Four Readings Related to Vera Rubin:

1. Obituary – Published in Nature in February 2017

2. Paper by Vera Rubin and collaborators on galactic rotation curves published in 1978 in the leading journal in the astrophysics research, The Astrophysical Journal. This is rather technical, but try to work through it, avoiding getting hung up on any jargon. It summarizes a great deal of the early work by Rubin and colleagues. Figure 1, in particular, shows how the emission lines (dark lines) shift a little in wavelength when looking at one side of the spiral galaxy and the other. Sections III and IV get quite technical (don't worry if you get lost there).

3. Paper 3 is a review paper by Vera Rubin and published in the Proceedings of the National Academy of Science in 1993. It is a more accessible paper, but also descends into jargon in points. Look at Figure 2 and 3 in particular, and try to sort through what is being presented.

4. This article was written by Scott Tremaine and published in 1992 in Physics Today, a more general magazine for physicists. I suggest reading the introduction and then up to page 34.

Vera Rubin

(1928–2016)

Observational astronomer who confirmed the existence of dark matter.

CARNEGIE INSTITUTION OF WASHINGTON

Vera Cooper Rubin was a pioneering astronomer, an admired role model and a passionate champion of female scientists. Her groundbreaking work confirmed the existence of dark matter and demonstrated that galaxies are embedded in dark-matter halos, which we now know contain most of the mass in the Universe.

In the 1970s, Rubin showed that the speed at which stars orbit around the centres of spiral galaxies remains high even at the outskirts. This contradicts the Newtonian theory of gravitation, which predicts that the speeds of distant stars should fall off as the pull of gravity declines, just as the farthest planets in the Solar System orbit more slowly around the Sun than do closer ones. The discrepancy is striking — if Jupiter moved at the same rate as Earth, for instance, it would orbit the Sun every 5 years, rather than every 12.

The only plausible explanation for these galactic 'flat rotation curves' (named for their shape on a graph) was that the mass of the galaxies must extend invisibly beyond the most distant stars and gas clouds. That excess mass is known as dark matter. Its existence was first suggested in 1933 by astronomer Fritz Zwicky, who saw that galaxies in clusters moved more quickly than would be predicted from observable mass.

Dark matter, which neither emits nor absorbs light, makes up 85% of the mass-density of the Universe. We don't know what it is for sure, but most astrophysicists think that it must be a new kind of particle, different from the familiar baryons (such as protons and neutrons) in stars, planets and people. The nature of dark matter is one of science's great unsolved mysteries. The answer will change how we think about the Universe.

Rubin made her discovery by working closely with her colleague Kent Ford, who built the sophisticated optical spectrograph needed for such accurate measurements. In 1970, they published their first rotation curve for the nearby Andromeda galaxy (V. C. Rubin and K. Ford *Astrophys. J.* **159,** 379; 1970). Instead of observing hydrogen gas in galactic disks at 21-centimetre radio wavelengths, as others had done, Rubin and Ford chose to use spectral signatures at optical wavelengths. This avoided the difficulties of interpreting gas dynamics that might affect other studies. Their confirmation in 1978 of flat rotation curves in a large sample of ten spiral galaxies confirmed the existence of dark matter (V. C. Rubin *et al. Astrophys. J. Lett.* **225,** L107–L111; 1978).

Vera Rubin passed away on 25 December 2016, aged 88. She was born in Philadelphia, Pennsylvania, in 1928 and moved with her family to Washington DC when she was 10. Her fascination with the night sky was evident at an early age. She would peer out of her bedroom window and wonder what caused the stars to move across the sky, or why the Moon seemed to follow her as she rode in the back seat of her parents' car.

Her insatiable desire to understand the Universe led her to study at Vassar College in New York state, which she chose because the first professional female astronomer in the United States, Maria Mitchell, had taught there in 1865–88. Rubin graduated in 1948 as the only astronomy major in her class, and was married the same year. She earned her master's degree at Cornell University in Ithaca, New York, then moved with her husband to Washington DC, where she received her PhD from Georgetown University in 1954. Her PhD thesis demonstrated that galaxies are clumped rather than evenly distributed in space — a surprising and crucial finding, the importance of which was not recognized until many years later.

After teaching at Georgetown, Vera accepted a research position in 1965 at the Carnegie Institution's Department of Terrestrial Magnetism in Washington DC, where she remained for the rest of her career. There she conducted her groundbreaking research with Ford, and mentored generations of young astronomers in the investigation of galaxy dynamics and dark matter. She was deeply loved and admired by all who knew her.

