THE ASTROPHYSICAL JOURNAL, 419:47–51, 1993 December 10

A TECHNIQUE FOR DETECTING STRUCTURE IN CLUSTER VELOCITY DISTRIBUTIONS[1]

ANN I. ZABLUDOFF,[2] MARIJN FRANX,[3] AND MARGARET J. GELLER
Harvard-Smithsonian Center for Astrophysics, 60 Garden Street, Cambridge, MA 02138;
E-mail: aiz@libra.ociw.edu; franx@astro.rug.nl; mjg@cfa.harvard.edu
Received 1993 February 22; accepted 1993 June 16

ABSTRACT

For clusters with $\gtrsim 100$ measured galaxy redshifts, we develop a new method for detecting departures from a Gaussian velocity distribution. We decompose the cluster velocity distribution into a sum of the orthogonal Gauss-Hermite functions. This method quantifies the asymmetric third-order (h_3) and symmetric fourth-order (h_4) terms of the distribution while minimizing the effect of interlopers in the tails.

We apply the method to a sample of eight rich clusters: A426, A548, A1060, A1656, A2151, A2670, DC 2048−52, and Centaurus. Each cluster has more than 100 members with known redshifts. The clusters A548, A1656, DC 2048−52, and Centaurus significantly depart from Gaussian. A1656 and DC 2048−52 have asymmetric velocity distributions; A548 and Centaurus have broad, symmetric distributions. The most likely interpretation is that these systems have substructure. This result is consistent with findings by other authors.

Subject headings: galaxies: clustering — galaxies: distances and redshifts

1. INTRODUCTION

Many well-sampled clusters are complex. Structure is present in both the galaxy and X-ray surface brightness distributions. There are at least three tests for dynamical complexity; it is possible to investigate the distribution of galaxies or X-ray emission on the sky, the distribution of galaxy velocities, or a combination of the two. For example, the cluster A119 has three distinct components in the galaxy distribution projected on the sky (Fabricant et al. 1993), Centaurus has a bimodal velocity histogram (Lucey et al. 1986a, b), and A2151 has three major subclumps identified from correlations of galaxy position with velocity (Dressler & Shectman 1988b). Here we demonstrate a new method of searching for substructure and orbital anisotropy in the velocity distributions of clusters.

The notorious problem of interlopers hinders analyses of cluster velocity distributions. Statistics which deweight the contribution of the tails of a velocity distribution include the first-order "location" and second-order "scale" (Beers, Flynn, & Gebhardt 1988). Methods which characterize the *shape* of the distribution are potentially more powerful if their sensitivity to outliers can be reduced (standard skewness and kurtosis tests may depend strongly on the wings of the distribution [Yahil & Vidal 1977]). In this paper, we describe a measure of the velocity distribution which depends upon an expansion in Gauss-Hermite polynomials. These functions are relatively insensitive to galaxies far from the mean of the distribution.

Our method decomposes the cluster line-of-sight velocity distribution into odd and even functions. We obtain two parameters, one measuring third-order asymmetric deviations from a Gaussian and the other measuring fourth-order symmetric deviations (in contrast, the standard Kolmogorov-Smirnov test does not estimate higher-order moments). The third-order moment may be significant if a system is not well-mixed (i.e., has substructure). A large fourth-order term might indicate substructure or orbital anisotropy (Lucey et al. 1986a, b; Merritt et al. 1988).

By quantifying departures from a Gaussian distribution, which are independent of the number of galaxies measured, we can compare the observations with the dynamical properties of clusters produced in *n*-body models. The observed distribution of second-order moments, the cluster velocity dispersions, places constraints on cold dark matter models of large-scale structure formation and evolution (Frenk et al. 1990; Zabludoff, Huchra, & Geller 1990; Weinberg & Cole 1993; Zabludoff & Geller 1993). Similarly, the distribution of third- and of fourth-order moments obtained from many clusters may provide useful tests of the models.

We describe the Gauss-Hermite expansion method we adopt from van der Marel & Franx (1993) in § 2. Section 3 discusses tests of the method on models of kinematically complex systems. We apply the method to the velocity distributions of eight-rich clusters in § 4. Section 5 gives our conclusions.

