EARLY-TYPE GALAXIES IN THE SLOAN DIGITAL SKY SURVEY. II.
CORRELATIONS BETWEEN OBSERVABLES

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,[14,10] MAMORU DOI,[15,16]
MASATAKA FUKUGITA,[17,18] JOSHUA FRIEMAN,[1,3] 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. YORK[1]

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 using morphological and spectral criteria. The sample was used to study how early-type galaxy observables, including luminosity L, effective radius R_o, surface brightness I_o, color, and velocity dispersion σ, are correlated with one another. Measurement biases are understood with mock catalogs that reproduce all of the observed scaling relations and their dependences on fitting technique. At any given redshift, the intrinsic distribution of luminosities, sizes, and velocity dispersions in our sample are all approximately Gaussian. A maximum likelihood analysis shows that $\sigma \propto L^{0.25\pm0.012}$, $R_o \propto L^{0.63\pm0.025}$, and $R_o \propto I^{-0.75\pm0.02}$ in the r^* band. In addition, the mass-to-light ratio within the effective radius scales as $M_o/L \propto L^{0.14\pm0.02}$ or $M_o/L \propto M_o^{0.22\pm0.05}$, and galaxies with larger effective masses have smaller effective densities: $\Delta_o \propto M_o^{-0.52\pm0.03}$. These relations are 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 little, with more evolution in the bluer bands. The luminosity function is consistent with weak passive luminosity evolution and a formation time of about 9 Gyr ago.

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

1. INTRODUCTION

This is the second 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, effective radius 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.

Section 2 of the present paper shows that the luminosity function of the galaxies in our sample, when expressed as a function of absolute magnitude, is well described by a

Gaussian form, and that the luminosities in the population as a whole appear to be evolving passively. Section 3 studies the distribution of (the logarithm of) velocity dispersion, size, surface brightness, effective mass, and effective density at fixed luminosity; all of these are quite well described by Gaussian forms, suggesting that the intrinsic distributions of log (size) and log (velocity dispersion) are, like the distribution of log (luminosity), approximately Gaussian. Maximum likelihood estimates of these and other correlations, which include the Faber-Jackson relation, the mass-to-light ratio, the Kormendy relation and a mass-density relation are presented in § 4. Appendix A describes a method for generating accurate mock complete and magnitude-limited galaxy catalogs, which are useful for assessing the relative importance of evolution and selection effects. The procedure used to estimate errors on our results is discussed in Appendix B.

Paper III (Bernardi et al. 2003b) of this series places special emphasis on the fundamental plane relation between size, surface brightness, and velocity dispersion. It shows how the fundamental plane depends on wave band, color, redshift, and environment. Paper IV (Bernardi et al. 2003c) uses the colors and spectra of these galaxies to provide information about the chemical evolution of the early-type population.

Except where stated otherwise, we write the Hubble constant as $H_0 = 100 \ h$ km s^{-1} 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 $t_0 = 9.43 \ h^{-1}$ Gyr. For comparison, an Einstein–de Sitter model has $(\Omega_M, \Omega_\Lambda) = (1, 0)$ and $t_0 = 6.52 \ h^{-1}$ 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} kpc and velocity dispersions in kilometers per second, respectively.

2. THE LUMINOSITY FUNCTION

Our sample is magnitude-limited (Table 1 of Paper I gives the magnitude limits in the different bands). Therefore, we measure the luminosity function of the galaxies in our sample using two techniques. The first uses volume-limited catalogs, and the second uses a maximum likelihood procedure (Sandage, Tammann, & Yahil 1979; Efstathiou, Ellis, & Peterson 1988).

In the first method, we divide our parent catalog into many volume-limited subsamples; this was possible because the parent catalog is so large. When doing this, we must decide what size volumes to choose. We would like our volumes to be as large as possible so that each volume represents a fair sample of the universe. On the other hand, the volumes must not be so large that evolution effects are important. In addition, because our catalog is cut at the bright, as well as the faint, end, large-volume subsamples span only a small range in luminosities. Therefore, we are forced to compromise: we have chosen to make the volumes about $\Delta z = 0.04$ thick, because $c\Delta z/H \approx 120 \ h^{-1}$ Mpc is larger than the largest structures seen in numerical simulations of the cold dark matter family of models (e.g., Colberg

et al. 2000). The catalogs are extracted from regions that cover a very wide angle on the sky, so the actual volume of any given volume-limited catalog is considerably larger than $(120 \ h^{-1}$ Mpc)^3. Therefore, this choice should provide volumes that are large enough in at least two of the three coordinate directions that they represent fair samples, but not so large in the redshift direction that the range in luminosities in any given catalog is small, or that evolution effects are washed out.

The volume-limited subsamples are constructed as follows. First, we specify the boundaries in redshift of the catalog: $z_{\min}$ and $z_{\max} = z_{\min} + 0.04$. In the context of a world model, these redshift limits, when combined with the angular size of the catalog, can be used to compute a volume. This volume depends on $z_{\min}$, $z_{\max}$ and the world model: as our fiducial model, we set $\Omega_M = 0.3$ and $\Omega_\Lambda = 1 - \Omega_M$. (Our results hardly change if we use an Einstein–de Sitter model instead.) We then compute the K-corrected limiting luminosities $L_{\max}(z_{\min})$ and $L_{\min}(z_{\max})$ given the apparent magnitude limits, the redshift limits, and the assumed cosmology. A galaxy i is included in the volume-limited subsample if $z_{\min} \le z_i \le z_{\max}$ and $L_{\min} \le L_i \le L_{\max}$. The luminosity function for the volume-limited subsample is obtained by counting the number of galaxies in a luminosity bin and dividing by the volume of the subsample.

The top panels in Figure 1 show the result of doing this in the g^* and r^* bands. Stars, circles, diamonds, triangles, squares, and crosses show measurements in volume-limited catalogs that have $z_{\min} = 0.04, 0.08, 0.12, 0.16, 0.20$, and 0.24 and $z_{\max} = z_{\min} + 0.04$. Each subsample contains more than 500 galaxies, except for the two most distant, which each contain about 100. As one would expect, the nearby volumes provide the faint end of $\phi(M)$, and the more distant volumes show the bright end. The extent to which the different volume-limited catalogs all trace out the same curve is a measure of how little the luminosity function at low and high redshifts differs from that at the median redshift.

The bottom panels in Figure 1 show evidence that, in fact, the galaxies in our data show evidence for a small amount of evolution: at fixed comoving density, the higher redshift population is slightly brighter than that at lower redshifts. Although volume-limited catalogs provide model-independent measures of this evolution, the test is most sensitive when a large range of luminosities can be probed at two different redshifts. Because the SDSS catalogs are cut at both the faint and the bright ends, our test for evolution is severely limited. Nevertheless, the small trends we see are both statistically significant, and qualitatively consistent with what one expects of a passively evolving population. (Note that our sample contains only early-type galaxies. Blanton et al. 2001 study the luminosity function in an SDSS sample that contains all galaxy types, but they ignore evolution effects. Since late-type galaxies are expected to evolve more rapidly than early types, it is important to redo Blanton et al.'s analysis after allowing for evolution.)

Before we make more quantitative conclusions, notice that a bell-like Gaussian shape would provide a reasonable description of the luminosity function. Although early-type galaxies are expected to have red colors, our sample was not selected using any color information. It is reassuring, therefore, that the Gaussian shape we find here also provides a good fit to the luminosity function of the redder objects in the SDSS parent catalog (see the curves for the two reddest
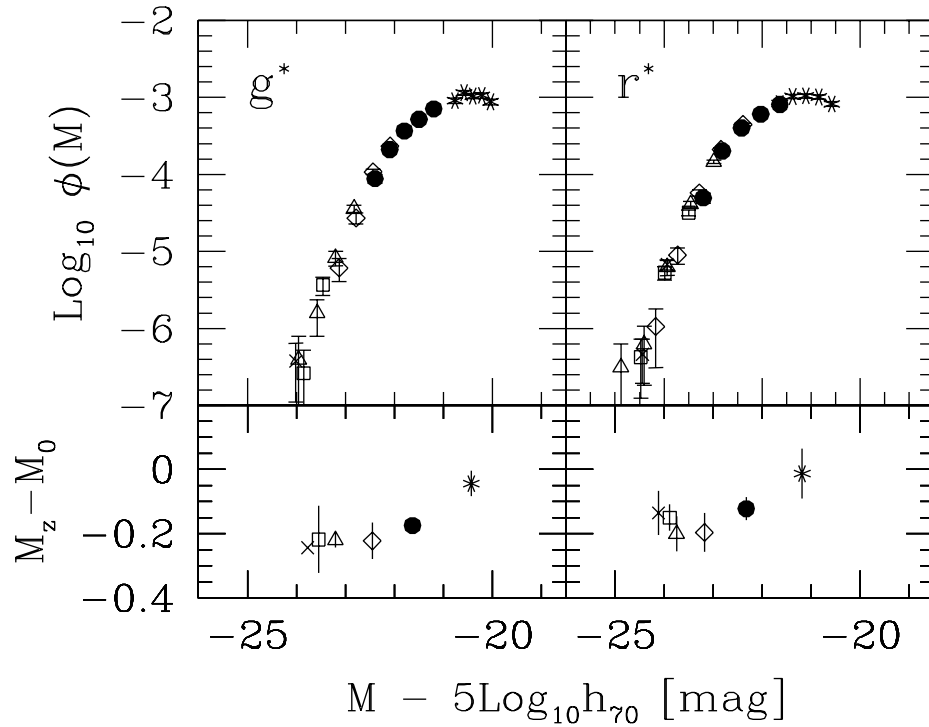
Fig. 1.—Luminosity functions in the $g*$ and $r*$ bands. Stars, circles, diamonds, triangles, squares, and crosses show measurements in volume-limited catalogs that are adjacent in redshift of width $\Delta z = 0.04$, starting from a minimum of $z_{min} = 0.04$. Top panels show that the higher redshift catalogs contribute at the bright end only. At the same comoving density, the symbols that represent the higher redshift catalogs tend to be displaced slightly to the left of the those that represent the lower redshift catalogs. Bottom panels show this small mean shift toward increasing luminosity with increasing redshift.

galaxy bins in Fig. 14 of Blanton et al. 2001). A Gaussian form also provides a reasonable description of the luminosity function of early-type galaxies in the CNOC2 survey (Lin et al. 1999, even though they actually fit a Schechter function to their measurements). The 2dF Galaxy Redshift Survey (2dFGRS) galaxies classified as being of type 1 by Madgwick et al. (2002) should be similar to early types. Their type 1 galaxies extend to considerably fainter absolute magnitudes than our sample and the shape of the luminosity function they report is quite different from ours. This is probably because the population of early-type galaxies at faint absolute magnitudes is quite different from the brighter ones (e.g., Sandage & Perelmuter 1990). In any case, their Schechter function fits underestimate the number density of luminous type 1 galaxies—a Gaussian tail would provide a significantly better fit.

