



Early-Type Galaxies in the Sloan Digital Sky Survey. III. The Fundamental Plane

Citation

Bernardi, Mariangela, Ravi K. Sheth, James Annis, Scott Burles, Daniel J. Eisenstein, Douglas P. Finkbeiner, David W. Hogg, et al. 2003. "Early-Type Galaxies in the Sloan Digital Sky Survey. III. The Fundamental Plane." *The Astronomical Journal* 125 (4) (April): 1866–1881. doi:10.1086/367794.

Published Version

10.1086/367794

Permanent link

<http://nrs.harvard.edu/urn-3:HUL.InstRepos:33461902>

Terms of Use

This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at <http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA>

Share Your Story

The Harvard community has made this article openly available. Please share how this access benefits you. [Submit a story](#).

[Accessibility](#)

EARLY-TYPE GALAXIES IN THE SLOAN DIGITAL SKY SURVEY. III. THE FUNDAMENTAL PLANE

MARIANGELA BERNARDI,^{1,2} RAVI K. SHETH,^{3,4} JAMES ANNIS,³ SCOTT BURLES,³ DANIEL J. EISENSTEIN,⁵
DOUGLAS P. FINKBEINER,^{6,7,8} DAVID W. HOGG,⁹ ROBERT H. LUPTON,⁷ DAVID J. SCHLEGEL,⁷ MARK SUBBARAO,¹
NETA A. BAHCALL,⁷ JOHN P. BLAKESLEE,¹⁰ J. BRINKMANN,¹¹ FRANCISCO J. CASTANDER,^{12,13} ANDREW J. CONNOLLY,⁴
ISTVÁN CSABAI,^{10,14} MAMORU DOI,^{15,16} MASATAKA FUKUGITA,^{17,18} JOSHUA FRIEMAN, TIMOTHY HECKMAN,¹⁰
GREGORY S. HENNESSY,¹⁹ ŽELJKO IVEZIĆ,⁷ G. R. KNAPP,⁷ DON Q. LAMB,¹ TIMOTHY MCKAY,²⁰ JEFFREY A. MUNN,¹⁹
ROBERT NICHOL,² SADANORI OKAMURA,^{16,21} DONALD P. SCHNEIDER,²² ANIRUDDHA R. THAKAR,¹⁰ AND DONALD G. YORK¹

Received 2002 September 27; accepted 2002 December 3

ABSTRACT

A magnitude-limited sample of nearly 9000 early-type galaxies in the redshift range $0.01 \leq z \leq 0.3$ was selected from the Sloan Digital Sky Survey (SDSS) using morphological and spectral criteria. The fundamental plane relation in this sample is $R_o \propto \sigma^{1.49 \pm 0.05} I_o^{-0.75 \pm 0.01}$ in the r^* band. It is approximately the same in the g^* , i^* , and z^* bands. Relative to the population at the median redshift in the sample, galaxies at lower and higher redshifts have evolved only a little. If the fundamental plane is used to quantify this evolution, then the apparent magnitude limit can masquerade as evolution; once this selection effect has been accounted for, the evolution is consistent with that of a passively evolving population that formed the bulk of its stars about 9 Gyr ago. One of the principal advantages of the SDSS sample over previous samples is that the galaxies in it lie in environments ranging from isolation in the field to the dense cores of clusters. The fundamental plane shows that galaxies in dense regions are slightly different from galaxies in less dense regions.

Key words: galaxies: elliptical and lenticular, cD — galaxies: evolution —
galaxies: fundamental parameters — galaxies: photometry — galaxies: stellar content

1. INTRODUCTION

This is the third of four papers in which the properties of ~ 9000 early-type galaxies in the redshift range $0.01 \leq z \leq 0.3$ are studied. Paper I (Bernardi et al. 2003a) describes how the sample was selected from the Sloan Digital Sky Survey (SDSS) database. The sample is essentially magnitude-limited, and the galaxies in it span a wide range of environments. Each galaxy in the sample has measured values of luminosity L , size R_o , and surface brightness $I_o = (L/2)/R_o^2$ in four bands (g^* , r^* , i^* , and z^*), a velocity dispersion σ , a redshift, and an estimate of the local density. Paper II (Bernardi et al. 2003b) shows that the joint distribution of early-type galaxy luminosities, radii, and velocity dispersions is reasonably well fitted by a trivariate Gaussian. It also shows various correlations between pairs of variables, such as the luminosity–velocity dispersion relation, the luminosity–size relation, and the relation between radius and surface brightness. Paper IV (Bernardi et al. 2003c) uses the spectra of these galaxies to provide information on the chemical evolution of the early-type population.

This paper places special emphasis on the fundamental plane (FP) relation between size, surface brightness, and velocity dispersion. It shows how the FP depends on wave band, color, redshift, and environment. In § 2.1, we compare the results of a maximum likelihood analysis, which can account for evolution and selection effects, as well as correlations between errors (e.g., Saglia et al. 2001; Paper II of this series) and standard regression estimates, which cannot. Section 2.2 checks if the residuals from the plane correlate with any other observables. A discussion of the mass-to-light ratio is provided in § 2.3. Evidence for weak evolution is presented in § 2.4, and weak trends with environment are found in § 2.5. The distribution of the galaxies in our sample in κ -space (Bender, Burstein, & Faber 1992) is shown in § 2.6. We summarize our findings in § 3.

¹ Department of Astronomy and Astrophysics, University of Chicago, 5640 South Ellis Avenue, Chicago, IL 60637.

² Department of Physics, Carnegie Mellon University, 5000 Forbes Avenue, Pittsburgh, PA 15213.

³ Fermi National Accelerator Laboratory, P.O. Box 500, Batavia, IL 60510.

⁴ Department of Physics and Astronomy, University of Pittsburgh, 3941 O'Hara Street, Pittsburgh, PA 15620.

⁵ Steward Observatory, University of Arizona, 933 North Cherry Avenue, Tucson, AZ 85721.

⁶ Department of Astronomy, 601 Campbell Hall, University of California, Berkeley, Berkeley, CA 94720.

⁷ Princeton University Observatory, Peyton Hall, Princeton, NJ 08544.

⁸ Hubble Fellow.

⁹ Department of Physics, New York University, 4 Washington Place, New York, NY 10003.

¹⁰ Department of Physics and Astronomy, Johns Hopkins University, 3400 North Charles Street, Baltimore, MD 21218-2686.

¹¹ Apache Point Observatory, P.O. Box 59, Sunspot, NM 88349-0059.

¹² Department of Physics, Yale University, P.O. Box 208101, New Haven, CT 06520.

¹³ Departamento de Astronomía, Universidad de Chile, Casilla 36-D, Santiago, Chile.

¹⁴ Department of Physics of Complex Systems, Eötvös Loránd University, Pf. 32, H-1117 Budapest, Hungary.

¹⁵ Institute of Astronomy, School of Science, University of Tokyo, Mitaka, Tokyo 181-0015, Japan.

¹⁶ Research Center for the Early Universe, School of Science, University of Tokyo, Tokyo 113-0033, Japan.

¹⁷ Institute for Cosmic Ray Research, University of Tokyo, Kashiwa 277-8582, Japan.

¹⁸ Institute for Advanced Study, Einstein Drive, Princeton, NJ 08540.

¹⁹ US Naval Observatory, 3450 Massachusetts Avenue, NW, Washington, DC 20392-5420.

²⁰ Department of Physics, University of Michigan, 500 East University, Ann Arbor, MI 48109.

²¹ Department of Astronomy, University of Tokyo, Tokyo 113-0033, Japan.

²² Department of Astronomy and Astrophysics, 525 Davey Laboratory, Pennsylvania State University, University Park, PA 16802.

Except where stated otherwise, we write the Hubble constant as $H_0 = 100 h \text{ km s}^{-1} \text{ Mpc}^{-1}$, and we perform our analysis in a cosmological world model with $(\Omega_M, \Omega_\Lambda, h) = (0.3, 0.7, 0.7)$, where Ω_M and Ω_Λ are the present-day scaled densities of matter and cosmological constant. In such a model, the age of the universe at the present time is $t_0 = 9.43 h^{-1} \text{ Gyr}$. For comparison, an Einstein–de Sitter model has $(\Omega_M, \Omega_\Lambda) = (1, 0)$ and $t_0 = 6.52 h^{-1} \text{ Gyr}$. We frequently use the notation h_{70} as a reminder that we have set $h = 0.7$. In addition, we will frequently be interested in the logarithms of physical quantities. Our convention is to set $R \equiv \log R_o$ and $V \equiv \log \sigma$, where R_o and σ are effective radii in $h_{70}^{-1} \text{ kpc}$ and velocity dispersions in kilometers per second, respectively.

2. THE FUNDAMENTAL PLANE

In any given band, each galaxy in our sample is characterized by three numbers: its luminosity, L , its size, R_o , and its velocity dispersion, σ . Correlations between these three observables are expected if early-type galaxies are in virial equilibrium, because

$$\sigma_{\text{vir}}^2 \propto \frac{GM_{\text{vir}}}{2R_{\text{vir}}} \propto \left(\frac{M_{\text{vir}}}{L}\right) R_{\text{vir}} \left(\frac{L/2}{R_{\text{vir}}^2}\right). \quad (1)$$

If the size parameter R_{vir} that enters the virial theorem is linearly proportional to the observed effective radius of the light, R_o , and if the observed line-of-sight velocity dispersion σ is linearly proportional to σ_{vir} , then this relates the observed velocity dispersion to the product of the observed surface brightness and effective radius. Following Djorgovski & Davis (1987), correlations involving all three variables are often called the fundamental plane. In what follows, we will show how the surface brightness, R_o , and σ are correlated. Because both $\mu_o \propto -2.5 \log[(L/2)/R_o^2]$ and σ are distance-independent quantities (this assumes that cosmological dimming and K -corrections have been computed correctly), it is in these variables that studies of early-type galaxies are usually presented.

2.1. Finding the Best-Fitting Plane

The FP is defined by

$$\log R_o = a \log \sigma + b \log I_o + c, \quad (2)$$

where the coefficients a , b , and c are determined by minimizing the residuals from the plane. There are a number of ways in which this is usually done. Let

$$\begin{aligned} \Delta_1 &\equiv \log R_o - a \log \sigma - b \log I_o - c, \\ \Delta_o &\equiv \frac{\Delta_1}{(1 + a^2 + b^2)^{1/2}}. \end{aligned} \quad (3)$$

Then summing Δ_1^2 over all N galaxies and finding that set of a , b , and c for which the sum is minimized gives what is often called the direct fit, whereas minimizing the sum of Δ_o^2 instead gives the orthogonal fit. Although the orthogonal fit is, perhaps, the more physically meaningful, the direct fit is of more interest if the FP is to be used as a distance indicator.

A little algebra shows that the direct-fit coefficients are

$$\begin{aligned} a &= \frac{\sigma_{\text{II}}^2 \sigma_{\text{RV}}^2 - \sigma_{\text{IR}}^2 \sigma_{\text{IV}}^2}{\sigma_{\text{II}}^2 \sigma_{\text{VV}}^2 - \sigma_{\text{IV}}^4}, & b &= \frac{\sigma_{\text{VV}}^2 \sigma_{\text{IR}}^2 - \sigma_{\text{RV}}^2 \sigma_{\text{IV}}^2}{\sigma_{\text{II}}^2 \sigma_{\text{VV}}^2 - \sigma_{\text{IV}}^4}, \\ c &= \overline{\log R_o} - a \overline{\log \sigma} - b \overline{\log I_o}, \\ \langle \Delta_1^2 \rangle &= (\sigma_{\text{II}}^2 \sigma_{\text{RR}}^2 \sigma_{\text{VV}}^2 - \sigma_{\text{II}}^2 \sigma_{\text{RV}}^4 - \sigma_{\text{RR}}^2 \sigma_{\text{IV}}^4 - \sigma_{\text{VV}}^2 \sigma_{\text{IR}}^4 \\ &\quad + 2\sigma_{\text{IR}}^2 \sigma_{\text{IV}}^2 \sigma_{\text{RV}}^2) / (\sigma_{\text{II}}^2 \sigma_{\text{VV}}^2 - \sigma_{\text{IV}}^4), \end{aligned} \quad (4)$$

where $\overline{\log X} \equiv \sum_i \log X_i / N$ and $\sigma_{xy}^2 \equiv \sum_i (\log X_i - \overline{\log X})(\log Y_i - \overline{\log Y}) / N$, and X and Y can be I_o , R_o , or σ . For what follows, it is also convenient to define $r_{xy} = \sigma_{xy}^2 / (\sigma_{xx} \sigma_{yy})$. The final expression above gives the scatter around the relation. If surface brightness and velocity dispersion are uncorrelated (we will show below that, indeed, $\sigma_{\text{IV}} \approx 0$), then a equals the slope of the relation between velocity dispersion and the mean size at fixed velocity dispersion, b is the slope of the relation between surface brightness and the mean size at fixed surface brightness, and the rms scatter is $\sigma_{\text{RR}}(1 - r_{\text{RV}}^2 - r_{\text{IR}}^2)^{1/2}$. Errors in the observables affect the measured σ_{xy}^2 and thus will bias the determination of the best-fit coefficients and the intrinsic scatter around the fit. If ϵ_{xy} is the rms error in the joint measurement of $\log X$ and $\log Y$, then subtracting the appropriate ϵ_{xy}^2 from each σ_{xy}^2 before using them provides estimates of the error-corrected values of a , b , and c . Expressions for the orthogonal fit coefficients can be derived similarly, although, because they require solving a cubic equation, they are lengthy, so we have not included them here.