2. THE METHOD

Here we modify the method introduced by van der Marel & Franx (1993, hereafter vMF) and Gerhard (1993) for the analysis of the line-of-sight velocity distribution of stars in elliptical galaxies. Given a set of velocity measurements v_i, $1 \le i \le N$, we calculate the Gauss-Hermite moments

$$h_j = \frac{2\sqrt{\pi}}{NS} \sum_{i=1,N} H_j(x_i) \tag{1a}$$

where

$$x_i = \frac{(v_i - V)}{S}, \quad H_j(x) = \frac{e^{-x^2/2}}{\sqrt{2\pi}} \mathscr{H}_j(x) . \tag{1b}$$

The functions $\mathscr{H}_j(x)$ are the Hermite polynomials as defined by vMF. The function H_0 is a Gaussian with velocity dispersion S and mean velocity V. The higher order terms measure the deviations from the Gaussian. The functions H_{2j+1} are odd

[1] Work reported here based partly on observations obtained by the Multiple Mirror Telescope, a joint facility of the Smithsonian Astrophysical Observatory and the University of Arizona.

[2] Current Address: Observatories of the Carnegie Institution of Washington, 813 Santa Barbara Street, Pasadena, CA 91101.

[3] Current address: Kapteyn Laboratory, P.O. Box 800, 9700 AV Groningen, The Netherlands.

and characterize asymmetric deviations from the Gaussian. The functions H_{2j+2} are even and characterize symmetric deviations. The functions $H_j(x)$ form an orthogonal set of functions. Realistic distributions of galaxy velocities can be written as a weighted sum of these functions. The velocity distribution can then be approximated by the series

$$L = \sum_j h_j H_j(x) . \quad (2)$$

The expansion defined by equations (1) and (2) works for any choice of V and S. We have considerable freedom in deriving V and S from the data. After some experimentation, we chose a definition very similar to the one used by vMF (see Gerhard 1983 for an alternative). The parameters V and S are chosen so that h_1 and h_2 are zero. The choice is based on the fact that a small change in V and/or S of a Gaussian is identical to an addition of small terms H_0, H_1, H_2. Thus any expansion with nonzero h_1 and h_2 can be reduced into an expansion with zero h_1 and h_2 by changing V and S. We iterate to achieve this expansion. The resulting parameters S and V are the velocity dispersion and mean velocity of the Gaussian function H_0 (and not necessarily the dispersion and mean of the cluster velocity distribution).

In the particular application of the Gauss-Hermite expansion to discrete galaxy data, our definition of V and S has the advantage that it is not strongly influenced by interlopers. The details of the cluster membership selection do not affect our results strongly, because the weight of a galaxy at a velocity very different from the mean is reduced by the exponential in equation (1b).

We use this method to quantify the deviations of velocity distributions from a Gaussian. In practice, much of the signal is satisfactorily described by the h_3 and h_4 terms alone. It is straightforward to determine the standard error for h_3 and h_4 if the underlying velocity distribution is Gaussian, and if V and S of the Hermite expansion are equal to the mean velocity and velocity dispersion of the parent distribution. We obtain $\sigma(h_3) \approx 0.69/(N)^{1/2}$ and $\sigma(h_4) \approx 0.64/(N)^{1/2}$. For example, these values imply that $\sim$200 velocities are required for a 2 σ detection of $h_3 = 0.10$, a 10% rms deviation from Gaussian.

Beacuse these error estimates are approximate, Monte Carlo simulations are necessary to test the significance of any observed h_3(obs) and h_4(obs). Because the distribution of observed h_4(obs) is not symmetric around zero for a Gaussian parent distribution, we use a two-tailed test. We analyze random realizations of a Gaussian distribution with $N = N_{\rm obs}$ with the same method as for the data. The probability $P[h_i < h_i(\text{obs})]$ is determined from 10,000 simulations. The significance of h_i(obs) is $P(h_i) = 2 \min \{P[h_i < h_i(\text{obs})], P[h_i > h_i(\text{obs})]\}$.

The advantages of the Gauss-Hermite expansion are that (1) it is parametric and yields h_3 and h_4, two physical properties which help characterize the dynamical state of a cluster. The asymmetric h_3 term is a measure of substructure. The symmetric h_4 term may indicate substructure or orbital anisotropies. (2) Both h_3 and h_4 are largely insensitive to the velocity cutoff used to define cluster members. We discuss both these points further in § 4.