Given that the Gaussian form provides a good description of our data, we use the maximum likelihood method outlined by Sandage et al. (1979) to estimate the parameters of the best-fitting luminosity function. For magnitude-limited samples that are small and shallow, this is the method of choice. For a sample such as ours, which spans a sufficiently wide range in redshifts that evolution effects might be important, the method requires a model for the evolution. We parametrize the luminosity evolution similarly to Lin et al. (1999). That is to say, if we were solving only for the luminosity function, then the likelihood function we maximize would be

$$\mathscr{L} = \prod_i \frac{\phi(M_i, z_i | Q, M_*, \sigma_M)}{S(z_i | Q, M_*, \sigma_M)} , \qquad (1)$$

where

$$\phi(M_i, z_i | Q, M_*, \sigma_M) = \frac{\phi_*}{\sqrt{2\pi\sigma_M^2}}$$

$$\times \exp\left(-\frac{[M_i - M_* + Qz_i]^2}{2\sigma_M^2}\right) ,$$

$$S(z_i | Q, M_*, \sigma_M) = \int_{M_{min}(z_i)}^{M_{max}(z_i)} dM \, \phi(M, z_i | Q, M_*, \sigma_M) ,$$

$M_{min}(z_i)$ and $M_{max}(z_i)$ denote the minimum and maximum absolute magnitudes at z_i that satisfy the apparent magnitude limits of the survey, and i runs over all the galaxies in the catalog. (At small z, this parametrization of the evolution in absolute magnitude implies that the luminosity evolves as $L_*(z)/L_*(0) \approx (1 + z)^q$, with $q = Q \ln (10)/2.5$. Note that, in assuming that only M_* evolves, this model assumes that there is no differential evolution in luminosities, i.e., that luminous and not so luminous galaxies evolve similarly.

Figure 2 shows the result of estimating the luminosity function in this way in the $g*$, $r*$, $i*$, and $z*$ bands. Later in this paper, we will solve simultaneously for the joint distribution of luminosity, size, and velocity dispersion; it is the parameters that describe the luminosity function of this joint solution that are shown in Figure 2. The dashed lines in each panel show the Gaussian shape of the luminosity function at redshift $z = 0$. For comparison, the symbols show the measurements in the same volume-limited catalogs as before, except that now we have subtracted the maximum
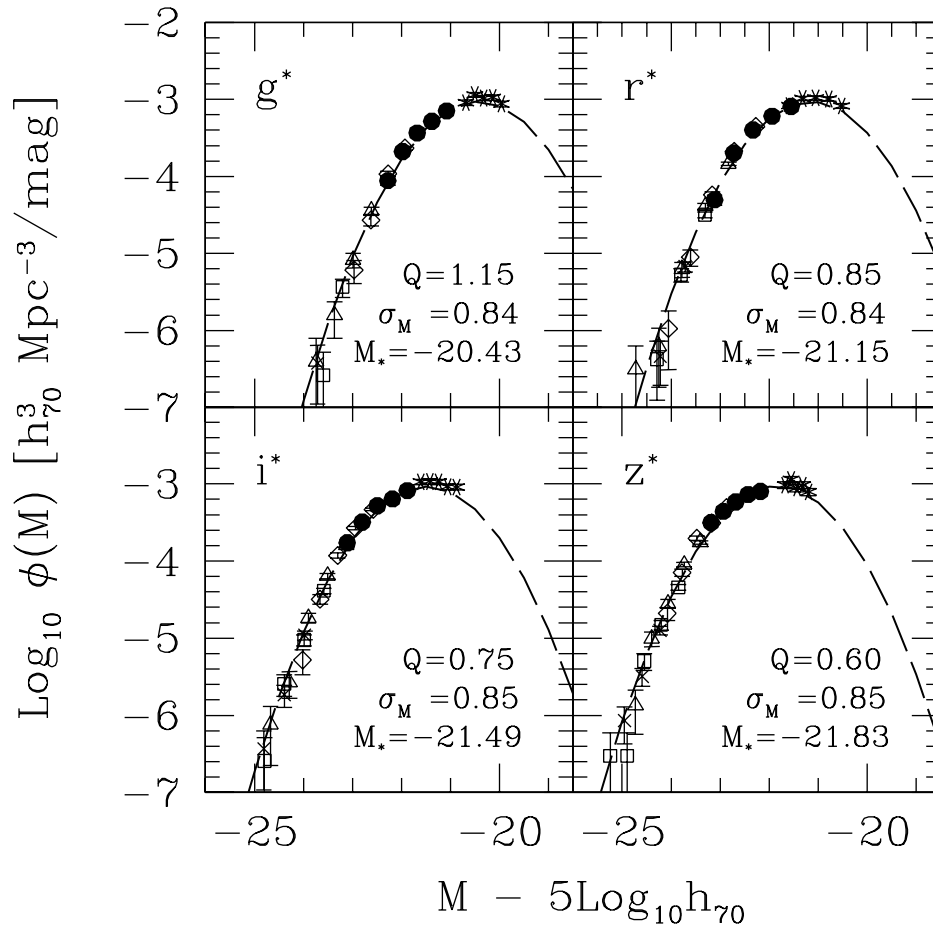
FIG. 2.—Luminosity functions in the g^*, r^*, i^*, and z^* bands, corrected for pure luminosity evolution. Symbols with error bars show the estimates from our various volume-limited catalogs; the higher redshift catalogs contribute at the bright end only. Dashed curves show the Gaussian-shaped luminosity function, which maximizes the likelihood of seeing this data at redshift $z = 0$.

likelihood estimate of the luminosity evolution from the absolute magnitudes M before plotting them. If the model for the evolution is accurate, then the different symbols should all trace out the same smooth dashed curve.

The comoving number density of the galaxies in this sample is $\phi_* = (5.8 \pm 0.3) \times 10^{-3} \, h^3 \, \mathrm{Mpc}^{-3}$ in all four bands. Because the different bands have different apparent magnitude limits, and they were fitted independently of each other, it is reassuring that the same value of ϕ_* works for all the bands. For similar reasons, it is reassuring that the best-fit values of M_* imply rest-frame colors at $z = 0$ of $g^*-r^* = 0.72$, $r^*-i^* = 0.34$, and $r^*-z^* = 0.68$, which are close to those of the models that we used to compute our K-corrections (Appendix A of Paper I), even though no a priori constraint was imposed on what these rest-frame colors should be.

The histograms in each of the four panels of Figure 3 show the number of galaxies observed as a function of redshift in the four bands. The peak in the number counts at $z \sim 0.08$ is also present in the full SDSS sample, which includes late types, and, perhaps more surprisingly, an overdensity at this same redshift is also present in the 2dFGRS. (The second bump at $z \sim 0.13$ is also present in the 2dFGRS counts.) The solid curves show what we expect to see for the evolving Gaussian function fits—the curves provide a reasonably good fit to the observed counts, although they slightly overestimate the numbers at high redshift in the red-

der wave bands. For comparison, the dashed curves show what is expected if the luminosities do not evolve and the no-evolution luminosity function is given by the one at the median redshift (i.e., a Gaussian with mean $M_* - 0.1Q$). Although the fit to the high-redshift tail is slightly better, this no-evolution model cannot explain the trends shown in the bottom panel of Figure 1. Moreover, a Bruzual & Charlot (2003) passive evolution model with a formation time of 9 Gyr ago predicts that the rest-frame luminosities at redshift $z = 0.2$ should be brighter than those at $z = 0$ by 0.3, 0.26, 0.24, and 0.21 mag in g^*, r^*, i^*, and z^*, respectively—not far off from what we estimate.

The bottom panels in Figure 1 suggest two possible reasons why our model of pure luminosity evolution overestimates dN/dz at higher z. One possibility is that the comoving number densities are decreasing slightly with redshift. A small amount of density evolution is not unexpected, because early-type galaxy morphologies may evolve (van Dokkum & Franx 2001), and our sample is selected on the basis of a fixed morphology. If we allow a small amount of density, as well as luminosity evolution, and we use $\phi_*(z) = 10^{0.4Pz}\phi_*(0)$ with $P \approx -2$, as suggested by the results of Lin et al. (1999), then the resulting dN/dz curves are also well fitted by the dashed curves. A second possibility follows from the fact that we only observe the most luminous part of the higher redshift population. If the most luminous galaxies at any given time are also the oldest, then
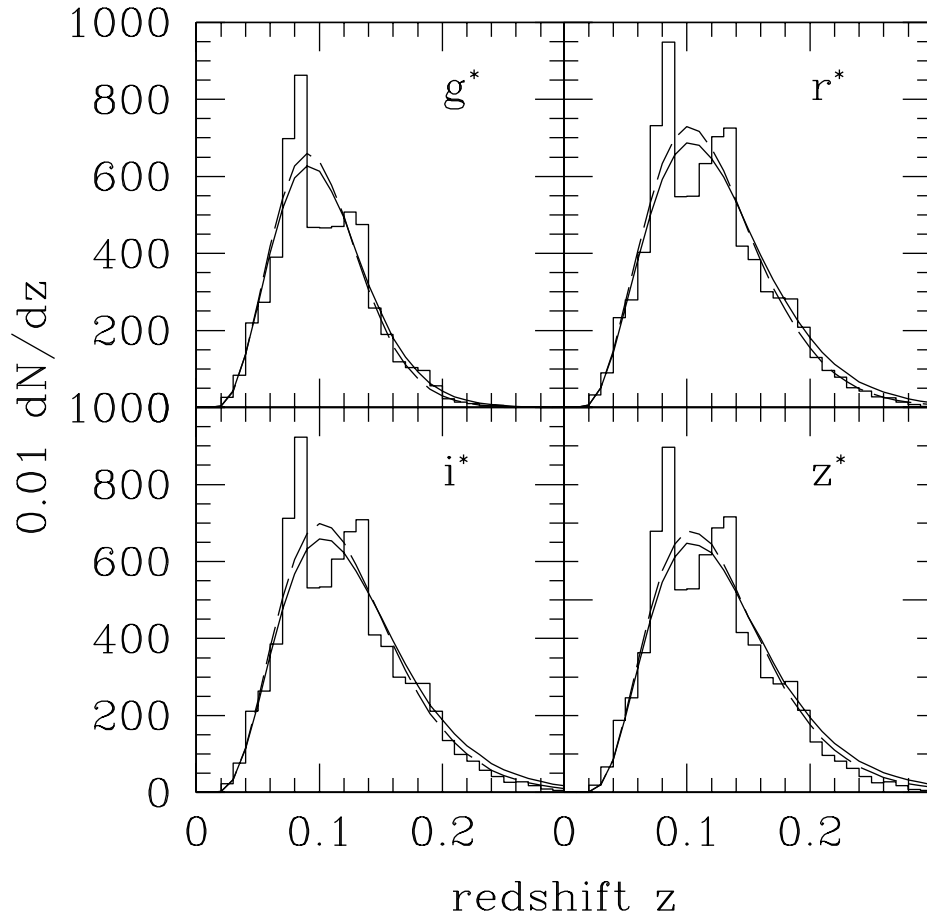
FIG. 3.—Number of galaxies as a function of redshift in our sample. Solid curves show the predicted counts if the comoving number densities are constant, but the luminosities brighten systematically with redshift: $M_*(z) = M_*(0) - Qz$ with Q given by the previous figure. Dashed curves show what one predicts if there is no evolution whatsoever, and the luminosity function is fixed to the value it has at the median redshift of our sample ($z = 0.1$).

one might expect the bright end of the luminosity function to evolve less rapidly than the fainter end. The curvature seen in the bottom panel of Figure 1 suggests that although the evolution of the fainter objects in our sample (which we only see out to low redshifts) is consistent with formation times of 9 Gyr ago, the brighter objects are not. Models of differential evolution in the luminosities also predict dN/dz distributions that are in better agreement with the observed counts at high redshift. Since the evolution of the luminosity function is small, we prefer to wait until we are able to make more accurate K-corrections before accounting for either of these other possibilities more carefully. Therefore, in what follows, we will continue to use the model with pure luminosity evolution.

Repeating the exercise described above but for an Einstein–de Sitter model yields qualitatively similar results, although the actual values of M_* and ϕ_* are slightly different. At face value, the fact that we see so little evolution in the luminosities argues for a relatively high formation redshift: the Bruzual & Charlot (2003) models indicate that $t_{\rm form} \sim 9\,{\rm Gyr}$.