Neither minimization procedure above accounts for the fact that the sample is magnitude-limited and has a cut at small velocity dispersions. In addition, because our sample spans a wide range of redshifts, we must worry about effects that may be due to evolution. The magnitude limit means that we cannot simply divide our sample up into small redshift ranges (over which evolution is negligible), because a small redshift range probes only a limited range of luminosities, sizes, and velocity dispersions. To account for all these effects, we use the maximum likelihood approach (e.g., Saglia et al. 2001) described in Paper II. This method is the natural choice given that the joint distribution of $M \equiv -2.5 \log L$, $R \equiv \log R_o$, and $V \equiv \log \sigma$ is quite well described by a multivariate Gaussian. The maximum likelihood estimates of the mean values of these variables and the parameters of the covariance matrix \mathcal{C} that describes the correlations between these variables are shown in Table 1 of Paper II. What remains is to write down how the covariance matrix changes when we change variables from (M, R, V) to (μ, R, V) . Because $(\mu_o - \mu_*) \equiv (M - M_*) + 5(R - R_*)$, the covariance matrix becomes

$$\begin{aligned} \mathcal{F} &\equiv (\lambda_1, \lambda_2, \lambda_3), \\ \lambda_1 &= \begin{pmatrix} \sigma_M^2 + 10\sigma_M\sigma_R\rho_{RM} + 25\sigma_R^2 \\ \sigma_R\sigma_M\rho_{RM} + 5\sigma_R^2 \\ \sigma_V\sigma_M\rho_{VM} + 5\sigma_R\sigma_V\rho_{RV} \end{pmatrix}, \\ \lambda_2 &= \begin{pmatrix} \sigma_R\sigma_M\rho_{RM} + 5\sigma_R^2 \\ \sigma_R^2 \\ \sigma_R\sigma_V\rho_{RV} \end{pmatrix}, \\ \lambda_3 &= \begin{pmatrix} \sigma_V\sigma_M\rho_{VM} + 5\sigma_R\sigma_V\rho_{RV} \\ \sigma_R\sigma_V\rho_{RV} \\ \sigma_V^2 \end{pmatrix}, \end{aligned} \quad (5a)$$

TABLE 1
MAXIMUM LIKELIHOOD ESTIMATES OF THE JOINT DISTRIBUTION OF LUMINOSITIES, SIZES, AND VELOCITY DISPERSIONS

Band	N_{gals}	μ_*	σ_μ	R_*	σ_R	V_*	σ_V	$\rho_{R\mu}$	$\rho_{V\mu}$	ρ_{RV}	Q
g^*	5825	20.74	0.654	0.520	0.254	2.197	0.113	0.801	0.005	0.536	1.15
r^*	8228	19.87	0.610	0.490	0.241	2.200	0.111	0.760	0.000	0.543	0.85
i^*	8022	19.40	0.600	0.465	0.241	2.201	0.110	0.753	-0.001	0.542	0.75
z^*	7914	18.99	0.604	0.450	0.241	2.200	0.110	0.759	-0.001	0.543	0.60

NOTES.—The mean values of the variables at redshift z are μ_* , Qz , R_* , V_* , and the elements of the covariance matrix \mathcal{F} defined by the various pairwise correlations between the variables are shown. These coefficients were obtained from the coefficients of the covariance matrix \mathcal{C} shown in Table 1 in Paper II.

so

$$\mathcal{F} \equiv \begin{pmatrix} \sigma_\mu^2 & \sigma_R \sigma_\mu \rho_{R\mu} & \sigma_V \sigma_\mu \rho_{V\mu} \\ \sigma_R \sigma_\mu \rho_{R\mu} & \sigma_R^2 & \sigma_R \sigma_V \rho_{RV} \\ \sigma_V \sigma_\mu \rho_{V\mu} & \sigma_R \sigma_V \rho_{RV} & \sigma_V^2 \end{pmatrix}. \quad (5b)$$

The coefficients of \mathcal{F} are given in Table 1; they were obtained by inserting the values shown in Table 1 of Paper II into the first of the equalities above. Note that $\mu = -2.5 \log I_o$.

This matrix is fundamentally useful because it describes the intrinsic correlations between the sizes, surface brightnesses, and velocity dispersions of early-type galaxies—the effects of how the sample was selected and observational errors have been accounted for. For example, the coefficients in the top right and bottom left of \mathcal{F} are very close to zero, indicating that surface brightness and velocity dispersion are uncorrelated. In addition, the eigenvalues and vectors of \mathcal{F} give information about the shape and thickness of the FP. For example, in r^* , the eigenvalues are 0.639, 0.179, and 0.052; the smallest eigenvalue is considerably smaller than the other two indicating that, when viewed in the appropriate projection, the plane is quite thin. The associated eigenvector gives the coefficients of the “orthogonal” fit, and the rms scatter around this orthogonal fit is given by the (square root of the) smallest eigenvalue (e.g., Saglia et al. 2001).

If we wish to use the FP as a distance indicator, then we are more interested in finding those coefficients that minimize the scatter in R_o . This means that we would like to find that pair (a , b) that minimizes $\langle \Delta_1^2 \rangle$, where Δ_1 is given by equation (3). A little algebra shows that the solution is given by inserting the maximum likelihood estimates of the scatter in surface brightnesses, sizes, and velocity dispersions into equation (4).

The maximum likelihood \mathcal{F} can be used to provide estimates of the direct and orthogonal fit coefficients, as well as the intrinsic scatter around the mean relations (orthogonal to the plane, as well as in the direction of R_o). These are given in Table 2. Although b is approximately the same both for the “orthogonal” and the “direct” fits, a from the direct fit is always about 25% smaller than from the orthogonal fit. In either case, note how similar a and b are in all four bands. This similarity, and the fact that the thickness of the FP decreases slightly with increasing wavelength, can be used to constrain models of how different stellar populations (which may contribute more or less to the different bands) are distributed in early-type galaxies. If the direct fit is used as a distance indicator, then the thickness of the FP

translates into an uncertainty in derived distances of about 20%.

Table 2 also shows results from the more traditional χ^2 -fitting techniques, which were obtained as follows. (These fits were not weighted by errors, and the intrinsic scatter with respect to the fits was estimated by subtracting the measurement errors in quadrature from the observed scatter.) Ignoring evolution and selection effects when minimizing $\langle \Delta_1^2 \rangle$ and $\langle \Delta_o^2 \rangle$ results in coefficients a that are about 10% larger than those we obtained from the maximum likelihood method. We have not shown these in the table for the following reason. If the population at high redshift is more luminous than that nearby, as expected if the evolution is passive, then the higher redshift population would have systematically smaller values of μ_o . Since the higher redshift population makes up most of the large R_o part of our sample, this could make the plane appear steeper, i.e., it could cause the best-fit a to be biased to a larger value. If we use the maximum likelihood estimate of how the luminosities brighten with redshift, then we can subtract off the brightening from μ_o before minimizing $\langle \Delta_1^2 \rangle$ and $\langle \Delta_o^2 \rangle$. This reduces the best-fit value of a so that it is closer to that of the maximum likelihood method. The coefficients obtained in this way are labeled “ χ^2 -Evolution” in Table 2; they are statistically different from the maximum likelihood estimates, presumably because they do not account for selection effects or for the effects of observational errors. If we weight each galaxy by the inverse of $S(z|M_*, Q)$ (the selection function defined in Paper II) when minimizing, then this should at least partially account for selection effects. The resulting estimates of a , b , and c are labeled “ χ^2 -Evolution Selection effects” in Table 2. The small remaining difference between these and the maximum likelihood estimates is likely due to the fact that the likelihood analysis accounts more consistently for errors.

Figure 1 shows the FP in the four SDSS bands. We have chosen to present the plane using the coefficients, obtained using the maximum likelihood method, which minimize the scatter orthogonal to the plane. (In all cases, the evolution of the luminosities has been subtracted from the surface brightnesses.) The results to follow regarding the shape of the FP, and estimates of how the mean properties of early types depend on redshift and environment, are independent of which fits we use. A fair number of the galaxies in our sample have velocity dispersion measurements with small signal-to-noise ratio (S/N; see, e.g., Fig. 19 of Paper I). The FP is relatively insensitive to these objects: removing objects with $S/N < 15$ had little effect on the best-fit values of a , b . Removing objects with small axis ratios also had little effect on the maximum likelihood coefficients.

TABLE 2
COEFFICIENTS OF THE FP IN THE FOUR SDSS BANDS

Band	a	b	c	$\text{rms}_{\text{orth}}^{\text{int}}$	$\text{rms}_{R_o}^{\text{int}}$
Orthogonal Fits					
Maximum Likelihood:					
g^*	1.45 ± 0.06	-0.74 ± 0.01	-8.779 ± 0.029	0.056	0.100
r^*	1.49 ± 0.05	-0.75 ± 0.01	-8.778 ± 0.020	0.052	0.094
i^*	1.52 ± 0.05	-0.78 ± 0.01	-8.895 ± 0.021	0.049	0.091
z^*	1.51 ± 0.05	-0.77 ± 0.01	-8.707 ± 0.023	0.049	0.089
χ^2 -Evolution Selection Effects:					
g^*	1.43 ± 0.06	-0.78 ± 0.01	-9.057 ± 0.032	0.058	0.101
r^*	1.45 ± 0.05	-0.76 ± 0.01	-8.719 ± 0.020	0.052	0.094
i^*	1.48 ± 0.05	-0.77 ± 0.01	-8.699 ± 0.024	0.050	0.090
z^*	1.48 ± 0.05	-0.77 ± 0.01	-8.577 ± 0.025	0.049	0.089
χ^2 -Evolution:					
g^*	1.35 ± 0.06	-0.77 ± 0.01	-8.820 ± 0.033	0.058	0.100
r^*	1.40 ± 0.05	-0.77 ± 0.01	-8.678 ± 0.023	0.053	0.092
i^*	1.41 ± 0.05	-0.78 ± 0.01	-8.688 ± 0.024	0.050	0.090
z^*	1.41 ± 0.05	-0.78 ± 0.01	-8.566 ± 0.026	0.048	0.089
Direct Fits					
Maximum Likelihood:					
g^*	1.08 ± 0.05	-0.74 ± 0.01	-8.033 ± 0.024	0.061	0.092
r^*	1.17 ± 0.04	-0.75 ± 0.01	-8.022 ± 0.020	0.056	0.088
i^*	1.21 ± 0.04	-0.77 ± 0.01	-8.164 ± 0.018	0.053	0.085
z^*	1.20 ± 0.04	-0.76 ± 0.01	-7.995 ± 0.021	0.053	0.084
χ^2 -Evolution Selection Effects:					
g^*	1.05 ± 0.05	-0.79 ± 0.01	-8.268 ± 0.026	0.063	0.094
r^*	1.12 ± 0.04	-0.76 ± 0.01	-7.932 ± 0.020	0.057	0.088
i^*	1.14 ± 0.04	-0.76 ± 0.01	-7.904 ± 0.019	0.054	0.085
z^*	1.14 ± 0.04	-0.76 ± 0.01	-7.784 ± 0.021	0.053	0.084
χ^2 -Evolution:					
g^*	0.99 ± 0.05	-0.76 ± 0.01	-7.921 ± 0.026	0.065	0.093
r^*	1.06 ± 0.04	-0.75 ± 0.01	-7.775 ± 0.020	0.059	0.088
i^*	1.09 ± 0.04	-0.77 ± 0.01	-7.823 ± 0.018	0.056	0.085
z^*	1.09 ± 0.04	-0.78 ± 0.01	-7.818 ± 0.020	0.053	0.083

NOTES.—For each set of coefficients, the scatter orthogonal to the plane and in the direction of R_o are also given.

