THE MAGNESIUM–VELOCITY DISPERSION RELATION AND THE GENESIS OF EARLY-TYPE GALAXIES

GUY WORTHLEY AND MAELA COLLOBERT
Washington State University, 1245 Webster Hall, Pullman, WA 99163-2814; gworthey@wsu.edu
Received 2002 June 21; accepted 2002 November 26

ABSTRACT
Available data on the magnesium–velocity dispersion (Mg–σ) relation for ~2000 early-type galaxies is analyzed. As noted previously, the Mg residuals from a fitted line are roughly Gaussian near the median but have an asymmetric blue tail, probably from subpopulations of relatively young stars. We define statistics for scatter and asymmetry of scatter in the Mg dimension and find impressive uniformity among data sets. We construct models of galaxy formation built to be as unbiased as possible toward the question of the importance of mergers in the formation of early-type galaxies. The observational constraints (Mg–σ width, asymmetry, and mean Mg strength, plus mean age and width of abundance distribution) are severe enough to eliminate almost all models. Eliminated are models with merger rates proportional to \((1 + z)^n\) with \(n > 0\), models that assume early formation followed by recent drizzling of new stars, merger-only models in which the number of mergers exceeds ~80, merger-only models with less than ~20 mergers, and models with a cold dark matter power spectrum (at least within our approximations). The most successful models were those with merger probability constant or mildly declining with time, with the number of mergers needed to form the galaxy around 50 and gas fractions of ~0.2–0.35. These models are characterized by mean light-weighted ages of 8–9 Gyr (consistent with spectroscopic studies), an abundance distribution that does not exceed local constraints, and a look-back time behavior nearly indistinguishable from passive evolution of old stellar populations. Our simulations suggest that the evolution of median Mg index strength is not a good discriminator between mergers and passive evolution and that better discriminators such as Mg–σ scatter and asymmetry require \(N > 1000\) sample sizes with accuracies similar to today’s local measurements.

Subject headings: galaxies: abundances — galaxies: elliptical and lenticular, cD — galaxies: evolution — galaxies: formation — galaxies: stellar content

1. INTRODUCTION
The process of formation of early-type galaxies is still largely unknown. Under study since Larson (1975), the choice between formation of elliptical galaxies by a single brief collapse or via merging is still disputed in the astronomical literature.

Some properties of early-type galaxies are very uniform, suggesting a homogeneous and probably ancient origin. Elliptical and S0 galaxies have smooth light profiles and very low gas and dust content, like overgrown globular clusters. As a class, they display scaling relations such as a tight fundamental plane (FP; Djorgovsky & Davies 1987; Dressler et al. 1987) among the variables luminosity, size, and velocity dispersion, a narrow Mg–σ relation between integrated starlight absorption line strength and velocity dispersion (Bender, Burstein, & Faber 1993; Ziegler & Bender 1997), and an orderly color-magnitude relation (Bower, Lucey, & Ellis 1992). Many authors note that elliptical galaxies nearby and at intermediate redshift have properties consistent with old, passively evolving systems (Stanford, Eisenhardt, & Dickinson 1998; van Dokkum & Franx 1996; Kelson et al. 1997; Ellis et al. 1997; Kodama & Arimoto 1997; Kodama et al. 1998; Bender et al. 1998; van Dokkum et al. 1998).

Confusingly, there is also a great deal of evidence that favors a messier, ongoing formation process for early-type galaxies. Some galaxy mergers are caught in the act (Toomre 1977), with the theoretical expectation that the remnant will soon relax to resemble an elliptical with an exponential light profile. The tails, ripples, shells, and other morphological aftereffects of merging are seen in many nearby elliptical galaxies (Schweizer et al. 1990; Schweizer & Seitzer 1992; Goudfrooij et al. 2001; Bardelli et al. 2002; Markevitch et al. 2002). Measurements of the integrated stellar absorption features compared with stellar population models indicate light-weighted mean ages for the near-nuclear regions of elliptical galaxies that range from ancient through a median of \(\sim 7\) Gyr to a few that appear less than 1 Gyr old (González 1993; Worthey 1997; Terlevich & Forbes 2002). Young stellar populations are much brighter than old populations, so that a relatively modest burst of young stars may be able to skew the mean age of an essentially old galaxy to appear much younger than a mass-weighted mean age, but the presence of any youthful subpopulation contradicts the uniform-and-old hypothesis.

With this paper, we attempt to bring clarity to the issue. To simplify, we consider only the relation between the central velocity dispersion (σ) and the strength of the integrated stellar Mg and MgH features around 5100 Å. The advantages of this relation are its distance independence, its almost total insensitivity to dust extinction, and its small scatter (Jørgensen, Franx, & Kjøgstad 1996; Bender et al. 1998). One disadvantage is that both age and metallicity have a similar effect on the Mg strength (Worthey 1994; Forbes et al. 2001; Colless et al. 1999); an older galaxy has a stronger Mg feature, but Mg strength can also be increased

\(^1\)Observatory of Paris, Meudon, Place J. Janssen, 92195 Meudon Cedex, France.
by Mg abundance. It is clear that the origin of the relation itself is primarily one of abundance: larger galaxies have more heavy elements. Intrinsic scatter exists in the $\text{Mg}_2$-$\sigma_0$ relation ($\sigma_0$ refers to an aperture-corrected central velocity dispersion). This scatter is not correlated with the appearance of the galaxies, with the degree of velocity anisotropy, with their environment (Dressler et al. 1987; Burstein, Faber, & Dressler 1990), or with deviation of the objects within (or perpendicular) to the FP. Bender et al. (1993) interpreted this as a combination of age and/or metallicity differences among elliptical galaxies with similar luminosity, and Colless et al. (1999) find that both mean age and metallicity vary, probably in a correlated manner with younger-appearing galaxies tending to be somewhat more metal-rich. This correlation is also seen with Balmer-metal indicators (Worthey, Trager, & Faber 1995; Terlevich & Forbes 2002).

The residuals of $\text{Mg}_2$ have a Gaussian core, but the galaxies with weaker $\text{Mg}_2$ at given $\sigma_0$ form a tail; i.e., the residual distribution is skewed with an excess of Mg-weak galaxies (Bender et al. 1993). This is suggestive of late star formation: a youthful galaxy will have weak Mg strength from the presence of bright, hot stars with weak Mg strength and will appear as a low outlier.

We test two opposing hypotheses. In both cases, the Mg- $\sigma$ relation is a mass-abundance correlation. In the ancient formation hypothesis, elliptical galaxies formed at the epoch of galaxy formation at high redshift (we assume $z_f = 5$). The Mg- $\sigma$ relation was put in place at that time. Furthermore, the symmetric portion of the scatter was set at formation. (We could be even more extreme and suggest that the asymmetric part of scatter was also imprinted at formation by an asymmetric abundance distribution, but the implication of this is trivial: except for the effect of passive stellar evolution, the Mg- $\sigma$ relation would remain unchanged back to $z_f$.) In the ancient formation hypothesis, mergers, as inconsequential as possible, cause occasional bursts of star formation that drive a few galaxies to weak Mg strength and cause the asymmetry of the observed relation. In the maximum merger hypothesis, the underlying Mg- $\sigma$ relation is very tight and all scatter—symmetric or not—is caused by merging events (convolved with observational error).

To compare these two views, we first collect and analyze published Mg- $\sigma$ data sets of E and S0 galaxies from field, group, and cluster environments. We use these data to find robust statistics to help us evaluate the expected asymmetrical scatter. In § 3, we describe special-purpose models that predict Mg- $\sigma$ relations given input parameters such as the number of mergers and gas fraction of each merger. Results are discussed in § 4, and a summary of conclusions is given in § 5.

### 2. OBSERVATIONAL DATA

We analyze the Mg- $\sigma$ relation in seven different data sets for E and S0 galaxies. There are three Lick/IDS system Mg indices: $\text{Mg}_b$, $\text{Mg}_2$, $\text{Mg}_1$ (Worthey et al. 1994), plus the variant ($\text{Mg}_2^*$) index introduced by Davies et al. (1987) and defined by

$$\langle \text{Mg}_2^* \rangle = 0.6 \text{ Mg}_2 + 0.4 \text{ Mg}_1^*,$$

with $\text{Mg}_1^* = 0.03 + 2.1 \text{ Mg}_1 - 62 \text{ Mg}_1^*$. The variant $\langle \text{Mg}_2^* \rangle$ uses the additional signal available in $\text{Mg}_b$ to improve the precision of the $\text{Mg}_2$ index. There does not appear to be any substantial mismatch between $\langle \text{Mg}_2^* \rangle$ and $\text{Mg}_2^*$ definitions in ultimate index behavior, although differences are detectable (Trager et al. 1998). We summarize the literature data that we collected with a few words and a list of characteristics in Table 1.