3. OBSERVED CORRELATIONS: DISTRIBUTIONS AT FIXED LUMINOSITY

This section presents scatter plots between different observables X and luminosity. This is done because, except

for a cut at small velocity dispersions, our sample was selected by luminosity alone. This means that the distributions of X at fixed luminosity are not biased by the selection cut (e.g., Schechter 1980). The distribution of X at fixed L is shown to be reasonably well described by a Gaussian for all the choices of X we consider. This simplifies the maximum likelihood analysis described in § 4 that we use to estimate a number of observed correlations (it is also used in Paper III to estimate the parameters of the fundamental plane).

The best way to think of any absolute magnitude M versus X scatter plot is to imagine that, at fixed absolute magnitude M, there is a distribution of X values. The scatter plot then shows the joint distribution

$$\phi(M, X|z)dM\,dX = dM\,\phi(M|z)p(X|M,z)dX\ ,\quad (2)$$

where $\phi(M, X|z)$ denotes the density of galaxies with X and M at z, and $\phi(M|z)$ is the luminosity function at z, which we computed in § 2. One of the results of this section is to show that the shape of $p(X|M, z)$ is simple for most of the relations of interest.

The mean value of X at fixed M is independent of the fact that our catalogs are magnitude-limited. Therefore, we estimate the parameters of linear relations of the form:

$$(X - X_*) = [-0.4(M - M_*)]/S \quad (3)$$

where $M = -2.5 \log L$ is the absolute magnitude and X is

the observable (for example, we will study $X = \log \sigma$, $\log R_o$ or $\mu_o = -2.5 \log I_o$). For each volume-limited catalog, we fitted for the slope S and zero point of the linear relation. If there really were a linear relation between M and X, and neither X nor M evolved, then the slopes and zero points of the different volume-limited catalogs would be the same.

To illustrate, the different symbols in Figure 4 show $\langle \log \sigma | M \rangle$, the Faber-Jackson relation (Faber & Jackson 1976), in our data set. Most data sets in the literature are consistent with the scaling $\langle \sigma | L \rangle \propto L^{1/4}$, approximately independent of wave band. For example, Forbes & Ponman (1999), using a compilation of data from Prugniel & Simien (1996) report $L \propto \sigma^{3.92}$ in the B band. At longer wavelengths, Pahre, Djorgovski, & de Carvalho (1998) report $L_K \propto \sigma^{4.14 \pm 0.22}$ in the K band, with a scatter of 0.93 mag.

Stars, circles, diamonds, triangles, squares, and crosses show the relation measured in volume-limited catalogs of successively higher redshift (redshift limits are the same as in Fig. 1). The galaxies in each subsample were further divided into two equal-sized parts based on luminosity. The symbols with error bars show the mean $\log \sigma$ for each of these small bins in M, and the rms spread around it (note that the error on the mean is smaller than the size of the symbols in all but the highest redshift catalogs). The solid line shows the maximum likelihood estimate of the slope of this relation at $z = 0$, which we

describe in § 4. Comparison with this line shows that the higher redshift population is slightly brighter. The slope of this line is shown in the top of each panel: $\sigma \propto L^{1/4}$, approximately, in all the bands, consistent with the literature. The zero point, however, is different; at fixed luminosity, the objects in our sample have velocity dispersions that are smaller than those reported in the literature by about $\log \sigma = 0.05$.

We have enough data that we can actually do more than simply measure the mean X at fixed M; we can also compute the distribution around the mean. If we do this for each catalog, then we obtain distributions that are approximately Gaussian in shape, with dispersions that depend on the range of luminosities that are in the subsample. Rather than showing these, we created a composite catalog by stacking together the galaxies from the nonoverlapping volume-limited catalogs, and we then divided the composite catalog into five equal-sized bins in luminosity. The histograms in the bottom of the plot show the shapes of the distribution of velocities in the different luminosity bins. Except for the lowest and highest redshift catalogs for which the statistics are poorest, the different distributions have almost the same shape; only the mean changes.

One might have worried that the similarity of the distributions is a signature that they are dominated by measurement
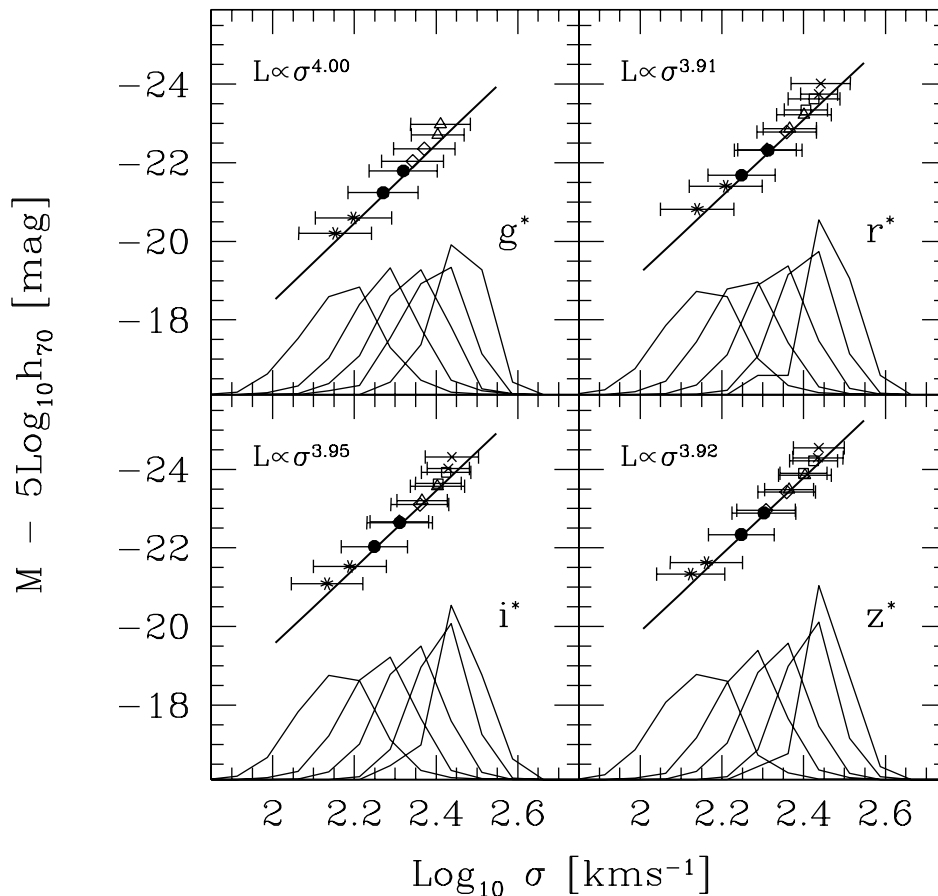
Fig. 4.—Relation between luminosity L and velocity dispersion σ. Stars, circles, diamonds, triangles, squares, and crosses show the error-weighted mean value of $\log \sigma$ for a small range in luminosity in each volume-limited catalog (see text for details). (Only catalogs containing more than 100 galaxies are shown.) Error bars show the rms scatter around this mean value. Solid line shows the maximum likelihood estimate of this relation, and the label in the top left shows the scaling it implies. Histograms show the distribution of $\log \sigma$ in small bins in luminosity. They were obtained by stacking together nonoverlapping volume-limited catalogs to construct a composite catalog, and then dividing the composite catalog into five equal-sized bins in luminosity.

error. This is not the case: the typical measurement error is about a factor of 2 smaller than the rms of any of these distributions. If we assume that the measurement errors are Gaussian distributed, then the distributions we see should be the true distribution broadened by the Gaussian from the measurement errors. The fact that the observed distributions are well approximated by Gaussians suggests that the true intrinsic distributions are also Gaussian. The fact that the width of the intrinsic distribution is approximately independent of M considerably simplifies the maximum likelihood analysis presented in the next section.

It is well known that color is strongly correlated with velocity dispersion (Paper IV of this series shows the color-σ relation in our sample). One consequence of this is that residuals from the σ-L relation shown in Figure 4 correlate strongly with color: at fixed magnitude, the redder galaxies have the highest velocity dispersions. In addition, as a whole, the reddest galaxies populate the high-σ part of the relation. Forbes & Ponman (1999) reported that residuals from the Faber-Jackson relation correlate with age. If color is an indicator of age and/or metallicity, then our finding is qualitatively consistent with theirs: the typical age/metallicity varies along the Faber-Jackson relation.

A similar study of the relation between the luminosities and sizes of galaxies is shown in Figure 5. Schade, Barientos, & López-Cruz (1997) find $L_B \propto R_o^{4/3}$ in the B band, whereas, at longer wavelengths, Pahre et al. (1998) find $L_K \propto R_o^{7/4}$ with an rms of 0.88 mag. This suggests that the relation depends on wavelength. We find $\langle R_o|L \rangle \propto L^{2/3}$ in g^*, but

$\langle R_o|L \rangle \propto L^{3/5}$ in the other bands. The distribution $p(\log R_o|M)$ is also reasonably well fitted by a Gaussian, with a mean that increases with luminosity and a dispersion that is approximately independent of M. The rms around the mean is about one and a half times larger than the rms around the mean σ-L relation. We argue in Paper IV that the color-magnitude and color-size relations are a consequence of the color-σ correlation. If this is correct, then residuals from the R_o-L relation should not correlate with size or magnitude. We have checked that this is correct, although we have not included a plot showing this explicitly.

There is an interesting correlation between the residuals of the Faber-Jackson and R_o-L relations. At fixed luminosity, galaxies which are larger than the mean $\langle R_o|L \rangle$ tend to have smaller velocity dispersions. This is shown in Figure 6, which plots the residuals from the σ-L relation versus the residuals from the R_o-L relation. The short dashed lines show the forward and inverse fits to this scatter plot. The long-dashed line in between the other two shows $\Delta_{R|M}/\sigma_{R|M} = -\Delta_{V|M}/\sigma_{V|M}$, where $\Delta_{X|M}$ denotes the residual from the mean relation at fixed M and $\sigma_{X|M}$ denotes the rms of this residual. The anticorrelation is approximately the same for all L.