In spite of the numerous obstacles she faced as a female scientist, Rubin triumphed. She was always cheerful, passionate and persistent. She had wanted to attend graduate school at Princeton University in New Jersey, but was denied because the university did not accept women at the time (Princeton awarded Rubin an honorary degree in 2005). She wanted to use the Hale telescope at the Palomar Observatory in California, but was denied because the facility did not allow women to do so until the 1960s. She paved a path for women not only by encouraging and inspiring them, but also by pressing for them to be hired for faculty positions, to be awarded honours and to be invited to speak at conferences. If too few women were listed as speakers, she would demand that organizers add more. As Vera liked to say, "Worldwide, half of all brains are in women."

Vera's accomplishments have been recognized by numerous honours, including the US National Medal of Science, the Gold Medal of the Royal Astronomical Society in London, the Gruber Foundation Cosmology Prize, the James Craig Watson Medal of the US National Academy of Sciences and many honorary degrees. She was widely seen as deserving of a Nobel prize for her pivotal observations. Vera's legacy as the 'mother' of flat rotation curves and dark matter will be forever remembered, as will her role in mentoring and inspiring generations of scientists, male and female. ■

Neta A. Bahcall *is a professor of astrophysics at Princeton University, New Jersey, and chair of the astronomy section of the US National Academy of Sciences. She was Vera Rubin's close friend and colleague for many years.*
e-mail: neta@astro.princeton.edu

THE ASTROPHYSICAL JOURNAL, 225:L107–L111, 1978 November 1

EXTENDED ROTATION CURVES OF HIGH-LUMINOSITY SPIRAL GALAXIES. IV. SYSTEMATIC DYNAMICAL PROPERTIES, Sa→Sc

VERA C. RUBIN,*† W. KENT FORD, JR.,* AND NORBERT THONNARD
Department of Terrestrial Magnetism, Carnegie Institution of Washington
Received 1978 June 7; accepted 1978 July 18

ABSTRACT

For a sample of 10 high-luminosity spiral galaxies, Sa through Sc, we have obtained accurate rotation curves which extend to about 80% of the de Vaucouleurs radii. For this sample: (1) All rotation curves are approximately flat, to distances as great as $r = 50$ kpc. Secondary velocity undulations indicate rotational velocities lower by about 20 km s^{-1} on the inner edges than on the outer edges of spiral features. (2) $V_{\max}$ is correlated with Hubble type, and total mass M is a function of both $V_{\max}$ and radius. At equal radii, $M(\text{Sa}) > M(\text{Sc})$. Hence, surface mass density decreases systematically along the Hubble sequence. (3) $V_{\max}$ is not correlated with luminosity or with radius. Galaxies with similar $V_{\max}$ have radii and luminosities which differ by factors of 2 and 3. This implies a large intrinsic scatter in the Tully-Fisher relation. (4) Masses are a few times $10^{11} M_{\odot}$ out to the de Vaucouleurs radius, and some masses approach $10^{12} M_{\odot}$ out to the Holmberg radius. M/L_B ratios are low, near 3.5. There is a weak suggestion that M/L_B is higher for early-type galaxies. While this is contrary to the accepted result of Roberts, masses for early-type galaxies are systematically low in his sample, due to an extrapolation procedure based on falling rotation curves.

Subject headings: galaxies: individual — galaxies: internal motions — galaxies: structure

I. INTRODUCTION

In the 50 years since Hubble (1926) introduced his classification sequence for galaxies, few systematic observational programs have attempted to study dynamical properties of galaxies as a function of Hubble type (HT). The constraints have been principally instrumental. Optical rotation curves (e.g., Burbidge and Burbidge 1975) have furnished valuable dynamical information, but generally only for the inner regions of late-type spirals. Neutral hydrogen observations have revealed integral properties of gas-rich systems (Roberts 1975), but with limited spatial resolution.

Available optical instrumentation now permits the detection of emission across a very large portion of the disks of spirals, well beyond the "turnover point" in the rotation curves. We have initiated a program to obtain spectra of spiral galaxies at high velocity resolution and at a large spatial scale, in order to study their properties as a function of HT. We now have velocities for 10 spirals, Sa through Sc, of high intrinsic luminosity. The galaxies, chosen with extreme care, have: angular diameters near 3′ or 4′ to match the KPNO/CTIO spectrograph slit lengths; high inclination so that uncertainties in inclinations produce little effect on rotational velocities and hence masses; high luminosity as indicated by the widths of their 21 cm profiles (if available); and large linear diameters. Two of the galaxies, NGC 4378 and NGC 7217, were studied originally for other reasons. These, plus NGC 3672, have been discussed separately (NGC 3672, Rubin, Thonnard, and Ford 1977 [Paper I]; NGC 4378, Rubin *et al.* 1978 [Paper II]; NGC 7217, Peterson *et al.* 1978 [Paper III]). Distances are determined from the corrected velocities, and we adopt $H = 50$ km s^{-1} Mpc^{-1} throughout.