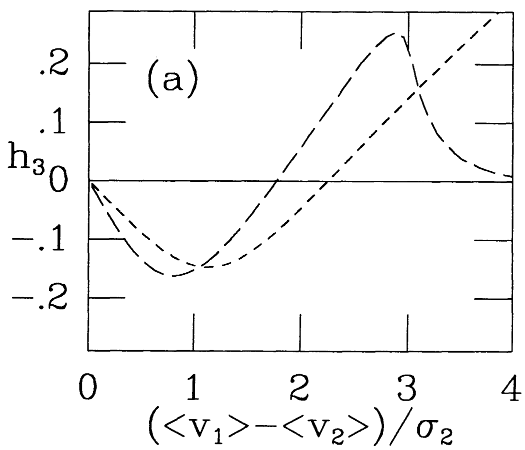
The disadvantages of the method are that (1) it is a parametric test which measures departures from a specified distribution. A Gaussian distribution, however, is a natural and historical choice for a comparison. (2) Once the h_j terms are determined, the velocity distribution reconstructed with equation (2) can have negative wings. In this case, the reconstructed distribution cannot be used in models as a probability distribution, but the terms h_j remain useful.

3. TESTING THE METHOD

To gain a physical understanding of how h_3 and h_4 behave when substructure is present, we examine a two-subclump model. We superpose two Gaussian distributions with populations n_1, n_2, mean velocities $\langle v_1 \rangle$, $\langle v_2 \rangle$, and velocity dispersions σ_1, σ_2. Figures 1*a* and 1*b* show h_3 and h_4 as a function of the velocity separation of the two clumps for three different cases. The solid line is for two clumps with the same velocity dispersion and number of galaxies, the short-dashed line is for two clumps with a velocity dispersion ratio of $\sigma_1/\sigma_2 = 0.5$ and populated in the ratio $n_1/n_2 = 0.5$. The long-dashed line is for clumps with $\sigma_1/\sigma_2 = 0.2$ and $n_1/n_2 = 0.2$.

In the case of two identical, superposed clumps (solid line), there is no skewness in the combined distribution at any velocity separation (h_3 is always zero). The h_4 signal grows as the increasing separation of the clumps broadens the distribution.

As we move apart the two clumps with $\sigma_1/\sigma_2 = 0.5$ and $n_1/n_2 = 0.5$ (short-dashed line), h_3 and h_4 change as the smaller clump augments different parts of the larger clump's Gaussian



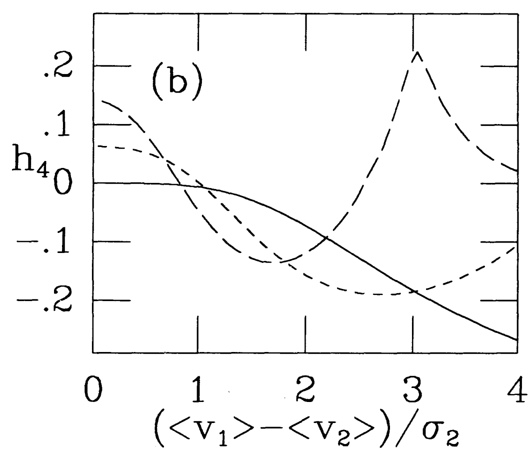


FIG. 1.—(*a*) Variation in the h_3 term with velocity separation for two superposed clumps with Gaussian velocity distributions. The solid line is for two clumps with the same velocity dispersions and number of galaxies, the short-dashed line is for two clumps with a velocity dispersion ratio of $\sigma_1/\sigma_2 = 0.5$ and populated in the ratio $n_1/n_2 = 0.5$. The long-dashed line is for clumps with $\sigma_1/\sigma_2 = 0.2$ and $n_1/n_2 = 0.2$. (*b*) The same as in (*a*) except for the h_4 term.

distribution. At velocity separations $\lesssim 1\ \sigma_2$, the smaller clump 1 enhances one side of the peak of clump 2; the skewness increases and the combined distribution becomes less peaked. At separations between ~ 1–$2\ \sigma_2$, the distribution is more symmetric and broad. At ~ 2–$3\ \sigma_2$, clump 1 moves outside of the peak and into the tail of clump 2; the sign of the skewness is reversed and the broadening of the combined peak reaches a maximum. Once clump 1 reaches the far tail of clump 2 (at separations exceeding $\sim 4\ \sigma_2$), both h_3 and h_4 fall quickly to zero. At these separations, the determination of V and S is dominated by the clump 2, and therefore deviations from H_0 are small.