This suggests that a plot of L versus some combination of R_o and σ should have considerably less scatter than either of the two individual relations. To illustrate, Figure 7 shows the distribution of the combination $R_o\sigma^2$ at fixed L. The scatter in L is significantly reduced, making the mean trend of increasing $R_o\sigma^2$ with increasing L quite clean. (The
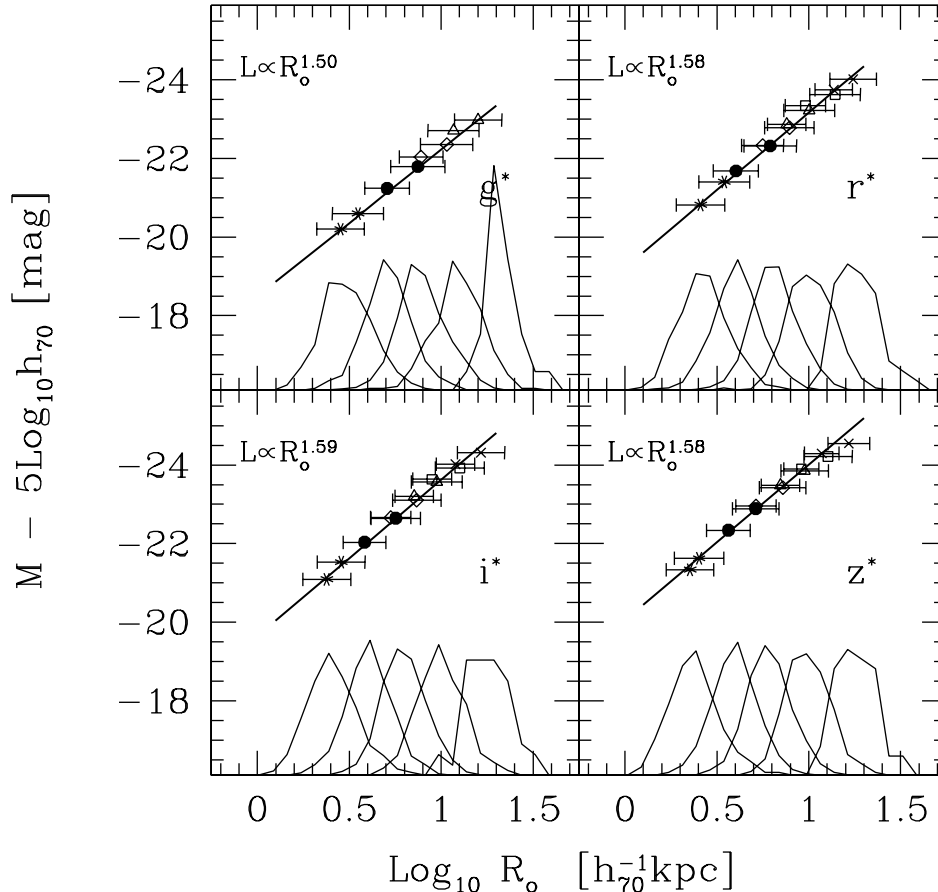

FIG. 5.—Same as Fig. 4, but for the relation between luminosity L and effective radius R_o
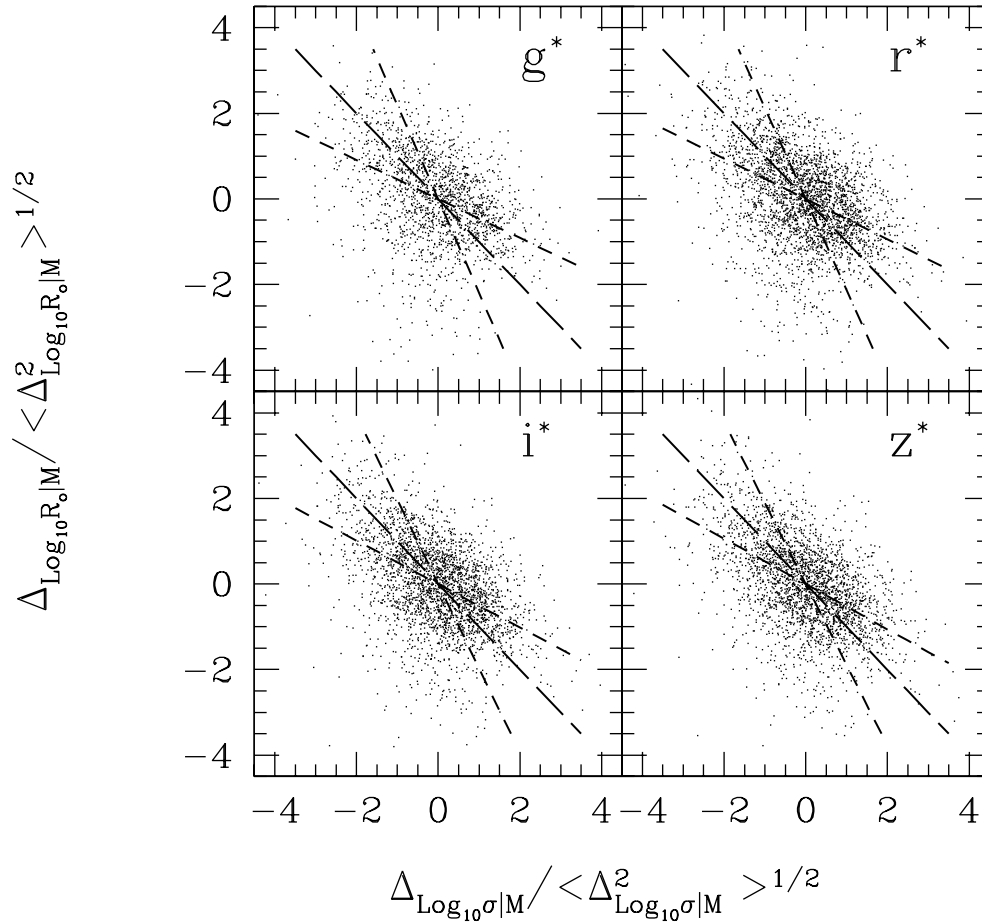
Fig. 6.—Residuals of the R_o-L relation are anticorrelated with residuals of the σ-L relation; galaxies of the same luminosity that are smaller than expected have larger velocity dispersions than expected. Plot shows the residuals normalized by their rms value. Short-dashed lines show forward and inverse fits to the scatter plots, and long-dashed line in between the other two shows $\Delta_{R_o|M}/\sigma_{R|M} = -\Delta_{\sigma-M}/\sigma_{V|M}$.

combination of observables for which the scatter is minimized is discussed in § 4.) This particular combination defines an effective mass: $M_o \equiv 2R_o\sigma^2/G$. In slightly more convenient units, this mass is

$$\left(\frac{M_o}{10^{10}\ h^{-1}\ M_\odot}\right) = 0.465\left(\frac{R_o}{h^{-1}\ \mathrm{kpc}}\right)\left(\frac{\sigma}{100\ \mathrm{km\ s^{-1}}}\right)^2. \quad (4)$$

(Because many of our galaxies are not spherical, some of their support must come from rotation, and so ignoring rotation as we are doing is likely to misestimate the true mass. See Bender, Burstein, & Faber 1992 for one way to account for this. This quantity will also misestimate the mass if some of the support comes from anisotropic velocity dispersions.)

Fiducial values of the effective mass-to-light ratio can be obtained by inserting the maximum likelihood values from Table 1 into this relation. This yields $M_o \approx 10^{10.56}\ h_{70}^{-1}\ M_\odot$ (we used the parameters for the r^* band, for which $R_* + 2V_* \approx 4.89$). The corresponding total absolute magnitude is $M_* - 5\log h_{70} \approx -21.15$. The luminosity of the Sun in r^* is 4.62 mag, so $L_* \approx 10^{10.31}\ h_{70}^{-2}\ L_\odot$. The luminosity within the effective radius is half this value, so that the effective mass-to-light ratio within the effective radius of an L_* object is $2h_{70} \times 10^{10.56-10.31} \approx 3.57h_{70}$ times that of the

Sun. Figure 7 shows that the effective mass-to-light ratio depends on luminosity:

$$\left\langle\frac{M_o}{L}\Big|L\right\rangle = 3.57h_{70}\left(\frac{L}{L_*}\right)^{0.15}\frac{M_\odot}{L_\odot}. \quad (5)$$

in r^*. At larger radii, the luminosity can double at most, whereas, if the galaxy is embedded in a dark matter halo, the mass at large radii may continue to increase. For this reason, one might expect the mass-to-light ratios to be significantly larger at larger radii.

Since the ratio $(2R_o\sigma_o^2/G)/(L/2)$ above is the mass-to-light ratio at the radius that encloses half the light, it is tempting to associate it with the mass-to-light ratio at the half-mass radius. Because both the numerator and the denominator are projected quantities, this is incorrect. For example, if the mass-to-light ratio is independent of distance from the galaxy center, then the three-dimensional half-mass radius is about 30% larger than the projected half-light radius (e.g., Hernquist 1990). If the velocity dispersion does not change substantially over the range in radii that contribute light, then a fairer estimate of the mass-to-light ratio within the half-mass radius would be about 30% larger than the value given above.

We can define an effective density by setting $3M_o/4\pi R_o^3 = (2R_o\sigma^2/G)/(4\pi R_o^3/3) = \Delta_o\rho_{\mathrm{crit}}$, with $\rho_{\mathrm{crit}} \equiv 3H^2/8\pi G$,
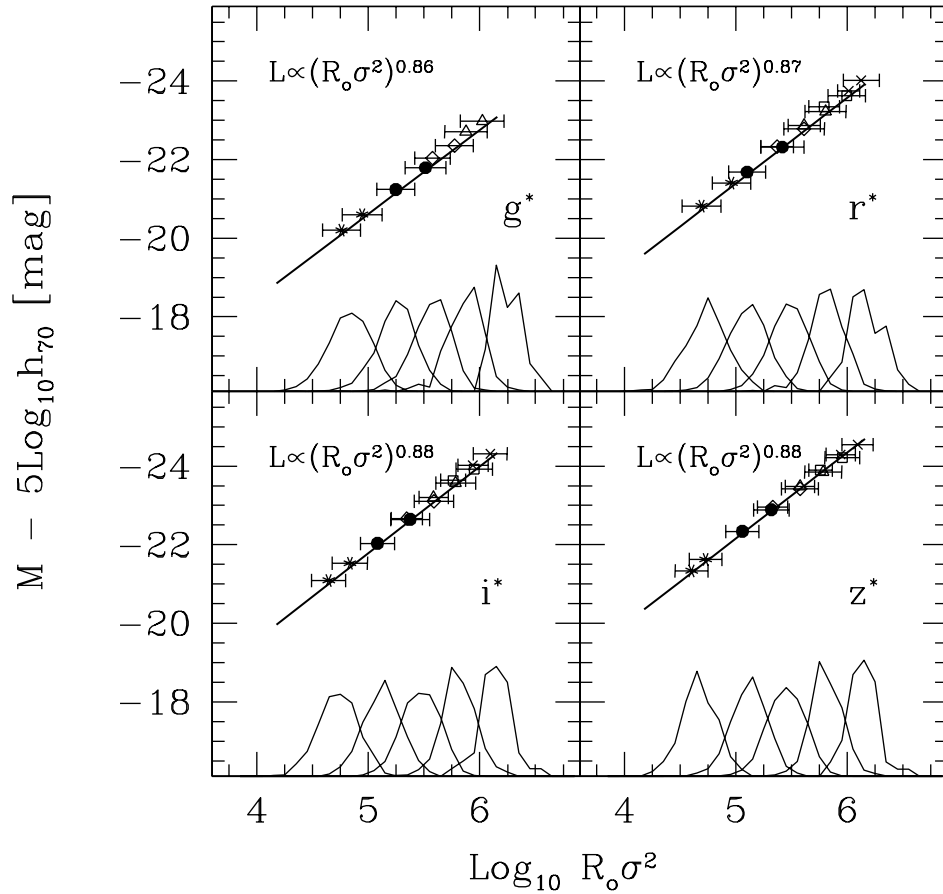
Fig. 7.—Same as Fig. 4, but for the relation between luminosity L and the combination $R_o\sigma^2$, which is supposed to be a measure of mass

then

$$\Delta_o = 4\left(\frac{\sigma}{HR_o}\right)^2 = (4 \times 10^6)0.7^2\left(\frac{\sigma}{100 \text{ km s}^{-1}}\frac{h_{70}^{-1} \text{ kpc}}{R_o}\right)^2$$

$$(6)$$

Figure 8 shows that this effective density decreases with increasing luminosity, although the scatter in densities at fixed luminosity is quite large ($\sim$0.32 dex). Inserting mean values for σ and R_o yields

$$\Delta_o = (5.16 \times 10^5)\left(\frac{\sigma/\sigma_*}{R_o/R_*}\right)^2 = (5.16 \times 10^5)\left(\frac{L}{L_*}\right)^{-3/4}$$

$$(7)$$

in r^*. Such a trend is qualitatively similar to that seen in numerical simulations of dissipationless gravitational clustering: the central densities of virialized halos in such simulations are smaller in the more massive halos (Navarro, Frenk, & White 1997).

