reduce the available parameter space. But it appears that the implications of dust-lane geometry and kinematics become statistical. Unless further work sheds new light, we still do not know how to measure the shapes of individual ellipticals.
4.4 *Dust and the Distinction Between E and S0 Galaxies*

The presence of dust contributes to a blurring of the distinction between E and S0 galaxies. This is partly just a practical problem of classification. When ellipticals were dust free by definition, dusty galaxies were easy to classify. If we now adopt as the main classification criterion the presence (S0) or absence (E) of a disk, then it is difficult to distinguish ellipticals from S0s with faint disks [see K82 and Capaccioli (1987) for reviews]. A significant number of galaxies must be misclassified in the literature. If the E–S0 sequence is continuous, this makes little difference for an individual object (Schweizer 1987). But it can systematically affect galaxy samples selected for physical studies.

There is also a more difficult problem of principle. Since dust lanes and gas may form stars, a galaxy can change our perception of its morphological type. For example, a slowly rotating, bright elliptical may, through judicious cannibalism, grow a disk and come to look like an S0. This would contribute noise to correlations between physical properties and type. For example, even if real bulges rotate rapidly, there would be apparent exceptions because some S0s started life as ellipticals. This is an example of how secular evolution can obscure a physical correlation that was set up during an earlier phase of galaxy formation. Since far-reaching conclusions are often based on a few galaxies with surprising behavior, we need to be careful to understand and allow for secular evolution.

5. SHELLS AND RIPPLES


Schweizer (1980, 1982, 1983) was the first to suggest that shells result from the accretion of small galaxies. Numerical simulations by Quinn (1982, 1984), Toomre (1983), Dupraz & Combes (1985, 1986), Quinn & Hernquist (1987), Hernquist & Quinn (1987, 1988), and others make a convincing case that they form through phase and spatial wrapping of cold material sloshing back and forth in the gravitational potential of an elliptical. Supporting this interpretation are the observations that (a) shells are made of stars similar in color to or slightly bluer than the underlying elliptical (Schweizer 1980, Carter et al. 1982, 1988, Bosma et al. 1985, Fort et al. 1986, Pence 1986, Clark et al. 1987, Schombert & Wallin 1987); (b)
shells at successive increasing radii alternate on opposite sides of the center [e.g. NGC 3923; Quinn 1982, 1984, Malin & Carter 1983, Fort et al. 1986, Pence 1986 (but see Prieur 1988)]; and (c) their outer edges are sharp and often edge brightened, like folded sheets (e.g. Malin & Carter 1980, 1983). Such structures form when a small accreted galaxy falls almost radially into a smooth and stationary potential.

Shell detection frequencies show that accretion of small companions is a normal event in the life of a galaxy. Surveys by Malin & Carter (1983) and by Schweizer & Ford (1985) find shells in 17% and 44% of field ellipticals, respectively. Not all encounter geometries and viewing angles produce visible shells, so the real percentage is larger. This suggests that a typical elliptical has experienced one or more accretion events. Disk galaxies also contain shells, albeit less often than ellipticals (Schweizer & Ford 1985, Schweizer & Seitzer 1988). There is no reason to believe that disk galaxies do not accrete companions. However, if the victim is too massive, the disk is destroyed. If the disk is very robust, the resulting flattened potential is not gentle enough to form ordered shells (Quinn 1984, Dupraz & Combes 1986).

At first it was hoped that shells could be used to measure galaxy mass distributions, because the number of shells and their radial distribution depend on the gravitational potential (Quinn 1984, Dupraz & Combes 1986, Quinn & Hernquist 1987, Hernquist & Quinn 1987). The steep potential gradient of an $r^{1/4}$-law mass distribution predicts a large number (100–200) of shells. The reason is that stars in the innermost shells have much shorter orbital periods than those at large $r$; one new shell forms every time inner stars complete an extra half-oscillation with respect to outer ones. But galaxies typically contain $\lesssim$20 shells. This can be understood if dark matter is added at large radii to reduce orbital periods there. However, this argument is oversimplified. The predictions were based on the assumption that the accreted galaxy is disrupted instantaneously. If not, it sinks toward the center through dynamical friction. Then inner shell stars have spent less time at small $r$ than we thought, so fewer shells are predicted. Dark matter is no longer required. This also solves the problem that some inner shells are surprisingly close to the center. All of this is discussed by Dupraz & Combes (1987), Hernquist & Quinn (1988), and Prieur (1988), who now conclude that shells cannot be used to measure mass distributions. However, Piran & Villumsen (1987) find that shells are not formed at all if stars are stripped too slowly from the victim. It remains to be demonstrated that we know how to construct regular shell systems over the largest observed radius range.

Shells may tell us something about galaxy shapes. Dupraz & Combes (1985, 1986) suggest that (a) when shells are short arcs bisected by the
major axis, the elliptical is likely to be prolate and edge-on; (b) when shells align with the minor axis, the elliptical is oblate and edge-on; and (c) when the shells are randomly distributed in azimuth and the elliptical is nearly round, it is likely to be oblate and face-on (e.g. 0422−476; Wilkinson et al. 1987). As in Section 4.3, these results are statistical, except perhaps in the most regular cases. However, they suggest independently that ellipticals span a wide range of shapes from oblate triaxial to prolate triaxial.

6. SUMMARY: SECULAR EVOLUTION BY ACCRETION

The observations reviewed in Sections 3–5 suggest that galaxy structure may be altered significantly by the accretion of gas and small companions. In the past, galaxies showing obvious effects of mergers were regarded as peculiar. Now we know that accretion happens often enough in a typical galaxy that even modest events can add up to a significant effect. However, this is secular evolution and not like the more violent mergers that may completely destroy a disk and convert two spirals into one elliptical. The present results are independent of the debate about what fraction of ellipticals formed in this way. Accreted material continues to trickle in long after dissipational collapse or merger formation is complete.

The amount of material added and its effects on the structure of a typical galaxy are unknown. However, typical masses of dust and gas ($10^7$−$10^8 \, M_\odot$; Section 4.1) are not much less than the mass of stars in a core. Substantial changes in core properties could result; this may increase the scatter in core parameter relations (cf. Lauer 1988b). Also, enough material can be accreted to make disks that would change a galaxy’s apparent morphological type (Section 4.4). And many shell galaxies have blue colors and early-type spectra implying recent bursts of star formation (Carter et al. 1988).

These results form part of a gradually changing picture of the formation and evolution of elliptical galaxies. Traditionally, galaxies were thought to form on a gravitational collapse time scale, with little subsequent evolution. Recent results suggest a picture in which formation is more gradual. As in disk galaxies, where gas infall and disk building are still going on today (Gunn 1982), accretion and the rearranging of mass in ellipticals may be a significant evolutionary process that is far from over (Schweizer 1983, 1986, Schweizer & Seitzer 1988).

7. BRIGHTNESS PROFILES AND TIDAL EFFECTS

The shapes of galaxy brightness profiles and their dependence on luminosity contain information about galaxy formation. Systematic trends with
environment tell us about tidal effects. And the functional form of the profile determines the best way to derive size and density scale parameters (Section 8).