**Data set 1.**—is a subset of data extracted from Trager et al. (1998). It consists of 255 galaxies observed at the Lick

---

**Table 1**

| Paper                        | Number of Galaxies in Sample | Number of Galaxies Used for Statistics | Value of Mg Index ($\sigma = 300 \text{ km s}^{-1}$) | Slope of Relation (mag/dex) | Standard Deviation | Skewness | Average Deviation | Average Deviation Ratio | $\sigma_x$ (gauss) |
|------------------------------|------------------------------|----------------------------------------|---------------------------------------------------|-----------------------------|--------------------|----------|------------------|------------------------|-------------------|
| Prugniel & Simien 1996........| 644                          | 488                                    | 0.314                                              | 0.181                       | 0.023              | -1.017   | 0.017            | 1.120                  | 0.013             |
| Trager et al. 1998 (Lick).....| 255                          | 184                                    | 0.339                                              | 0.183                       | 0.032              | -1.750   | 0.022            | 1.544                  | 0.014             |
| Wegner et al. 1999............| 147                          | 130                                    | 0.331                                              | 0.208                       | 0.021              | -0.324   | 0.017            | 1.210                  | 0.013             |
| Dressler et al. 1991..........| 137                          | 107                                    | 0.318                                              | 0.170                       | 0.023              | -0.722   | 0.018            | 1.538                  | 0.011             |
| Bernardi et al. 1998 (all)....| 931                          | 645                                    | 0.329                                              | 0.219                       | 0.033              | -0.184   | 0.025            | 1.230                  | 0.020             |
| Bernardi et al. 1998 (field)  | 631                          | 458                                    | 0.331                                              | 0.236                       | 0.033              | -0.186   | 0.025            | 1.151                  | 0.020             |
| Bernardi et al. 1998 (group)  | 128                          | 91                                     | 0.333                                              | 0.266                       | 0.031              | -0.369   | 0.024            | 1.361                  | 0.015             |
| Bernardi et al. 1998 (cluster)| 151                          | 83                                     | 0.331                                              | 0.199                       | 0.032              | -0.052   | 0.023            | 1.128                  | 0.013             |
| Hudson 2001....................| 528                          | 391                                    | 0.332                                              | 0.172                       | 0.018              | -0.369   | 0.014            | 1.105                  | 0.013             |
| $\langle \text{Mg}_b \rangle$ |                              |                                        |                                                   |                             |                    |          |                  |                        |                  |
| Trager et al. 1998 (Lick).....| 255                          | 184                                    | 5.091                                              | 2.824$^a$                  | 0.526              | -0.971   | 0.378            | 1.220                  | 0.2               |
| Wegner et al. 1999.............| 164                          | 144                                    | 5.260                                              | 2.916$^a$                  | 0.422              | -0.573   | 0.314            | 0.990                  | 0.3               |
| $\langle \text{Mg}_2 \rangle$ |                              |                                        |                                                   |                             |                    |          |                  |                        |                  |
| Trager et al. 1998 (Lick).....| 255                          | 184                                    | 0.338                                              | 0.171                       | 0.032              | -2.472   | 0.021            | 1.788                  | 0.011             |
| Davies et al. 1987.............| 470                          | 362                                    | 0.324                                              | 0.172                       | 0.038              | -5.545   | 0.020            | 1.609                  | 0.015             |

---

$^a$ Slope of the relation in $\text{Å}$/dex for $\text{Mg}_b$.

$^b$ $\langle \text{Mg}_2 \rangle = 0.6 \text{ Mg}_2 + 0.4 \text{ Mg}_2^*$, with $\text{Mg}_2^* = 0.03 + 2.10 \text{ Mg}_1 - 62 \text{ Mg}_1^*$. 

---
Observatory between 1972 and 1984. Average errors for Mg\(_2\), Mg\(_b\), and (Mg\(_i\)) indices were 0.008 mag, 0.27 Å, and 0.007 mag, respectively. For velocity dispersion, the error is less than 10% (mean 8.9%). The advantage of these data is that they define the Lick system and are very homogeneous. The disadvantage is that no aperture corrections for distance have been included. This appears to have no measurable consequence (see Table 1). Propagating velocity dispersion error to increase index error yields errors of 0.010 mag, 0.29 Å, and 0.009 mag for Mg\(_2\), Mg\(_b\), and (Mg\(_i\)), respectively.

Data set 2.—is part of the ENEAR database from Bernardi et al. (1998). It is an independent sample. It is subdivided into three data sets according to whether the galaxies were in field, group, or cluster environments. Measurements of \(\sigma\) are accurate to 5%–13% (mean 10%), Mg\(_2\) to 0.005–0.011 mag (mean 0.008 mag).

Data set 3.—Dressler, Faber, & Burstein (1991) have presented 136 cluster and group elliptical and S0 galaxies in the direction of the large-scale streaming flow attributed to the great attractor. The published errors are 0.05 dex for log \(\sigma\) and 0.017 mag for Mg\(_2\), implying total errors in the Mg\(_2\) direction of 0.019 mag.

Data set 4.—The 528 galaxies available with the Mg\(_2\) index in Hudson et al. (2001) are from 56 galaxy clusters. Uncertainties in \(\sigma\) are 5% and 0.009 mag for Mg\(_2\), increasing to 0.010 mag if \(\sigma\) error is propagated.

Data set 5.—We culled 147 (Mg\(_3\)) and 164 (Mg\(_b\)) galaxies taken from the EFAR data set (Wegner et al. 1999). The whole sample is much larger, but we considered only those galaxies that had errors of 7% or less in \(\sigma\) (mean error 5%). Errors for Mg\(_2\) were 0.010 mag and for Mg\(_b\), 0.20 Å. Propagating velocity dispersion error, we find total errors of 0.011 for Mg\(_3\) and 0.21 Å for Mg\(_b\).

Data set 6.—Davies et al (1987, one of the Seven Samurai) computed \(\sigma\) and lines indices for 600 galaxies from all environments. The scatter of measurements indicates a mean \(\sigma\) uncertainty of 10% and 0.009 mag for (Mg\(_3\)). We use an updated Seven Samurai data set kindly provided by D. Burstein (1999, private communication). This data set includes much of the Lick data (data set 1, above) as well as data taken with other telescopes. The effective (Mg\(_3\)) error is 0.011 mag.

Data set 7.—This data set of 644 elliptical galaxies and S0 galaxies was extracted from Prugniel & Simien (1996). It is a literature compilation and nominally a superset of both the Seven Samurai data and the Lick data but with different procedures for homogenization. The data are in \(\sigma\) (mean error 6.7%) and in Mg\(_2\) (mean error 0.008 mag, total error 0.009 mag).

With the data in hand, we then sought suitable statistics to describe the scatter and asymmetry of the residuals from a linear fit to the data. The linear fit itself was tried several ways. A line fit via least-squares method was too sensitive to the presence of outliers. Even when we clipped the outliers in a rejection loop, we judged that the resulting line was too unstable. Finally, we binned the data in log \(\sigma\), found the median Mg index value of each bin, weighted each datum by 1/(\(N^3\))\(^{1/2}\), and performed a least-squares fit on the array of medians. This method is outlier-insensitive and very stable. Unless specified otherwise in this paper, the fit was computed only for 150 < \(\sigma\) < 320 km s\(^{-1}\), medium-sized and large galaxies. In Table 1, we report the number of galaxies really used for our fit.

For convenience, we would prefer to pretend that all of the scatter and error is in the Mg direction if we can. Propagation of velocity dispersion errors indicates that approximately 0.002 mag of additional error is added to the Mg error (which is about 0.009 mag), so velocity dispersion is a minor component in the total error budget. One remaining concern is that \(\sigma\) errors grow faster than Mg index errors for smaller galaxies. We have eliminated this problem by simply dropping all of the very small galaxies from our analysis. We therefore quote errors in the Mg dimension only, modestly inflated by velocity dispersion error. This paragraph applies only to observational errors: intrinsic structural differences in galaxies may be a very important part of the intrinsic scatter in the Mg-\(\sigma\) relation.