If the two clumps are very different ($\sigma_1/\sigma_2 = 0.2$ and $n_1/n_2 = 0.2$; long-dashed line), the h_3 term behaves similarly to the second case, but diminishes more quickly at clump separations $\gtrsim 3\ \sigma_2$. For separations $\lesssim 1.5$–$2\ \sigma_2$, the behavior of the h_4 term also resembles the second case. However, if clump 1 is at ~ 2–$3\ \sigma_2$, it makes the combined distribution leptokurtic by extending the tail. Clump 1 ceases to contribute to h_4 at separations $\gtrsim 3\ \sigma_2$.

The two-clump model gives us a feeling for the behavior of h_3 and h_4 in systems ranging from a superposition of two similar clumps to a superposition of a poor group and a rich cluster. Yet real clusters are often more complex. Our model does not include orbital anisotropies, multiple clumps, or infall from the field. Numerical simulations are required to produce estimates of h_3 and h_4 in such cases.

Although our simple model cannot be applied to individual, extremely complex systems, we can still make some statistical arguments. For a large sample of clusters consistent with a two-clump model, it is possible to turn the question around and to use the measured distributions of h_3 and h_4 to place limits on the relative populations of the clumps (Franx & Zabludoff 1993). For now, we can further evaluate the method by applying it to well-studied clusters and comparing our results to those obtained with other tests for substructure.

4. APPLICATION TO DATA: EIGHT RICH CLUSTERS

We test the power of the method to measure departures from a Gaussian in the velocity distributions of eight rich clusters. We choose clusters with more than 100 members with known redshifts. Figure 2 shows the Gauss-Hermite expansions superposed on the redshift data for A426, A548, A1060, A1656, A2151, A2670, DC 2048−52, and Centaurus. Previous studies have revealed substructure in all of these systems, except A2670 (Fitchett & Smail 1993; Dressler & Shectman 1988b; Fitchett & Merritt 1988; Fitchett & Webster 1987; Mellier et al. 1988; Davis & Mushotzky 1992; Merritt 1988; Lucey et al. 1986a, b). A2670 has an unremarkable velocity distribution, but the large peculiar velocity of its cD galaxy with respect to the cluster mean suggests some kinematic complexity (Sharples, Ellis, & Gray 1988; Zabludoff, Huchra, & Geller 1990).

Table 1 lists the cluster (col. [1]) and its coordinates (cols. [2] and [3]), radius sampled in h^{-1} Mpc (col. [4]), velocity range Δv (col. [5]) number of galaxies (col. [6]), mean velocity V (col. [7]), and velocity dispersion S of the Gaussian (col. [8]). We use the velocity limits determined by previous authors or by the membership procedure described in Zabludoff, Huchra, & Geller (1990). Because the tails of the velocity distribution are deweighted by our fitting procedure, varying these limits has little effect on the calculation of h_3 and h_4.

The h_3 and h_4 terms and their errors for all eight systems are in columns (9) and (11) of Table 1. We test the significance of these terms with a two-tailed test; the significance of the observed h_3 and h_4 are in columns (10) and (12). Daggers mark the significant values of h_3 and h_4. In columns (13) and (14), we show the "D-statistic" (D_{KS}) and confidence level (P_{KS}) of the Kolmogorov-Smirnov test (Lilliefors 1967) which measures the departure from a Gaussian with the mean and dispersion of the data. The redshift sources are in column (15).

Table 1 shows that three different choices of membership in A1656 ($\Delta v = 0$–$20{,}000$ km s^{-1} is an extreme test case) do not affect the results for h_3 and h_4, verifying that the method is insensitive to the tails of the distribution. The Kolmogorov-Smirnov test is more sensitive to outliers and therefore yields changing results.

Because the redshift sample in A2670 is not uniform over 40′ (Sharples et al. 1988), we test the velocity distribution for galaxies within both 40′ and 20′. We do not find significant h_3 or h_4 terms for either radius.

Four clusters deviate significantly from a Gaussian. Two clusters, A1656 and DC 2048−52, have significant h_3 terms (at the >95% level). In two other clusters, A548 and Centaurus, the symmetric h_4 terms are significant. For comparison, the K-S test identifies significant deviations in three of these systems, A548, A1656, and Centaurus; the 108 redshifts measured in DC 2048−52 are too few to yield a detection. The significances of the K-S detections and those of our method are comparable (e.g., 98% for the D-statistic vs. 99% in either h_3 or h_4). Our method, however, provides a quantitative evaluation of the symmetric and asymmetric properties of the distribution.