Figure 9 shows a final relation at fixed luminosity: the surface brightness–L relation. In such a plot, luminosity evolution moves objects upward and to the right (larger luminosities and surface brightnesses at high redshift), so that the higher redshift population should be obviously displaced from the zero-redshift relation. The plot shows the distribution of μ_o and M after subtracting the maximum likelihood estimate of the evolution from both quantities. The solid line shows the maximum likelihood value of the slope of this relation. This differs slightly from the $I \propto L^{-0.45}$

TABLE 1

MAXIMUM LIKELIHOOD ESTIMATES, IN THE FOUR SDSS BANDS, OF THE JOINT DISTRIBUTION OF LUMINOSITIES, SIZES, AND VELOCITY DISPERSIONS

Band	N_{gals}	M_*	σ_M	R_*	σ_R	V_*	σ_V	ρ_{RM}	ρ_{VM}	ρ_{RV}	Q
g^*..........	5825	−20.43	0.844	0.520	0.254	2.197	0.113	−0.886	−0.750	0.536	1.15
r^*..........	8228	−21.15	0.841	0.490	0.241	2.200	0.111	−0.882	−0.774	0.543	0.85
i^*...........	8022	−21.49	0.851	0.465	0.241	2.201	0.110	−0.886	−0.781	0.542	0.75
z^*..........	7914	−21.83	0.845	0.450	0.241	2.200	0.110	−0.885	−0.782	0.543	0.60

NOTES.—Mean values of the variables at redshift z, $M_* - Qz$, R_*, V_*, and the elements of the covariance matrix $\mathscr{C}$ defined by the various pairwise correlations between the variables are shown. These coefficients are also used in computing the matrix $\mathscr{F}$ in Paper III.
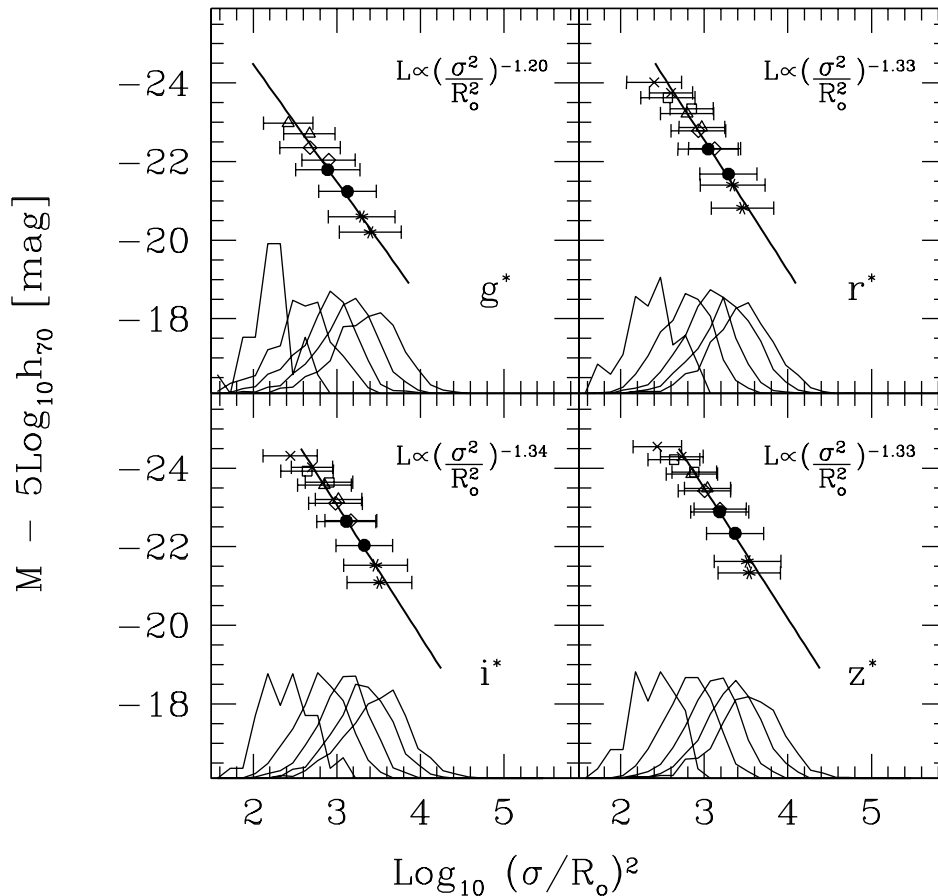
FIG. 8.—Same as Fig. 7, but for the relation between luminosity L and the combination $(\sigma/R_o)^2$, which is supposed to be a measure of density

0.45 scaling Sandage & Perelmuter (1990) find for giant galaxies with $M_B < -20$, although the scatter around the mean relation of ~ 0.58 mag is similar. This relation is considerably broader than any of the others we have studied so far, which may account for some of the difference. However, a careful inspection of the figure suggests that the relation is becoming shallower at high redshift; whether or not this is a signature of differential evolution in the luminosities is the subject of work in progress.

4. A PARAMETRIC MAXIMUM LIKELIHOOD ANALYSIS

Section 2 showed that, after accounting for the fact that the SDSS sample is magnitude-limited, the distribution of $M = -2.5 \log L$ is quite well described by a Gaussian. In principle, by extending the Efstathiou et al. (1988) method (along the lines described by Sodré & Lahav 1993), we could derive nonparametric maximum likelihood estimates of the three-dimensional distribution of L, R_o, and σ. The virtue of this approach is that it accounts for the fact that the observed sample is magnitude-limited, that there is also a cut at small velocity dispersions, and that there are correlated measurement errors associated with the luminosities, sizes, and velocity dispersions. Once the shape of the three-dimensional distribution has been estimated, it is straightforward to obtain estimates of the various correlations with luminosity we studied in the previous section. This is the

subject of § 4.2. However, the real benefit of the maximum likelihood analysis is that it also yields estimates of correlations between observables that do not include luminosity—some examples of these are shown in § 4.3.

We chose not to make a nonparametric estimate of the joint distribution because just 10 bins in each of L, R_o, and σ yields 10^3 free parameters to be determined from 10^4 galaxies. Moreover, § 3 showed that, in each of the SDSS wave bands, the distributions of $\log R_o$ and $\log \sigma$ at fixed absolute magnitude are quite well described by Gaussian forms. Therefore, the joint distribution of early-type galaxy luminosities, sizes, and velocity dispersions should be well described by a trivariate Gaussian distribution in the variables $M = -2.5 \log L$, $R = \log R_o$, and $V = \log \sigma$. Saglia et al. (2001) describe a maximum likelihood analysis of early-type galaxy correlations in which they *assume* that a trivariate Gaussian is a reasonable description of their data: we have the luxury of knowing that this is indeed a reasonable description of our data set. Thus, we have a simple parametrization of the joint distribution for which, in each wave band, nine numbers suffice to describe the statistical properties of our sample: three mean values, M_*, R_*, and V_*, three dispersions, σ_M^2, σ_R^2, and σ_V^2, and three pairwise correlations, $\sigma_R \sigma_M \rho_{RM}$, $\sigma_V \sigma_M \rho_{VM}$, and $\sigma_R \sigma_V \rho_{RV}$.

In addition, we will also allow for the possibility that the luminosities are evolving—a 10th parameter to be estimated from the sample. The maximum likelihood technique allows us to estimate these 10 numbers as follows. We define the
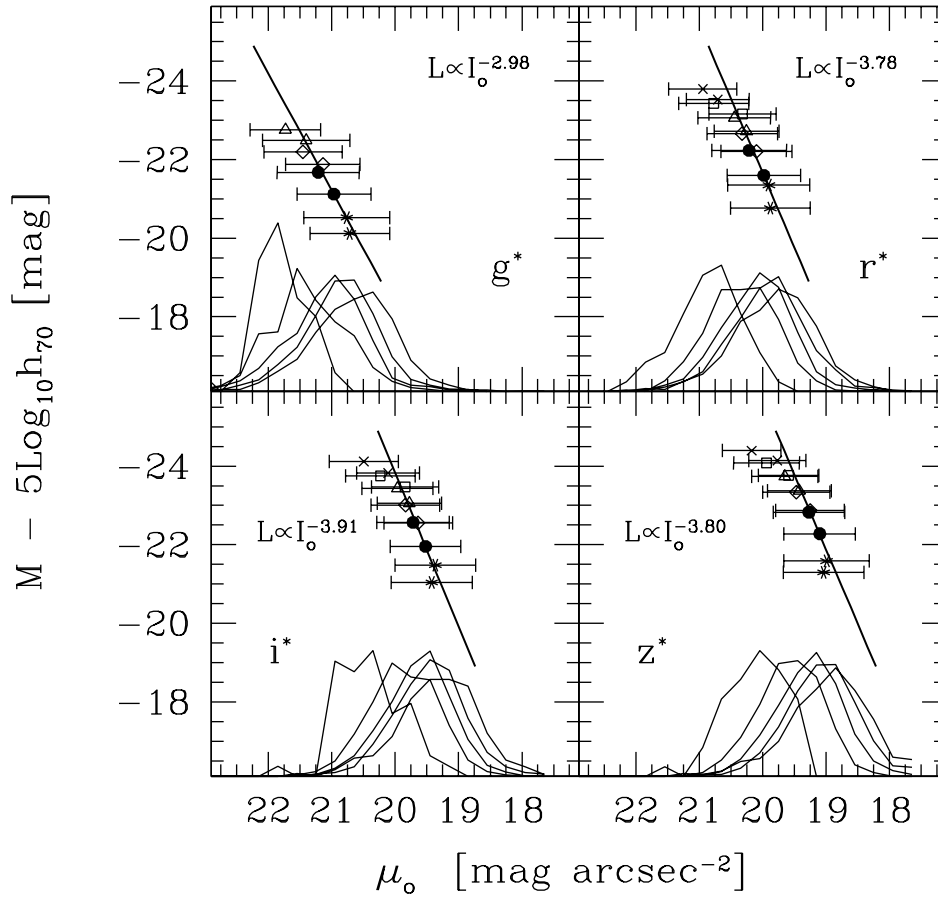
FIG. 9.—Relation between luminosity L and surface brightness μ_o in different volume-limited catalogs (higher redshift catalogs contribute points to the top-left corners of each plot). Passive evolution of luminosities would shift points upward and to the right of the zero-redshift relation, but the slope of the relation should remain unchanged. This shift has been subtracted.

likelihood function

$$\mathcal{L} = \prod_i \frac{\phi(\mathcal{X}_i, \mathcal{C}, \mathcal{E}_i)}{S(z_i)} , \qquad (8)$$

where

$$\mathcal{X} = (M - M_* + Qz, R - R_*, V - V_*) ,$$

$$\mathcal{E} = \begin{pmatrix} \epsilon_{MM}^2 & \epsilon_{RM}^2 & \epsilon_{VM}^2 \\ \epsilon_{RM}^2 & \epsilon_{RR}^2 & \epsilon_{RV}^2 \\ \epsilon_{VM}^2 & \epsilon_{RV}^2 & \epsilon_{VV}^2 \end{pmatrix} ,$$

$$\mathcal{C} = \begin{pmatrix} \sigma_M^2 & \sigma_R \sigma_M \rho_{RM} & \sigma_V \sigma_M \rho_{VM} \\ \sigma_R \sigma_M \rho_{RM} & \sigma_R^2 & \sigma_R \sigma_V \rho_{RV} \\ \sigma_V \sigma_M \rho_{VM} & \sigma_R \sigma_V \rho_{RV} & \sigma_V^2 \end{pmatrix} ,$$

$$\phi(\mathcal{X}, \mathcal{C}, \mathcal{E}) = \frac{\phi_*}{(2\pi)^{3/2} |\mathcal{C} + \mathcal{E}|^{-1/2}} \exp\left[-\frac{1}{2} \mathcal{X}^T (\mathcal{C} + \mathcal{E})^{-1/2} \mathcal{X} \right] .$$

Similarly to when we discussed the luminosity function, $S(z_i)$ is defined by integrating over the range of absolute magnitudes, velocities, and sizes at z_i, which make it into the catalog. Here $\mathcal{X}$ is the vector of the observables, and $\mathcal{E}$ describes the errors in the measurements.