There is a small amount of drift in Mg strength among the data sets. The Mg 300 column of Table 1 is the value of the Mg index fit for \(\sigma = 300\) km s\(^{-1}\). This quantity is 0.327 ± 0.039 for Mg\(_3\), 5.1 ± 0.085 for Mg\(_b\), and 0.333 ± 0.007 for (Mg\(_3\)) averaged over the data sets. For illustration, we plot three data sets and the computed median line in Figure 1.

Once the line was fitted, we sought statistics to characterize the distribution of Mg index residuals in terms of both

Fig. 1.—Observational Mg-\(\sigma\) relation for three of our data sets: Bernardi et al. (1998; all galaxies), Hudson et al. (2001), and Davies et al (1987). The median fit line is shown within the limits we chose for calculating statistics: 150 km s\(^{-1}\) < \(\sigma\) < 320 km s\(^{-1}\).
width and asymmetry. For width, the standard deviation has the advantage that everyone is familiar with it, but it is quite sensitive to outliers due to its \((x_j - \bar{x})^2\) dependence. We prefer the average deviation or ADev (mean absolute deviation). It is an estimator of width less sensitive to outliers than the standard deviation. It is defined by

\[
\text{ADev} = \frac{1}{N} \sum_{j=1}^{N} |x_j - \bar{x}| ,
\]

but in this case, we substitute the median-fit line value for \(\bar{x}\) instead of the traditional mean.

For measuring asymmetric tails, we need another statistic. We considered the skewness, a dimensionless number that describes only the shape of the distribution. We take the usual definition

\[
\text{Skewness} = \frac{1}{N} \sum_{j=1}^{N} \left( \frac{x_j - \bar{x}}{\sigma} \right)^3 ,
\]

where \(\sigma\) is the distribution’s standard deviation and \(\bar{x}\) is the fitted line. A positive value of skewness signifies a distribution with an asymmetric tail extending out toward more positive \(x\). A negative value signifies a distribution whose tail extends out toward more negative \(x\). We expect a negative skewness if galaxies scatter preferentially toward weaker Mg index values.

Because we wish to be outlier insensitive, we also define an average deviation ratio (ADR) of the ADev's above and below the fitted line that avoids the \((x_j - \bar{x})^2\) dependence of traditional skewness. Since we expect more weak-lined galaxies than strong-lined, we define the ADR so that it will be greater than 1 for asymmetries toward weak Mg index values. As for the ADev, the \(x\) is the median, not the mean, i.e.,

\[
\text{ADR} = \frac{\sum_{j=1}^{N} |x_j - \bar{x}|_{\text{bottom}}}{\sum_{j=1}^{N} |x_j - \bar{x}|_{\text{top}}} .
\]

The results are summarized in Table 1. The value of the standard deviation is relatively stable at 0.027 ± 0.003 for Mg2, 0.474 ± 0.052 for Mg b, and 0.035 ± 0.003 for \(\langle \text{Mg}_2 \rangle\) among the different data sets. The skewness is definitely negative, between −0.052 and −1.75 for Mg2, around −0.7 for Mg b, and −4 for \(\langle \text{Mg}_2 \rangle\) but shows too much scatter to be very useful.

The ADev and ADR are more stable because of their relative insensitivity to outliers (ADev: 0.0206 ± 0.0014, 0.346 ± 0.032, and 0.02 ± 0.001 and ADR: 1.265 ± 0.058, 1.105 ± 0.12, and 1.7 ± 0.09 for Mg2, Mg b, and \(\langle \text{Mg}_2 \rangle\), respectively). In the rest of the paper, we use ADev and ADR.

Bender et al. (1993) noted that Mg residuals could be described by a Gaussian core with extra galaxies at weak Mg strength. The width of this Gaussian core is also of crucial interest for the ancient formation hypothesis because it represents the amount of scatter imprinted on the Mg−σ relation at formation. In order to find its width from the various data sets, we compute the number of galaxies within bins of 0.01 mag in both Mg2 and \(\langle \text{Mg}_2 \rangle\) residual and within bins of 0.1 A in Mg b.

We fit with a Gaussian the principal peak from Mg residual distributions, weighting by \(N^{1/2}\) in each bin (see Figs. 2, 3, and 4). These fits yielded widths that were very close for all the data sets except for the Bernardi et al. (1998) field sample (a bit large) and the Dressler et al. (1991) sample (a bit small). It is possible that the larger Bernardi et al. (1998) width is intrinsic (and a possible environmental effect since this sample is the only one that consists solely of field galaxies). The Dressler et al. (1991) residuals were difficult to fit with a Gaussian since they display a second peak toward the blue side (see Fig. 2). These data also had the largest observational errors of our collection. On average, we found \(\sigma_g\) equal to 0.015 ± 0.002 for Mg2 and 0.013 ± 0.002 for \(\langle \text{Mg}_2 \rangle\). Because of the difference of units, \(\sigma_g\) for the Mg b index is broader: \(\sigma_g = 0.25 ± 0.05\).

All data sets give remarkably similar characteristics. We see no statistically significant difference in width or asymmetry among E and S0 galaxies in different environments (field, group, or cluster). Also note that our method of line-fitting yields Mg300 and Mg500. field-group-cluster offsets even smaller than those computed by Bernardi et al. (1998) from the same data.

For the remainder of the paper, we use the ADev and ADR statistics for providing the first-order discriminant for the synthetic Mg−σ diagrams, whose generation we discuss in the following section.

### 3. MERGER MODELS

In order to understand the effects of galaxy merging on the scatter in the Mg−σ relation, we constructed hybrid models that contain simple galaxy merging trees and rudimentary chemical evolution but a relatively rigorous treatment of the effect of stellar ages on Mg feature strength. The aim of the resultant Monte Carlo simulation is to make a synthetic Mg−σ diagram in which the stretch along the \(\sigma\) axis is heuristic and approximate but where the scatter in the Mg strength direction is modeled well. Also, relative zero point changes in Mg strength are reliable; e.g., the weakening of Mg feature strength with increasing look-back time is well modeled (if one can assume that elliptical galaxies today were also elliptical galaxies in the past).

These models are simple in the sense that they involve neither rigorous chemical evolution, N-body simulations, nor hydrodynamics, but nevertheless go several steps beyond what has been attempted in the past. For instance, Bender et al. (1993) use the scatter in Mg−σ, filtered through stellar population models, to estimate a maximum allowed scatter in age of 15% (although Worthey, Trager, & Faber 1995 point out that if young subpopulations tend to be more metal-rich, then the allowed age scatter increases greatly). Colless et al. (1999) and Trager et al. (2000) confirm this age-metallicity anticorrelation. Trager et al. (2000) also use Monte Carlo techniques within the principle components formalism they develop to show that random galaxies with a factor of 10 spread in age nevertheless form an Mg−σ relation with rms scatter of only 0.007 mag (substantially more narrow than observed) if the age-metallicity anticorrelation as derived from Balmer-metal diagrams is correct. Also, Kauffmann & Charlot (1998) showed that cold dark matter (CDM) semianalytical models produce a narrow Mg−σ relation. Another modeling effort that has considerable parallel to our effort is that of Bower et al. (1992), who studied the color-magnitude diagram (CMD). They explore the scatter in the CMD with (1) burst models with artificial scatter in time of burst and (2) two-burst models with the major burst happening 10 Gyr ago and a secondary burst happening 5 Gyr ago. Our effort is of roughly the same philosophy as the Bower et al. (1992) exploration but (1) includes enough
chemical evolution so that younger subpopulations of stars will tend to be more metal-rich and (2) explores a much more vast parameter space in terms of when bursts happen and how much gas and stellar material is involved.

There are several simplifying assumptions in the models. Chemical enrichment proceeds in a schematic way. A mass-[Mg/H] relation is assumed, and a premerger galaxy fragment is assumed to have the [Mg/H] appropriate for its mass. The fragment is part gas and part stars, the gas fraction being a parameter in the models. Merging occurs in the simplest of trees: we follow a trunk protogalaxy, into which a fragment will merge at a random time but with a constant (parameterized) mass and gas fraction. Upon merging, newly incorporated gas instantly becomes stars, slightly metal-enriched with [Mg/H] derived from the postmerger mass via the mass-[Mg/H] relation. Note that while total mass in gas and stars is rigorously accounted for, neither mass in heavy elements in gas phase nor gas reinjected into the interstellar medium through stellar mass loss is tracked, so these are not true chemical evolution models. Only gas enriches: the stellar portion of the merger fragments retain their original abundance (and age) as they are incorporated into the postmerger object. The protoelliptical to which fragments are added is always purely stellar. Using this scheme, the Mg abundance is driven toward higher values in larger galaxies. This is one of the more heuristic elements of the models, the chief intent being to produce a spread of galaxy Mg strengths. Almost any mechanism that accomplishes

Fig. 2.—Number of galaxies vs. Mg$_2$ residual. Curve: Gaussian fit for the peak at zero. Bar at upper left: Typical uncertainty on the measurements.
this is sufficient for the purposes of this exercise since it is the spread of Mg strength from stellar age effects that is under study, not the origin of the relationship itself.