Depending on the population of each peak and their location relative to the sample mean, two or more clumps in velocity space can produce either an asymmetric velocity distribution or a broadened symmetric distribution. In the sample, the clusters with significant asymmetric terms (h_3), A1656 and DC 2048−52, have substructure. On the basis of our test alone, we do not know whether the h_4 departures from a Gaussian in A548 and Centaurus are produced by substructure, orbital anisotropies, or both. Other authors, however, have argued that substructure exists in A548 and Centaurus. A548 has two obvious clumps on the sky (Dressler & Shectman 1988b); the h_4 term is small for each clump, indicating that the h_4 signal for the entire system results from the superposition of the two clumps in velocity space. Centaurus has a symmetric, bimodal velocity distribution, and Lucey et al. (1986a, b) identify the peaks which produce our signal as subclumps.

Our method alone cannot rule out the contribution of orbital anisotropy to the h_4 term. Theoretically, a cluster with predominantly radial orbits will have a centrally peaked velocity distribution with extended tails; a system dominated by circular orbits will have a broad peak with truncated wings (Merritt 1987). Although we measure a broadening of the velocity distribution in A548 and Centaurus, simulations of collapsed systems suggest that circular orbits are unlikely to dominate (e.g., Dubinski 1992).

At present, there are no observational signatures of orbital anisotropy in the sample clusters. If most cluster members followed radial orbits, we might observe a large population of tidally-stripped galaxies (Pryor & Geller 1984) or a decrease in the cluster velocity dispersion with radius more rapid than expected from the system's steeply falling mass profile (Rood et al. 1972; Dressler & Shectman 1988b). In clusters without obvious substructure, a significant h_4 detection might indicate that these other tests will yield interesting results.

5. CONCLUSIONS

For clusters with $\gtrsim 100$ measured galaxy redshifts, we develop a new method for detecting departures from a Gaussian velocity distribution. We decompose the cluster velocity distribution into a sum of the orthogonal Gauss-Hermite functions. This method quantifies the asymmetric third-order (h_3) and symmetric fourth-order (h_4) terms of the distribution while minimizing the effect of interlopers in the tails.

The clusters A548, A1656, DC 2048−52, and Centaurus depart significantly from Gaussian. A1656 and DC 2048−52 have asymmetric velocity distributions, A548 and Centaurus have broad, symmetric distributions. The most likely interpretation is that these systems have substructure. This result is consistent with findings by other authors.