II. THE ROTATION CURVES

Optical spectra were obtained with the Kitt Peak and Cerro Tololo 4 m spectrographs plus Carnegie image tube, usually at 25 Å mm^{-1}. Errors in the rotational velocities (measuring errors plus projection uncertainties) are generally less than ± 8 km s^{-1} per point. Reproductions of eight spectra are shown in Figure 1 (Plate L3) (Papers II and III show the other two). Many striking features can be observed directly from the spectra. (1) *All rotation curves are approximately flat*, with only a slight rise or fall following the initial steep gradient. (2) Secondary velocity undulations occur at positions of spiral features. Rotational velocities are ~20 km s^{-1} lower on the inner edges of the arms than on the outer edges. This is especially noticeable in NGC 2998, whose velocities are plotted in Figure 2. While this confirms a prediction of the density wave theory, velocity gradients will exist also in any gravitational model with a mass concentration in the arms. (3) Nuclear emission is often strong, often exhibits a large velocity gradient, and often remains intense several kiloparsecs beyond the nucleus. There follows

* Visiting Astronomer, Kitt Peak National Observatory, which is operated by the Association of Universities for Research in Astronomy, Inc., under contract with the National Science Foundation.

† Visiting Astronomer, Cerro Tololo Inter-American Observatory, which is supported by the National Science Foundation under contract AST 74-04128.