The choice of \([\text{Mg}/\text{H}]\) rather than \([\text{Fe}/\text{H}]\) is mandated by observation. At high metallicities, elliptical galaxies are seen to have approximately constant Fe-feature strength as a function of velocity dispersion (Trager et al. 1998). Mg-feature strengths increase with velocity dispersion, and the tight relation between the two variables is the cause of this investigation. Scaled-solar stellar population models predict that both species should increase together with increasing metallicity (Worthey 1994), so it is mostly Mg (and presumably other light metals) that drives the increase of abundance strengths. Index strengths are computed as a function of age, \([\text{Mg}/\text{H}]\), and mass for each tagged star parcel within the final galaxy assemblage. The part of the models that is not approximate is the treatment of the stellar populations. Quick interpolation tables were summarized from Worthey (1994) models using both the original Yale/VandenBerg evolution and the Bertelli et al. (1994) isochrones for younger ages. Lookups can be done for \(V\) luminosity and the three Mg indices. The simulation delivers a galaxy formed at several (or many) epochs. These are collected as a list of ages, masses, and \([\text{Mg}/\text{H}]\) abundances. Index strengths are computed as a function of age and \([\text{Mg}/\text{H}]\) and weighted by \(V\) luminosity computed as a function of age, \([\text{Mg}/\text{H}]\), and mass for each tagged star parcel within the final galaxy assemblage.

The main parameters in the models number four: the number of equal-mass merger fragments that will eventually combine to form the final galaxy, the fraction of the final galaxy that formed at the redshift of formation \(z = z_f\), the gas fraction of the premerger fragments \(F_{\text{gas}} = m_{\text{gas}}/m_{\text{total}}\), and a probability envelope that is either flat with time between \(z_f\) and \(z = 0\) or proportional to \((1 + z)^n\), where \(n\) is

\[
\frac{\text{Number of Galaxies}}{-0.15 \text{ to } 0} \quad \begin{array}{c}
\text{Hudson} \\
\text{Prugniel} \\
\text{Trager}
\end{array}
\]

\[
\text{Mg}_2 \text{ Residual}
\]
a free parameter. Other parameters control cosmology, set $z_f$, set the number of galaxies to be generated and their mass range, introduce artificial scatter into the computed Mg-$\sigma$ relation, and provide for extra (or not as much) Mg enrichment at each star formation episode.

Throughout this paper, we adopt $H_0 = 60$ km s$^{-1}$ Mpc$^{-1}$, $\Omega = 0.2$, $\Omega_R = 0.8$, and $\Omega_k = 0$. With these parameters, the universe is 13.8 Gyr old, and $z_f = 5$ corresponds to a lookback time of 12.3 Gyr. We flagged galaxies that experienced star formation within the last 50 Myr as "star forming" in the sense that they probably show emission lines, and they are omitted from all statistical calculations. As the majority of data sets feature the Mg$_2$ index, we use this index for all of the following discussion even though the models predict all index varieties.

4. MERGER-MODEL DISCUSSION

4.1. Basic Model Behavior

As an introduction to the model output, we focus on examples of the effect of our four main parameters. Figures 5, 6, 7, and 8 show the effect of varying the number of equal-mass mergers that eventually form the galaxy. With a lot of mergers ($\sim 80$), the asymmetry becomes low since all the galaxies have almost the same history; i.e., they all suffer a peppering of small amounts of star formation over the age of the universe. But with very few mergers ($\sim 5$), the dispersion is larger and the asymmetry is much more pronounced. Each galaxy experiences a distinct timing fingerprint for mergers and, thus, each is more individualistic, some having had recent star formation, others having had none.

Fig. 4.—Number of galaxies vs. Mg $b$ or $\langle$Mg$_2$$\rangle$ residual. Curve: Gaussian fit for the peak at zero. Bar at upper left: Typical uncertainty on the measurements.
The shape of the probability envelope has a strong impact as well. The frequency of mergers is often parameterized in the literature as being proportional to \((1 + z)^n\), where \(n\) is a free parameter. Figure 9 translates some of these choices from redshift to time. One can see that \(n = -2.2\) roughly corresponds to a uniform probability with time. (Note that when we refer to uniform probability in this paper, we mean a perfectly flat [with time] curve, avoiding altogether the \((1 + z)^n\) formalism). The parameter \(n = -1.7\) corresponds to a present-day merger frequency about half of that at

![Fig. 5.—Influence of the number of equal-mass mergers; 5, 40, and 80 merger models are shown in the different panels. All models have a uniform probability envelope, \(P_{\text{gas}} = 0.2\), and a fraction of galaxy formed at \(z = z_f\) of \(F_{\text{ini}} = 0\).](image1)

![Fig. 6.—Influence of variable merger probability with time. We show uniform probability, \(n = -1.5\) and \(n = -0.5\) cases with other parameters held fixed at 40 mergers, \(P_{\text{gas}} = 0.2\), and \(F_{\text{ini}} = 0\). Lower panel: Y-axis scaling differs to accommodate a more peaked histogram.](image2)

![Fig. 7.—Influence of the fraction of gas in each merging event. Gas fractions \(P_{\text{gas}} = 0.1, 0.2,\) and \(0.5\) are shown with other parameters held fixed at 40 mergers, a uniform probability envelope, and a fraction of galaxy formed at \(z = z_f\), \(F_{\text{ini}} = 0\). Top panel: Y-axis scaling differs to accommodate a more peaked histogram.](image3)

![Fig. 8.—Influence of the fraction of the galaxy formed at \(z = z_f\). Fractions \(F_{\text{ini}} = 0.0, 0.3,\) and \(0.6\) are shown with other parameters held fixed at 40 mergers, a uniform probability envelope, and \(P_{\text{gas}} = 0.2\). Lower panel: Y-axis scaling differs to accommodate a more peaked histogram.](image4)
$z = z_f$, and $n > 0$ corresponds to merger rates strongly skewed to high redshift. These numbers change somewhat with cosmology.

As Figure 6 shows, it is plain that with $n = -0.5$, the obtained distribution is very thin. Values of $n$ larger than $-0.5$ produce even thinner relations. Since mergers are concentrated in the past, few are lucky enough to have had recent star formation. We should point out that the narrowness of the diagrams at higher $n$ has a subtle effect on the ADR: any galaxies that do deviate from the relation do so relatively dramatically compared to the width so that the ADR is magnified. This effect can be seen later, in Figures 10–15.

The parameter $F_{\text{gas}}$ is the fraction of gas present in each premerger fragment. In Figure 7, we see that more gas leads to more star formation and hence more young galaxies. Less gas leads to weak starbursts and a narrow Mg-$\sigma$ relation.

The fraction of galaxy formed at $z = z_f$, $F_{\text{ini}}$, has less of a dramatic effect (see Fig. 8), except for the trivial case of $F_{\text{ini}} \approx 1$. There is little difference between forming half of the galaxy at $z_f$ and forming the galaxy entirely by mergers, rather to the initial surprise of the authors. Apparently, it is the last few Gyr of history that are the most important.

### 4.2. Maximum Merger Models

We now search quantitatively for solutions under the maximum merger hypothesis. For fair statistical comparison, the artificial data were further randomized by an artificial Gaussian scatter $\sigma_a = 0.01$ mag, representing the observational error only, i.e., $\sigma_a < \sigma_g$, the width of the Gaussian core of the residual distribution. This assumption is modified for the ancient formation hypothesis described below.

Because we compare to the mean of all of the observational data, we analyzed data from a similar number of artificial galaxies—roughly 2000 Monte Carlo galaxies. The

---

2 This is actually the median error of the data sets, where we combine errors due to velocity dispersion with errors due to index observation as $\sigma^2 = \sigma_M^2 + (\text{slope} \times \sigma_g)^2$. 