In principle, the likelihood analysis provides an estimate of the error on each of the derived coefficients. However, this estimate assumes that the parametric Gaussian form is indeed a good fit. Although we present evidence in Paper II that the Gaussian form is indeed good, we emphasize that, when the data set is larger a nonparametric fit should be performed. Therefore, we have estimated errors on the numbers quoted in Table 2 as follows. The large size of our sample allows us to construct many random subsamples, each of which is substantially larger than most of the samples available in the literature. Estimating the elements of the covariance matrix presented in Table 1, and then transforming to get the FP coefficients in Table 2, in each of these subsamples provides an estimate of how well we have determined a , b , and c . (Note that the errors we find in this way are comparable to those sometimes quoted in the literature, even though each of the subsamples we generated is an order of magnitude larger than any sample available in the literature.) Because each subsample contains fewer galaxies than our full sample, this procedure is likely to provide an overestimate of the true formal error for our sample. However, the formal error does not account for the uncertainties in our K -corrections and velocity dispersion aperture corrections (discussed in more detail in Paper I), so an overestimate is probably more realistic.

As a check on the relative roles of evolution and selection effects, we simulated complete and magnitude-limited samples (with a velocity dispersion cut) following the procedures outlined in Appendix A of Paper II. We then estimated the coefficients of the FP in the simulated catalogs using the different methods. The results are summarized in Table 3. When applied to the complete simulations, the χ^2 -minimization method yields estimates of a that are biased high; it yields the input FP coefficients only after evolution has been subtracted from the surface brightnesses. However, in the magnitude-limited simulations, once evolution has been subtracted, it provides an estimate of a that is biased low, unless selection effects are also accounted for. Note that this is similar to what we found with the data. The maximum likelihood method successfully recovers the same intrinsic covariance matrix and evolution as the one used to generate the simulations, both for the complete and the magnitude-limited mock catalogs, and so it recovers the same correct coefficients for the FP in both cases. (We have not shown these estimates in the table.)

A selection of results from the literature is presented in Table 4. Many of these samples were constructed by combining new measurements with previously published photometric and velocity dispersion measurements, often made by other authors. (Exceptions are Jørgensen, Franx, &

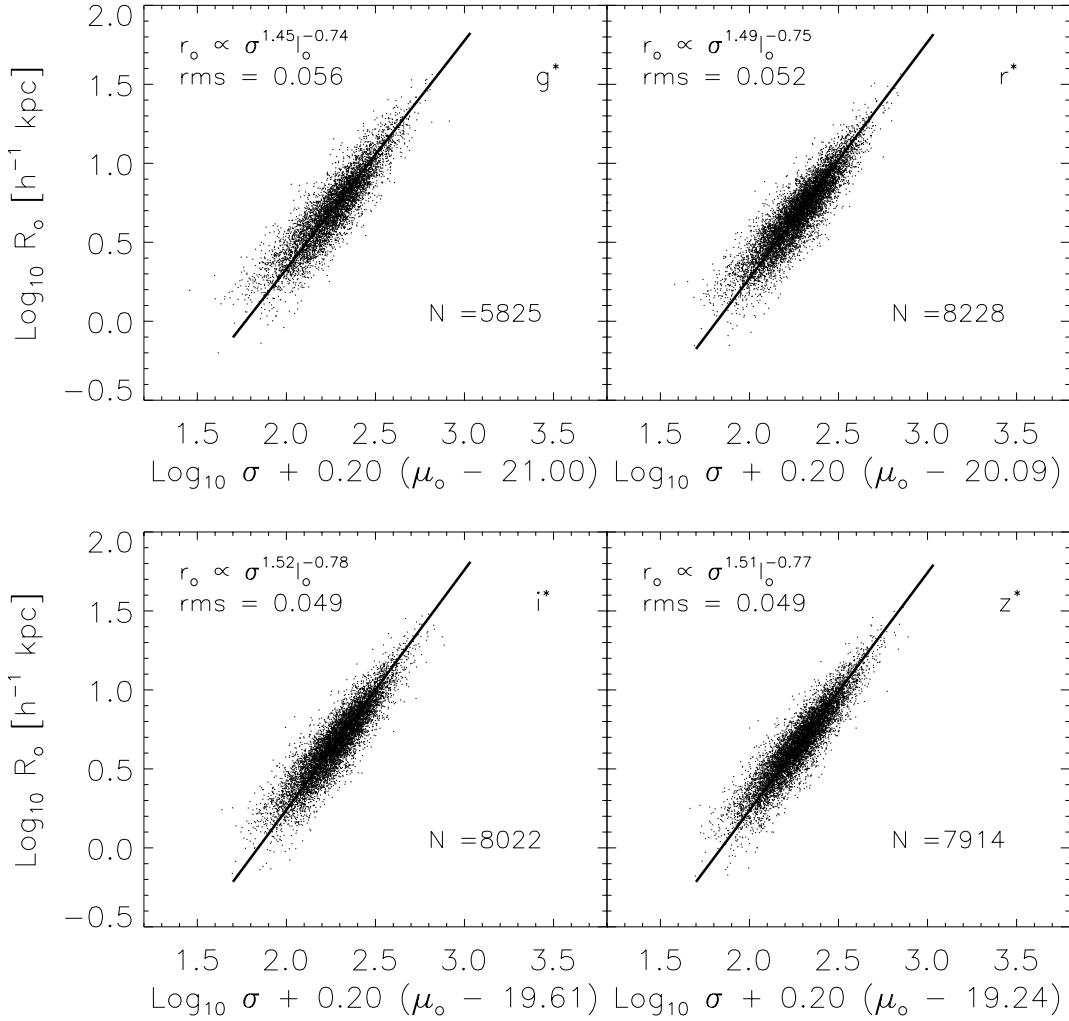


FIG. 1.—FP in the four SDSS bands. Coefficients shown are those that minimize the scatter orthogonal to the plane, as determined by the maximum likelihood method. Surface brightnesses have been corrected for evolution.

TABLE 3
COEFFICIENTS OF THE FP IN THE COMPLETE AND MAGNITUDE-LIMITED SIMULATED CATALOGS

Band	a	b	c	rms_{orth}	rms_{R_o}
Orthogonal Fits					
Complete:					
g^*	1.44 ± 0.05	-0.74 ± 0.01	-8.763 ± 0.028	0.056	0.100
r^*	1.48 ± 0.05	-0.75 ± 0.01	-8.722 ± 0.020	0.052	0.094
Magnitude-Limited:					
g^*	1.39 ± 0.06	-0.74 ± 0.01	-8.643 ± 0.028	0.056	0.100
r^*	1.43 ± 0.05	-0.76 ± 0.01	-8.721 ± 0.021	0.052	0.093
Direct Fits					
Complete:					
g^*	1.09 ± 0.04	-0.74 ± 0.01	-7.992 ± 0.023	0.061	0.091
r^*	1.16 ± 0.04	-0.75 ± 0.01	-8.005 ± 0.020	0.056	0.088
Magnitude-Limited:					
g^*	1.04 ± 0.05	-0.74 ± 0.01	-7.817 ± 0.025	0.061	0.090
r^*	1.11 ± 0.04	-0.75 ± 0.01	-7.895 ± 0.020	0.056	0.087

NOTES.—Coefficients of the FP in the complete and magnitude-limited simulated catalogs have been obtained by minimizing a χ^2 in which evolution in the surface brightnesses has been removed and which weights objects by the inverse of the selection function.

TABLE 4
SELECTION OF FP COEFFICIENTS FROM THE LITERATURE

Source	Band	N_{gal}	a	b	Δ_{R_o}	Fit Method
Dressler et al. 1987	B	97	1.33 ± 0.05	-0.83 ± 0.03	20%	Inverse
Lucey, Bower, & Ellis 1991.....	B	26	1.27 ± 0.07	-0.78 ± 0.09	13%	Inverse
Guzmán, Lucey, & Bower 1993.....	V	37	1.14	-0.79	17%	Direct
Kelson et al. 2000	V	30	1.31 ± 0.13	-0.86 ± 0.10	14%	Orthogonal
Djorgovski & Davis 1987	r_G	106	1.39 ± 0.14	-0.90 ± 0.09	20%	Two-step inverse
Jørgensen et al. 1996.....	r	226	1.24 ± 0.07	-0.82 ± 0.02	19%	Orthogonal
Hudson et al. 1997.....	R	352	1.38 ± 0.04	-0.82 ± 0.03	20%	Inverse
Gibbons, Fruchter, & Bothun 2001	R	428	1.37 ± 0.04	-0.825 ± 0.01	20%	Inverse
Colless et al. 2001	R	255	1.22 ± 0.09	-0.84 ± 0.03	20%	ML
Scodreggio 1997	I	294	1.55 ± 0.05	-0.80 ± 0.02	22%	Orthogonal
Pahre et al. 1998	K	251	1.53 ± 0.08	-0.79 ± 0.03	21%	Orthogonal

Kjærgaard 1996, Scodreggio 1997, and Colless et al. 2001.) With respect to previous samples, the SDSS sample presented here is both extremely large and homogeneous.

Note the relatively large spread in published values of a , and the fact that a is larger at longer wavelengths. In contrast, the FP we obtain in this paper is remarkably similar in all wave bands—although our value of b is consistent with those in the literature, the value of a we find in all wave bands is close to the largest published values. In addition, the eigenvectors of our covariance matrix satisfy the same relations presented by Saglia et al. (2001). Namely, $\hat{v}_1 = \hat{R}_o - a\hat{V} - b\hat{I}_o$, $\hat{v}_2 \approx -\hat{R}_o/b - \hat{V}(1 + b^2)/(ab) + \hat{I}_o$, and $\hat{v}_3 \approx \hat{R}_o + \hat{I}_o/b$. In addition, when used as a distance indicator, the FP we find is as accurate as most of the samples containing more than ~ 100 galaxies in the literature. Unfortunately, at the present time, we have no galaxies in common with those in any of the samples listed in Table 4, so it is difficult to say why our FP coefficients appear to show so little dependence on wavelength, or why a is higher than it is in the literature.

The fact that $a \neq 2$ means that the FP is tilted relative to the simplest virial theorem prediction $R_o \propto \sigma^2/I_o$. One of the assumptions of this simplest prediction is that the kinetic energy that enters the virial theorem is proportional to the square of the observed central velocity dispersion. Busarello et al. (1997) argue that, in fact, the kinetic energy is proportional to $\sigma^{1.6}$ rather than to σ^2 . Since this is close to the $\sigma^{1.5}$ scaling we see, it would be interesting to see if the kinetic energy scales with σ for the galaxies in our sample similarly to how it does in Busarello et al.'s sample. This requires measurements of the velocity dispersion profiles of (a subsample of) the galaxies in our sample and has yet to be done.

Correlations between pairs of observables, such as the Faber & Jackson (1976) relation between luminosity and velocity dispersion, and the Kormendy (1977) relation between the size and the surface brightness, can be thought of as projections of the FP. They are studied in Paper II. The κ -space projection of Bender et al. (1992) is presented in § 2.6 below.

2.2. Residuals and the Shape of the FP

Once the FP has been obtained, there are at least two definitions of its thickness that are of interest. If the FP is to be used as a distance indicator, then the quantity of interest is the scatter around the relation in the R_o direction only. On the other hand, if the FP is to be used to constrain models of

stellar evolution, then one is more interested in the scatter orthogonal to the plane. We discuss both of these below.

The thickness of the FP is some combination of residuals that are intrinsic and those coming from measurement errors. We would like to verify that the thickness is not dominated by measurement errors. The residuals from the FP in the different bands are highly correlated; a galaxy that scatters above the FP in g^* also scatters above the FP in, say, z^* . Although the errors in the photometry in the different bands are not completely independent, this suggests that the scatter around the FP has a real, intrinsic component. It is this intrinsic thickness that the maximum likelihood analysis is supposed to have estimated. The intrinsic scatter may be somewhat smaller than the maximum likelihood estimates because there is a contribution to the scatter that comes from our assumption that all early-type galaxies are identical when we apply the K -correction, for which we have not accounted.