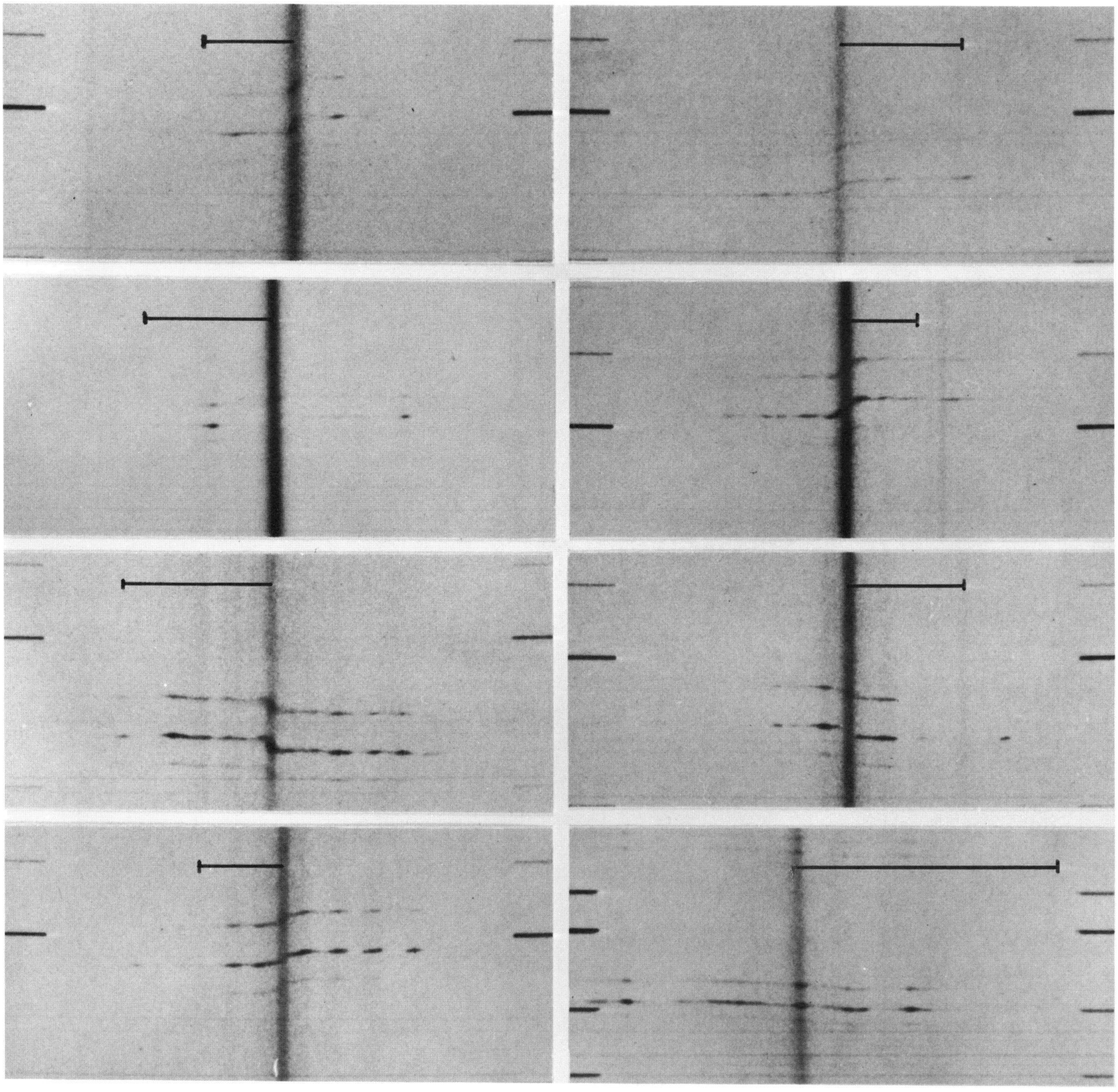


FIG. 1.—Hα region major axis spectra for galaxies of different Hubble types, taken with the 4 m RC spectrograph plus Carnegie image tube plus preflashed IIIa-J plate. Plates are (*) H_2 treated, 26 Å mm^{-1}, KPNO; or (†) N_2 baked, 52 Å mm^{-1}, CTIO. For all spectra, scale perpendicular to the dispersion is 24″ mm^{-1}, and transfer optics are f/2. (*a*)*NGC 2590, Sb, exposure 120 minutes. (*b*)*NGC 1620, Sbc, exposure 129 minutes. (*c*)†NGC 3145, Sbc I, exposure 90 minutes. (*d*)*NGC 801, Sbc–Sc, exposure 150 minutes. (*e*)*NGC 7541, Sbc–Sc, exposure 114 minutes. (*f*)*NGC 7664, Sbc–Sc, exposure 119 minutes. (*g*)*NGC 2998, Sc I, exposure 200 minutes. (*h*)†NGC 3672, Sc I–II, exposure 120 minutes. On each spectrum, Hα is strongest emission line; [N II] λ6583 is at longer λ (*up* on print). Vertical stripe is continuum from stars in nucleus. Solid horizontal line on each spectrum indicates 20 kpc in plane of galaxy. Linear extent of spectra varies from a radius of r = 17.4 kpc (NGC 2590) to r = 49 kpc (NGC 801). Note that velocity is often not constant across emission regions (spiral features) but is lower at inner edge and higher at outer edge, especially apparent in NGC 2998. The letters (*a*) and (*b*) refer to the upper left and upper right, respectively.

RUBIN *et al.* (*see* page L107)

a region of weaker emission with a shallow velocity minimum. Beyond this minimum the emission increases in intensity, with a corresponding velocity increase where the first spiral feature is encountered. (4) Star formation is proceeding vigorously over great distances in a single galaxy. In NGC 801, emission extends to a radius of 49 kpc. (5) Nuclear spectra come in a wide variety; emission lines can be narrow with a large velocity gradient (NGC 801), or broad, with the entire gradient within the nucleus (NGC 7541).

Parameters for these galaxies discussed are listed in Table 1. Classifications come from de Vaucouleurs, de Vaucouleurs, and Corwin (1975) (hereafter RC2) and/or Sandage (1978). Three galaxies not listed in

TABLE 1

DATA FOR PROGRAM GALAXIES

			Heliocentric Velocity								
NGC (1)	Hubble Class (2)	Source (3)	Opt. (km s^{-1}) (4)	21 cm (km s^{-1}) (5)	V_0 Opt. (km s^{-1}) (6)	Distance (Mpc) (7)	i (degrees) (8)	Radius r_{deV} (kpc) (9)	Nuclear Distance of Last Velocity (kpc) (10)	Fraction r_{deV} Observed (11)	Rotation Curve at Large r (12)
4378..	Sa I	deV;S	2540±12	2536	2428	48.6	35	23.4	22.0	0.94	Falling
4594..	Sa	deV	1076±10		911	18.2	84	23.6	15.	0.64	Rising
7217..	Sb–Sab III	deV;S	955±6	954±10	1236	24.7	35	13.8	11.0	0.80	Falling
2590..	Sb	N	4990±25	4992±20	4793	95.9	75	34.8	17.4	0.50	Rising
1620..	Sbc	deV	3496±10	3509±8	3424	68.5	70	30.8	21.9	0.71	Rising
3145..	Sbc I	deV;S	3650±15	3648±20	3414	68.3	61	30.8	25.3	0.82	Flat
801...	Sbc–Sc	N	5763±10	5764±20	5948	119.	86	57.1	49.1	0.86	Flat or rising
7541..	Sbc–Sc III	deV;S	2685±10	2665:	2873	57.5	72	28.6	23.2	0.81	Rising
7664..	Sbc–Sc	N	3464±10	3481±15	3709	74.2	58	35.6	28.1	0.79	Flat or rising
2998..	Sc I	deV;S	4767±10	4777±5	4781	95.6	62	39.1	34.0	0.87	Flat
3672..	Sc I–II	deV;S	1857±10	1867±10	1655	33.1	70	19.8	17.6	0.89	Flat