---

Fig. 11.—ADev (top) and ADR (bottom) for a $(1 + z)^{-1.5}$ probability vs. the number of mergers. An artificial observed scatter of $\sigma_a = 0.01$ mag was included. Symbols and lines are the same as Fig. 10. Some models for $N_{\text{merges}} = 5$ are beyond the plot limits in the ADR panel.
scatter from limited sample size, of course, decreases with an increased number of samples; for 250 galaxies, the errors were 0.0009 and 0.2 for the ADev and the ADR, respectively (computed from multiple realizations of the same simulation). With 2000 galaxies, we obtained 0.0002 for the ADev and 0.04 for the ADR. These errors are not too different from those seen in the observational data sets, so it is all but certain that sampling error causes most of the statistical deviations between data sets. This leads to important guidelines for future work at higher redshift: large sample size is crucial if the scatter is to be adequately characterized, and errors may not creep too much higher than their nearby brethren.

Figures 10–12 show our main diagnostic statistics, ADev and ADR, for models that cover the grid of input parameters. Data for a uniform probability are shown in Figure 10, for a \((1 + z)^{-0.5}\) envelope in Figure 11, and for a \((1 + z)^{-1.5}\) probability vs. the number of mergers. An artificial observed scatter of \(\sigma_a = 0.01\) mag was included. Symbols and lines are the same as Fig. 10.

Fig. 12.—ADev (top) and ADR (bottom) for a \((1 + z)^{-0.5}\) probability vs. the number of mergers. An artificial observed scatter of \(\sigma_a = 0.01\) mag was included. Symbols and lines are the same as Fig. 10.

Fig. 13.—ADev (top) and ADR (bottom) for a uniform probability vs. the number of mergers, with an artificial scatter \(\sigma_a = 0.015\) mag. Filled circles, diamonds, and squares: \(F_{\text{ini}} = 0.0\), \(F_{\text{ini}} = 0.1\), 0.2, and 0.5, respectively. Open circles, diamonds, and squares: \(F_{\text{ini}} = 0.5\), \(F_{\text{ini}} = 0.1\), 0.2, and 0.5, respectively. Lines: Situated at \(\pm \sigma\) and \(\pm 2.5 \sigma\) around the observed mean.

Fig. 14.—ADev (top) and ADR (bottom) for a \((1 + z)^{-1.5}\) probability vs. the number of mergers with an artificial scatter \(\sigma_a = 0.015\) mag. Symbols and lines are the same as Fig. 13.

Fig. 15.—ADev (top) and ADR (bottom) for a \((1 + z)^{-0.5}\) probability vs. the number of mergers with an artificial scatter \(\sigma_a = 0.015\) mag. Symbols and lines are the same as Fig. 13.
envelope in Figure 12. Different $F_{\text{ini}}$ and $F_{\text{gas}}$ possibilities are plotted as various symbol types as a function of the number of mergers. The trend with increasing number of mergers is clear: a tighter, more symmetrical relation is found. If this seems counterintuitive—because with many mergers one should expect a more volatile and bursty collection of galaxies—consider instead that each galaxy has a more similar history if the number of mergers grows large (see Fig. 5). To compare directly with the observational data, we plot the mean observational values found in $\S$ 2 with $\pm 1.0$ and $\pm 2.5$ $\sigma$ errors.

Other trends are also apparent. At a given $F_{\text{ini}}$, the ADev increases with the gas fraction since, with more gas available, more stars will form, making the mean stellar population more volatile. Similarly, at fixed gas fraction, the ADev decreases with higher $F_{\text{ini}}$. Indeed, if $F_{\text{ini}}$ is set to 1, the relation becomes razor thin since the whole galaxy formed at $z = z_f$. Except for values approaching 1, the authors were surprised at the ineffectiveness of $F_{\text{ini}}$ to modify the Mg-$\sigma$ relation by much. It is the weakest of the four parameters.

As for the influence of when in time the mergers are likely to happen (the probability envelope), we compare the three figures. It is rather difficult to find many points that land simultaneously in the $2.5 \sigma$ regions of both ADev and ADR statistics, but to the extent that some do, most of them are in the uniform probability figure or, in the mildest case of $(1 + z)^{-0.3}$, around 40–60 mergers. No solutions are found for $n = 0$; the slope is too extreme, and galaxy formation is concentrated too far in the past to generate sufficiently large ADev or ADR. Any probability exponent $n$ greater than zero yields Mg-$\sigma$ relations that are even thinner than the $n = 0$ case, and we conclude that elliptical galaxy formation that was very concentrated toward $z_f$ is ruled out under the maximum merger hypothesis.

### 4.3. Ancient Formation Models

We now relax our assumption that the Mg-$\sigma$ relation is inherently thin. We suppose instead that there is an intrinsic, symmetric scatter $\sigma_g$ equal to that which is observed today but imprinted at formation. The mean width derived in $\S$ 2 is $\sigma_g = 0.014$ mag, and the median is 0.013 mag. This is close to the observational scatter, median 0.010 mag. Clearly, most of the Gaussian core of the Mg-$\sigma$ relation is due to observational error. At a maximum (assuming $\sigma_g = 0.014$ mag in Mg), the intrinsic symmetric galaxy scatter is 0.010 mag, but this is an upper limit. A higher precision data set is required to sort this problem out with certainty.

With a generous artificial scatter of 0.015 mag to simulate both observational and intrinsic scatter, we plot the three different probability cases exactly as in the previous section: uniform probability (Fig. 13), $(1 + z)^{-1.5}$ envelope (Fig. 14), and $(1 + z)^{-0.3}$ envelope (Fig. 15). We derived the statistics using 2000 galaxies and show the possibilities allowed by $F_{\text{gas}} = 0.1, 0.2, 0.5$, and $F_{\text{ini}} = 0.0$ or 0.5.

We find that most of the remarks from the maximum merger section are still valid. The simulations with 5 or 10 mergers do not fit both statistics simultaneously, and the most numerous solutions still occur for uniform probability and $N_{\text{merger}} = 40$ or an $n = -1.5$ envelope and $N_{\text{merger}} = 40–60$. The main difference is that the additional artificial scatter calms the asymmetry slightly, and this allows a few more successful models.

The close similarity between maximum-merger and minimum-merger cases is driven by the high quality and consistency of the observational data. With the low and high bounds of allowed artificial scatter set at 0.010 and 0.015 mag, tight reins are held on our freedom to interpret the results.

### 4.4. Additional Visualization

Figure 16 shows the models from the previous two sections in ADev, ADR space to more easily illustrate trends in the model tracks. For instance, it is clearly seen that any models with $n > -0.5$ will miss the observational locus independent of the other parameters. The extra artificial scatter in the right-hand ancient formation panels increases ADev without affecting ADR very much, and this improves the fit for Mg-$\sigma$ relations that are too thin, such as those with $n \approx -0.5$. The uniform-probability case offers more near misses than other probability envelopes, although the $n = -1.5$ case is similar. The very best matches appear in panels 10, 13, and 14, with models between 40 and 60 mergers and $F_{\text{ini}} = 0.2$.

We can explore model fits a little more by considering a goodness-of-fit parameter. Including Mg$_{500}$, the fitted Mg strength at $\sigma = 300$ km s$^{-1}$ along with ADR and ADev, we define a goodness parameter to evaluate the relative merit of the various models as follows:

$$
\sigma^2_g = \frac{(\text{ADev} - \bar{\text{ADev}})^2}{\sigma^2_{\text{ADev}}} + \frac{(\text{ADR} - \bar{\text{ADR}})^2}{\sigma^2_{\text{ADR}}} + \frac{(\text{Mg}_{500} - \bar{\text{Mg}}_{500})^2}{\sigma^2_{\text{Mg}_{500}}}.
$$

![Figure 16](image-url)
We adopted the observational averages $\Delta$Dev = 0.0206, ADR = 1.265, and Mg300 = 0.329. The different scatters are $\sigma_{\Delta \text{Dev}} = 1.4 \times 10^{-3}$, $\sigma_{\text{ADR}} = 5.8 \times 10^{-2}$, and $\sigma_{\text{Mg300}} = 0.04$. This last value is less well motivated than the others since we expect considerable uncertainty in model Mg zero point so that cannot adopt the observed uncertainty for $\sigma_{\text{Mg300}}$. On the other hand, the overall Mg level is an important diagnostic and too-weak overall Mg is clearly to be discouraged, so our value of 0.04 is an approximate compromise.