All our estimates of the scatter around the FP show that the FP appears to become thicker at shorter wavelengths. Presumably, this is because the light in the redder bands, being less affected by recent star formation and extinction by dust, is a more faithful tracer of the dynamical state of the galaxy. The orthogonal scatter in our sample, which spans a wide range of environments, is comparable to the values given in the literature obtained from cluster samples (e.g., Pahre, Djorgovski, & de Carvalho 1998); this constrains models of how the stellar populations of early-type galaxies depend on environment. If the direct fit to the FP is used as a distance indicator, then the intrinsic scatter introduces an uncertainty in distance estimates of $\sim \ln(10) \times 0.09 \sim 20\%$.

Our next step is to check that the FP really is a plane, and not, for example, a saddle. To do this, we should show the residuals from the orthogonal fit as a function of distance along the long axis of the plane. Specifically, if $X \equiv \log \sigma + (b/a) \log I_o + (c/a)$, then

$$\begin{aligned}
 X_{\text{FP}} &\equiv X \sqrt{1 + a^2} + (\log R_o - aX) \frac{a}{\sqrt{1 + a^2}} \\
 &= \frac{X + a \log R_o}{\sqrt{1 + a^2}}, \quad (6)
 \end{aligned}$$

and we would like to know if the residuals Δ_o defined earlier correlate with X_{FP} . A scatter plot of these residuals versus X_{FP} is shown in Figure 2 (we have first subtracted off the weak evolution in the surface brightnesses). The symbols

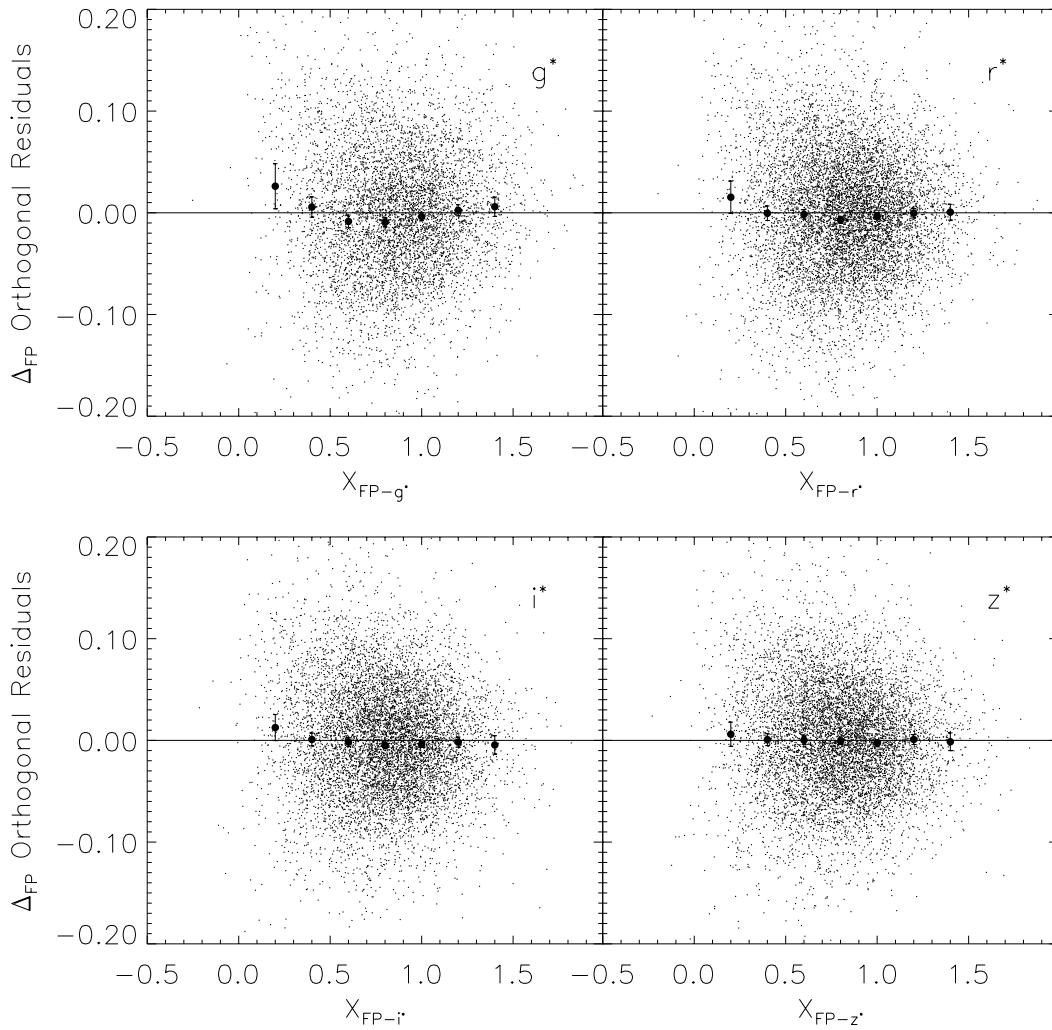


FIG. 2.—Residuals orthogonal to the maximum likelihood FP fit as a function of distance along the fit (the long axis of the plane). Error bars show the mean ± 3 times the error in the mean in each bin. Galaxies with low/high velocity dispersions populate the top left/bottom right of each panel, but the full sample shows little curvature.

superposed on the scatter plot show the mean value of the residuals and ± 3 times the error in the mean, for a few small bins along X_{FP} . The figure shows that the FP is reasonably flat; it is slightly more warped in the shorter wavelengths.

Given that the FP is not significantly warped, we would like to know if deviations from the plane correlate with any of the three physical parameters used to define it. When the plane is defined by minimizing with respect to $\log R_o$, there is little if any correlation of the residuals with absolute magnitude, surface brightness, effective radius, axis ratio, velocity dispersion, or color, so we have chosen to not present them here. Instead, Figure 3 shows the result of plotting the residuals orthogonal to the plane when the plane is defined by the orthogonal fit. The residuals show no correlation with M , μ_o , $\log R_o$, or axis ratio (we have subtracted the weak evolution in M and μ_o when making the scatter plots). The residuals are anticorrelated with $\log \sigma$ and slightly less anticorrelated with $g^* - r^*$ color. The correlation with color is due to the fact that velocity dispersion and color are tightly correlated (this correlation is studied in more detail in Paper IV). The correlation with velocity dispersion is not a selection effect, nor is it associated with evolution; we see a

similar trend with velocity dispersion in both the complete and the magnitude-limited simulated catalogs.

Figure 4 shows why this happens. The four panels show the FP in four subsamples of the full r^* sample, divided according to velocity dispersion. Note how the different scatter plots in Figure 4 show sharp cutoffs approximately perpendicular to the x -axis: lines of constant σ are approximately perpendicular to the x -axis. Whereas the direct fit is not affected by a cutoff that is perpendicular to the x -axis, the orthogonal fit is. Hence, the residuals with respect to the orthogonal fit show a correlation with velocity dispersion, whereas those from the direct fit do not. (Indeed, by using the coefficients provided in Tables 1 and 2, and the definition of the residuals Δ_1 and Δ_o , one can show that $\langle \Delta_1 | \log \sigma \rangle$ is proportional to $\log \sigma$, with a constant of proportionality that is close to zero when the parameters for the direct fit are inserted. However, when the parameters for the orthogonal fit are used, then the slope of the $\langle \Delta_o | \log \sigma \rangle$ versus $\log \sigma$ relation is significantly different from zero.)

To illustrate, the solid curves in Figure 4 (the same in each panel) show the maximum likelihood FP for the full sample. The dashed curves show the FP, determined by using the

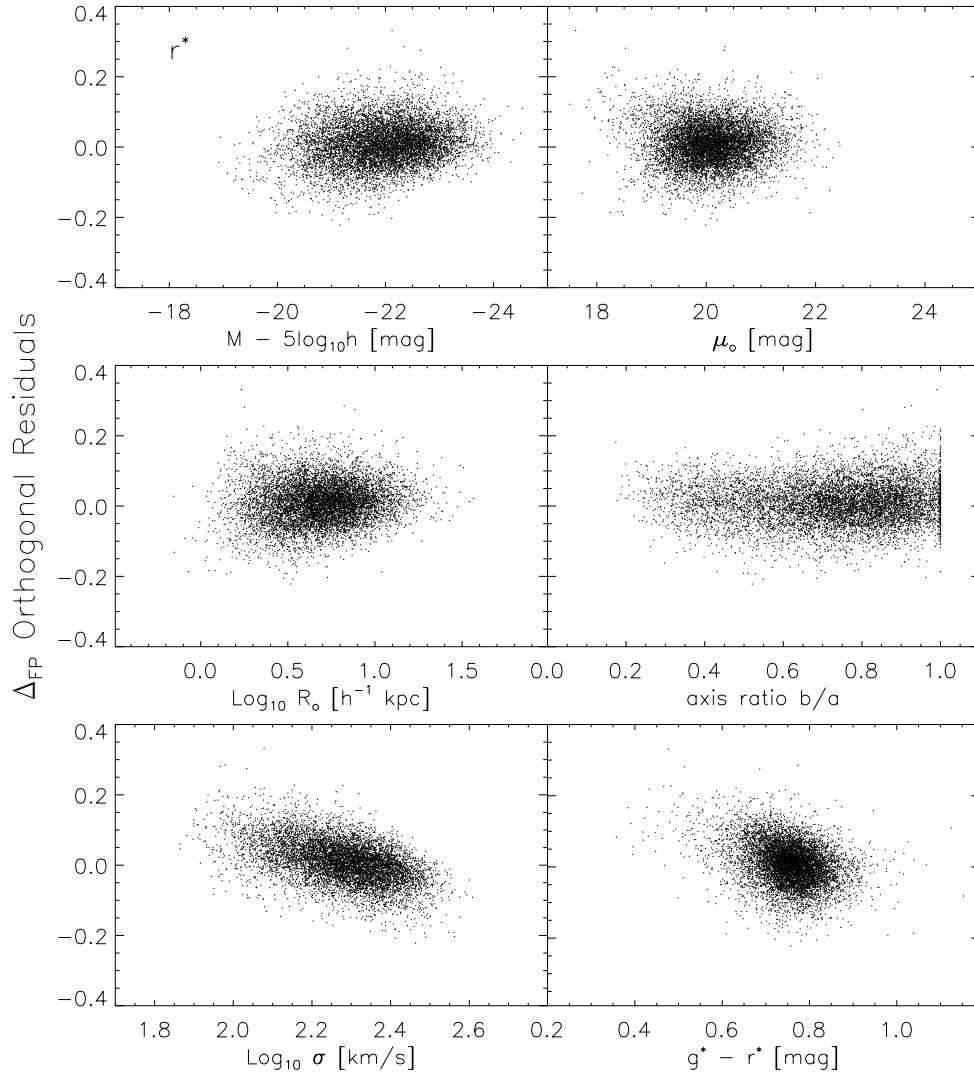


FIG. 3.—Residuals orthogonal to the FP in r^* vs. absolute magnitude M , surface brightness μ_o , effective radius $\log R_o$, axis ratio b/a , velocity dispersion $\log \sigma$, and $g^* - r^*$ color. Note the absence of correlation with all parameters other than velocity dispersion and color.

χ^2 -method to minimize the residuals orthogonal to the plane, in various subsamples defined by velocity dispersion. The panels for larger velocity dispersions show steeper relations. Evidence for a steepening of the relation with increasing velocity dispersion was seen by Jørgensen et al. (1996). Their sample was considerably smaller than ours, and so they ruled the trend they saw as only marginal. Our much larger sample shows this trend clearly. We have already argued that this steepening is an artifact of the fact that lines of constant velocity dispersion are perpendicular to the x -axis. The maximum likelihood fit to the subsamples is virtually the same as that for the full sample, provided we include the correct velocity dispersion cuts in the normalization of the likelihood. In other words, the maximum likelihood fit is able to account for the bias introduced by making a cut in velocity dispersion, as well as apparent magnitude.