NGC (1)	Hubble Class (2)	$V_{\max}$ Opt. (km s^{-1}) (13)	$V_{\max}$ 21 cm (km s^{-1}) (14)	Disk Mass to Last Velocity ($10^{11} M_\odot$) (15)	Mass to r_{deV} ($10^{11} M_\odot$) (16)	$\int S dV$ (Jy km s^{-1}) (17)	H I ($10^{10} M_\odot$) (18)	H I/ $M(r_{\mathrm{deV}})$ (19)	m_B (mag) (20)	$L_B{}^{b,i}$ ($10^{10} L_\odot$) (21)	$M(r_{\mathrm{deV}})/L_B{}^{b,i}$ (22)
4378..	Sa I	320	323	3.05	3.24/3.89	6.6	0.37	0.01	12.28 S	4.53±0.4	8.6±2
4594..	Sa	344		2.8							>5
7217..	Sb–Sab III	290	313	1.13	1.41/1.55	11.7	0.17	0.01	11.08 S	4.41±0.4	3.5±1
2590..	Sb	262	251	1.74	4.11/4.52	18.3	8.1	0.18	13.4 P	10.5±2	4.3±2
1620..	Sbc	248	231	1.95	2.96	27.7	4.7	0.16	12.62 P	9.20±2	3.2±1
3145..	Sbc I	251	273	2.55	3.11	37.7	4.1	0.13	12.35 S	10.5±1	3.0±1
801...	Sbc–Sc	248	224	3.73	4.40	15.8	11.	0.25	13.1 Z	23.8±9	1.8±2
7541..	Sbc–Sc III	238	236	1.92	2.46				12.45 S	7.05±0.7	3.5±1
7664..	Sbc–Sc	208	204	1.77	2.36	29.7	4.0	0.17	12.9 Z	7.38±3	3.2±1
2998..	Sc I	211	213	2.38	2.74	25.1	5.4	0.20	12.65 S	14.9±1.4	1.8±1
3672..	Sc I–II	208	223	1.07	1.22	66.3	2.7	0.22	11.66 S	4.45±0.4	2.7±1

NOTES TO TABLE 1

Col. (3).—N, Nilson 1973. S, Sandage 1978. deV, de Vaucouleurs *et al.* 1976.

Col. (5).—Thonnard *et al.* 1978 except: NGC 4378, Krumm and Salpeter 1976; NGC 7217, Paper III; NGC 7451, Shostak 1978, contaminated by companion, not included in optical–21 cm comparison.

Col. (6).—$V_0 = V_{\mathrm{H}} + 300 \sin l \cos b$.

Col. (7).—Distance = V_0/H; $H = 50$ km s^{-1} Mpc^{-1}.

Col. (8).—$\cos^2 i = [1.042\ b^2/a^2 - 0.042]$; b/a from RC2 or Nilson 1973.

Col. (9).—Radius to 25 mag arcsec^{-2} from de Vaucouleurs *et al.* 1976 or Nilson 1973.

Col. (13).—$V_{\max}$ from smooth rotation curve.

Col. (14).—$\Delta V/2 \sin i$; ΔV read at 25% of mean flux, corrected for filter width; see n. (5).

Col. (15).—Factor of 1.1 (Brandt 1960) included.

Col. (16).—/Mass × 1.2 (Sa). /Mass × 1.1 (Sab, Sb). $\epsilon_M = \pm 25\%$; except NGC 2590, $\epsilon_M = 50\%$.

Col. (17).—See n. (5).

Col. (18).—Corrected for optical depth effects at 21 cm after Roberts 1975.

Col. (19).—See n. (18).

Col. (20).—P, Peterson unpublished; $\sigma_m = \pm 0.4$ mag. S, Sandage 1978; $\sigma_m = \pm 0.1$ mag. Z, Zwicky *et al.* 1961–68, corrected by $\Delta m = -0.4$ mag. Correction established by 4 galaxies in common with Sandage 1978, $\sigma_m = \pm 0.4$ mag.

Col. (21).—Corrected for: $A_b = 0.13$ (csc b − 1); $A_i = 0.8 \log a/b$ (RC 2); σ_L comes from σ_m (n. 20).

Col. (22).—Error estimate from ϵ_M and σ_L.

NOTE.—All data for NGC 4594 are from Schweizer 1978.