We compute this $\sigma_G$ for three cases: uniform merger probability envelope, $(1 + z)^{-1}$; and $(1 + z)^{-3}$ probability and for both $F_{\text{ini}} = 0$ or 0.5. For each case, we make plots from 5 to 100 mergers, and from $F_{\text{gas}} = 0.0$ to 0.9 for 40 values using a linear distribution for the fraction of gas and a logarithmic distribution for the number of mergers. We plot isovales in $1/\sigma_G$ so that the best models show as maxima rather than minima. Figure 17 shows the results.

In this plot, there are elongated regions of high $\sigma_G$ values. They tend to slant, indicating significant degeneracy between gas fraction and number of mergers in determining a successful model. The amplitude of the band decreases with the $n$ of the envelope probability. We see that the valid possibilities shift to slightly higher values of gas fraction when $F_{\text{ini}} = 0.5$. The $n = -1.5$ envelope probability appears to be the best fit.

![Fig. 17.—Isovalue contours of $1/\sigma_G$ for a uniform probability (top), $n = -1.5$ probability (middle), and $n = 0$ probability (bottom) for $F_{\text{ini}} = 0$ (left) and 0.5 (right). An artificial observed scatter of $\sigma_G = 0.01$ mag was included. The contours are set at multiples of $(3)^{1/2}/4$ so that borderline acceptability corresponds approximately to the second contour level, good fits to the fourth.](image)

4.5. Constraints from Mean Age and Abundance Spread

We now check to see if the possibly successful models indicated by ADev, ADR, and Mg300 in Figure 17 are compatible with the abundance distribution and the $V$ band light-weighted age observed in nearby galaxies. The abundance distribution is observed to be narrow in all galaxies the size of M32 and larger (Worthey, Dorman, & Jones 1996), with a FWHM of less than 0.4 dex (Grillmair et al. 1996) in M32 itself and about the same in the Milky Way (Rana 1991). We can check the abundance distributions of the simulations as well. The chemical evolution of these models is primitive, but if the abundance distributions are wider than about 0.5 dex FWHM, then we must regard them with suspicion. Also, if the light-weighted mean age (the closest our present models can come to approximating a spectroscopic Balmer-metal feature mean age) falls below about 7 Gyr (the approximate median of such studies; Terlevich & Forbes 2002; González 1993), we likewise regard this model with much suspicion. Older ages are allowed because the age zero point of the Balmer-metal feature technique is still quite uncertain.

What we find is emphatic. For $n = 0$ or $n = -0.5$ models, no model passes the ADR/ADev test very well, but they also tend to have an abundance scatter greater than 0.5 dex. For uniform probability and its close cousin $n = -1.5$, the mean galaxy age becomes too young for a combination of high gas fraction and large numbers of mergers, but this never becomes a critical problem because the ADR/ADev test excludes these models as well. The best surviving models are those that fit the best in Figure 16: those with uniform probability or $n = -1.5$, $F_{\text{gas}} = 0.2$ (for $F_{\text{ini}} = 0$) to $F_{\text{gas}} = 0.35$ (for $F_{\text{ini}} = 0.5$), and around 50 mergers. The abundance histogram for these models is at the 0.5 dex width limit, mean age is 8–9 Gyr, and the ADR and ADev are within the observational error bars.

At large number of mergers, the $F_{\text{ini}} = 0.5$ models tend to have a bimodal abundance distribution caused by the initial metallicity of the galaxy to which additional material from the gaseous portion of the merger is added, plus a low-metallicity peak from the stellar component of the relatively low-mass merger fragments. In fact, all of the $F_{\text{ini}} = 0.5$ high points in Figure 17 are at the borderline of having a too-wide abundance distribution, and we should mentally penalize these models a couple of contour levels for this near transgression. The $F_{\text{ini}} = 0$ models do not suffer from this problem as much: they all have suitably narrow abundance distributions.

4.6. Model Variants

Being left with only one narrow family of successful models after trying so many hundreds left the authors scratching their heads. We decided to explore other options. Our first option was to abandon the constant merger-mass assumption by adopting a cold dark matter (CDM) power spectrum for galaxy merger fragments. We use the formulae of White & Frenk (1991) and Kauffmann & White (1993) that are based on Press & Schechter (1974) formalism. CDM clusters hierarchically with preferentially small masses in the early universe, building toward larger structures later on. We built a look-up table with 40 steps in redshift and 200 steps in mass in order to quickly invert the probability function for halos of a given mass and redshift. Within each step, a small randomization assigned the final redshift and
mass to the merger fragment. To calibrate the zero point between the circular velocity ($V_c$, which the White & Frenk formulae give) and mass, we normalized to the Milky Way. We took $V_c = 220$ km s$^{-1}$, mass equal to $10^{11} M_\odot$, and White & Frenk formulae to obtain $M_{\text{gal}} = 10^{11} M_\odot$ ($V_c/220$ km s$^{-1}$).

The only adjustable parameters in the CDM merger models were gas fraction and the CDM bias parameter, $b$. The ADR statistics from the CDM models were very high unless the gas fraction was set below 10%, at which point the ADev becomes too narrow. The $V$-weighted age and abundance scatter for the $F_{\text{gas}} = 0.1$ model is acceptable: 11.4 Gyr and 0.4 dex, respectively. Increasing the bias parameter helps a little, but no CDM models were found that matched all statistics simultaneously. Keep in mind that our treatment is less sophisticated than, for example, Kauffmann & Charlot (1998), and we are left with far fewer adjustable parameters than with our original scheme.

The CDM simulations were also interesting because they predict variations in Mg-σ as a function of galaxy size. For small galaxies compared to large, CDM predicts similar ADev, younger mean ages, and much higher asymmetry. Spectroscopic observations (Terlevich & Forbes 2002) indicate a younger mean age for smaller galaxies. Our collection of Mg-σ observations indicate higher ADev (which may reflect only increased observational error for these fainter galaxies) but ADR roughly the same.

A variant on the CDM models was to throw out the CDM power spectrum and adopt a constant power spectrum ($dN \propto M^{-1}$) and return to a merger probability uniform. These models have suitable ADR and ADev if the gas fraction is set at 2%–3%. The 2% model has a (V-weighted) mean age of 11.7 Gyr, with an abundance distribution FWHM scatter of 0.3 dex. This is the only model with a mass spectrum (as opposed to the original models) that fits all of the observational constraints. Consequently, even though this constant power spectrum model is not physical, the good fit seems to indicate that there will be some power spectrum that will satisfy the constraints equally well.

Other variants were tried. We tried to superpose two different schemes atop one another. For instance, suppose half of the galaxies were truly ancient and the other half was allowed to have mergers. These superposition schemes fail dramatically. Unless one is exquisitely artful, one ends up with Mg-σ diagrams that are clearly double.

We also checked the influence of implicit parameters. (1) We assigned a timescale of 50 Myr for nebular emission to die away after an accretion and omitted emission-line objects from statistical analysis. If this timescale is increased, overall goodness-of-fit increases mildly due to a decrease in asymmetry, but it is not a dramatic effect. (2) We also made our assumed [Mg/H]-mass relation shallower and steeper to see what effect this might have. If one makes large galaxies more metal-rich, newly formed stars are more metal-rich, redder, and postburst galaxies join the main Mg-σ relation faster. This decreases the ADR without affecting the ADev very much. If we steepen the relation to the point where large galaxies have [Mg/H] $\approx 1$, this decreases the ADR by about 0.2. This is enough, for instance, to make the CDM models with 10% gas fraction fit both ADR and ADev statistics. In the original models, this drives the best fits to slightly lower $N_{\text{merger}}$ or toward a more uniform merger probability.

Finally, we tried schemes of variable gas fraction, where the early universe was assumed to be gas-rich but later mergers are mostly stellar. Like the original four-parameter models, envelope parameters $n > -0.5$ are ruled out, but uniform-with-time probability gives good matches for $20 < N_{\text{merger}} < 40$. In fact, one such model, with $F_{\text{gas}}$ varying linearly with time from 0.9 at $z_f$ to 0.1 today matches the Mg-σ data as good as any model that we tried. Pursuing variable gas-fraction options promises to be a fruitful avenue for future work. This model is also intriguing because it shows strong Mg-σ asymmetry evolution at only modest redshift (see next section).