2.3. The Mass-to-Light Ratio

The FP is sometimes used to infer how the mass-to-light ratio depends on different observed or physical parameters. For example, the scaling required by the virial theorem, $M_o \propto R_o \sigma^2$, combined with the assumption that the mass-

to-light ratio scales as $M_o/L \propto M_o^\gamma$ yields an FP-like relation of the form

$$R_o \propto \sigma^{2(1-\gamma)/(1+\gamma)} I_o^{-1/(1+\gamma)}. \quad (7)$$

The observed FP is $R_o \propto \sigma^a I_o^b$. If the relation above is to describe the observations, then γ must simultaneously satisfy two relations: $\gamma = (2 - a)/(2 + a)$ and $\gamma = -(1 + b)/b$. The values of b in the literature are all about -0.8 ; setting γ equal to the value required by b and then writing a in terms of b gives $a = -2(1 + 2b)$. Most of the values of a and b in the shorter wave bands reported in the literature (see, e.g., Table 4) are consistent with this scaling, whereas the higher values of a found at longer wavelengths are not. Although the direct fits to our sample have small values of a , the orthogonal fits give high values in all four bands. These fits do not support the assumption that M_o/L can be parametrized as a function of M_o alone.

Another way to phrase this is to note that, when combined with the virial theorem requirement that $(M_o/L) \propto \sigma^2/(R_o I_o)$, the FP relation $R_o \propto \sigma^a I_o^b$ yields

$$(M_o/L)_{\text{FP}} \propto \sigma^{2+a/b} R_o^{-(1+b)/b} \quad (8)$$

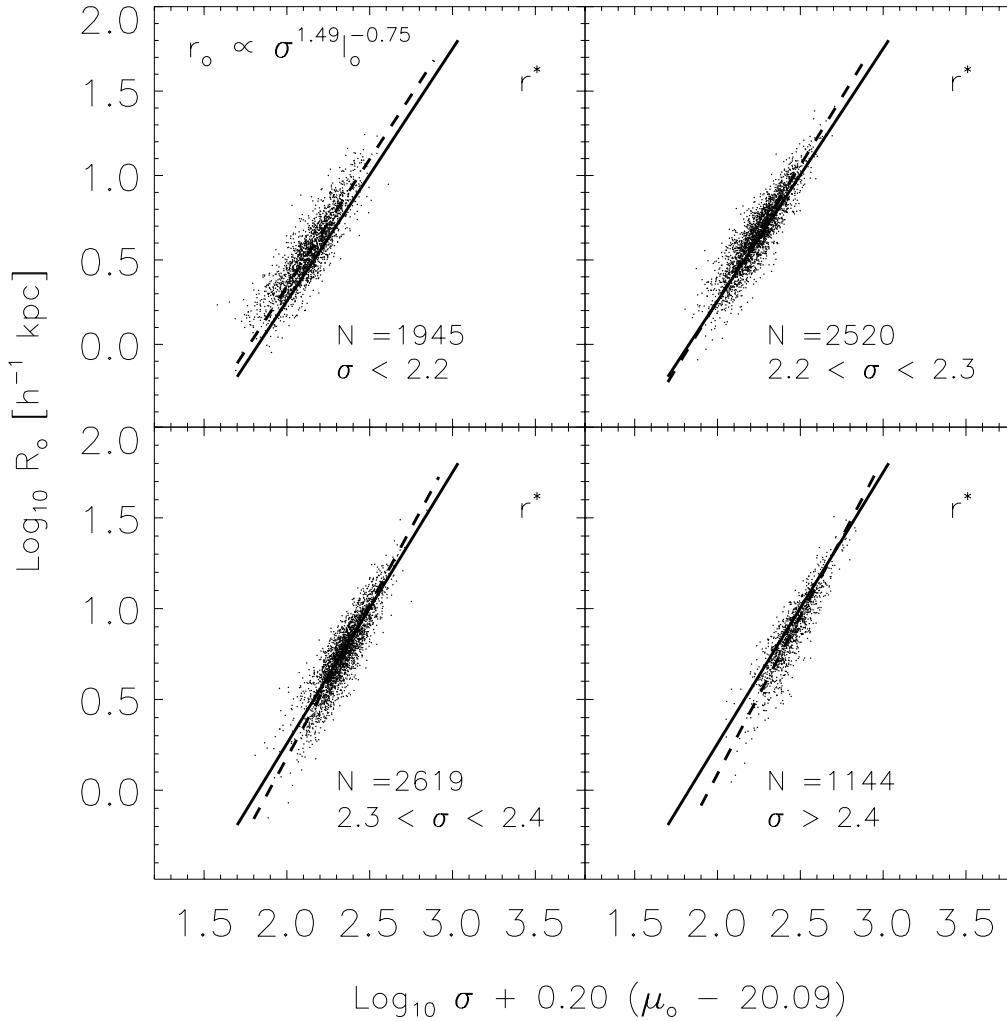


FIG. 4.—FP in four subsamples defined by velocity dispersion. Solid curve (same in all four panels) shows the maximum likelihood relation of the parent r^* sample and dashed lines show the best fit, obtained by minimizing the residuals orthogonal to the plane, using only the galaxies in each subsample. The slope of the minimization fit increases with increasing velocity dispersion, whereas maximum likelihood fits to the subsamples, which account for the cut on σ , give the same slope as for the full sample.

(e.g., Jørgensen et al. 1996; Kelson et al. 2000). The quantity on the right-hand side is the mass-to-light ratio “predicted” by the FP, if σ and R_o are given and the scatter in the FP is ignored. This is a function of M_o alone only if $a = -2(1 + 2b)$. Our orthogonal fit coefficients a and b are not related in this way. Rather, for our FP, the dependence on σ in equation (8) cancels out almost exactly: to a very good approximation, we find $(M_o/L)_{FP} \propto R_o^{-(1+b)/b} \propto R_o^{0.33}$.

Alternatively, a little algebra shows that mass-to-light ratio is determined by the combination $(\sigma^2/I_o)^{0.25}$. Whether there is a simple physical reason for this is an open question.

In contrast to the predicted ratio, $(M_o/L)_{FP}$, the combination $R_o\sigma^2/L$ is the “observed” mass-to-light ratio. The ratio of the observed value to the FP prediction of equation (8) is $(R_o/I_o^b\sigma^a)^{1/b}$. The scatter in the logarithm of this ratio is $1/b$ times the scatter in FP in the direction of R_o (i.e., it is the scatter in the quantity we called Δ_1 in the previous subsections, divided by b). Inserting the values from Table 2 shows that if the values of σ and the effective radius in r^* are used to predict the values of the mass-to-light ratio in r^* , then the uncertainty in the predicted ratio is 26%. This is larger than

the values quoted in the literature for early-type galaxies in clusters (e.g., Jørgensen et al. 1996; Kelson et al. 2000).

Unfortunately, this is somewhat confusing terminology, because the two mass-to-light ratios are not proportional to each other. This can be seen by using the maximum likelihood results of Table 1 to compute the mean of the observed mass to light ratio $R_o\sigma^2/L$ at fixed predicted $(M/L)_{FP}$, or simply by plotting the two quantities against one another. Figure 5 shows how $R_o\sigma^2/L$ correlates with luminosity, mass $R_o\sigma^2$, velocity dispersion, surface brightness, the ratio predicted by the FP, and color. The different panels show obvious correlations; the maximum likelihood predictions for these correlations can be derived from the coefficients in Table 1: $(R_o\sigma^2/L) \propto L^{0.14 \pm 0.02}$, $(R_o\sigma^2/L) \propto (R_o\sigma^2)^{0.22 \pm 0.05}$, $(R_o\sigma^2/L) \propto \sigma^{0.84 \pm 0.1}$, and $(R_o\sigma^2/L) \propto R_o^{0.27 \pm 0.06}$. These are shown as dashed lines in the top four panels. A linear fit to the scatter plot in the bottom left panel gives $(R_o\sigma^2/L) \propto (M/L)_{FP}^{0.80 \pm 0.05}$, with an rms scatter around the fit of 0.14: the ratio predicted by the FP is not proportional to the observed ratio. A scatter plot of $(M/L)_{FP}$ against all these quantities is tighter, of course (recall the scatter around the FP has been removed), although some of

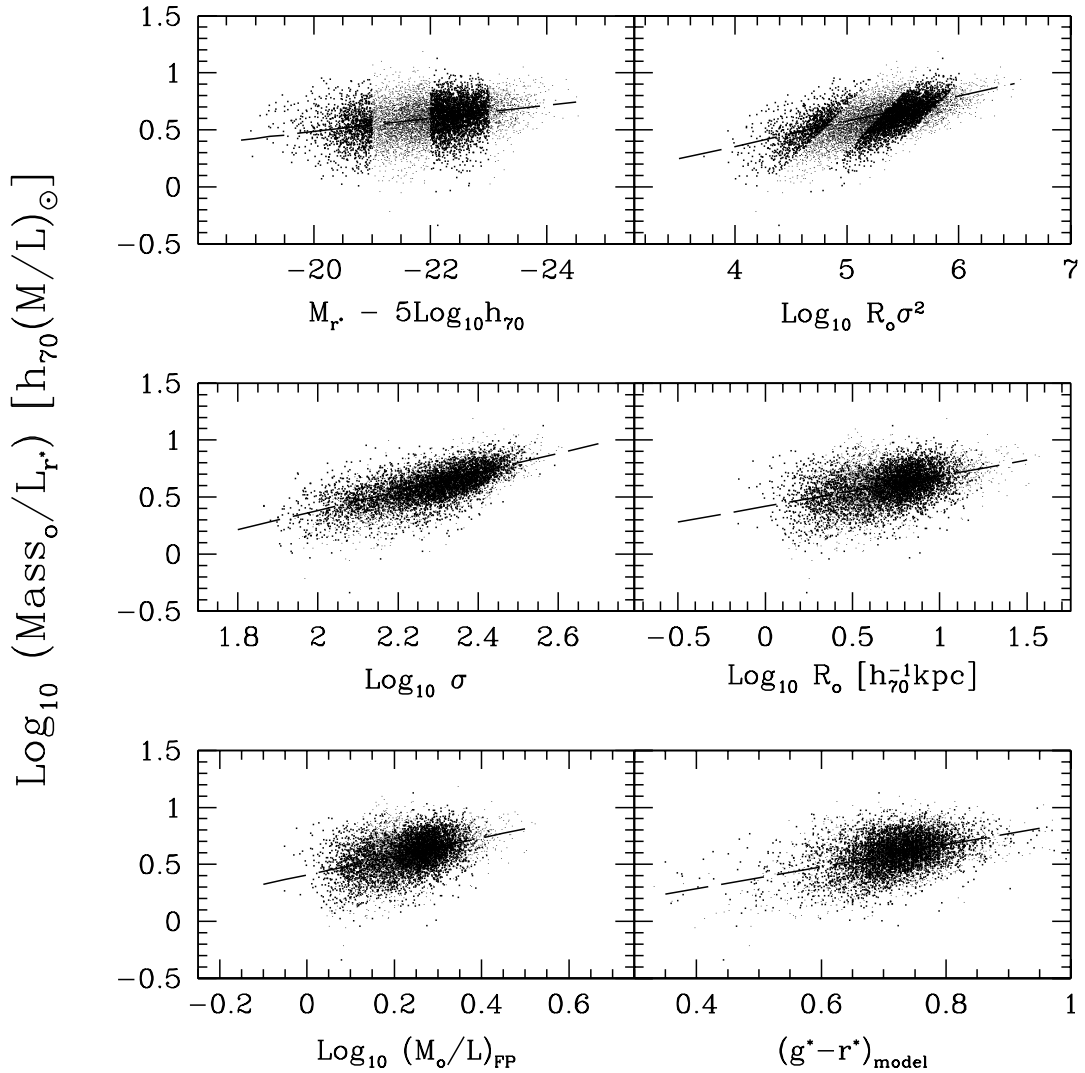


FIG. 5.—Ratio of effective mass $R_o\sigma^2$ to effective luminosity ($L/2$) as a function of luminosity (top left), mass (top right), velocity dispersion (middle left), surface brightness (middle right), the combination of velocity dispersion and size suggested by the FP (bottom left), and color (bottom right). Note the substantial scatter around the best fit linear relation in the bottom left panel, the slope of which is shallower than unity.

the slopes are significantly different. For example, $(M/L)_{FP} \propto L^{0.16 \pm 0.04}$, $(M/L)_{FP} \propto (R_o\sigma^2)^{0.13 \pm 0.03}$, and $(M/L)_{FP} \propto \sigma^{0.21 \pm 0.03}$: the “observed” and “predicted” slopes of the mass-to-light ratio versus σ relations are very different. For this reason, one should be careful in interpreting what is meant by the “predicted” mass-to-light ratio. Our own view is that the observed ratio, $R_o\sigma^2/L$ is to be preferred, as it is directly related to observables, and is independent of the fitting procedure used to fit the FP.