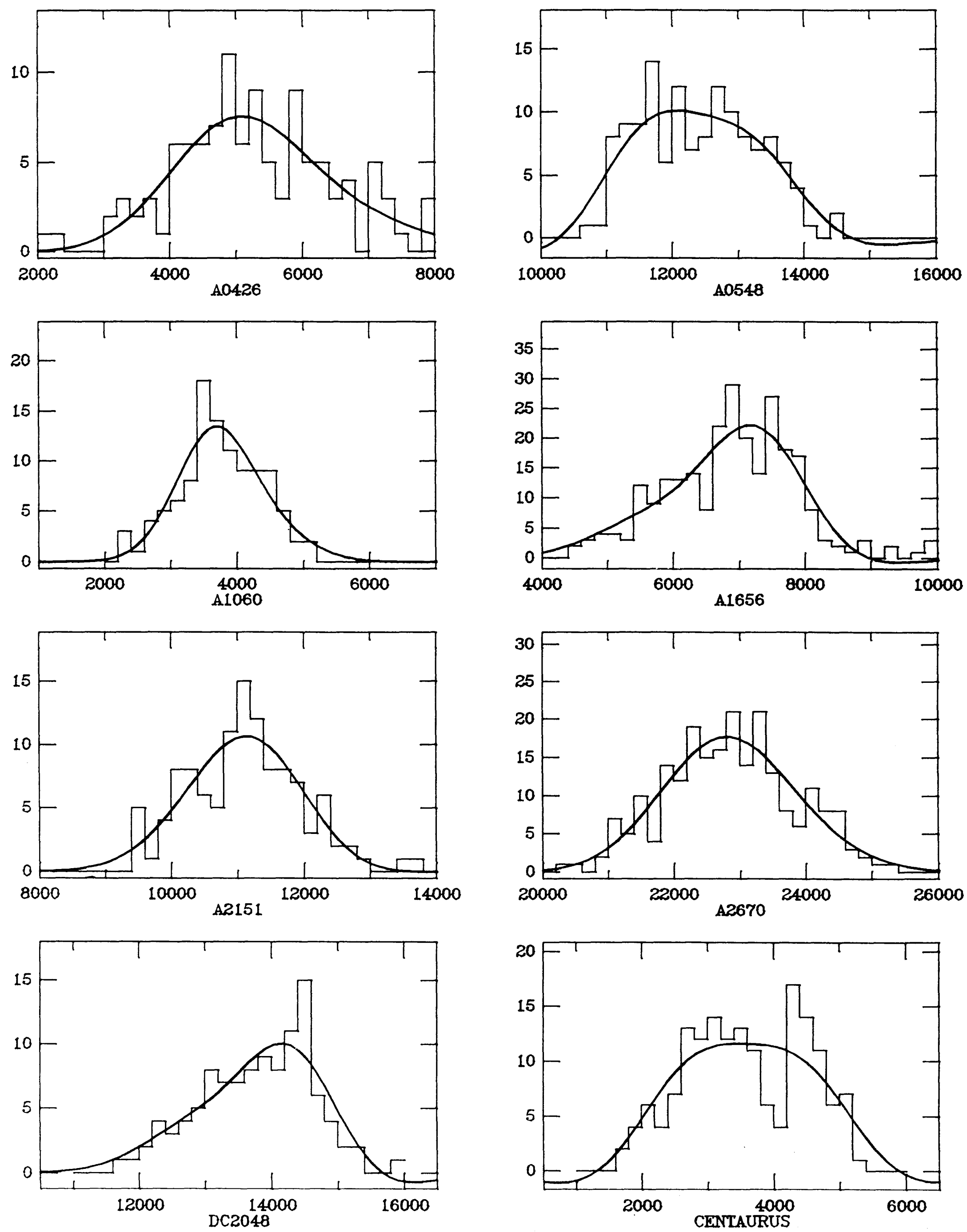


FIG. 2.—The fourth-order Gauss-Hermite expansions (*thick lines*) superposed on the line-of-sight velocity histograms for A426, A548, A1060, A1656, A2151, A2670, DC 2048−52, and Centaurus.

TABLE 1

CLUSTER DATA

Cluster	α	δ	R	Δv	N_c	V	S	h_3	$P(h_3)$	h_4	$P(h_4)$	D_{KS}	P_{KS}	Ref.
	1950.0		h^{-1} Mpc	10^3 km s^{-1}		km s^{-1}								
A426	3 15.3	41 20	1.5	2-8	110	5215	1182	0.07±0.07	0.26	0.02±0.06	0.73	0.054	0.71	1
A548	5 45.1	-25 39	2.2	10-16	133	12355	1021	0.06±0.06	0.31	-0.15±0.06†	≤ 0.005	0.087†	0.01	2
A1060	10 34.5	-27 16	1.5	2-6	106	3736	638	0.04±0.07	0.49	0.02±0.06	0.77	0.060	0.46	3,4
A1656	12 57.4	28 15	1.5	4-10	255	6975	935	-0.14±0.04†	≤ 0.005	0.02±0.04	0.67	0.061†	0.02	5
			"	3-13	261	6973	939	-0.14±0.04†	≤ 0.005	0.02±0.04	0.64	0.085†	≤ 0.005	
			"	0-20	315	6974	937	-0.11±0.04†	≤ 0.005	0.02±0.04	0.66	0.222†	≤ 0.005	
A2151	16 3.0	17 53	1.9	8-14	114	11091	855	-0.03±0.06	0.66	-0.01±0.06	0.90	0.052	0.63	2
A2670	23 51.6	-10 41	5.0	20-26	223	22834	1014	0.02±0.05	0.63	0.00±0.04	0.97	0.038	0.60	6
			2.5	"	183	22912	1107	-0.01±0.05	0.78	-0.05±0.05	0.24	0.045	0.50	
DC2048-52	20 47.9	-52 54	2.4	11-16	108	13918	886	-0.17±0.07†	0.01	0.02±0.06	0.68	0.070	0.21	2
Centaurus	12 47.0	-41 02	1.5	1-6	164	3587	1073	0.01±0.05	0.87	-0.16±0.05†	≤ 0.005	0.100†	≤ 0.005	7,8

† Significant.