Studies of profile shapes are affected by a variety of problems, some of which are worse for CCDs than for photographic observations. (a) Some data are not accurate at large radii. CCDs have a reputation for omnipotence that does not apply to measurements of galaxy halos. CCDs are small, so sky estimates are often uncertain (Capaccioli 1987). Observers know this, but poor sky subtraction nevertheless plagues even the best CCD photometry of halos. Capaccioli et al. (1988) and Peletier et al. (1988a) cite examples; errors of 0.2–0.5 mag arcsec$^{-2}$ are common. (b) Most CCD data reach out to only a few de Vaucouleurs (1948) effective radii $r_e$; the systematic departures from $r^{1/4}$ laws that are discussed below begin at about these radii. (c) Seeing is a problem, especially for low-luminosity galaxies. These have such tiny cores that seeing can completely change their brightness profiles. For example, if M32 were in the Virgo cluster, we would know nothing about its inner power-law profile (Tonry 1984b, 1987). No one has studied enough nearby cases; for example, Schombert's (1986, 1987) low-luminosity galaxies are in the Coma cluster. (d) Tidal effects modify outer profiles upward or downward (Section 7.2). (e) Some "ellipticals" are misclassified S0s. (f) Finally, many ellipticals contain unrecognized dust (Section 4.1).

7.1 Do Elliptical Galaxies Have $r^{1/4}$-Law Brightness Profiles?

De Vaucouleurs' (1948, 1953) $r^{1/4}$ law fits bright elliptical galaxies reasonably well except where tidal effects are important. We do not attach physical significance to this choice of function, although Binney (1982c) and Bertin & Stiavelli (1984, 1989) find reasonable distribution functions whose density profiles are similar to it. The $r^{1/4}$ law is a convenient parameterization that extracts all of the scaling information that we are entitled to derive, given the similarity of profiles to power laws (Kormendy 1980, K82). But how well does it work?

No definitive study has been published. Based on large photometric surveys, Michard (1985), Djorgovski et al. (1985), Djorgovski (1985), Schombert (1986, 1987), Kodaira et al. (1986), Jedrzejewski (1987b), Capaccioli et al. (1988), and de Carvalho & da Costa (1988) conclude that ellipticals have a wide variety of profile shapes. A corollary is that fitting functions with two scale parameters but no shape parameter are not particularly useful. However, these conclusions are undermined by the problems discussed above.

Much of what we know about galaxy halos still comes from photo-
graphic data. These show that profiles of isolated ellipticals vary with luminosity (e.g. Kormendy 1980, Michard 1985, Schombert 1986, 1987), although with significant scatter. The $r^{1/4}$ law fits best near $M_B = -21$. Even at this luminosity, profiles are slightly concave upward when plotted against $r^{1/4}$; typical deviations are $\pm 0.1$--0.2 mag arcsec$^{-2}$ over $\gtrsim 6$ mag arcsec$^{-2}$ (Kormendy 1977b, Capaccioli 1985). Galaxies much brighter than $M_B = -21$ have more light at large radii than the extrapolation of $r^{1/4}$ laws fitted further in, and fainter galaxies have less.

We believe that a good approach for future investigation of profile shapes is one suggested by Schombert (1986, 1987). For each luminosity bin, Schombert constructs template profiles by averaging many observed profiles. Two further improvements are needed. First, total luminosities should be used. Schombert's 16-kpc metric absolute magnitudes measure different fractions of the total light in giant and dwarf galaxies: They are total magnitudes for dwarfs but contain only $\sim 50\%$ of the luminosity of first-ranked galaxies (see Figure 8 in Schombert 1986). Second, we need to use isolated galaxies to minimize tidal effects. The resulting templates can then be compared with profiles of galaxies that have companions to study tidal effects.

It remains true that characteristic sizes and densities are well measured using two-parameter fitting functions. These are basically equivalent. None has a special physical interpretation, but among formulas explored so far, the $r^{1/4}$ law is most convenient and fits best. Profile fits can be improved by adding a third parameter, but then the parameters are too coupled to be useful. All this has been reviewed by Kormendy (1980, K82) and Capaccioli (1988b). Parameters can also be derived without using fitting functions. For example, the actual half-light radius and surface brightness can be used. Or scale radii can be derived using dimensionless monotonic functions like Petrosian's (1976) $\eta$ function, i.e. the ratio of the surface brightness at a given radius to the mean surface brightness within that radius.

7.2 Tidal Effects

Tidal effects are reviewed in K82. In the cores of rich clusters, galaxies are observed to have abnormally small sizes (Strom & Strom 1978, and subsequent papers; see K82). This is particularly true of faint ellipticals, which moreover have outer cutoffs in their profiles (see also Schombert 1986, 1987). These observations are convincingly interpreted as truncation by the mean gravitational field of the cluster, especially during virialization (Merritt 1984, and references therein).

Whether small galaxies are truncated by large ones is less clear. Suggestions that M32 (King 1962) and NGC 4486B (Rood 1965, Kormendy 1977a) are truncated conflict with recent photometry [see Nieto & Prugniel
(1987a,b) and N88 for reviews]. On the other hand, King & Kiser (1973) find that NGC 5846A has an outer cutoff. Examples of both "truncated" and untruncated small companions are given by Prugniel et al. (1987, 1988). Some photometry is uncertain because the galaxies are embedded in the halos of companions (N88). Also, some close pairs must be optical doubles. Thus the implications of these observations are unclear.

Despite possible truncation, it is clear that ellipticals like M32 are not dwarfs only because of tidal effects. They are genuinely the low-$L$ end of the luminosity function of elliptical galaxies (Section 8.1; see also Nieto & Prugniel 1987a,b, N88).

Encounters between ellipticals of nearly the same mass cannot by symmetry produce truncation if the total mass lost to the system is small (Aguilar & White 1985, and references therein). Kormendy (1977b, 1982a) concluded that ellipticals with companions of comparable sizes have distended outer profiles, and he interpreted this as tidal stretching or heating. The effect needs checking: Schombert (1988) and de Carvalho & da Costa (1988) did not see it in their samples. However, distension is seen in $N$-body simulations. Aguilar & White (1986) find that encounters produce transient tidal waves in the density distribution. An encounter heats each galaxy. Strong encounters steepen the profiles (i.e. make the galaxies smaller) because mass is lost; weak encounters make the profiles shallower. In either case, the final profile is set up first at small radii. As time passes, the transition between the old and new profile moves outward until only the final profile is left. An $r^{1/4}$-law profile shape is approximately preserved; there is no truncation.


8. PARAMETER CORRELATIONS AND SCALING LAWS

One of the main astrophysical uses of surface photometry is for the study of parameter correlations and scaling laws. These contain valuable information about galaxy formation and evolution. Also, correlations between distance-dependent and distance-independent quantities are vital for the mapping of large-scale structure and velocity fields.

8.1 Families of Ellipsoidal Stellar Systems

A fundamental application of parameter correlations has been the demonstration that diffuse dwarf spheroidal (dSph) galaxies are a family of
objects unrelated to ellipticals. Baade (1944) long ago noted that NGC 147, NGC 185, NGC 205, and the Galactic dwarf spheroidals form a low-
surface-brightness sequence quite unlike ordinary ellipticals. Until the mid-
1980s, most people believed that the transition between these sequences is
continuous [e.g. Binggeli et al. (1984), but contrast Michard (1979) and
Farouki et al. (1983)]. Then, in important and somewhat neglected papers,
Saito (1979a,b) pointed out that dSphs have anomalously low binding
energies compared to giant ellipticals. He suggested the now-favored
explanation that dwarfs have low densities because supernova-driven
winds have removed large amounts of gas. Later, Wirth & Gallagher
(1984) were the first to emphasize that there are two unrelated
sequences of early-type galaxies: the diffuse dwarfs, and an E-galaxy sequence whose
low-luminosity end consists of galaxies like M32. They pointed out that
the extreme properties of M32-like dwarfs are not due to tidal truncation
but are intrinsic to low-luminosity ellipticals (see also Section 7.2). Also,
they found additional examples in the Fornax cluster, showing that M32
is not a fluke. The case was further strengthened by Kormendy (1985b,
1987c), who demonstrated a clear separation into two families overlapping
in luminosity. The key to this was CFHT seeing good enough to define
the low-\(L\) end of core parameter scaling laws for ordinary ellipticals. The
differences between the families are also seen in global properties\(^3\) (Saito
1979a, Okamura 1985, Dekel & Silk 1986, Ichikawa et al. 1986, 1988,
Kormendy 1987c). These results are not due to selection effects, as sug-
gested by Phillipps et al. (1988). The distribution of parameters for diffuse
dwarfs is undoubtedly biased by selection; remarkably low-surface-
brightness galaxies are still being discovered (Sandage & Binggeli 1984,
Impey et al. 1988). But luminosity “icebergs” hidden under the sky bright-
ness (Disney & Phillipps 1987, 1988) only contribute to the distinction
between E and dSph galaxies. Further evidence for this distinction includes
a large difference in luminosity functions: Ellipticals have a nearly Gauss-
ian luminosity function that peaks at \(M_B \simeq -18\), while dSph galaxies