2.4. The FP: Evidence for Evolution?

The FP is sometimes used to test for evolution. This is done by plotting R_o versus the combination of μ_o and σ , which defines the FP at low redshift, and then seeing if the high-redshift population traces the same locus as the low-redshift population. Figure 6 shows this test for our g^* -band sample: solid lines (same in each panel) show the relation that fits the zero-redshift sample; dashed lines show a line with the same slope that best fits the higher redshift sample. The population at higher redshift is displaced slightly to the left of the low-redshift population; the text in the bottom of

each panel shows this shift, expressed as a change in the surface brightness μ_o . The plot appears to show that, on average, the higher redshift galaxies are brighter, with the brightening scaling approximately as $\Delta\mu_o \approx -2z$.

How much of this apparent brightening is really due to evolution, and how much is an artifact of the fact that our sample is magnitude-limited? To address this, we generated complete and magnitude-limited mock galaxy catalogs, as described in Appendix A of Paper II, and then performed the same test for evolution. Comparing the shifts in the two simulations allows us to estimate how much of the shift is due to the selection effect. Figure 7 shows the results in our simulated g^* (left) and r^* (right) catalogs. The solid lines in each panel show the zero-redshift relation, and the dotted and dashed lines show lines of the same slope that best fit the points at low and high redshift, respectively. The text in the bottom shows how much of the shift in μ_o is due to the magnitude limit, and how much to evolution. The sum of the two contributions is the total shift seen in the magnitude-limited simulations. Note that this sum is similar to that seen in the data (Fig. 6), both at low and high redshifts, suggesting that our simulations describe the varying roles

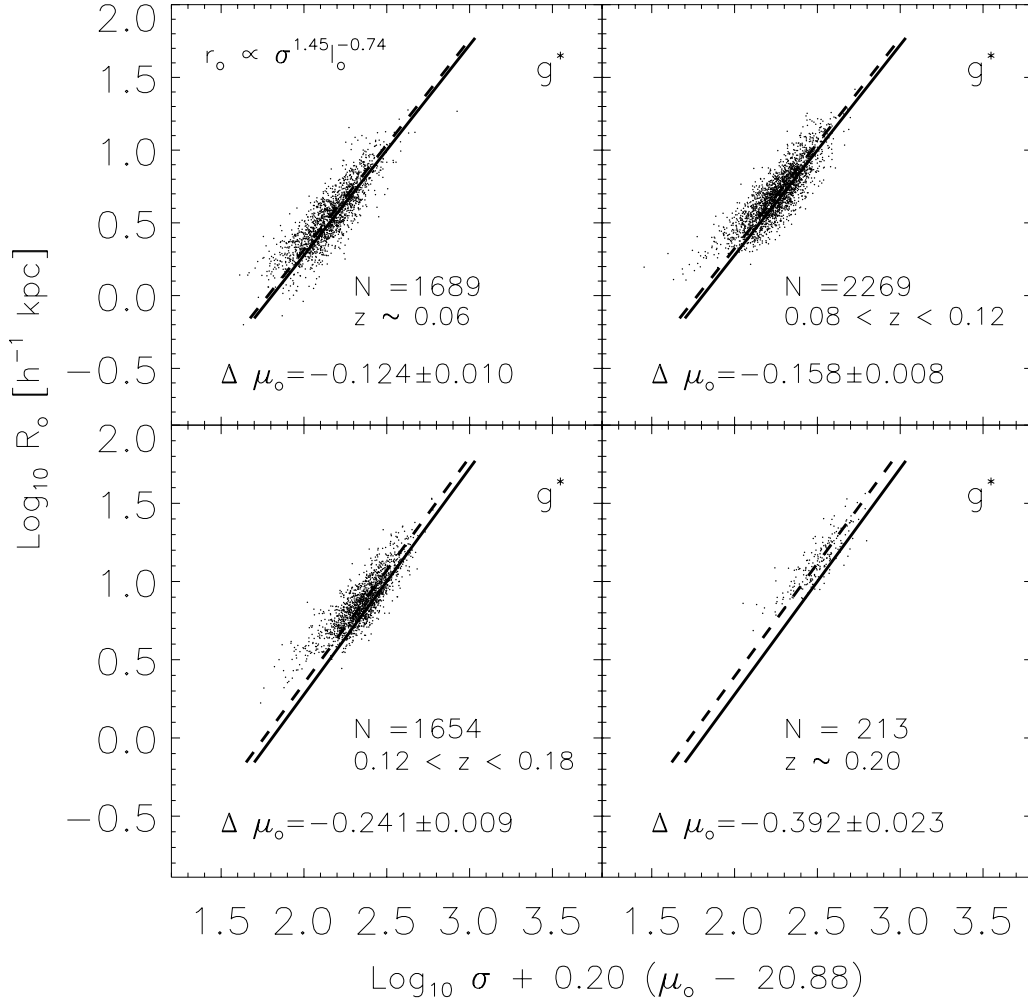


FIG. 6.—The g^* FP in four redshift bins. The slope of the FP is fixed to that at zero redshift; only the zero point is allowed to vary. The zero-point shifts systematically with redshift. The same plot for r^* shows similar but smaller shifts.

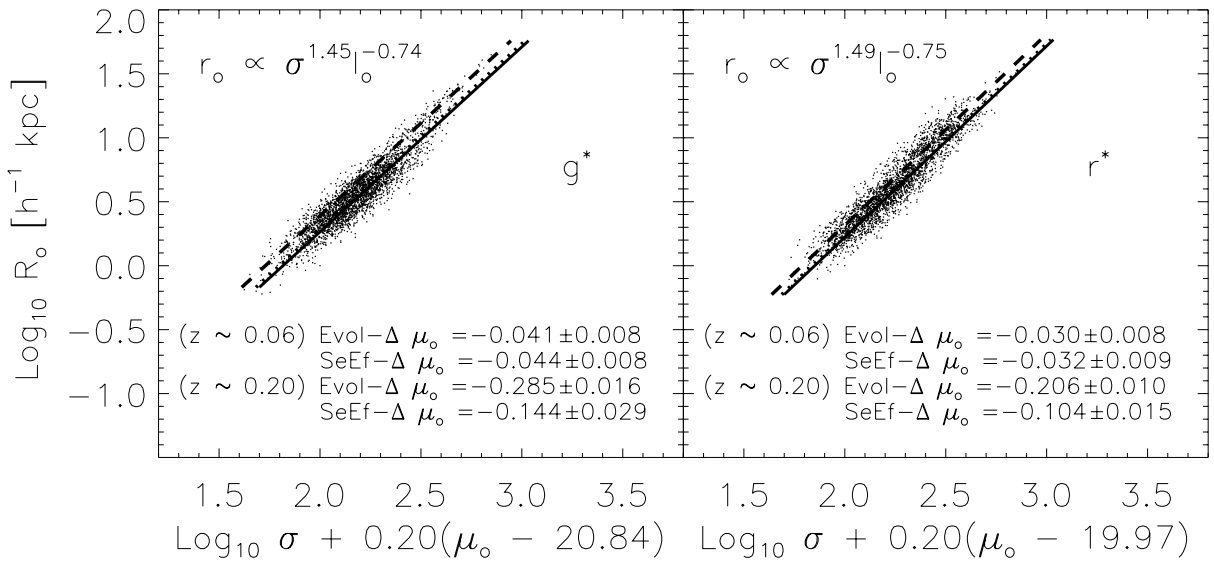


FIG. 7.—FP in the g^* (left) and r^* (right) magnitude-limited mock catalogs. The solid line shows the FP at $z = 0$. The dotted and dashed lines show fits using a low- and high-redshift subsample only. For these fits, the slope of the FP is required to be the same as the solid line; only the zero point is allowed to vary. The shift seen in the complete simulations is labeled “Evol- $\Delta\mu_o$,” whereas the shift seen in the magnitude-limited simulations is the sum of this and the quantity labeled “SeEf- $\Delta\mu_o$.” This sum is similar to the shift seen in the SDSS data, suggesting that selection effects are not negligible.

played by evolution and selection effects accurately. Since the parameters of the simulations were set by the maximum likelihood analysis, we conclude that the likelihood analysis of the evolution in luminosities is reasonably accurate ($\Delta\mu_o \approx -1.15z$) in g^* , but we note that this evolution is less than one would have inferred if selection effects were ignored ($\Delta\mu_o \approx -2z$).

The importance of selection effects in our sample has implications for another way in which studies of evolution are presented. If galaxies do not evolve, then the FP can be used to define a standard candle, so the test checks if residuals from the FP in the direction of the surface brightness variable, when plotted versus redshift, follow Tolman's $(1+z)^4$ cosmological dimming law. If Friedmann-Robertson-Walker models are correct, then departures from this $(1+z)^4$ dimming trend can be used to test for evolution. This can be done if one assumes that the main effect of evolution is to change the luminosities of galaxies. If so, then evolution will show up as a tendency for the residuals from the FP, in the μ direction, to drift away from the

$(1+z)^4$ dimming (e.g., Sandage & Perlmutter 1990; Pahre, Djorgovski, & de Carvalho 1996).

Figure 8 shows this trend in our data set. The lowest dashed lines in all panels show the expected $(1+z)^4$ dimming; panels on the left/right show results in g^*/r^* . Consider the top two panels first. The points show residuals with respect to the zero-redshift FP in our sample. The crosses show the median residual in a small redshift bin. The galaxies do not quite follow the expected $(1+z)^4$ dimming. The similarity to the $(1+z)^4$ dimming argues in favor of standard cosmological models, whereas the small difference from the expected trend is sometimes interpreted as evidence for evolution (e.g., Jørgensen et al. 1999; van Dokkum et al. 1998, 2001; Treu et al. 1999, 2001a, 2001b).

Of course, to correctly quantify this evolution, we must account for selection effects. The dashed lines that lie between the $(1+z)^4$ scaling and the data (i.e., *crosses*) show how the surface brightness should scale if there was passive evolution of the form suggested by the maximum likelihood analysis, but there was no magnitude limit. That is, if