REFERENCES.—(1) Kent & Sargent 1983; (2) Dressler & Schectman 1988; (3) Richter, Materne & Huchtmeier 1982; (4) Richter 1987; (5) Huchra et al. 1992; (6) Sharples, Ellis, & Gray 1988, (7) Dawe, Dickens, & Peterson 1977; (8) Dickens, Curry, & Lucey 1986.

The method allows us to characterize the shape of cluster velocity distributions with the parameters h_3 and h_4. These physical properties can be determined in the same way for simulated clusters in cosmological models. Like the distribution of cluster velocity dispersions (Frenk et al. 1990; Zabludoff et al. 1990; Weinberg & Cole 1993; Zabludoff & Geller 1993), the distributions of h_3 and h_4 are dynamical measures which may prove effective at discriminating among models for the evolution of large-scale structure with and without biased galaxy formation.

The method could be generalized to measure the departures from particular distributions in three dimensions; including the positions of galaxies would produce a more powerful test.

We thank V. McIntyre for his assistance and the anonymous referee for a thoughtful reading of this manuscript. A. I. Z. acknowledges the support of a NASA Graduate Student Research Fellowship (NGT-50667). M. F. acknowledges the support of a Hubble Fellowship (1016.01.91A). This research was made possible in part by NASA Grant NAGW-201 and by the Smithsonian Institution.

REFERENCES

Beers, T. C., Flynn, K., & Gebhardt, K. 1988, ApJ, 100, 32
Davis, M., & Mushotzky, R. F. 1992, preprint
Dawe, J. A., Dickens, R. J., & Peterson, B. A. 1977, MNRAS, 178, 675
Dickens, R. J., Currie, M. J., & Lucey, J. R. 1986, MNRAS, 220, 679
Dressler, A., & Shectman, S. A. 1988a, AJ, 95, 284
———. 1988b, AJ, 95, 985
Dubinski, J. 1992, ApJ, 401, 441
Fabricant, D. G., Kurtz, M., Geller, M. J., Zabludoff, A. I., Mack, P., & Wegner, G. 1993, AJ, 105(3), 788
Fitchett, M., & Merritt, D. 1988, ApJ, 335, 18
Fitchett, M., & Smail, I. 1993, in preparation
Fitchett, M. J., & Webster, R. L. 1987, ApJ, 317, 653
Franx, M., & Zabludoff, A. I. 1993, in preparation
Frenk, C. S., White, S. D. M., Efstathiou, G., & Davis, M. 1990, ApJ, 351, 10
Gerhard, O. E. 1993, MNRAS, in press
Huchra, J. P., Geller, M. J., Clemens, C. M., Tokarz, S. P., & Michel, A. 1992, Bull. Inf. Centre de Données Stellaises, 41
Kent, S. M., & Sargent, W. L. W. 1983, AJ, 88, 697
Lilliefors, H. W. 1967, J. Amer. Statist. Assoc., 62, 399
Lucey, J. R., Currie, M. J., & Dickens, R. J. 1986a, MNRAS, 221, 453
———. 1986b, MNRAS, 222, 427
Mellier, Y., Mazure, A., Mathez, G., Chauvineau, B., & Proust, D. 1988, A&A, 199, 67
Merritt, D. 1987, ApJ, 313, 121
———. 1988 in The Minnesota Lectures on Clusters of Galaxies and Large-Scale Structure, ed. J. M. Dickey (San Francisco: ASP), 175
Pryor, C., & Geller, M. J. 1984, ApJ, 278, 457
Richter, O. G. 1987, A&AS, 67, 237
Richter, O. G., Materne, J., & Huchtmeier, W. K. 1982, A&A, 111, 193
Rood, H. J., Page, T. L., Kintner, E. C., & King, I. R. 1972, ApJ, 175, 627
Sharples, R. M., Ellis, R. S., & Gray, P. M. 1988, MNRAS, 231, 479
van der Marel, R. F., & Franx, M. 1993, ApJ, 407, 525
Weinberg, D. H., & Cole, S. 1993 MNRAS, 259, 652
Yahil, A., & Vidal, N. 1977, ApJ, 214, 347
Zabludoff, A. I., & Geller, M. J. 1993, preprint
Zabludoff, A. I., Huchra, J. P., & Geller, M. J. 1990, ApJS, 74, 1