Appendix D of Paper I describes how the elements of the error matrix $\mathcal{E}$ were obtained. Briefly, the error in the

absolute magnitude assumes that there are no errors in the redshift or the K-correction, so all the error comes from the error on the apparent magnitude $m_{\rm dev}$; the error on the circularly averaged radius R_o is given by adding the error on the angular length of the longer axis $r_{\rm dev}$ to those that come from the error on the axis ratio b/a. We assume that the errors in b/a are neither correlated with those in $\log r_{\rm dev}$ nor with those in the absolute magnitude. However, because both $m_{\rm dev}$ and $r_{\rm dev}$ come from the same fitting procedure, the errors in M and R_o are correlated. Finally, we assume that errors in magnitudes are not correlated with those in velocity dispersion, so $\epsilon_{\rm VM}^2$ is set to zero, and that errors in size and velocity dispersion are only weakly correlated because of the aperture correction we apply.

The covariance matrix $\mathcal{C}$ contains six of the 10 free parameters we are seeking. It is these parameters, along with the three mean values, M_*, R_*, and V_*, and the evolution parameter Q, which are varied until the likelihood is maximized. The maximum likelihood estimates of these parameters in each band are given in Table 1. Notice that although the luminosity and size distributions differ from band to band, the velocity distributions do not. This is reassuring, because the intrinsic distribution of velocity dispersions, estimated from the spectra, should not depend on the band in which the photometric measurements were made. As an additional test, we also computed maximum likelihood

estimates of the 2×2 covariance matrices of the bivariate Gaussians for the pairs (M, R) and (M, V). These estimates of, e.g., ρ_{RM} and ρ_{VM} were similar to those in Table 1.

The remainder of this paper uses $\mathscr{C}$ to estimate various pairwise correlations. In Paper III, we transform the covariance matrix $\mathscr{C}$ into one which describes the fundamental plane variables of size, surface brightness, and velocity dispersion.

4.1. The Intrinsic Distributions of Sizes and Velocity Dispersions

Before we present maximum likelihood estimates of various correlations, it is worth remarking that because the trivariate Gaussian is a good description of the data, our results indicate that, in addition to the intrinsic distribution of absolute magnitudes, the intrinsic distributions of (the logarithms of) early-type galaxy sizes and velocity dispersions are also well fitted by Gaussian forms. The means and dispersions of these Gaussians are given by (R_*, σ_R^2) and (V_*, σ_V^2) in Table 1. Note that the width of the distribution of $\log \sigma$ is about half that of $\log R_o$. This is consistent with earlier work (e.g., it is one of the motivations for the κ-space parametrization of Bender et al. 1992).

4.2. Correlations with Luminosity

As we describe below, appropriate combinations of the coefficients in Table 1 provide maximum likelihood estimates of various linear regressions between pairs of observables that are often studied; these are summarized in Table 2. Plots comparing some of these linear regressions with the maximum likelihood estimates are shown in § 3.

In the Gaussian model, the mean of $\log \sigma$ at fixed M is

$$\left\langle V - V_* \middle| M - M_* \right\rangle = \frac{(M - M_*)}{\sigma_M} \sigma_V \rho_{VM}$$
$$\equiv \frac{(M - M_*)}{-2.5 S_{VL}} , \qquad (9)$$

where the second equality defines S_{VL}, for ease of comparison with equation (3). The dispersion around this mean is

$$\sigma_{V|M}^2 \equiv \sigma_V^2 (1 - \rho_{VM}^2) . \qquad (10)$$

Inserting the values in Table 1 into these expressions for S_{VL} and $\sigma_{V|M}^2$ provides the maximum likelihood estimate of the slope and thickness of this relation. These are shown in the second column of Table 2, and the fit itself is shown in Figure 4. The errors we quote on the slopes of this, and the other relations in the table, were obtained using subsamples as described in Appendix B. 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 selected is an order of magnitude larger than any sample available in the literature.

The mean size at fixed absolute luminosity M, and the dispersion around this mean, are obtained by replacing all V's with R's in equation (9). The third column in Table 2 gives the maximum likelihood value of the slope S_{RL}, of the size-at-fixed-luminosity relation in the four bands. This fit is shown in Figure 5.

Similarly, one can show that the slopes of the mean L-mass and L-density relations shown in Figures 7 and 8 are $S_{ML} = (2/S_{VL} + 1/S_{RL})^{-1}$ and $S_{DL} = 1/(2/S_{VL} - 2/S_{RL})^{-1}$. These are the fourth and fifth columns of Table 2. The dispersions around these mean mass-L and density-L relations can be written in terms of the elements of $\mathscr{C}$, although we have not included the expressions here. Even though these relations are made from linear combinations of R and V, they may be tighter than either the L-σ or L-R_o relations because the correlation coefficients ρ_{RM}, ρ_{VM}, and ρ_{RV} are different from zero.

The surface brightnesses of the galaxies in our sample are defined by $(\mu_o - \mu_*) \equiv (M - M_*) + 5(R - R_*)$, so the dispersion in μ is $\sigma_\mu^2 = \sigma_M^2 + 10\sigma_M \sigma_R \rho_{RM} + 25\sigma_R^2$. The mean surface brightness at fixed luminosity is obtained by replacing all V's with μ's in equations (9) and (10) above. This means that we need $\rho_{\mu M}$, which we can write in terms of σ_M, σ_R, and ρ_{RM}. The sixth column in Table 2 gives the slope of the surface brightness I at fixed luminosity relation, $\langle I_o | L \rangle \propto L^{1/S_{IL}}$, in the four bands. These fits are shown in Figure 9.

4.3. Inverse Relations and Other Correlations

So far, we have shown that the maximum likelihood analysis provides estimates of correlations that are in good agreement with quantities that can also be estimated by a more straightforward regression technique. However, with the coefficients of the correlation matrix $\mathscr{C}$ in hand, it is straightforward to obtain estimates of correlations that, because of selection effects, cannot be reliably estimated using simple regressions. For example, the mean luminosity given the velocity dispersion is

$$\left\langle M - M_* \middle| V - V_* \right\rangle = \frac{(V - V_*)}{\sigma_V} \sigma_M \rho_{VM} , \qquad (11)$$

with dispersion $\sigma_M^2 (1 - \rho_{VM}^2)$ (compare eqs. [9] and [10]). Inserting the coefficients in Table 1 yields $\langle L | \sigma \rangle \propto \sigma^{2.34}$. Similarly, one can show that $\langle L | R_o \rangle \propto R_o^{1.23}$ and $\langle M_o/L | M_o \rangle \propto M_o^{0.22}$ in r^*.

TABLE 2
MAXIMUM LIKELIHOOD ESTIMATES OF THE SLOPES S_{VL}, S_{RL}, S_{ML}, S_{DL}, S_{IL}, AND S_k

Band	S_{VL}	S_{RL}	S_{ML}	S_{DL}	S_{IL}	S_k
g^*	4.00 ± 0.25	1.50 ± 0.06	0.86 ± 0.02	-1.20 ± 0.08	-2.98 ± 0.16	-0.73 ± 0.02
r^*	3.91 ± 0.20	1.58 ± 0.06	0.87 ± 0.02	-1.33 ± 0.07	-3.78 ± 0.17	-0.75 ± 0.02
i^*	3.95 ± 0.15	1.59 ± 0.06	0.88 ± 0.02	-1.34 ± 0.08	-3.91 ± 0.18	-0.76 ± 0.02
z^*	3.92 ± 0.15	1.58 ± 0.06	0.88 ± 0.02	-1.33 ± 0.07	-3.80 ± 0.17	-0.76 ± 0.01

NOTES.—Maximum likelihood estimates of the slopes S_{VL}, S_{RL}, S_{ML}, S_{DL}, and S_{IL}, of the relations between luminosity and the mean velocity dispersion, effective radius, effective mass, effective density, and effective surface brightness at fixed luminosity, as a function of luminosity. The slope of the relation between surface brightness and the mean size at fixed surface brightness is S_k.

We can also study correlations that do not involve luminosity. The best studied of these is the Kormendy (1977) relation: the surface brightnesses of early-type galaxies decrease with increasing effective radius. The mean size at fixed surface brightness in our sample is

$$\langle R - R_* | \mu - \mu_* \rangle = [(\mu - \mu_*)/\sigma_\mu]\sigma_R\rho_{\mu R}$$
$$\equiv -0.4 S_k(\mu - \mu_*) , \qquad (12)$$

where $\rho_{\mu R}$ can be written in terms of σ_M, σ_R, and ρ_{RM}, and the final equality defines S_k. The seventh column in Table 2 gives the slope of this relation in the four bands. For comparison, Kormendy (1977) found that $\log I_o \propto 1.29 \log R_o$ in the B band, and Pahre et al. (1998) find $R_o \propto I_o^{-0.61}$ in the K band.

For the reasons described in § 3, when presented with a magnitude-limited catalog, correlations at fixed luminosity are useful because they are unbiased by the selection. When luminosity is not one of the variables, then forward and inverse correlations may be equally interesting and equally biased. For example, in the Kormendy (1977) relation, $\langle R - R_* | \mu - \mu_* \rangle$ may be just as interesting as $\langle \mu - \mu_* | R - R_* \rangle$. The slopes of the two relations are, of course, simply related to each other. In fact, it may be preferable to study the relations that are defined by the principle axes of the ellipse in (R, μ) space that the galaxies populate. The directions of these axes are obtained by computing the eigenvalues and vectors of the covariance matrix associated with the sizes and surface brightnesses. To illustrate, the eigenvalues of the 2×2 covariance matrix associated with the Kormendy relation are

$$\sigma_\pm^2 = \left(\sigma_R^2 + \sigma_\mu^2 \pm \sqrt{D_{\mu R}}\right)\Big/ 2 ,$$

where we have set $D_{\mu R} = (\sigma_R^2 - \sigma_\mu^2)^2 + (2\sigma_R\sigma_\mu\rho_{\mu R})^2$. The $+/-$ eigenvalues give the dispersions along and perpendicular to the major axis of the ellipse. The long axis of the ellipse describes the mean relation, $(R - R_*) = S_K(\mu_o - \mu_*)$, where

$$S_K = \frac{\sigma_R^2 - \sigma_\mu^2 + \sqrt{D_{\mu R}}}{2\sigma_R\sigma_\mu\rho_{\mu R}} .$$

With obvious changes of variables, analogous expressions can be derived for all the correlations presented earlier, although we do not show them here.