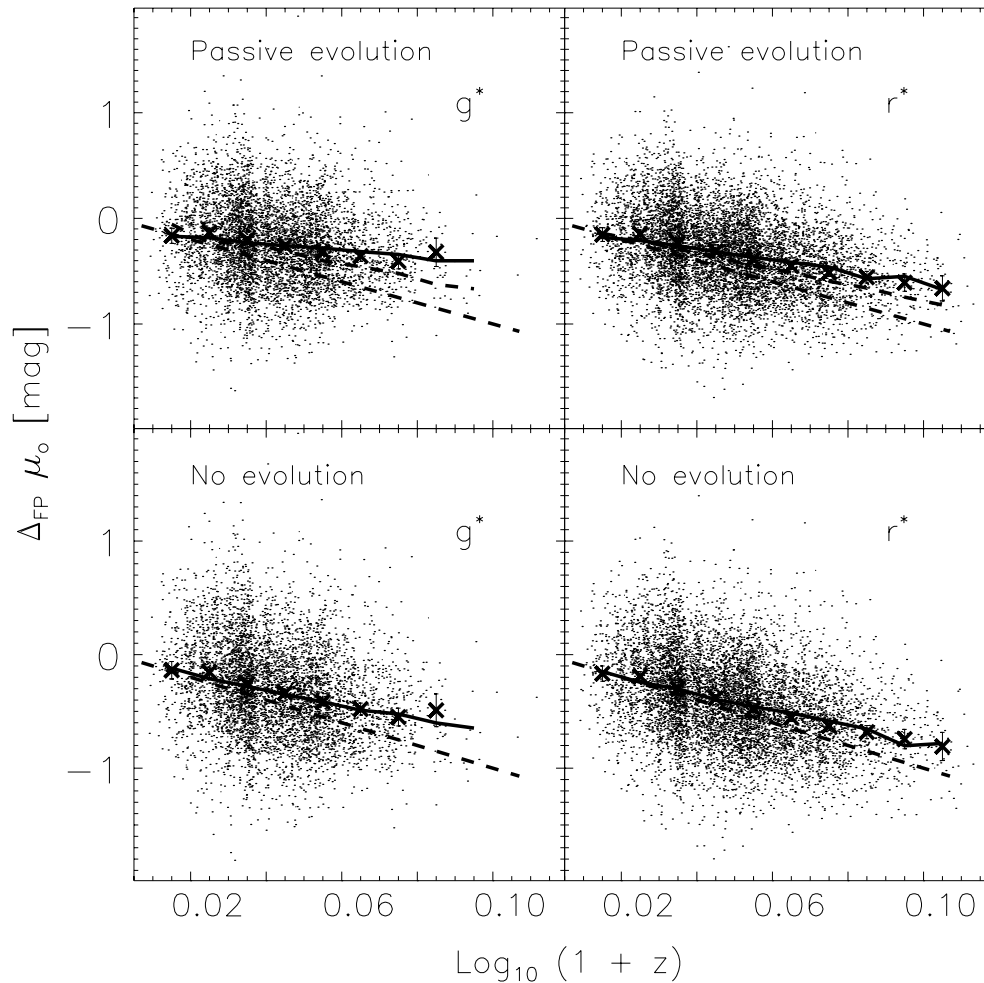


FIG. 8.—Residuals of the zero-redshift FP with respect to the surface brightness, before correcting for cosmological dimming, vs. redshift in the four bands. Lowest dashed line in all panels shows the $(1+z)^4$ dimming expected if there is no evolution. Solid curves in top panels show the same measurement in mock simulations of a magnitude-limited sample of a passively evolving population. The dashed lines in between show the actual evolution in surface brightness—the difference between these and the solid curves is an artifact of the magnitude limit. The bottom panel shows the same test applied using the parameters of the FP that best describe the data if there is required to be no evolution whatsoever. In this case, the entire trend away from the $(1+z)^4$ dimming is a selection effect. Note how, once the magnitude limit has been applied, both the evolving (*top*) and nonevolving populations (*bottom*) appear very similar to our observed sample.

$M_*(z) = M_*(0) - Qz$, then the surface brightnesses should scale as $(1+z)^{4-0.92Q}$. The solid curves show the result of making the measurement in simulated magnitude-limited catalogs that include this passive evolution. Note how different these solid curves are from the dashed curves (they imply Q about twice the correct value), but note how similar they are to the data. This shows that about half of the evolution one would naively have inferred from such a plot is a consequence of the magnitude limit.

To further emphasize the strength of this effect, we constructed simulations in which there was no evolution whatsoever. We did this by first making maximum likelihood estimates of the joint luminosity, size, and velocity dispersion distribution in which no evolution was allowed. (For the reasons discussed earlier, the associated no-evolution FP coefficient a is steeper by about 10%.) This was then used to generate mock catalogs in which there is no evolution. The crosses in the bottom panels show the result of repeating the same procedure as in the top panels, but now using the parameters of the no-evolution FP, and the solid line shows the measurement in the no-evolution simulations in which, by construction, the population of galaxies at all redshifts is identical. Therefore, the shifts from the $(1+z)^4$ dimming we see in the magnitude-limited no-evolution catalogs (*bottom, solid curves*) are entirely due to the magnitude limit. Note how similar the solid lines from our no-evolution simulations are to the actual data. If we believed there really was no evolution, then the results shown in the bottom panel would lead us to conclude that much of the trend away from the $(1+z)^4$ dimming is a selection effect—it is not evidence for evolution. (The fact that we were able to find a nonevolving population that mimics the observations so well suggests that the population of early-type galaxies at the median redshift of our sample must be rather similar to the population at lower and at higher redshifts. This, in turn, can constrain models of when the stars in these galaxies must have formed.)

We view our no-evolution simulations as a warning about the accuracy of this particular test of evolution. If the evolution is weak, then it appears that the results of this test depend critically on how the catalog was selected, and on what one uses as the fiducial FP. To make this second point, we followed the procedure adopted by many other recent publications. Namely, we assumed that the zero-redshift FP has the shape reported by Jørgensen et al. (1996) for Coma, for which a is about 15% smaller than what we find in g^* . If no account is taken of selection effects, then the inferred evolution in μ_o results in a value of Q that is about a factor of 4 times larger than the one we report in Table 1!

Our results indicate that inferences about evolution that are based on this test depend uncomfortably strongly on the strength of selection effects, and on what one assumes for the fiducial shape of the FP. In this respect, our findings about the role of, and the need to account for, selection effects are consistent with those reported by Simard et al. (1999). While we believe we have strong evidence that the early-type population is evolving, we do not believe that the strongest evidence of this evolution comes from either of the tests presented in this subsection. Nevertheless, it is reassuring that the evolution we see from these FP tests is consistent with that which we estimated using the likelihood analysis in Paper II and is also consistent with what we use to make our K -corrections. Namely, a passively evolving population that formed the bulk of its stars about 9 Gyr ago

appears to provide a reasonable description of the evolution of the surface brightnesses in our sample.

2.5. The FP: Dependence on Environment

This section is devoted to a study of if and how the properties of early-type galaxies depend on environment. Paper I describes our working definition of environment—essentially, we use the number of galaxies that are nearby in color, angular, and redshift space as an indication of the local density. Our procedure for assigning neighbors is least secure in the lowest redshift bin (typically $z \leq 0.08$).

Paper I shows that when the number of near neighbors is small, the luminosities, sizes, and velocity dispersions all increase slightly as the local density increases, whereas the surface brightnesses decrease slightly, although all these trends are very weak. A more efficient way of seeing if the properties of galaxies depend on environment is to show the residuals from the FP. As we argue below, this efficiency comes at a cost: if the residuals correlate with environment, it is difficult to decide if the correlation is due to changes in luminosity, size, or velocity dispersion.

Figure 9 shows the differences between galaxy surface brightnesses and those predicted by the zero-redshift maximum likelihood FP given their sizes and velocity dispersions, as a function of local density. Stars, circles, diamonds, triangles, squares, and crosses show averages over galaxies in the redshift ranges $z \leq 0.075$, $0.075 < z \leq 0.1$, $0.1 < z \leq 0.12$, $0.12 < z \leq 0.14$, $0.14 < z \leq 0.18$, and $z > 0.18$. Error bars show the error in determining the mean. (For clarity, the symbols have been offset slightly from each other.) The plot shows that the residuals depend on redshift—we have already argued that this is a combination of evolution and selection effects. Note that in all redshift bins the residuals tend to increase as local density increases. This suggests that the mean residual from the FP depends on environment. If the offset in surface brightness is interpreted as evidence that galaxies in denser regions are slightly less luminous than their counterparts in less dense regions, then this might be evidence that they formed at higher redshift. While this is a reasonable conclusion, we should be cautious: because $\mu_o - \mu_{\text{FP}}(R_o, \sigma) = -\Delta_1/b$, what we have really found is that the residuals in the direction of R_o correlate with environment. Because $\sigma - \sigma_{\text{FP}}(R_o, \mu_o) = -\Delta_1/a$, we might also have concluded that the velocity dispersions of galaxies in dense regions are systematically different from those of galaxies that have the same sizes and luminosities but are in the field. For similar reasons, a plot of the mean residual orthogonal to the plane shows a dependence on environment. (However, the rms scatter in the orthogonal direction of the residuals around the mean residual in each density bin, when plotted as a function of density, shows no trend.) Thus, while the FP suggests that the properties of galaxies depend on environment, it does not say how.

2.6. The κ -Space Projection

Bender et al. (1992) suggested three simple combinations of the three observables:

$$\begin{aligned}\kappa_1 &= \log(R_o \sigma^2) / \sqrt{2}, \\ \kappa_2 &= \log(I_o^2 \sigma^2 / R_o) / \sqrt{6}, \\ \kappa_3 &= \log(I_o^{-1} \sigma^2 / R_o) / \sqrt{3},\end{aligned}\tag{9}$$

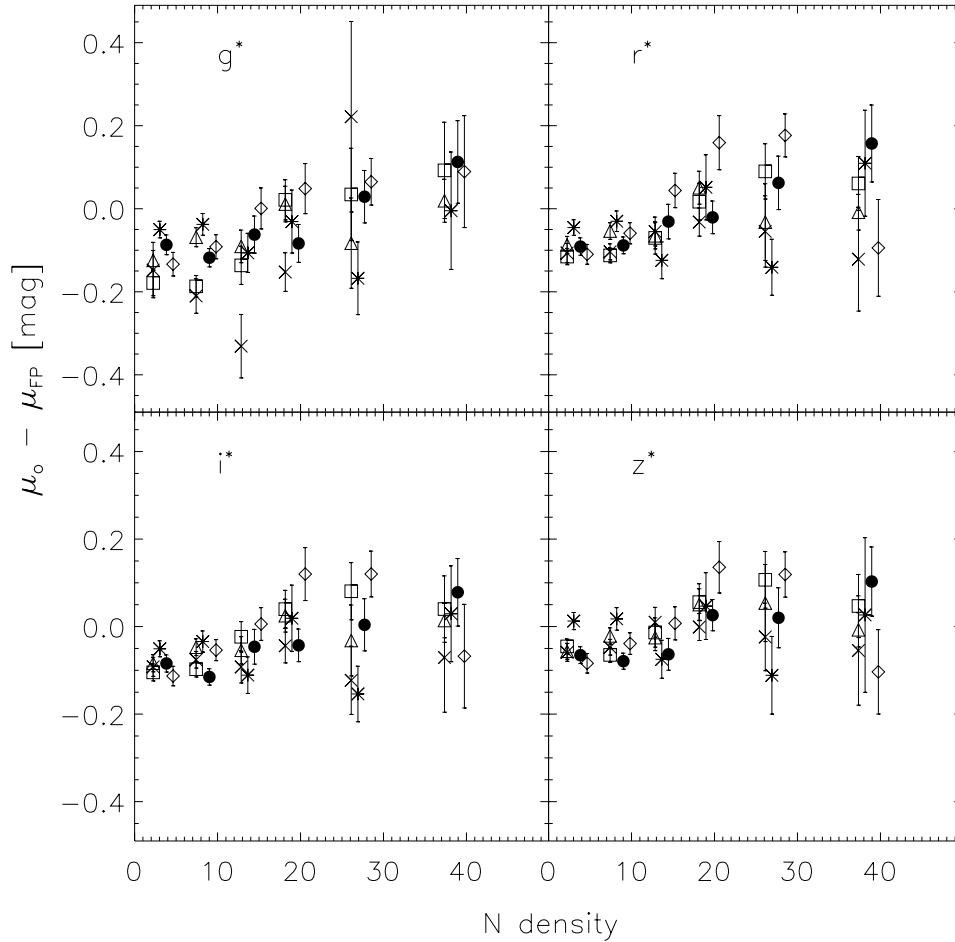


FIG. 9.—Residuals from the FP as a function of number of nearby neighbors. Stars, circles, diamonds, triangles, squares, and crosses show averages over galaxies in the redshift ranges $z \leq 0.075$, $0.075 < z \leq 0.1$, $0.1 < z \leq 0.12$, $0.12 < z \leq 0.14$, $0.14 < z \leq 0.18$, and $z > 0.18$.

which, they argued, correspond approximately to the FP viewed face-on ($\kappa_2 - \kappa_1$), and the two edge-on projections ($\kappa_3 - \kappa_1$ and $\kappa_3 - \kappa_2$). They also argued that their parametrization was simply related to the underlying physical variables. For example, κ_1 is proportional to mass and κ_3 is proportional to the mass-to-light ratio. The $\kappa_1 - \kappa_2$ projection would view the FP face-on if $R_o \propto \sigma^2 / I_o$; recall that we found $R_o \propto (\sigma^2 / I_o)^{0.75}$.

Bender et al.’s choice of parameters was criticized by Pahre et al. (1998) on the grounds that if the effective radius R_o is a function of wavelength, then the “mass” becomes a function of wavelength, which is unphysical. On average, the effective radii of the galaxies in our sample do increase with decreasing wavelength (Paper I), so one might conclude that Pahre et al.’s objections are valid. However, recall that we do not use the measured velocity dispersion directly; rather, σ represents the value the dispersion would have had at some fixed fraction of R_o . If R_o depends on wavelength, and we wish to measure the velocity dispersion at a fixed fraction of R_o , then one might argue that we should also correct the measured velocity dispersion differently in the different bands. The velocity dispersion decreases with increasing radius. So, if R_o is larger in the blue band than the red, then the associated velocity dispersion we should use in the blue band should be smaller than in the red. One might imagine that the combination $R_o \sigma^2$ remains approximately constant

after all. Whether or not it does, we have chosen to present the SDSS early-type sample in the κ -space projection introduced by Bender et al.