\(^3\) Uncertainty about whether there is a discontinuity (Binggeli et al. 1984, Sandage et al.
1985a, Binggeli 1985, Caldwell & Bothun 1987) is based mainly on three problems. (a) Some
global parameters used (e.g. isophotal mean surface brightnesses) are insensitive structure
indicators. (b) Bright dSph galaxies, which are close to the E sequence, often contain both
an exponential component and (apparently) a bulge (e.g. NGC 5206; Caldwell & Bothun
1987). Inclusion of bulges guarantees convergence with the E sequence; disks and bulges
should be plotted separately in these diagrams. (c) Seeing effects are so large that it is
very difficult to define the faint end of the E sequence using ground-based observations of
galaxies as far away as 20 Mpc (Kormendy 1987c). Effects (a)–(c) appear sufficient to explain
the apparent merging of the E and dSph sequences seen by the above authors. Nevertheless,
it is important that the galaxies they cite as transition objects be measured with Space
Telescope for inclusion in the parameter diagrams.
begin to appear at $M_B \approx -18$ and in Virgo then become more numerous at least as rapidly as $L^{-1.35}$ [Wirth & Gallagher 1984, Sandage et al. 1985a,b, Impey et al. 1988; see Binggeli (1987) for a review]. The distinction between the elliptical and diffuse dwarf galaxy families points to a fundamental difference in formation history.

A clue to the origin of dSph galaxies is provided by the observation that they are structurally similar to dwarf spiral and irregular (dS + I) galaxies (Faber & Lin 1983, Lin & Faber 1983, Caldwell 1983, Wirth & Gallagher 1984, Sandage & Binggeli 1984, Binggeli et al. 1985, Kormendy 1985b, 1987c, Okamura 1985, Binggeli 1985, Ichikawa et al. 1986, 1988, Karachentseva et al. 1987, Impey et al. 1988). They are not merely dS + I galaxies seen between bursts of star formation, because they contain virtually no gas (Bothun et al. 1985, Impey et al. 1988). Two formation mechanisms are discussed at length in the literature. First, the basic low-density structure of dwarf galaxies is probably due at least in part to supernova-driven galactic winds; these can turn some dS + I galaxies into dwarf spheroidals (e.g. Larson 1974, Saito 1979b, Silk 1983, Dekel & Silk 1986, Vader 1986a, 1987, Silk et al. 1987, Yoshii & Arimoto 1987). In addition, there is strong evidence, at least in clusters, that some dwarf spheroidals formed from dS + I galaxies by ram-pressure stripping of their gas [see Lin & Faber (1983), Binggeli (1985), and Kormendy (1987c) for reviews]. There are other possibilities too. Dwarf spheroidals could be dS + I galaxies that turned all of their gas into stars (Kormendy 1985b, Davies & Phillipps 1988, Binggeli et al. 1989). In certain circumstances, dSph galaxies could even turn back into dS + I galaxies by accreting gas (Silk et al. 1987). In the search for simple, unique explanations, we should not forget that all of these things may happen.

8.2 Correlations With Galaxy Luminosity

The results of the previous section were based on correlations of various physical scale parameters with total luminosity $L$. The best known of these is the Faber-Jackson (1976) relation $L \propto \sigma^n$. The slope is $n \approx 4$, but with a real variation, depending on the sample definition (Faber & Jackson 1976, Tonry & Davis 1981, Tonry 1981, Terlevich et al. 1981, de Vaucouleurs & Olson 1982, Kormendy & Illingworth 1983, Dressler 1984a).

Many of these authors combined their data with estimates of effective radii and found a weak correlation between mass-to-light ratio and luminosity, $M/L \propto L^{0.35 \pm 0.15}$. This is also seen in core mass-to-light ratios (K87, Kormendy 1987c).

A correlation between the de Vaucouleurs (1948) effective radius $r_e$ and surface brightness $I_e$ was found by Kormendy (1977b, 1980, K82); modern data give $r_e \propto I_e^{-0.83 \pm 0.08}$ (e.g. Hoessel & Schneider 1985, Hamabe &
Kormendy 1987, Djorgovski & Davis 1987). This implies that more luminous galaxies have larger \( r_e \) and fainter \( I_e \), although with large scatter. Hamabe & Kormendy show that these relations are not significantly affected by the coupling of measurement errors in the parameters. They are also largely independent of how the parameters are defined; e.g. similar relations hold for core parameters (K82, Kormendy 1984, 1985b, 1987c, K87, Lauer 1985b, 1988b).

An important correlation between luminosity and the dynamical importance of rotation was discovered by Davies et al. (1983). They found that low-luminosity ellipticals and bulges rotate rapidly, have nearly isotropic velocity dispersions, and are flattened by rotation. In contrast, bright ellipticals rotate slowly, are pressure supported, and owe their shapes to velocity anisotropy. Let \( V/\sigma \) be the ratio of the maximum rotation velocity to a suitable mean velocity dispersion (see Davies et al. 1983). Then the level of rotational support can be parametrized by the ratio \((V/\sigma)^*\) of \( V/\sigma \) to the value expected for an isotropic oblate spheroid. The result that \((V/\sigma)^* \approx 1\) for faint galaxies and \(<1\) for bright ones could arise if proto-ellipticals acquired angular momenta through tidal torques, and if mergers then produced brighter ellipticals in which rotation got scrambled.

8.3 Multiparameter Correlations: the "Fundamental Plane" of Elliptical Galaxies

Multiparameter correlations were discovered through studies of correlated residuals from relations like those of Section 8.2. A breakthrough in our understanding of scaling laws required the appearance of large, homogeneous data sets based mainly on CCD photometry and long-slit spectroscopy. Also, this work has benefited from the application of statistical tools like principal component analysis (PCA) (e.g. Brosche 1973, Bujarrabal et al. 1981, Brosche & Lentes 1983, Lentes 1983, Efstatthiou & Fall 1984, Whitmore 1984, Murtagh & Heck 1987). However, in PCA, the astrophysics can get lost in too many eigenvectors. Therefore, simple techniques like bilinear least-squares fits remain useful to provide physical insight.

The presence of intrinsic scatter in the Faber-Jackson relation was correctly interpreted as an indication of a "second parameter." In an important paper, Terlevich et al. (1981) proposed that this second parameter is metallicity, measured by the \( \text{Mg}_2 \) index, and possibly axial ratio. Their results were challenged by Tonry & Davis (1981) and then readdressed by Efstatthiou & Fall (1984). However, relatively poor data sets available at the time did not permit a resolution of the problem. Authors agreed that elliptical galaxies are at least a two-parameter family, but the second parameter could not clearly be identified.