### 4.7. Predictions for Look-back Studies

The recent proliferation of 8–12 m telescopes will greatly accelerate the rate at which Mg-σ data will become available for distant galaxies. The result of studies to date, that the drift of the Mg zero point is “consistent with passive evolution” (Ziegler & Bender 1997), will eventually be broadened to study the shape of Mg residuals as well. We therefore present some model predictions for high-redshift studies. There is a major caveat to consider here: model galaxies are assumed to be “star forming” for 50 Myr after a burst, but “early type” at all other times. If merging is a dominant process, then premerger fragments are often likely to be spiral galaxies. If semianalytic models such as Kauffmann & White (1993) are correct, then, over time, some elliptical galaxies can accrete a gas disk and become spirals. So the predictions we show in this section are of questionable validity for merger-dominated models (i.e., all of those that match the $z = 0$ data!) but nevertheless provide a useful first cut for the variety of behavior we might expect for Mg-σ relations at significant look-back times.

Figure 18 shows a collection of near-best models as they would appear if observed at large look-back times. The merging trees of the model galaxies are computed to redshift zero, but only the partially complete history of each galaxy is considered for nonzero redshift. So the further back in time we look, the more incomplete the galaxies are. Figure 18 at a glance shows that Mg-σ relations survive intact even under fairly severe merging scenarios back to when the universe was a quarter of its present age.

The bottom panel shows $M_{\text{gas}}$ as a function of redshift. Within a zero-point shift, all models show Mg evolution very similar to the ancient, passive evolution case shown in large diamond symbols at strong Mg strength. It is easy to shift any given model track by adjusting the assumed metallicity scale, so the overall Mg level by itself is not a strong constraint, but the relative time evolution should be fairly well modeled. The most shallow slope, i.e., the most constant Mg level with time, is given by the $F_{\text{gas}} = 0.25$, 40 merger, uniform probability, $F_{\text{ini}} = 0.5$ model (small open diamonds). This is a merger-heavy model. We think that some researchers will be surprised by the fact that a merger-driven model is redder, relatively speaking, at $z = 1.5$ than the pure passive evolution scenario because one usually thinks of mergers as causing more star formation activity and hence bluer colors and weaker line strengths. In reality, it can go either way, depending on the model. The steepest Mg track is from a very similar model but with the variable $F_{\text{gas}}$ option and $F_{\text{ini}} = 0$ (filled circles). Neither the steepest nor the shallowest models differ by more than 0.05 mag in Mg$_{300}$ from the passive evolution case if they are normalized.
to match at one end or the other of the illustrated redshift range.

The tightness and asymmetry of the Mg-σ relation also do not show large changes with look-back time. Both the ADR and ADev can increase or decrease according to the model, but at a very modest level that will be hard to measure without a very large sample of distant galaxies.

5. SUMMARY AND SYNTHESIS

In this paper, we analyze the Mg-σ relation in elliptical and S0 galaxies by comparing data and models. We adapted an ADR statistic for measuring residual asymmetry and we used the ADev statistic to measure the width of the scatter. For the combination of all data sets and for 150 km s$^{-1}$, ADev $= 0.0206 \pm 0.0014$, ADR $= 1.265 \pm 0.058$, and Mg$_{300} = 0.329 \pm 0.0039$. The width of the distribution increases at low $σ$. For 50 < $σ$ < 150 km s$^{-1}$, the ADev $= 0.0282 \pm 0.0035$. The errors of the ADev and ADR statistics are roughly consistent with the random errors one would expect from finite number counting as estimated from running simulations multiple times. (The observed ADR scatter is even somewhat lower than what we would expect.)

We also fit the central peak of the Mg residual distribution with a Gaussian, thinking that this would be a useful number to know under an ancient formation hypothesis where all of this scatter was imprinted at formation. We found a width for the central peak of the Mg residual $\sigma_g = 0.014 \pm 0.002$ mag for Mg$_2$, while the observational error was $\sigma_{\text{obs}} = 0.010$. So if total scatter is a quadrature sum of observational and intrinsic scatter, i.e., $\sigma^2 = \sigma^2_{\text{intrinsic}} + \sigma^2_{\text{obs}}$, we find $\sigma_{\text{intrinsic}} = 0.01$ mag. After performing this exercise, we realized that there would be little difference for models that began with the entire scatter or just the observational portion of the scatter, and this turned out to be true: the ancient formation models were not dramatically different from the maximum merger models.

Merger models were constructed to simulate the Mg-σ relation. The merger fragments that join to form the final galaxy were assumed to be equal in mass. Mergers occur randomly in time (uniform probability distribution) or randomly under a probability envelope of the form $(1+z)^n$, where $n$ is a free parameter. The other three main parameters are the number of mergers, the gas fraction of each event, and the fraction of the galaxy that formed in the distant past, at $z = z_f = 5$ for this paper. Also included was an artificial scatter in Mg strength. This scatter was set equal to the observational error for the maximum merger hypothesis and equal to the entire symmetric scatter (plus a bit) for the ancient formation hypothesis. The two hypotheses generated similar results since observational scatter appears to account for the bulk of the observed scatter.

The models predict a simulated Mg-σ diagram, from which we compute Mg$_{300}$, ADev, and ADR in the same way we did for the observational data. We also output $V$-weighted mean age and an abundance distribution since observational constraints also exist on these quantities. We found this avenue of investigation unexpectedly fruitful in the sense that very strong observational constraints allow only very select interpretations to survive. Some immediate conclusions are as follows:

1. A small number (5–20) of (severe) mergers leads to scatter and asymmetry larger than observed.
2. To the contrary, with many mergers (60–100), a large number of small events leads to small scatter and small asymmetry because merging histories are similar. Many flavors of these models showed too tight of a relation, and the gas-rich ones also tended toward unacceptably young mean ages.
3. For the merger probability envelope, a positive $n$ in $(1+z)^n$ concentrates galaxy formation to the early universe, making the Mg-σ relation thin. Because we find no successful models for $n > 0$, it may appear that we clash with many published merger rate results such as $(1+z)^{0.12 \pm 0.05}$, $(1+z)^{0.28 \pm 0.09}$, $(1+z)^{0.45}$, or $(1+z)^{0.11 \pm 0.05}$ (Carlberg et al. 2000).

The $n < 0$ result is very robust, so we presume that the difference between this result and others stems from differences in treatment. For instance, our redshift range is 0–5, but observational studies tend to fit to a much more narrow range. Also, galaxy Hubble types may change with look-back time, other authors have definitions of merger rates that differ, and the observable under study also has a profound influence. For example, Kaufmann, Charlot, & White (1996) find a $(1+z)^{1.6}$ dependence for early-type galaxies, but this is for number evolution, a quantity that we do not model because it requires transformations to other Hubble types.
4. An increased gas fraction increases the number of new stars formed in each merger. This increases the scatter in Mg strength but increases asymmetry only on some models.
5. If a larger fraction of the galaxies form at \( z = z_f \), both scatter and asymmetry decrease.

6. Monte Carlo scatter indicates that several thousand high-redshift galaxies need to be observed at nearby galaxy accuracy to measure ADR to within \( \lesssim 0.05 \), good enough for stringent tests on the evolution of the Mg-\( \sigma \) relation. Such measurements are needed because measurement of Mg\( _{500} \) with look-back time is predicted to be model-degenerate.

Finally, let us briefly discuss the meaning of these results. Our most successful models involved a moderate number of mergers (40–60) with mergers equally probable over cosmic time. A model with a constant gas fraction that fit nicely was with \( F_{\text{gas}} = 0.2 \), but a model with gas fraction that varies linearly with time from 0.9 at \( z_f \) to 0.1 at the present epoch also fit well. One other model, a variable-mass scenario with a flat power spectrum and merger probability constant with time fit the constraints if the gas fraction was 2%–3%. The overall yield of successful models was far below our initial expectations: we thought many solutions would be possible. This is good news for understanding galaxy evolution since tight constraints are better than loose ones.