RC2 have Nilson (1973) classifications. Sources for other data and corrections applied are indicated in the notes. Differences between the optical and the 21 cm velocities are small: $\langle V_{21} - V_{opt} \rangle = 5$ km s^{-1}. However, the mean velocity difference for four galaxies with velocities in RC2 is (ours−catalog) = 77 km s^{-1}. Velocity errors this large still permeate the best available catalogs. Columns (9), (10), and (11) give the galaxy radius (25 mag arcsec^{-2}; RC2), the radius of the last measured velocity, and the ratio of the two. In the mean, our velocities extend over 80% of the galaxy radius. V_{max}, the peak velocity of the optical rotation curve (in the plane of the galaxy), is listed in column (13). Data for NGC 4594 obtained by Schweizer (1978) with the same CTIO equipment are also included.

Rotation curves are plotted in Figures 3 and 4. The general flatness of the curves, and the pronounced increase in V_{max} with earlier HT, are notable. A plot of V_{max} versus HT, Figure 5, shows this tight correlation. A correlation between V_{max} and HT found earlier by Brosche (1971) lies about 50 km s^{-1} below that indicated in Figure 5, and is defined principally from galaxies with types later than Sbc. The correlation indicated in Figure 5 may represent an upper envelope defined by high-luminosity galaxies.

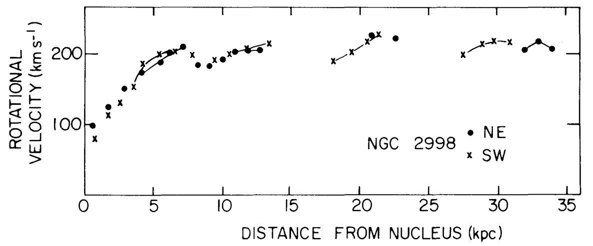


FIG. 2.—Rotational velocities in NGC 2998, as a function of distance from nucleus. Velocities for strongest emission regions are connected with lines. Note fairly good velocity agreement between velocities from NE and SW major axes, and positive velocity gradient across each arm.

III. MASSES AND MASS-TO-LIGHT RATIOS

Masses and mass distributions have been determined from the rotation curves by two procedures: (1) disk-modeled galaxies, which give lower limits, and (2)

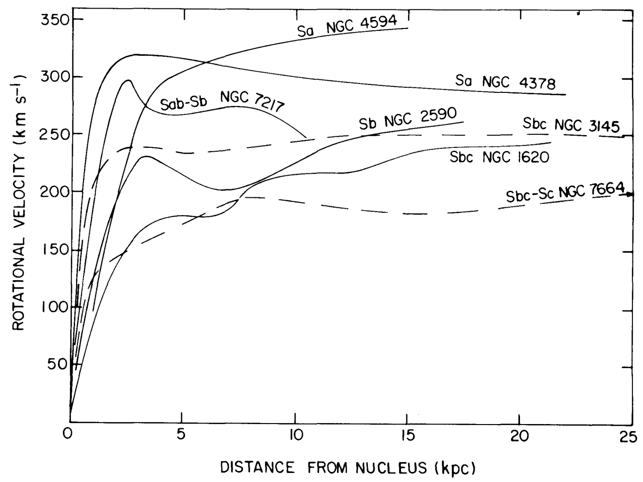


FIG. 3.—Rotational velocities for seven galaxies, as a function of distance from nucleus. Curves have been smoothed to remove velocity undulations across arms and small differences between major-axis velocities on each side of nucleus. Early-type galaxies consistently have higher peak velocities than later types.

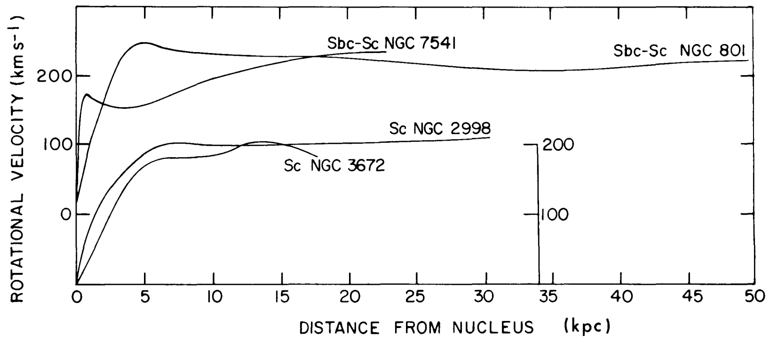


FIG. 4.—Rotation curves for two pairs of galaxies, which illustrate the lack of Tully-Fisher relation. NGC 7541 and NGC 801, both Sbc–Sc, have V_{max} values of 238 and 248 km s^{-1}. However, their luminosities (7.05 ± 0.7 and 23.8 ± 9 × 10^{10} $L_{\odot}$) and radii (23.2 and 49.1 kpc) differ by factors of 3 and 2. Similarly, the Sc galaxies NGC 2998 and NGC 3672 have V_{max} of 211 and 208 km s^{-1}, but luminosities 14.9 ± 1.4; 4.45 ± 0.4 × 10^{10} $L_{\odot}$ and radii 34.0 and 17.6 kpc.