More accurate data confirm that the variance of global properties is exhausted almost entirely by two variables (Tonry & Davis 1981, Lauer
1985b, 1987, Burstein et al. 1986, Djorgovski & Davis 1986, 1987, Dressler et al. 1987, Faber et al. 1987, Djorgovski 1987a, Dressler 1987, de Carvalho & da Costa 1989). These data show that bulges and ellipticals lie in an inclined “fundamental plane” in the space of observed parameters (Figure 2),

\[ R \propto \sigma^{1.4 \pm 0.15} f^{-0.9 \pm 0.1}. \]

Here \( R \) can be any consistently defined radius derived from surface brightness profiles, such as the core or effective radius, but not an isophotal radius. An equivalent relation is obtained for luminosity. The old Faber-Jackson and radius-surface brightness relations are projections of the fundamental plane. The luminosity-color and mass-metallicity relations are also contained in the fundamental plane. Its tilt with respect to the

![Figure 2](image-url)

*Figure 2* Projections of the fundamental parameter plane of elliptical galaxies. *Top panels:* the one-parameter scaling relations discussed in Section 8.2, i.e. (left) the relation between radius and mean surface brightness, and (right) that between luminosity and velocity dispersion (the Faber-Jackson relation). *Bottom left:* the surface brightness–velocity dispersion correlation is the fundamental plane seen almost face-on. This is an observer’s version of the cooling diagram from theories of galaxy formation. *Bottom right:* this relation between the radius and a combination of surface brightness and velocity dispersion is the fundamental plane seen edge-on. The data are from Djorgovski & Davis (1987). All photometric quantities are in the Lick \( r_c \) band and are measured at or within the \( r_c \) elliptical isophote. The crosses are median error bars for all points in each panel.

© Annual Reviews Inc. • Provided by the NASA Astrophysics Data System
planes of observed parameters produces the correlated intrinsic scatter seen in the projected relations.

An alternate form of Equation 1 is the relation between the modified isophotal diameter and velocity dispersion, \( D_n \propto \sigma^{4/3} \) (Burstein et al. 1986, Dressler et al. 1987). Here \( D_n \) is defined as the circular diameter within which the mean surface brightness reaches a certain fiducial value, e.g. \( \mu_B = 20.75 \) mag arcsec\(^{-2} \) in the case of Dressler et al. (1987). The above authors show that the \( D_n-\sigma \) relation is equivalent to Equation 1, provided that all elliptical galaxies have brightness profiles of the same shape.

Another alternative representation of the fundamental plane—a relation between radius, surface brightness, and a metallicity indicator (color or \( M_{g_2} \) index)—has been obtained by de Carvalho & Djorgovski (1989). The two-dimensional nature of the manifold of elliptical galaxies implies that there must be a second parameter in the relation between mass and metallicity; this is identified as the luminosity density (S. Djorgovski & R. R. de Carvalho, in preparation).

The residual scatter about the fundamental plane is \( \sim 20\% \) per galaxy (given as the relative error of distance or radius). It is mostly or entirely due to measurement errors. Any cosmic scatter cannot be larger than a few percent.

A group at the Tokyo Astronomical Observatory has obtained and analyzed photographic surface photometry of galaxies of all Hubble types (Kodaira et al. 1983, Okamura et al. 1984, Watanabe et al. 1985). Based on PCA, they also conclude that there are two dominant dimensions in the parameter space of luminosity, isophotal diameter, surface brightness, and central light concentration. Kodaira (1988) proposes that one of the principal components is phase space density. These results are in agreement with work described above when ellipticals are treated separately. They also agree with results obtained separately for spirals (e.g. Whitmore 1984). Spirals and ellipticals have similar principal-component solutions. Nevertheless, it is not clear whether it is meaningful to lump together galaxies of different Hubble types, since different dynamical subsystems and stellar populations (young disks and old spheroids) contribute to the measured quantities.

In retrospect, the fundamental plane was already implicit in papers by Michard (1979), de Vaucouleurs & Olson (1982), Brosche & Lentes (1983), and Lentes (1983), although they did not recognize the full implications of their results.

There is some controversy about whether luminosity is the "first" parameter (i.e. whether it accounts for the greater part of the variance in other parameters). The present authors disagree on this point. SD believes that it is misleading to consider luminosity as the first parameter. The
axis perpendicular to the luminosity in the fundamental plane does not correspond to any direct observable. Even if subsystems of galaxies have a first parameter, this does not prove that the same parameter has the same controlling effect on galaxies as a whole. SD therefore believes that it is most profitable to think of the velocity dispersion and surface brightness as the principal variables from which one can derive luminosity, radius, and other quantities of interest. JK is unconvincing. Even though the \( \langle \mu \rangle - \sigma \) diagram in Figure 2 shows no correlation, he worries that subtle problems may have enlarged the scatter. Large samples were required to explore these issues; then many of the galaxies are far away and suffer from problems like seeing and sample selection. Also, cores of nearby galaxies do show a \( \mu_0 - \sigma \) correlation (J. Kormendy, in preparation). JK believes that more work is needed on the question of whether one first parameter is more fundamental than the others.

8.4 Uses and Interpretation of the Fundamental Plane

The fundamental plane is a powerful new distance indicator for early-type galaxies. Using it, Lynden-Bell et al. (1988) have discovered large-scale galaxy streaming motions toward the Hydra-Centaurus Supercluster (the "Great Attractor" model).

The plane also contains valuable clues about galaxy formation. Its solutions are very robust, and the residual scatter is very low. The solution has the same form for ellipticals and bulges, it spans about three orders of magnitude in luminosity, it varies little (if at all) in different environments, and it does not depend on how the parameters are measured. It must reflect an important regularity in the process of elliptical galaxy formation or transformation. One useful representation of the fundamental plane is its projection on the log \( \sigma \)-log \( I \) plane of observables. This is the "cooling diagram" in theories of galaxy formation [i.e. virial temperature vs. density (e.g. Rees & Ostriker 1977, Faber 1982, Silk 1983, 1985, 1987, Blumenthal et al. 1984)]. The position of a galaxy in this diagram is related to the amount of dissipation during its formation.

The fundamental plane can be understood using the following simple argument (Faber et al. 1987, Djorgovski et al. 1989). The virial theorem implies that galaxies must satisfy a relation that is very similar to Equation 1:

\[
R \sim k_S k_E \sigma^2 I^{-1} (M/L)^{-1}.
\]

The parameter \( k_S \) reflects the density, luminosity, and kinematic structure of a galaxy; it would be a constant if all galaxies considered had the same dynamical structure. The parameter \( k_E \) is the ratio of absolute potential energy to kinetic energy for a galaxy: \( k_E > 1 \) for a bound system, and
\( k_F = 2 \) for a virialized one. The deviations in Equation 1 of the coefficients of \( \sigma \) and \( I \) from 2 and \( -1 \), respectively, reflect the dependence of \( k_SK(F)(M/L)/L^{-1} \) on galaxy mass or other fundamental plane variables. If all of the variation is in mass-to-light ratio, this implies the scaling relation \( M/L \propto M^{0.2} \) (Faber et al. 1987, Djorgovski 1987b; cf. Section 8.2). A more complete discussion is given by Djorgovski et al. (1989).

The parameters \( k_S, k_F, \) and \( M/L \) depend on the formation and evolutionary histories of galaxies (Djorgovski et al. 1989). Our present understanding of galaxy formation is that it consists of a series of dissipative merging and infall processes, most of which are affected by environment (e.g. Silk & Norman 1981, Silk 1987). In fact, Vader (1986b) found a marginal but systematic difference between the \( L-\sigma-M_g \) relations in the Virgo and Coma clusters. Also, Djorgovski et al. (1989) find that the \( D_n-\sigma \) relation in different clusters varies with cluster richness class. Within clusters, it varies with distance from the cluster center. Further investigation of the fundamental plane and its dependence on environment is desirable but will require large bodies of high-quality data.