We also find some clarity for the basic question, “Are elliptical galaxies old or young?” To arrive at age information, one must look at the most sensitive indicators. A brief perusal of Worthey (1994) or other population-model papers weighted by the inverse of the observational error orders the list from most sensitive to least sensitive: Balmer-metal diagrams, Mg-\( \sigma \) relation, FP, and color-magnitude diagram. Thus, Balmer-metal diagrams show very large age scatter that is harder to detect with other methods. The present study of the Mg-\( \sigma \) relation finds that the models that fit every observable have an age range of 7–10 Gyr. This does not force a revision of the Balmer-based age scale, but it keeps alive the long-held suspicion that the Balmer ages do not force a revision of the Balmer-based age scale, but imply considerable age scatter and a mean age (7–10 Gyr) within the uncertainty of the Balmer method. The tightness of the Mg-\( \sigma \) relation is compatible with the less-sensitive FP and color-magnitude diagram tightness, and both are compatible with the Balmer-metal diagrams. In short, we see no barriers to fairly strong merger hypotheses. In contrast, we find little support for the ancient merger hypothesis: the tamest merger scenarios that we could find are still much wilder than quiescent evolution since \( z_f \). The least active model (2%–3% gas fraction and flat mass spectrum) still involves massive mergers (albeit without much gas) up to the present time. On the other hand, this model’s mean age was 11 Gyr, and it was still able to match the width and asymmetry of the Mg-\( \sigma \) relation, so there may be a related model that can satisfy a relaxed version of the ancient merger hypothesis.

Our “sweet spot” of 50 mergers on average is an interestingly large number. We should really replace the term “merger” with “accretion” to describe such numerous, small-coalescence events. This number also suggests that elliptical galaxies share considerable kinship with spiral galaxies in the sense that 50 accretions over a Hubble time differs little from a constant drizzle of star formation over a Hubble time. Elliptical galaxies must simply be better than spirals at turning gas into stars, for reasons that either are circumstantial, structural, or environmental. This study strongly suggests that near-equal-mass spiral-spiral merging events are not a dominant mode of elliptical galaxy formation. Rather, many accretions of satellite galaxies or gas clouds appear to be a much more plausible way to explain elliptical galaxies as a class.

Future theoretical work should involve different dark matter schemes and better chemical evolution, although many aspects of the modeling could be improved. Observational work at intermediate redshift is also needed, although our results regarding high quality and large \( N \) of such studies make this job more challenging.

We thank M. Bernardi, D. Burstein, M. J. Hudson, and P. Prugniel for machine-readable Mg-\( \sigma \) data, S. C. Trager for good advice, and an anonymous referee for a thorough critique. Washington State University provided major funding for this effort.

REFERENCES

Bardelli, S., De Grandi, S., Ettori, S., Molendi, S., Zucca, E., & Cofalferccos, S. 2002, A&A, 382, 17
Bender, R., Burstein, D., & Faber, S. M. 1992, ApJ, 399, 462
Bender, R., Kagawa, R. P., Ziegler, B., Belloni, P., Greggio, L., & Hopp, U. 1998, ApJ, 493, 529
Bernardi, M., Renzini, A., Da Costa, L. N., Wegner, G., Alonso, V., Pellegrini, P. S., Rité, C., & Willmer, C. N. A. 1998, ApJ, 508, 143
Bertelli, G., Bressan, A., Chiosi, C., Fagotto, F., & Nasi, E. 1994, A&A, 285, 12
Bower, R. G., Lucey, J. R., & Ellis, R. S. 1992, MNRAS, 254, 601
Bower, R. G., Lucey, J. R., & Ellis, R. S. 1992, MNRAS, 254, 601
Burstein, D., Dressler, A., & Faber, S. M. 1987, ApJ, 315, 46
Carlberg, R. G., et al. 2000, ApJ, 532, L1
Colless, M., Bressan, A., Chiosi, C., Fagotto, F., & Nasi, E. 1994, A&A, 285, 12
Davies, R. L., Burstein, D., Davies, R. L., McManus, R. K., Saglia, R. P., & Wegner, G. 1999, MNRAS, 303, 813
Conselice, C. J., Bershady, M. A., & Dickinson, M. 2001, BAAS, 32, 34
Davies, R. L., Burstein, D., Dressler, A., Faber, S. M., Lynden-Bell, D., Terlevich, R. J., & Wegner, G. 1987, ApJS, 64, 581
Djorgovski, S., & Davis, M. 1987, ApJ, 313, 59
Dressler, A., Faber, S. M., & Burstein, D. 1991, ApJ, 368, 54
Dressler, A., Lynden-Bell, D., Burstein, D., Davies, R. L., Faber, S. M., Terlevich, R. J., & Wegner, G. 1987, ApJ, 313, 42
Ellis, R. S., Small, I., Dressler, A., Couch, W. C., Oemler, A., Butcher, H., & Sharples, R. M. 1997, ApJ, 483, 582
Faber, S. M., & Jackson, R. E. 1976, ApJ, 204, 668
Forbes, D. A., Beasley, M. A., Brodie, J. P., & Kissler-Patig, M. 2001, ApJ, 563, 143
González, J. J. 1993, Ph.D. thesis, Univ. California, Santa Cruz
Goudfrooij, P., Mack, J., Kissler-Patig, M., Meylan, G., & Minniti, D. 2001, MNRAS, 322, 643
Governato, F., Gardner, J. P., Stadel, J., Quinn, T., & Lake, G. 1999, AJ, 117, 1651
Grillmair, C. J., et al. 1996, AJ, 112, 1975
Hudson, M. J., Lucey, J. R., Smith, R. J., Schlegel, D. J., & Davies, R. L. 2001, MNRAS, 327, 265
Jorgensen, I., Franx, M., & Kjærgaard, P. 1996, MNRAS, 280, 167
Kauffmann, G., & Charlot, S. 1998, MNRAS, 294, 705
Kauffmann, G., Charlot, S., & White, S. D. M. 1996, MNRAS, 283, 117
Kauffmann, G., & White, S. D. M. 1993, MNRAS, 261, 921
Kelson, D. D., van Dokkum, P. G., Franx, M., Illingworth, G. D., & Fabricant, D. 1997, ApJ, 148, 27
Kodama, T., & Arimoto, N. 1997, A&A, 320, 41
Kodama, T., Arimoto, N., Barger, A. J., & Aragón-Salamanca, A. 1998, A&A, 334, 99
Larson, R. B. 1975, MNRAS, 173, 671
Markevitch, M., Gonzalez, A. H., Vikhlinin, A., Murray, S., Forman, W., Jones, C., & Tucker, W. 2002, ApJ, 567, 27
Okamoto, T., & Habe, A. 2000, PASJ, 52, 670
Patton, D. R., et al. 2002, ApJ, 565, 208
Press, W. H., & Schechter, P. L. 1974, ApJ, 187, 425
Prugniel, P., & Simien, F. 1996, BAAS, 309, 749
Rana, N. C. 1991, ARA&A, 29, 129
Schweizer, F., & Seitzer, P. 1992, AJ, 104, 1039
Schweizer, F., Seitzer, P., Faber, S. M., Burstein, D., Dalle Ore, C. M., &
Gonzalez, J. J. 1990, ApJ, 364, L33
Stanford, S., Eisenhardt, P., & Dickinson, M. 1998, ApJ, 492, 461
Terlevich, A. I., & Forbes, D. A. 2002, MNRAS, 330, 547
Toomre, A. 1977, in The Evolution of Galaxies and Stellar Populations, ed.
B. M. Tinsley & R. B. Larson (New Haven: Yale Univ. Obs.), 401
Trager, S. C., Faber, S. M., Worthey, G., & Gonzalez, J. J. 2000, AJ, 120,
165
Trager, S. C., Worthey, G., Faber, S. M., Burstein, D., & Gonzalez, J. J.
1998, ApJS, 116, 1
van Dokkum, P. G., & Franx, M. 1996, MNRAS, 281, 985
van Dokkum, P. G., Franx, M., Kelson, D. D., & Illingworth, G. D. 1998,
ApJ, 504, 17

Wegner, G., Colless, M., Saglia, R. P., McMathan, R. K., Jr., Davies, R. L.,
Burstein, D., & Bagley, G. 1999, MNRAS, 305, 259
White, S. D. M., & Frenk, C. S. 1991, ApJ, 379, 52
Worthey, G. 1994, ApJS, 95, 107
———. 1997, in Star Formation Near and Far, ed. S. S. Holt & L. G.
Mundy (Woodbury: AIP), 525
Worthey, G., Dorman, B., & Jones, L. A. 1996, AJ, 112, 948
Worthey, G., Faber, S. M., Gonzalez, J. J., & Burstein, D. 1994, ApJS, 94,
687
Worthey, G., Trager, S. C., & Faber, S. M. 1995, in ASP Conf. Ser. 86,
Fresh Views of Elliptical Galaxies, ed. A. Buzzoni, A. Renzini, &
A. Serrano (San Francisco: ASP), 203
Ziegler, B. L., & Bender, R. 1997, MNRAS, 291, 527