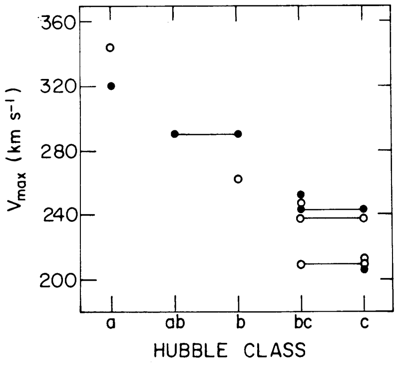


FIG. 5.—Maximum rotational velocity, in the plane of the galaxy, as a function of Hubble type. Galaxies with two classifications are entered twice with a connecting line. Open circles denote values of V_{max} which come from the last measured velocity (i.e., rising rotation curve).

spherical galaxies, which give upper limits to the masses. Both models are parameter independent. For the thin disk approximation (Kuz'min 1952), increased by 1.1 (Brandt 1960), integral masses as a function of radius (mass within a disk of radius r) are shown in Figure 6. The linear increase of mass with r is a consequence of the flat rotation curve: $M(r) = (2/G\pi) V^2(r)r$ for disk models, and V^2 is approximately constant. Because total mass varies both with r and V^2, and V is correlated with HT (Figure 5), total mass is a function of HT and of radius. While early-type galaxies have steeper mass gradients, dM/dr, large late-type galaxies (NGC 801, Figure 6) will have larger total masses than small early-type (NGC 4378). The *surface mass density decreases systematically along the Hubble sequence.*

The mass calculated to the last observed velocity, and the mass extrapolated to the RC2 radius, are listed in columns (15) and (16) of Table 1. If we had chosen to determine the masses from disk-plus-bulge models, then these masses would have been increased by about 1.2 (Sa) and 1.1 (Sab) times over the disk-molded masses. These refinements are included in column (16) and are discussed in Papers II and III. Masses to the RC2 radii lie in the range $1.2 < M(r_{\rm deV}) < 4.5 \times 10^{11} M_\odot$; random errors should not exceed 25%. The expression $M(r_{\rm deV}) = 1.5 \times 10^5 V^2r$ (V in km s^{-1}, r in kpc) reproduces all 11 masses to within 20%, and seven of them to within 10%. Total masses calculated from integrated 21 cm profiles (Roberts 1975) are in good agreement with those calculated from the detailed optical rotation curves, but are about 20% smaller, due to the falling rotation curves adopted in the 21 cm models (Roberts 1969). If the exponent in the adopted Brandt (1960) curve is changed from 3 to 2, the 21 cm masses will agree with the disk-modeled masses determined from optical rotation curves.

For galaxies modeled as spheres, $M(r) = G^{-1}V^2r$, or $\frac{1}{2}\pi$ (1/1.1) times that used above. The integrated mass out to any r is about 40% larger than the mass determined from the same rotation curve for a disk model. For spherical models, and flat rotation curves, $M_{\rm sph} = 2.1 \times 10^5 V^2r$ is a good approximation to the mass distribution. For our sample, spherical models give $2 < M(r_{\rm deV}) < 7 \times 10^{11} M_\odot$; out to their Holmberg radii (26.5 mag arcsec^{-2}), masses for some must approach $10^{12} M_\odot$. However, both disk and spherical mass distributions can reproduce the observed rotation curves.

By choosing to observe galaxies of high inclinations, we have maximized the accuracy of the rotational velocities and masses, but the uncertain internal extinction corrections pose a problem for the total magnitudes. The calculated blue luminosities (following RC2) and mass–to–blue-luminosity ratios are tabulated in columns (21) and (22). To the limits of the RC2 radii, integrated mass-to-luminosity ratios, M/L_B, are consistently small for this sample of high-luminosity spirals, near 3.5 for disk mass models, and about 5 for spherical mass models. Uncertainties in M/L_B are of the order of 1 or 2 in solar units, excluding systematic effects due to the choice of models and corrections applied.

There is a weak suggestion that M/L is higher for the earlier HT (col. [22]) a result in conflict with the widely quoted result (Roberts 1969) that M/L ratios are constant across the Hubble spiral sequence. However, five of the eight early-type galaxies in Robert's study had no 21 cm detection, so total masses come from limited optical measures, with extrapolation based on falling rotation curves. Hence, the masses for these early-type galaxies are too low by factors of from 2.5 to 5. Thus the earlier conclusions is probably not valid.