9. ISOPHOTE SHAPES

Fundamental new constraints on the structure of elliptical galaxies are emerging from studies of isophote shapes. This work may resolve a well-known shortcoming of the Hubble classification scheme: While the sequence S0–Im is one of changing physical properties, that from E0 to E6 is not a sequence of anything fundamental (Tremaine 1987). New observations suggest that ellipticals form a physical sequence that is continuous with S0s at one end. Along this sequence, rotation decreases in importance compared with anisotropic velocity dispersions. This subject is developing rapidly; we summarize it as of December 1988.

CCD photometry shows that the isophotes of elliptical galaxies are usually not perfect ellipses. Some are box shaped and others have disk-shaped distortions along their major axes (Carter 1979, 1987, Lauer 1985c, Jedrzejewski 1987a,b, Jedrzejewski et al. 1987, Michard & Simien 1987, 1988, Bender & Möllenhoff 1987, Bender et al. 1987, 1988a, Ebneter et al. 1988, Franx et al. 1988, Peletier et al. 1988a). It is convenient to parametrize these departures by the amplitude \( a(4) \) of the cos(4\( \theta \)) term in a Fourier expansion of the isophote radius in polar coordinates [see Carter (1978) and the above papers]. Along the major axis, the fractional radial departures from ellipses are typically \( a(4)/a \approx 1\% \). Positive values of \( a(4)/a \) describe diskly isophotes; negative values describe boxy isophotes.

Our discussion of \( a(4)/a \) measurements follows an excellent paper by
Bender et al. (1988b; hereafter B+88). Figure 3 shows correlations of various parameters with $a(4)/a$. The upper-left panel shows that rotation is dynamically less important in boxy ellipticals than in disky ellipticals (Lauer 1985c, Carter 1987, Bender 1987, 1988a, Nieto et al. 1988, Wagner et al. 1988, B+88, Nieto & Bender 1988). All disky ellipticals show significant rotation, and many are consistent with isotropic models. Boxy ellipticals have a variety of $(V/σ)^*$ values but include all of the galaxies with negligible rotation. They are also notable for showing minor-axis rotation (Davies & Birkinshaw 1986, Wagner et al. 1988). Bender, Nieto,
and collaborators suggest that ellipticals can be divided into two groups: boxy, slowly rotating, anisotropic ellipticals; and rapidly rotating, disky galaxies that connect with the S0 sequence.

Such a division also segregates ellipticals by other physical properties. For example, the upper-right panel shows that \( a(4)/a \) correlates with ellipticity. The distribution of points is V-shaped. Galaxies that appear almost round are almost elliptical. More flattened galaxies tend to be either box- or disk-shaped (see also Jedrzejewski 1987a). Since ellipticals have a preferred shape of E3.8 (Sandage et al. 1970, Binney & de Vaucouleurs 1981), these flattened galaxies are close to edge-on, and most round galaxies are close to face-on. B+88 therefore suggest that essentially all ellipticals are either boxy or disky when seen edge-on. This again suggests a dichotomy (but see below).

Bender et al. (1987) and B+88 also find that X-ray and radio luminosities of ellipticals show striking correlations with \( a(4)/a \) (e.g. Figure 3, bottom right). With few exceptions, only box-shaped ellipticals are strong radio or X-ray sources. Disky ellipticals have X-ray luminosities that are consistent with emission by compact sources only. Also, they are weak radio sources, like S0 galaxies (Hummel & Kotanyi 1982). These results are remarkably clear-cut, at least in the B+88 sample. We do not know what they mean.

B+88 also note that average mass-to-light ratios are higher in boxy than in disky ellipticals. This is consistent with Heckman's (1983) finding that powerful radio galaxies have higher \( M/L \) values than do other ellipticals. B+88 calculate global \( M/L \) values using central velocity dispersions and effective radii. However, slow rotation demonstrates that velocity anisotropies are important in ellipticals, and these affect the calculations (e.g. Binney & Tremaine 1987, Merritt 1988). We therefore checked the B+88 results using core \( M/L \) values from K87. These are not immune from anisotropy problems but should be more secure. Only 24 objects with \( a(4)/a \) values quoted in B+88 are in both galaxy samples, but they are a fair sample of Bender's \( M/L \) values. We confirm Bender's result: The boxy and disky galaxies have mean \( M/L_\nu \) values of \( 7.0 \pm 0.6 \) and \( 4.2 \pm 0.6 \), respectively. Since \( M/L_\nu \) depends on \( L \), this is not the best way to express the result (although these boxy and disky ellipticals happen to have the same mean luminosity). Rather, the bottom-left panel of Figure 3 suggests that \( a(4)/a \) may be a second parameter in the \( M/L_\nu-L \) correlation. Galaxies with low mass-to-light ratios for their luminosities tend to have larger \( a(4)/a \) values.

We believe that this is new evidence for velocity anisotropy. Figure 3 (upper left) shows that boxy ellipticals are especially anisotropic. If the radial component of \( \sigma \) is larger than the tangential component, then by
ignoring this we overestimate $M/L_V$. Similarly, we underestimate $M/L_V$ in disky ellipticals because we neglect rotational support. This suggests that there should be a correlation between $\delta \log (M/L_V)$ and $(V/\sigma)^*$, and one is observed, but it is not better than the one between $\delta \log (M/L_V)$ and $a(4)/a$. Therefore, anisotropy is not the whole story. A larger galaxy sample is needed to pursue these questions.

The distribution of points in Figure 3 and other similar correlations in B+88 suggest two alternative interpretations. First, it is possible that $a(4)/a$ measures the distribution of ellipticals along a continuous (but not necessarily uniformly populated) sequence that connects smoothly with S0s at one end. As rotation decreases, galaxies become intrinsically more spherical and anisotropic. However, the most anisotropic galaxies must be flattened and turn out to be boxy. Alternatively, perhaps only the disky ellipticals are the continuation of the Im–S0 sequence, and boxy ellipticals are a separate group with a different origin.

It is clear that at least two kinds of boxy structure are seen (e.g. Bender 1988a, Nieto & Bender 1988), because boxy ellipticals include the slowest rotators, whereas box-shaped bulges of disk galaxies rotate rapidly (K82). Interestingly, the few box-shaped ellipticals that rotate rapidly are small companions of much larger galaxies (Nieto & Bender 1988; see also Jedrzejewski 1987a, Peletier et al. 1988a). This suggests to these authors that the boxy structure is related to interactions (May et al. 1985). Accretion events also seem capable of leaving behind an excess of box or tube orbits (ex-polar rings?) that could create slowly or rapidly rotating boxy structure, respectively (Binney & Petrou 1985, Whitmore & Bell 1988, Hernquist & Quinn 1988, Statler 1988). The extreme case IC 3370 may be an example of the latter (Jarvis 1987). Of course, it is also possible that one or both kinds of boxy structure are primordial.

These developments have great potential for clarifying our picture of galaxy formation and structure. However, it is still early. Also, we have ignored complications like dust, variations of $a(4)/a$ with radius, and other Fourier components in the isophotes. The present discussion will undoubtedly prove inadequate; our main aim is to stimulate further work on these important issues.

10. COLORS AND COLOR GRADIENTS

In old stellar systems, colors are a complicated measure of metallicity (Burstein et al. 1984, Aragón et al. 1987) and age (O’Connell 1986). Colors, color gradients, and their correlations with other galaxy properties can therefore be used to test theories of galaxy formation. A few of the many
reviews of this subject include Faber (1977), Pagel & Edmunds (1981), Burstein (1985), Norman et al. (1986), and Thomsen & Baum (1988).