The Kormendy relation in our sample is shown in Figure 10. The dashed lines show forward and inverse fits to the data, i.e., the mean size at fixed surface brightness, and the mean surface brightness at fixed size. The parameters of the fits are affected by the magnitude limit of the catalog. To estimate the effect of the magnitude-limit cut on this relation, we compute the direct and inverse fits to the Kormendy relation in the simulated complete and magnitude-limited samples we describe in Appendix A. The dotted line in
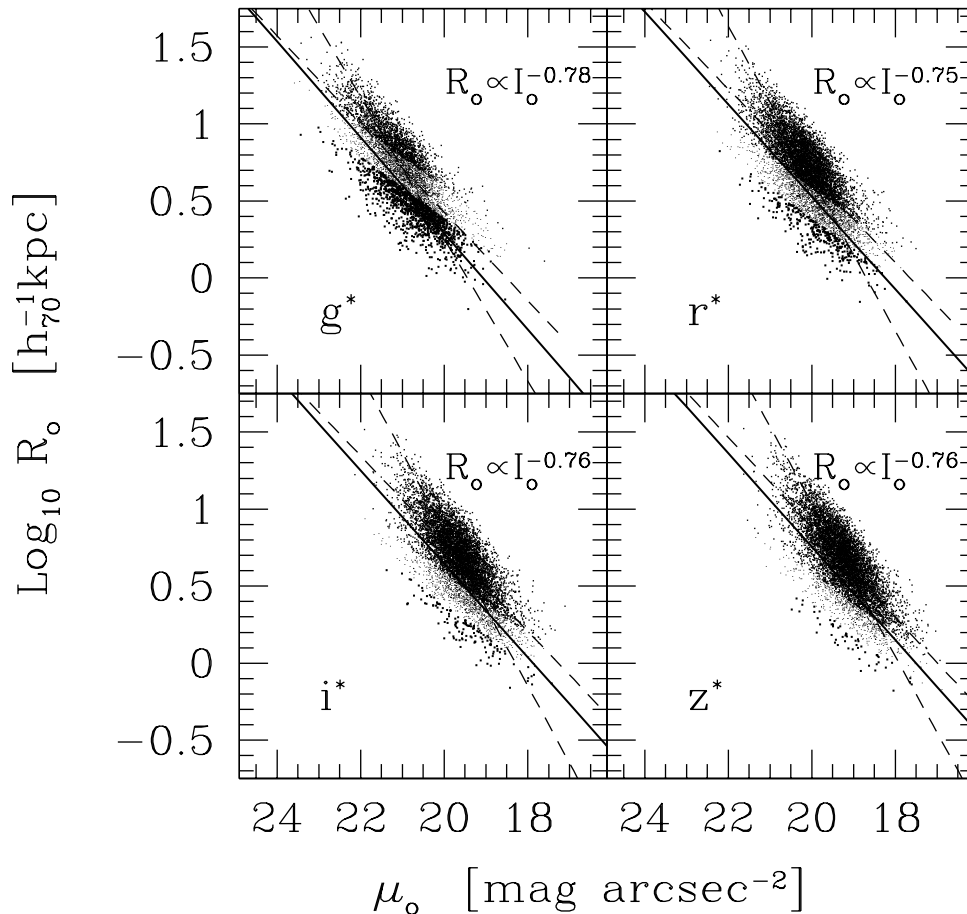

Fig. 10.—Relation between effective radius and surface brightness. Short dashed lines show forward and inverse fits to this relation. The zero points of these fits are strongly affected by the magnitude limit of our sample. To illustrate, solid line shows the maximum likelihood estimate of the relation in the simulated complete catalog from which the magnitude-limited catalog, shown by the dotted line, was drawn.

Figure 10 shows the direct fit to the magnitude-limited simulations (it can hardly be distinguished from the fit to the data).

In comparison, the maximum likelihood estimate of the true direct relation provides a very good description of the relation in the complete simulations in which there is no magnitude limit: it is shown as the solid line. Notice that the dashed and dotted lines have approximately the same slope as the solid line: the magnitude limit hardly affects the slope, although it changes the zero point dramatically. At fixed surface brightness, the typical R_o is significantly larger in the magnitude-limited sample than in the complete sample. This happens because lines of constant luminosity run downward and to the right with slope $-1/5$, so that changes in luminosity act approximately perpendicular to the relation.

This shows that although linear regression fits to the data provide good estimates of the true slope of the Kormendy relation, they provide bad estimates of the true zero point. In comparison, the maximum likelihood technique, which accounts for the selection on apparent magnitudes, is able to estimate the slope and the zero point correctly.

Another interesting correlation is that between the effective mass and density defined in equations (4) and (6). A little algebra shows that $\langle \Delta_o | M_o \rangle \propto M_o^{-0.52}$. Figure 11 shows forward and inverse fits to this relation. The characteristic density Δ_c of halos seen in numerical simulations of hierarchical clustering scales with halo mass: $\langle \Delta_c | m \rangle \sim 9^3 (m/m_*)^{-0.4}$ (Bullock et al. 2001), which is qualitatively similar to the scaling of effective density with effective mass in our sample. The scatter in characteristic densities at fixed halo mass, ~ 0.33 dex, is also rather similar to the scatter in effective densities at fixed mass. These coincidences may provide important clues to how early-type galaxies formed.

In contrast to the Faber-Jackson, radius-luminosity, Kormendy, and mass-density relations, the relations between luminosity and mass and luminosity and density involve three variables. Is there some combination of these variables that provides the least scatter? The eigenvectors and eigenvalues of the 3×3 matrix $\mathscr{C}$ give the directions of the principle axes of the ellipsoid in (M, R, V) space that the early-type galaxies populate. One of the eigenvalues of $\mathscr{C}$ is considerably smaller than the others, suggesting that the galaxies populate a two-dimensional plane in (M, R, V) space. The eigenvectors show that the plane is viewed edge-on in the projection

$$-0.4(M - M_*) = \alpha(R - R_*) + \beta(V - V_*), \quad (13)$$

where M_*, R_*, and V_* were given in Table 1, and the coefficients α and β, and the thickness of the plane in this projection, σ_{MRV}, are given in Table 3. Section 3 shows that a scatter plot of luminosity versus mass is considerably tighter than plots of M versus log R_o or log σ. The eigenvectors of $\mathscr{C}$ show that this is because the M versus $R + 2V$ projection is actually quite close to the edge-on projection. It is interesting that this plane is only about 10% thicker than the
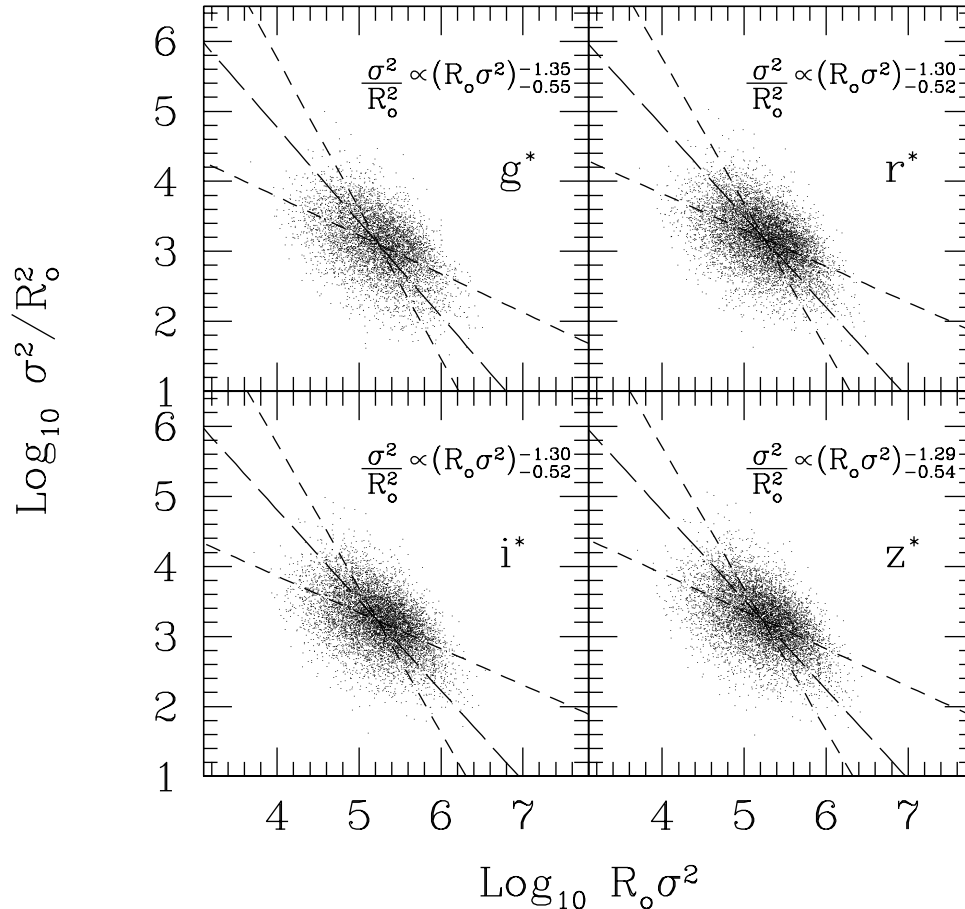

Fig. 11.—Relation between effective density and effective mass. Short dashed lines show forward and inverse fits to this relation. The more massive galaxies are less dense. Text in top right of each panel shows the slopes of the dashed lines.

TABLE 3

COEFFICIENTS α AND β THAT DEFINE THE
PROJECTION OF MINIMUM SCATTER, σ_{MRV},
IN THE SPACE DEFINED BY ABSOLUTE
MAGNITUDE, AND THE LOGARITHMS OF
THE SIZE AND VELOCITY DISPERSION

Band	α	β	σ_{MRV}
g^*...............	0.76	1.94	0.063
r^*...............	0.79	1.93	0.058
i^*...............	0.82	1.89	0.054
z^*...............	0.81	1.90	0.054

fundamental plane relation between R_o, I_o, and σ which is the subject of Paper III.

5. DISCUSSION AND CONCLUSIONS

We have studied the properties of $\sim$9000 early-type galaxies over the redshift range $0 \leq z \leq 0.3$ using photometric (the g^*, r^*, i^*, and z^* bands) and spectroscopic observations. The intrinsic distributions of luminosity, velocity dispersion, and half-light radius of the galaxies in our sample are each well described by Gaussians in absolute magnitude, $\log \sigma$, and $\log R_o$.

A maximum likelihood analysis of the joint distribution of luminosities, sizes, and velocity dispersions suggests that the population at higher redshifts is slightly brighter than the population nearby, and that the change with redshift is faster in the shorter wave bands: If $M_*(z) = M_*(0) - Qz$, then $Q = 1.15$, 0.85, and 0.75 in g^*, r^*, and i^*. This evolution is sufficiently weak that, relative to their values at the median redshift ($z \sim 0.15$) of our sample, the sizes, surface brightnesses, and velocity dispersions of the early-type galaxy population at lower and higher redshifts has evolved little. The fact that we see so little evolution in the luminosities argues for a relatively high formation redshift: Bruzual & Charlot (2003) single burst stellar population synthesis models indicate that $t_{form} \sim 9$ Gyr. This is consistent with the model we use to make K-corrections in Paper I and is also consistent with the formation time estimates based on the fundamental plane in Paper III and galaxy colors and spectral line indexes in Paper IV.

We find that $\langle \sigma | L \rangle \propto L^{1/4}$ and $\langle R_o | L \rangle \propto L^{3/5}$ (see Table 2 for the exact coefficients, and Figs. 4 and 5 for the fits). Galaxies that are slightly larger than expected (given their luminosity) have smaller velocity dispersions than expected (Fig. 6). This is expected if galaxies are in virial equilibrium.

A plot of luminosity versus effective mass $M_o = 2R_o\sigma^2/G$ is substantially tighter than either the L-σ or the L-R_o relations. It has a slope that is slightly shallower than unity. In particular, on scales of a few kiloparsecs, $L \propto M_o^{0.86}$, approximately independent of wave band (Fig. 7). This complements recent SDSS weak-lensing analyses (McKay et al. 2001) that suggest that mass is linearly proportional to luminosity in these same wave bands, but on scales that are 2 orders of magnitude larger ($\sim$260 h^{-1} kpc). Together, these two measurements of the mass-to-light ratio can be used to provide a constraint on the density profiles of dark matter halos.