Figure 10 shows the results for the four SDSS bands. Because the mean surface brightness depends on wave band, we set $\log I = 0.4(27 - \mu_o + \langle \mu_o \rangle - \langle \mu_g \rangle)$ when making the plots, so as to facilitate comparison with Bender et al. (1992) and Burstein et al. (1997). The dashed line in the top right corner of each panel shows $\kappa_1 + \kappa_2 = 8$, what Burstein et al. termed the “zone of avoidance.” Had we not accounted for the fact that the mean surface brightness is different in the different bands, then the galaxies would populate this zone.

The magnitude limit is clearly visible in the bottom right corner of the $\kappa_3 - \kappa_1$ projection; we have not made any correction for it.

When the sample is split by color, the redder galaxies appear to follow a tighter relation than the bluer galaxies, and they also tend to lie slightly closer to the zone of avoidance. We leave quantifying and interpreting these trends to future work.

3. DISCUSSION AND CONCLUSIONS

We have studied the FP populated by ~ 9000 early-type galaxies over the redshift range $0 \leq z \leq 0.3$ in the g^* , r^* ,

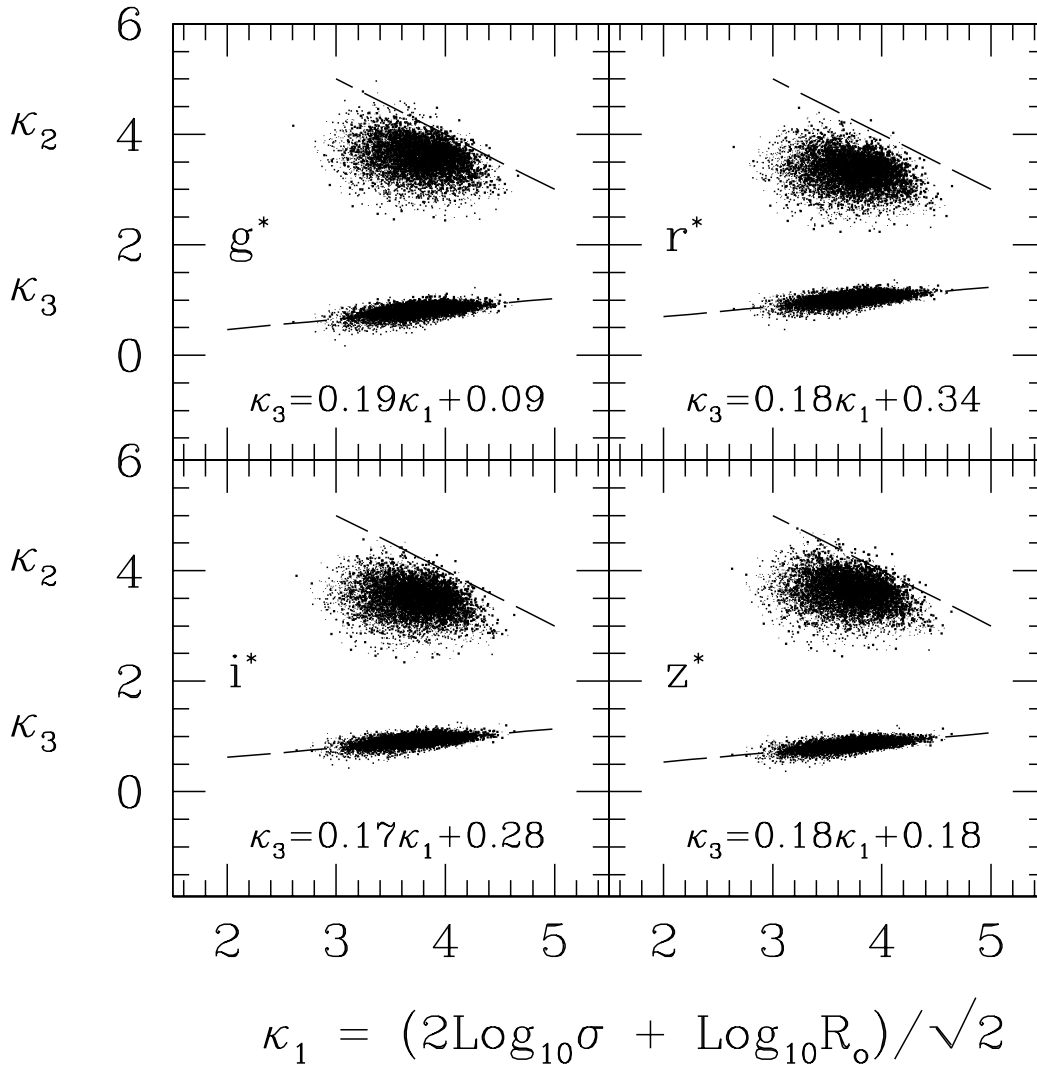


FIG. 10.—Early-type sample in the four SDSS bands viewed in the κ -space projection of Bender et al. (1992). The dashed line in the top right corner of each panel shows $\kappa_1 + \kappa_2 = 8$, what Burstein et al. (1997) termed the “zone of avoidance.”

i^* , and z^* bands. If this FP is defined by minimizing the residuals orthogonal to it, then $R_o \propto \sigma^{1.5} I_o^{-0.77}$ (see Table 2 for the exact coefficients). The FP is remarkably similar in the different bands (Fig. 1) and appears to be slightly warped in the shorter wave bands (Fig. 2). Residuals with respect to the direct fit (i.e., the FP defined by minimizing the residuals in the direction of $\log R_o$) do not correlate with either velocity dispersion or color, whereas residuals from the orthogonal fit correlate with both (Fig. 3). This correlation with σ is simply a projection effect (see Fig. 4 and related discussion), whereas the correlation with color is mainly due to the fact that color and σ are strongly correlated (Paper IV). The FP is intrinsically slightly thinner in the redder wave bands. This thickness is sometimes expressed in terms of the accuracy to which the FP can provide redshift-independent distance estimates—this is about 20%. If the thickness is expressed as a scatter in the mass-to-light ratio at fixed size and velocity dispersion, then this scatter is about 30%.

The simplest virial theorem prediction for the shape of the FP is that $R_o \propto \sigma^2 / I_o$. This assumes that the observed velocity dispersion σ^2 is proportional to the kinetic energy

σ_{vir}^2 that enters the virial theorem. Busarello et al. (1997) argue that in their data $\log \sigma = (1.28 \pm 0.11) \log \sigma_{\text{vir}} - 0.58$, so that $\sigma^{1.5} \propto \sigma_{\text{vir}}^{1.92}$. Since the coefficient of σ in the FP we find in all four bands is ~ 1.5 , it would be interesting to see if the kinetic energy for the galaxies in our sample scales as it did in Busarello et al.’s sample. To do this, measurements of the velocity dispersion profiles of (a subsample of) the galaxies in our sample are required.

Tests for passive luminosity evolution that use the FP are severely affected by selection effects and the choice of the fiducial FP against which to measure the evolution (Figs. 6–8). These tests suggest that the surface brightnesses of galaxies at higher redshifts in our sample are brighter than those of similar galaxies nearby. The amount of brightening is consistent with the luminosity evolution estimated in Paper II.

The way in which galaxies scatter from the FP correlates weakly with their local environment (Fig. 9). If this is caused entirely by differences in surface brightness, then galaxies in overdense regions are slightly fainter. If so, then single-age stellar population models suggest that early-type galaxies in denser regions formed at higher redshift. However, it may

be that the velocity dispersions are higher in denser regions (Paper II). A larger sample is necessary to make a more definitive statement.

By the time SDSS is complete, the uncertainty in the K -corrections, which prevent us at the present time from making more precise quantitative statements about the evolution of the luminosities and colors, will be better understood. In addition, the size of the sample will have increased by more than an order of magnitude. This will allow us to provide a more quantitative study of the effects of environment than we are able to at the present time.

We would like to thank S. Charlot for making his stellar population synthesis predictions for the SDSS filters available to the collaboration and N. Benitez for making his package available. We thank D. Kelson and M. Pahre for

helpful discussions. Funding for the creation and distribution of the SDSS Archive has been provided by the Alfred P. Sloan Foundation, the Participating Institutions, the National Aeronautics and Space Administration, the National Science Foundation, the US Department of Energy, the Japanese Monbukagakusho, and the Max Planck Society. The SDSS Web site is <http://www.sdss.org/>. The SDSS is managed by the Astrophysical Research Consortium for the Participating Institutions. The Participating Institutions are the University of Chicago, Fermilab, the Institute for Advanced Study, the Japan Participation Group, Johns Hopkins University, Los Alamos National Laboratory, the Max-Planck-Institute for Astronomy, the Max-Planck-Institute for Astrophysics, New Mexico State University, University of Pittsburgh, Princeton University, the US Naval Observatory, and the University of Washington.

REFERENCES

- Bender, R., Burstein, D., & Faber, S. M. 1992, *ApJ*, 399, 462
 Bernardi, M., et al. 2003a, *AJ*, 125, 1817 (Paper I)
 ———. 2003b, *AJ*, 125, 1849 (Paper II)
 ———. 2003c, *AJ*, 125, 1882 (Paper IV)
 Burstein, D., Bender, R., Faber, S. M., & Nolthenius, R. 1997, *AJ*, 114, 1365
 Busarello, G., Capaccioli, M., Capozziello, S., Longo, G., & Puddu, E. 1997, *A&A*, 320, 415
 Colless, M., Saglia, R. P., Burstein, D., Davies, R. L., McMahan, R. K., & Wegner, G. 2001, *MNRAS*, 321, 277
 Djorgovski, S., & Davis, M. 1987, *ApJ*, 313, 59
 Dressler, A., Lynden-Bell, D., Burstein, D., Davies, R. L., Faber, S. M., Terlevich, R. J., & Wegner, G. 1987, *ApJ*, 313, 42
 Faber, S. M., & Jackson, R. 1976, *ApJ*, 204, 668
 Gibbons, R. A., Fruchter, A. S., & Bothun, G. D. 2001, *AJ*, 121, 649
 Guzmán, R., Lucey, J. R., & Bower, R. G. 1993, *MNRAS*, 265, 731
 Hudson, M. J., Lucey, J. R., Smith, R. J., & Steel, J. 1997, *MNRAS*, 291, 488
 Jørgensen, I., Franx, M., Hjorth, J., & van Dokkum, P. G. 1999, *MNRAS*, 308, 833
 Jørgensen, I., Franx, M., & Kjørgaard, P. 1996, *MNRAS*, 280, 167
 Kelson, D. D., Illingworth, G. D., van Dokkum, P. G., & Franx, M. 2000, *ApJ*, 531, 184
 Kormendy, J. 1977, *ApJ*, 218, 333
 Lucey, J. R., Bower, R. G., & Ellis, R. S. 1991, *MNRAS*, 249, 755
 Pahre, M. A., Djorgovski, S. G., & de Carvalho, R. R. 1996, *ApJ*, 456, L79
 ———. 1998, *AJ*, 116, 1591
 Saglia, R. P., Colless, M., Burstein, D., Davies, R. L., McMahan, R. K., & Wegner, G. 2001, *MNRAS*, 324, 389
 Sandage, A., & Perlmutter, J.-M. 1990, *ApJ*, 361, 1
 Scodreggio, M. 1997, Ph.D. thesis, Cornell Univ.
 Simard, L., et al. 1999, *ApJ*, 519, 563
 Treu, T., Stiavelli, M., Bertin, G., Casertano, S., & Møller, P. 2001a, *MNRAS*, 326, 237
 Treu, T., Stiavelli, M., Casertano, S., Møller, P., & Bertin, G. 1999, *MNRAS*, 308, 1037
 Treu, T., Stiavelli, M., Møller, P., Casertano, S., & Bertin, G. 2001b, *MNRAS*, 326, 221
 van Dokkum, P. G., Franx, M., Kelson, D. D., & Illingworth, G. D. 1998, *ApJ*, 504, L17
 ———. 2001, *ApJ*, 553, L39