The correlation between color and luminosity for early-type galaxies is well known (Sandage 1972, Visvanathan & Sandage 1977, Sandage & Visvanathan 1978a,b, Strom et al. 1976, 1978, Frogel et al. 1978): More luminous galaxies are redder and thus more metal rich. For example, Sandage & Visvanathan find that \( \log L = 4.1 (u-V) + \) constant. Also, the centers of most early-type galaxies are redder than their envelopes; typical gradients are \( \Delta(b-V) \approx -0.03 \) and \( \Delta(u-V) \approx -0.10 \) magnitudes per decade in radius. Color gradients in bulges of spirals are generally stronger than those in ellipticals (Wirth 1981, Wirth & Shaw 1983). The same effects are seen in spectroscopic metallicity indicators (Faber 1973, Terlevich et al. 1981, Tonry & Davis 1981). From population synthesis models, Tinsley (1978) derives the mass-metallicity relation, \( Z \propto M^{0.25} \), where \( Z \) is the logarithm of the metallicity. Metallicity variations can partly, but not entirely, explain the observed dependence of \( M/L \) on luminosity (Smith & Tinsley 1976).

These results can be understood within the framework of dissipative galaxy formation (Larson 1974, 1975, Silk & Norman 1981, Carlberg 1984a,b, Arimoto & Yoshii 1987, Yoshii & Arimoto 1987, Matteucci & Tornambè 1987). Carlberg’s models in particular avoid some technical limitations of Larson’s pioneering work and make more detailed predictions. In these models, the removal of enriched gas by supernova-driven galactic winds is more efficient for less massive galaxies (see also Section 8.1). In this spirit, the color-luminosity relation is recast as a metallicity–escape velocity relation by Vigroux et al. (1981): \( Z \propto V_e^{0.9} \). Similarly, colors and metallicities correlate better with central velocity dispersions than with luminosities. Carlberg (1984b) also predicts the existence of a second parameter in the Faber-Jackson and mass-metallicity relations. Larson and Carlberg both predict that color and metallicity gradients should be stronger in more massive galaxies. Finally, they make testable predictions about the relative shapes of isophotes and isochromes (i.e. isometallicity contours). In Larson’s models, isochromes are considerably flatter than isophotes. However, Larson’s ellipticals are supported by rotation, which we now know is incorrect (e.g. Illingworth 1981, Davies et al. 1983). Carlberg’s models are generally supported by velocity anisotropy; then isochromes are only slightly flatter than isophotes.

CCD photometry has provided high-quality measurements of color gradients for large numbers of galaxies. Boroson et al. (1983), Davis et al. (1985), Cohen (1986), Boroson & Thompson (1987), and Bender & Möllenhoff (1987) present data on relatively small samples of ellipticals, mostly in the Virgo cluster. They conclude that color gradients are common.
in ellipticals. In the absence of nonthermal emission or recent star formation, colors always get redder toward the center. Interestingly, isophotes and isochromes generally have the same shape. In fact, isochromes are occasionally rounder than isophotes (Boroson et al. 1983). This is consistent with Carlberg's but not Larson's models.

The interpretation of color gradients in terms of stellar population gradients has been discussed recently by Efstathiou & Gorgas (1985), Gorgas & Efstathiou (1987), Davies & Sadler (1987), Couture & Hardy (1988), and references therein. They present extensive evidence for gradients in $Mg_2$ indices. Assuming the somewhat uncertain conversions between $Mg_2$ index and metallicity (Terlevich et al. 1981) and between color and metallicity (Strom et al. 1976, 1978, Tinsley 1978), they find that the $Mg_2$ and color gradients are mutually consistent and imply typical changes of $\Delta[Fe/H] \sim -0.2$ per decade in radius. In excellent papers, Baum et al. (1986) and Thomsen & Baum (1987, 1988) derive metallicity gradients from narrowband surface photometry. They also find that isochromes are not flatter than isophotes, in agreement with spectroscopic results. Similar photometric measurements of the $Mg_2$ index are reported by Vigroux et al. (1988). Also, Vader et al. (1988) find that $Mg_2$ gradients correlate well with broadband color gradients. Further constraints are obtained by Peletier and coworkers (Peletier et al. 1987, 1988a, b, c, Peletier & Valentijn 1988, Peletier 1988). They show that observed optical and near-infrared ($JHK$) color gradients are mutually consistent (i.e. one can be derived from the other using the separate optical and infrared color-luminosity relations). All this suggests that the same change in stellar population produces both the color-luminosity relation and the color gradients. Using the new Yale isochrones (Green et al. 1987), Peletier and coworkers conclude that most color variations and gradients are due to changes in metallicity. In typical ellipticals, these do not exceed a factor of 10 inside $r_e$. However, age gradients may be present as well; the fraction of young stars may increase at larger radii.

Large data sets are needed to investigate correlations of color gradients with other galaxy properties. The measurements are difficult because color gradients are weak and because differential magnitude measurements are sensitive to systematic errors. Nevertheless, important data for early-type galaxies have been obtained by Jedrzejewski (1987b), Vigroux et al. (1988), Franx (1988), Franx et al. (1988), and Peletier and collaborators (see above).

Vader et al. (1988) analyze data from Vigroux et al. (1988) and obtain several interesting results. Whereas inward reddening is the rule in elliptical galaxies and bulges, they find that dSph galaxies tend to become bluer toward the center. Particularly interesting is the observation that color gradients are correlated with the rotation parameter $(V/\sigma)^*$: Anisotropic,
pressure-supported ellipticals have smaller color gradients. We find the same effect, although with more scatter, in the Franx and Peletier et al. data (Figure 4).

The bright ellipticals in the Franx (1988) sample show weak correlations of color gradients with luminosity, velocity dispersion, integrated color, and Mg$_2$ index: Weaker gradients are seen in brighter, hotter, redder, and more metal-rich galaxies. Gorgas & Efstathiou (1987) also find a marginally significant anticorrelation between Mg$_2$ gradients and velocity dispersion. However, Peletier et al. (1988a) find no significant correlations with the above quantities. When we combine the Vader et al., Franx, and Peletier et al. samples (Figure 4), color gradients in E and S0 galaxies are weak or absent at low luminosities ($M_B > -20$) and largest near the peak of the luminosity function ($M_B \sim -20$). The scatter exceeds the measurement errors at all luminosities.

We also find a marginal correlation of color gradients with isophote

![Figure 4](image)

**Figure 4** Correlations of color gradients with other galaxy properties. The gradients are defined as $\Delta$(Color)/$\Delta$(log $r$), in magnitudes per decade in radius; positive values indicate reddening toward the center. The data are from Vader et al. (1988; circles), Franx (1988; squares), and Peletier et al. (1988a; triangles). The left panel shows the dependence of color gradients on luminosity for dSph galaxies (open circles) and for ellipticals and S0s (solid symbols). Ellipticals become redder toward the center, but most dSph galaxies have inverse gradients. The top-right panel shows the relation between color gradients and the level of rotational support; $(V/\sigma)^* = 1$ for an isotropic oblate rotator. Anisotropic galaxies tend to have smaller gradients. The bottom-right panel shows the correlation with isophote shape (measured by B + 88). More boxy galaxies [$a(4)/a < 0$] tend to have smaller color gradients. Similar trends are obtained using $(U - R)$ color measurements.
shapes (Figure 4). Color gradients in boxy ellipticals get smaller as boxyness increases [i.e. as $a(4)/a$ decreases further below 0]. There are not enough data to look for variations of color gradients in disky ellipticals.