A plot of luminosity versus effective density $\Delta_o \propto \sigma^2/R_o^2$ shows that $\langle \Delta_o | L \rangle \propto L^{-3/4}$ (Fig. 8). Moreover, a maximum likelihood analysis suggests that the more massive galaxies are less dense: $\langle \Delta_o | M_o \rangle \propto M_o^{-0.52}$ (Fig. 11). This is qualitatively similar to a trend seen in numerical simulations of hierarchical clustering: more massive halos tend to be less centrally concentrated (Navarro et al. 1997). This coincidence may provide an important clue to how early-type galaxies formed.

The Kormendy relation between size and surface brightness has approximately the same slope $\langle R_o | I_o \rangle \propto I_o^{-0.77}$ in all four SDSS bands (Fig. 10). Our maximum likelihood analysis, and measurements made in mock catalogs that reproduce all the observed scalings of the data set (a procedure for generating such catalogs is described in Appendix A), show that the zero point of this relation is strongly affected by the magnitude limit of the sample (§ 4.3).

APPENDIX A

SIMULATING A COMPLETE SAMPLE

This appendix describes how to use our knowledge of the covariance matrix $\mathscr{C}$ to simulate mock galaxy samples that have the same correlated observables as the data. We use these mock samples to estimate the effect of the magnitude-limit cut on the relations we wanted to measure in the main text.

The observed parameters L, R_o, and σ of each galaxy in our sample are drawn from a distribution, say, $\phi(M, R, V|z)$, where M is the absolute magnitude, $R = \log R_o$, and $V = \log \sigma$. We show in § 3 that $\phi(M, R, V|z) = p(R, V|M, z)\phi(M|z)$, where $\phi(M|z)$ is the luminosity function at redshift z, and the distribution of R and V at fixed luminosity is, to a good approximation, a bivariate Gaussian. The maximum likelihood estimates of the parameters of the luminosity function and of the bivariate distribution at fixed luminosity can be obtained from Table 1.

To make the simulations, we must assume that, when extrapolated down to luminosities that we do not observe, these relations remain accurate. Assuming this is the case, we draw M from the Gaussian distribution that we found was a good fit to $\phi(M|z)$ (§ 2). We then draw R from the Gaussian distribution with mean $\langle R|M \rangle$ and dispersion $\sigma^2_{R|M}$. Finally, we draw V from a Gaussian distribution with mean and variance that accounts for the correlations with both M and R. In practice, we draw three zero mean unit variance Gaussian random numbers: g_0, g_1, and g_2, and

then set

$$M = M_* + \sigma_M g_0 ,$$

$$R = R_* + \frac{(M - M_*)}{\sigma_M} \sigma_R \rho_{RM} + g_1 \sigma_R \sqrt{1 - \rho_{RM}^2}$$

$$V = V_* + \frac{(M - M_*)}{\sigma_M} \xi_{MV} + \frac{(R - R_*)}{\sigma_R} \xi_{RV} + g_2 \, \sigma_{V|RM} ,$$

where

$$\xi_{MV} = \sigma_V \frac{(\rho_{VM} - \rho_{RM}\rho_{RV})}{(1 - \rho_{RM}^2)} ,$$

$$\xi_{RV} = \sigma_V \frac{(\rho_{RV} - \rho_{RM}\rho_{VM})}{1 - \rho_{RM}^2} ,$$

$$\sigma_{V|RM} = \sigma_V \sqrt{\frac{1 - \rho_{RM}^2 - \rho_{RV}^2 - \rho_{VM}^2 + 2\rho_{RV}\rho_{VM}\rho_{RM}}{1 - \rho_{RM}^2}} .$$

Because each simulated galaxy is assigned a luminosity and size, its surface brightness is also fixed: $\mu = M + 5R +$ constant.

If we generate a catalog in r^*, then we can also generate colors using the parameters given in Table 1 of Paper IV. Specifically, generate a Gaussian variate g_3, and then set

$$C = C_* + \xi_{CM}(M - M_*)/\sigma_M + \xi_{CV}(V - V_*)/\sigma_V$$
$$+ g_3 \sigma_{C|MV} ,$$

where ξ_{CM}, ξ_{CV}, and $\sigma_{C|MV}$ are defined analogously to ξ_{MV}, ξ_{RV}, and $\sigma_{V|RM}$ above. Inserting the values from Table 1 of Paper IV shows that $\xi_{CM} \approx 0$, and $\sigma_{C|MV} \approx \sigma_{C|V} = \sigma_C (1 - \rho_{VM}^2)^{1/2}$: the mean color is determined by the velocity dispersion and not by the absolute magnitude.

Passive evolution of the luminosities and colors is incorporated by adding the required z-dependent shift to M and C after the sizes and velocity dispersions have been generated.

This complete catalog can be used to simulate a magnitude-limited catalog if we assign each mock galaxy a redshift, assuming a world model and homogeneity. Let m_{min} and m_{max} denote the apparent magnitude limits of the observed sample. Let M_{bright} denote the absolute magnitude of the most luminous galaxy we expect to see in our catalog. Because the luminosity function cuts off exponentially at the bright end, we can estimate this by setting $M_{bright} \approx M_* + 5\sigma_M$. This means that the most distant object that can conceivably make it into the magnitude-limited catalog lies at a luminosity distance of about $d_{Lmax} = 10^{(m_{max} - M_{Bright} - 25)/5}$, from which the maximum redshift z_{max} can be determined. If the comoving number density of mock galaxies is to be independent of redshift, we must assign redshifts as follows. Draw a random variate u_1 distributed uniformly between zero and one and set $d_{Com} = u_1^{1/3} d_{Lmax}/(1 + z_{max})$. The redshift z can be obtained by inverting the $d_{Com}(z; \Omega, \Lambda)$ relation. The apparent magnitude of this mock galaxy is $m = M + 5 \log d_L + 25 + K(z)$, where $K(z)$ is the K-correction. If $m_{min} \le m \le m_{max}$, then this galaxy would have been observed; we add it to the subset of galaxies from the complete catalog that would have been observed in the magnitude-limited catalog.

If our simulated catalogs are accurate, then plots of magnitude, size, surface brightness, and velocity dispersion versus redshift made using our magnitude-limited subset should look very similar to the SDSS data set shown in Figure 12 of Paper I. In addition, dN/dz in the simulated magnitude-limited subset should be similar to that in Figure 3. Furthermore, any correlations between observables in the magnitude-limited subset should be just like those in the actual SDSS data set. If they are, then one has good reason to assume that similar correlations measured in the complete, rather than the magnitude-limited simulation, represent the true correlations between the parameters of SDSS galaxies, corrected for selection effects. In this way, the simulations allow one to estimate the impact that the magnitude-limited selection has when estimating correlations between early-type galaxy observables.

We have verified that our simulated magnitude-limited catalogs have similar dN/dz distributions to those observed, and the simulated σ and R_o versus z plots show the same selection cuts at low velocities and sizes as do the observed data. The distribution of apparent magnitudes, angular sizes, and velocity dispersions in the magnitude-limited simulations are very similar to those in the real data. The simulated parameters also show the same correlations at fixed luminosity as the data. Maximum likelihood analysis on the simulations produces an estimate of the covariance matrix that is similar to that of the data. Therefore, we are confident that our simulated complete catalogs have correlations between luminosity, size, and velocity dispersion that are similar to the data.

APPENDIX B

COMPOSITE VOLUME-LIMITED CATALOGS

Our parent sample is magnitude-limited; unless accounted for, this will introduce a bias into a number of correlations we study in this series of papers. For this reason, we often present results measured in a few volume, limited subsamples. Because of the cuts at both the faint and the bright ends of the catalog, each volume-limited subsample used in the main text spans only a small range in luminosity. However, because the galaxies in our sample luminosity show little or no evolution relative to the values at the median redshift of the sample, we can extend this range in either of three ways.

One method is to construct a composite volume-limited catalog by stacking together smaller volume-limited subsamples that are adjacent in redshift and in luminosity, but that do not overlap at all. Let V_i denote the volume of the ith subsample, and let N_i denote the number of galaxies in it. A conservative approach is to randomly choose the galaxies in V_i with probability proportional to min $(V_i)/V_i$, where min (V_i) denotes the volume of the smallest of the subsamples. This has the disadvantage of removing much of the data, but, because our data set is so large, we can afford this luxury. A more cavalier approach is to choose all the galaxies in the largest V_i, all the galaxies in the other V_j, and to generate a set of additional galaxies by randomly choosing one of the N_j galaxies in V_j, adding to each of its observed parameters a Gaussian random variate with dispersion given by the quoted observational error and repeating this $N_j[\max (V_j)/V_j - 1]$ times. A final possibility is to weight all the galaxies in V_i (even those that were not in the volume-limited subsample) by the inverse of the volume in which they could have been observed ($V_{max} - V_{min}$). We chose the first, most conservative option.

By piecing together three volume-limited subsamples, we were able to construct composite catalogs of about 10^3 objects each. Because the completeness limits are different in the different bands, the composite catalogs are different for each band. In addition, because any one composite catalog is obtained by subsampling the set of eligible galaxies, by subsampling many times, we can generate many realizations of a composite catalog. This allows us to estimate the effects of sample variance on the various correlations we measure.

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, 1866 (Paper III)
———. 2003c, AJ, 125, 1882 (Paper IV)
Blanton, et al. 2001, AJ, 121, 2358
Bruzual A., G., & Charlot, S. 2003, in preparation
Bullock, J. S., Kolatt, T. S., Sigad, Y., Somerville, R. S., Kravtsov, A. V., Klypin, A. A., Primack, J. R., & Dekel, A. 2001, MNRAS, 321, 559
Colberg, J. M., et al. 2000, MNRAS, 319, 209
Efstathiou, G., Ellis, R. S., & Peterson, B. S. 1988, MNRAS, 232, 431
Faber, S. M., & Jackson, R. 1976, ApJ, 204, 668
Forbes, D. A., & Ponman, T. J. 1999, MNRAS, 309, 623
Hernquist, L. 1990, ApJ, 356, 359
Kormendy, J. 1977, ApJ, 218, 333
Lin, H., Yee, H. K. C., Carlberg, R. G., Morris, S. L., Sawicki, M., Patton, D. R., Wirth, G., & Shepherd, C. W. 1999, ApJ, 518, 533
Madgwick, D. S., et al. 2002, MNRAS, 333, 133
McKay, T. A., et al. 2001, ApJ, submitted (astro-ph/0108013)
Navarro, J. F., Frenk, C. S., & White, S. D. M. 1997, ApJ, 490, 493
Pahre, M. A., Djorgovski, S. G., & de Carvalho, R. R. 1998, AJ, 116, 1591
Prugniel, P., & Simien, F. 1996, A&A, 309, 749
Saglia, R. P., Colless, M., Burstein, D., Davies, R. L., McMahan, R. K., & Wegner, G. 2001, MNRAS, 324, 389
Sandage, A., & Perelmuter, J.-M. 1990, ApJ, 361, 1
Sandage, A., Tammann, G. A., & Yahil, A. 1979, ApJ, 232, 352
Schade, D., Barientos, L. F., & López-Cruz, O. 1997, ApJ, 477, L17
Schechter, P. L. 1980, AJ, 85, 801
Sodré, L., Jr., & Lahav, O. 1993, MNRAS, 260, 285
van Dokkum, P. G., & Franx, M. 2001, ApJ, 553, 90