These results are very preliminary. However, the trends are probably real: All of the parameters are measured independently, and measurement errors only diminish the correlations. If confirmed, these correlations will provide important new information about galaxy formation.

The absence of a strong correlation between the strengths of color gradients and $L$ or $\sigma$ is contrary to the predictions of the Larson and Carlberg models. However, these do not include the effects of postcollapse mergers, which are important at least for bright ellipticals. The observations suggest that the properties of early-type galaxies are determined by dissipative collapse and then modified by mergers (Vader et al. 1988). Dissipative formation could produce a mass-metallicity relation and metallicity gradients, which are then gradually erased by mergers. The trend with luminosity at $M_B \gtrsim -21$ would then be a fossil of the initial correlation predicted by Larson and Carlberg. Mergers may also produce correlated changes in $(V/\sigma)^*$ and $a(4)/a$ (cf. Figure 3). Normal color gradients are not produced in diffuse dwarfs because of galactic winds (Vader 1986a, Dekel & Silk 1986); their inverse color gradients could be due to recent star formation.

Finally, we discuss color gradients in cooling-flow galaxies. The existence of cooling flows in ellipticals and cluster cores is reasonably well established (e.g. Fabian 1988). Mass flow rates are uncertain, but $\dot{M} \sim 1$–1000 $M_\odot$ yr$^{-1}$ are believed to be deposited inside a few $r_e$, with a strong dependence on radius. The most plausible fate of the gas is star formation. Unless the initial mass function (IMF) is strongly biased toward low-mass stars, observable color gradients should result (Sarazin & O'Connell 1983, Silk et al. 1986, O'Connell 1988). Blueing toward the centers of some high-$\dot{M}$ cooling-flow galaxies is seen and interpreted as evidence for age gradients (e.g. Wirth et al. 1983, Romanishin 1986a, 1987, Maccagni et al. 1988). However, the extreme gradients reported by Valentijn (1983) and Valentijn & Moorwood (1985) are not confirmed by subsequent work. A detailed comparison of color gradients in normal and cooling-flow galaxies could provide strong constraints on the fate of gas in cooling flows.

11. cD AND BRIGHTEST CLUSTER GALAXIES

The most luminous galaxies in rich clusters are giant ellipticals. Some of these are classified as D (possessing a large, diffuse envelope) or cD (extra-large D) (Matthews et al. 1964, Morgan & Lesh 1965, Morgan et al. 1975, Albert et al. 1977). The distinction between cD and E galaxies is useful,
but we argue below that the term "D galaxy" is not. When we wish to discuss brightest cluster members without regard to their morphology, we will refer to them as BCMs.

The importance of BCMs is twofold. First, because of their large luminosities, they are used as standard candles for cosmological studies. Second, because of their large masses and special locations, they are believed to be the sites of interesting evolutionary phenomena (e.g. dynamical friction, galactic cannibalism, interactions with the intracluster medium, and cooling flows). These subjects are reviewed in detail by Dressler (1984b) and Tonry (1987); here we discuss them briefly in the context of recent photometry. Extensive photometric studies are reported by Oemler (1976), Thuan & Romanishin (1981), Schneider et al. (1983), Lugger (1984), Malumuth & Kirshner (1985), Hoessel & Schneider (1985), Schombert (1986, 1987, 1988), Lauer (1988a), and Porter (1988). Except for Valentijn (1983) and Morbey & Morris (1983), most authors are in fairly good agreement.

The classification of galaxies as D and cD is often done loosely and may be misleading. More luminous ellipticals have shallower brightness profiles (Section 7.1). Also, galaxies can have S0 disks and tidally stretched halos. All of these satisfy the definition of a D galaxy but are not new phenomena, nor even a single class of phenomena. For this reason, we recommend that the term D galaxy not be used (Kormendy 1987b). On the other hand, the cD classification is useful: cDs are physically different from ellipticals.

Finding an objective way to recognize cDs is nontrivial. The most objective way is to look for an inflection in the outer brightness profile that is independent of the way the profile is plotted. This is interpreted as the signature of a halo that is a distinct dynamical subsystem. In practice, not all halos are prominent enough to produce inflection points in the profiles. Then, less objective identification criteria are necessary, such as extra light compared with mean profiles of comparably bright ellipticals. We use the term cD only for bright ellipticals with such extra envelopes. It is not clear whether a cD envelope belongs to the galaxy or whether it was formed by the parent cluster independently of whether there happened to be a bright galaxy at the bottom of the cluster potential well.

BCM and cD galaxies are generally believed to have formed or been modified by mergers (e.g. V Zw 311; Schneider & Gunn 1982, and references therein). The luminosities of BCMs are weakly correlated with some properties of their parent clusters, like Abell richness class or Bautz-Morgan type (Sandage 1972, Sandage & Hardy 1973, Schneider et al. 1983, Morbey 1984, Hoessel & Schneider 1985, Schombert 1987). Schombert (1987) also finds weak correlations with cluster velocity dispersion and X-ray luminosity. Beers & Geller (1983) find that cD galaxies are always found in local density maxima, even if they are not the brightest or central cluster.
members. Such correlations suggest that environment-dependent processes are responsible for at least some of the luminosity of BCMs. Mergers are a natural candidate, but other options are possible. Examples include a gradual accumulation of cluster tidal debris (Malumuth & Richstone 1984, and references therein) or star formation in (now possibly extinguished) cooling flows (Sarazin 1986, and references therein). Or the large luminosities of BCMs may be a consequence of environment-dependent initial conditions. In the language of biased galaxy formation, BCMs may originate from unusually large primordial fluctuations, which are most likely to occur in dense environments.

An argument often cited in favor of mergers is the high frequency of secondary nuclei in BCMs. These may be semidigested cluster members. Schneider et al. (1983) and Hoessel & Schneider (1985) find that about half of the BCMs in their sample of Abell clusters are multiple-nucleus systems, considerably more than would be expected from chance projections if the clusters have cores. This argument is weakened by the conclusion of Beers & Tonry (1986) that rich clusters do not have cores, but instead have steep number density profiles. Then many nuclei are predicted to be near the central galaxy. Tonry (1984a, 1985; see also Hoessel et al. 1985) shows that many of the secondary nuclei move too quickly to be gravitationally bound to the BCM core. This effect was explained independently by Merritt (1984) and Tonry (1984a) as a natural outcome of the evolution of galaxy orbits in a rich cluster. A possible way to distinguish between bound and unbound secondary nuclei is to decompose BCM images into elliptically symmetric components and look for tidal distortions. Lauer (1986, 1988a) made such a study of 17 multiple-nucleus systems and found that \(\sim 50\%\) of the secondary nuclei show isophote distortions. However, these distortions do not correlate as expected with the kinematics. Even nuclei with large relative velocities show distortions. Therefore, the problem of which nuclei are currently being accreted is not settled. Based only on the observed distortions, Lauer estimates that material is being cannibalized at an average rate of \(\sim 4 L_\star \) (primary galaxy)\(^{-1}\) (10 Gyr)\(^{-1}\). [Here \(L_\star\) is the characteristic luminosity of the Schechter (1976) luminosity function.] This is in rough agreement with models by Merritt (1985), which imply accretion rates of \(\sim 1 L_\star \) (primary galaxy)\(^{-1}\) (10 Gyr)\(^{-1}\). Since the total luminosity of a cD galaxy is typically \(\sim 12 L_\star\), this argues that not all of a cD originates through cannibalism.

The structural properties of BCMs are often discussed in the framework of the homologous merger picture (e.g. Ostriker & Tremaine 1975, Ostriker & Hausman 1977, Hausman & Ostriker 1978, Malumuth & Richstone 1984, Merritt 1985; see also White 1982, and references therein). In this
picture, the kinetic energy per unit mass is preserved. Then, if the orbital structure of the cannibal galaxy stays the same, its projected central velocity dispersion does not change, even though the mass and luminosity increase. That is, merger products should deviate from the Faber-Jackson relation by being too luminous for their velocity dispersions. Such an effect was found by Malumuth & Kirshner (1985). Because of the conversion of galaxy orbital energy into internal random motions in the merger remnant, the envelope of the cannibal should get shallower after every merger [as measured, say, by the Gunn-Oke (1975) structure parameter $\alpha$]. Therefore, more luminous merger remnants should have shallower profiles. This is the sense of the luminosity dependence of profile shapes for all ellipticals (Section 7.1; Schneider et al. 1983, Hoessel & Schneider 1985). Finally, compared with the color-luminosity relation for normal galaxies, merger remnants should be too blue for their luminosities. This prediction is not confirmed: Gallagher et al. (1980), Lugger (1984), Lachièze-Rey et al. (1985), and Schombert (1988) find that integrated colors and color gradients in the envelopes of cD galaxies are consistent with those in normal ellipticals of comparable luminosities.

Of more interest is the radius–surface brightness relation (cf. Section 8.2). Kormendy (1980), Thomsen & Frandsen (1983), Lugger (1984), and Romanishin (1986b) find strong correlations consistent with those for normal ellipticals. With larger data sets, Schneider et al. (1983), Schombert (1987), and Hoessel et al. (1987) conclude that the relations between $r_e$ and $\mu_e$ or $L$ are steeper for BCM than for non-BCM ellipticals. Hoessel et al. (1987) find that $r_e \propto I_e^{-0.80}$ for BCMs and $r_e \propto I_e^{-0.55}$ for non-BCM in the Gunn $r$ band. BCMs are larger at a given surface brightness than normal ellipticals. However, since the relation for non-BCM ellipticals that are well resolved is $r_e \propto I_e^{-0.83}$ (Section 8.2), the above difference may be due partly to seeing effects. (The non-BCM ellipticals in the Hoessel et al. sample typically have $r_e \approx 2-5^\prime\prime$.)

An even more informative comparison of BCMs and other ellipticals can be made using fundamental plane solutions. Hoessel et al. (1987) find that the $R-\sigma-\mu$ solutions for the galaxies in their sample are consistent with solutions for normal ellipticals (Djorgovski & Davis 1987), but with a hint of a different slope. S. Djorgovski & R. R. de Carvalho (in preparation), using data from Malumuth & Kirshner (1985), obtain different solutions for BCMs and normal ellipticals. At a given effective radius or luminosity, the range of velocity dispersions is much smaller than for normal ellipticals; as a result, the scatter in the $r_e-\mu_e$ relation is smaller for BCMs than for non-BCM ellipticals. This can be understood in the homologous merger picture, because velocity dispersions are not changed much by mergers, while luminosities and radii increase. Different fun-
damental planes for BCMs and other ellipticals imply different formation histories.

The $r_e-\mu_e$ relation and especially the fundamental plane solutions for BCMs are promising distance indicators (Thomsen & Frandsen 1983, Hoessel et al. 1987; S. Djorgovski & R. R. de Carvalho, in preparation). They may also lead to an improved angular size–redshift cosmological test.

The origin of cD halos remains murky. They are purely a rich cluster phenomenon: Thuan & Romanishin (1981) and Schombert (1986) find that cD halos do not occur in poor clusters. Struble (1988) has discovered what appears to be a cD envelope without a central galaxy in the rich cluster Abell 545. It would be interesting to know whether more such cases exist. They are difficult to find because of their low surface brightnesses ($\mu_V \sim 24$ mag arcsec$^{-2}$) and unimposing luminosities ($L \sim L_*$ for Struble’s “star pile”).

The most systematic photometric study of cD envelopes to date is by Schombert (1988), building on work by Oemler (1976). He subtracted template brightness profiles of ellipticals from those of cD galaxies and measured the properties of the envelopes. He found that envelopes have brightness profiles $I(r) \propto r^{-1.6}$ similar to those of their galaxies and any X-ray halos. Envelope luminosities are comparable to those of the central galaxies ($\sim 10^{12} L_\odot$). Therefore, if the theoretical models are correct, mergers are an insufficient source of material to build either cD galaxies or their envelopes. Envelope luminosities correlate with parent galaxy luminosities, which suggests that similar processes may be responsible for both. They also correlate with cluster richness ($L_{\text{env}} \propto N^{1.6}$, where $N$ is the Abell galaxy count) and weakly with Bautz-Morgan type. Finally, there is a good correlation with the cluster X-ray luminosity ($L_{\text{env}} \propto L_X^{0.6 \pm 0.18}$).

Other connections between BCMs and their clusters include alignment effects. In clusters with well-defined orientations, BCM isophotes tend to align with cluster major axes (Carter & Metcalfe 1980, Binggeli 1982, Porter 1988, Rhee & Roos 1989; see also the review by Djorgovski 1987c). Cluster position angles are uncertain, but Porter (1988) also finds a tendency for alignment with cluster X-ray gas isodensity contours. Ellipticity tends to increase strongly with radius in BCMs. Porter finds that BCMs tend to have larger ellipticities and ellipticity gradients and smaller isophote twists than normal ellipticals.

All of these correlations suggest that cD envelopes are products of their clusters. They may be the accumulated debris of all tidal interactions. Further support for this interpretation comes from the kinematics: The projected velocity dispersions of cDs increase with radius and approach cluster velocity dispersions (Dressler 1979, Carter et al. 1981, 1985).
12. CONCLUSION

This paper is being written at a time of unusually rapid progress in galaxy photometry. Some tentative conclusions, especially in the latter parts of this paper, may even evolve by the time you read this. CCDs have made it possible to measure second-order structure parameters like isophote shape distortions, color gradients, and dust. More difficult physical questions have become accessible. Soon, near-infrared array detectors will provide important new kinds of data. Therefore, the next few years promise substantial progress in our understanding of the structure of elliptical galaxies.

ACKNOWLEDGMENTS

We are most grateful to R. Bender, R. Davies, R. de Carvalho, G. Illingworth, J.-L. Nieto, R. Peletier, F. Schweizer, A. Toomre, and especially T. de Zeeuw for very helpful discussions, comments on the manuscript, or data in advance of publication. Also, it is a pleasure to thank the many people who sent us preprints, including those on subjects that were ultimately not covered because of space limitations. During this work, SD was supported by Caltech and the Alfred P. Sloan Foundation.

Literature Cited


© Annual Reviews Inc. • Provided by the NASA Astrophysics Data System
Carter, D., Inglis, I., Ellis, R. S., Efstatthiou,
KORMENDY & DJORGOVSKI

de Carvalho, R. R., da Costa, L. N. 1989. Submitted for publication
de Vaucouleurs, G. 1953. MNRAS 113: 134

The Revised Yale Isochrones and Luminosity Functions. New Haven, Conn: Yale Univ. Obs.


Kormendy & Djorgovski

Kodaira, K. 1988. Submitted for publication
Kormendy, J. 1989. Submitted for publication

© Annual Reviews Inc. • Provided by the NASA Astrophysics Data System
Sarazin, C. L. 1986. Rev. Mod. Phys. 58: 1

Terlevich, R., Davies, R. L., Faber, S. M., Burstein, D. 1981. MNRAS 196: 381

© Annual Reviews Inc. • Provided by the NASA Astrophysics Data System