Previous correlations (Holmberg 1964) between high M/L and early HT have been questioned because of the procedure for obtaining masses (Sandage, Freeman, and

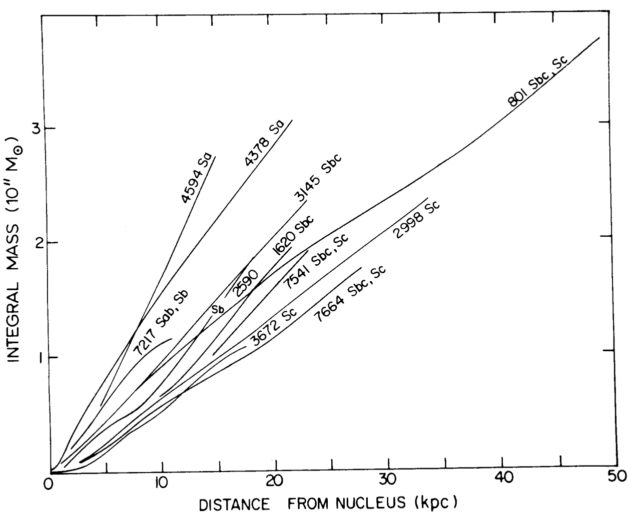


FIG. 6.—Integral mass within disk of radius r, as a function of r, for 11 galaxies, Sa through Sc, out to last measured velocity. Scale gives mass for disk models; masses for spherical galaxies are 1.4 times larger. Linear increase of mass with radius is a consequence of flat rotation curves. Mass gradient, dM/dr, is greater for early types, but total mass is a function both of V^2 and r. Surface mass density decreases systematically with Hubble type.

Stokes 1970). A correlation of M/L with HT might be expected from flat rotation curves ($M \propto V^2 r$), at least for galaxies of approximately equal radii, because of the relation between HT and V, and the tight correlation between luminosity and radius ($L \propto r^{2.4}$, Holmberg 1975). We prefer to keep an open mind on the question of whether M/L varies across the Hubble spiral sequence. At least some early-type galaxies have high M/L ratios.

IV. 21 CENTIMETER PROPERTIES

The 21 cm integrated flux densities, the hydrogen masses, and the hydrogen-mass–to–total-mass ratios are listed in columns (17), (18), and (19). H I constitutes 1% of the mass for our single Sa and Sab galaxies, but about 20% of the mass for the Sb and Sc galaxies.

From the 21 cm profile width, ΔV, we determine the peak velocity of the rotation curve, $V_{\max} = \Delta V/(2 \sin i)$ (col. [14]). This value is in good agreement with $V_{\max}$ (optical), found from the (resolved) optical rotation curve. The mean difference, $\langle V_{\max}$ (opt) $- V_{\max}$ (21 cm)$\rangle$, is 1 ± 6 km s^{-1}. A major surprise is the lack of a correlation of $V_{\max}$ (optical or 21 cm) with luminosity or with radius, i.e., a Tully-Fisher (1977) relation. The pairs of rotation curves in Figure 5 show that we do not observe such a correlation. Galaxies of the same HT and same $V_{\max}$ have radii which differ by factors of 2, and luminosities which differ by factors of 3.

V. CONCLUSIONS

The major result of this work is the observation that rotation curves of high-luminosity spiral galaxies are flat, at nuclear distances as great as $r = 50$ kpc. Roberts and his collaborators (Roberts 1976) deserve credit for first calling attention to flat rotation curves. Recent 21 cm observations by Krumm and Salpeter (1976, 1977) have strengthened this conclusion. These results take on added importance in conjunction with the suggestion of Einasto, Kaasik, and Saar (1974), and Ostriker, Peebles, and Yahil (1974) that galaxies contain massive halos extending to large r. Such models imply that the galaxy mass increases significantly with increasing r which in turn requires that rotational velocities remain high for large r. The observations presented here are thus a necessary but not sufficient condition for massive halos. As shown above, mass distributions from disk models or spherical models adequately reproduce the observed velocities. The choice between spherical and disk models is not constrained by these observations.

We thank the directors of Kitt Peak National Observatory and Cerro Tololo Inter-American Observatory for observing time, Drs. J. Ostriker, M. S. Roberts, R. J. Rubin, and F. Schweizer for valuable comments, and A. Sandage for making available his galaxy classifications before publication. V. R. thanks the staff of CTIO for their hospitality during her stay as an invited Visiting Astronomer, where these results were compiled.

REFERENCES

Brandt, J. C. 1960, *Ap. J.*, **131**, 293.
Brosche, P. 1971, *Astr. Ap.*, **13**, 293.
Burbidge, E. M., and Burbidge, G. R. 1975, in *Galaxies and the Universe*, ed. A. and M. Sandage and J. Kristian (Chicago: University of Chicago Press).
de Vaucouleurs, G., de Vaucouleurs, A., and Corwin, H. G. 1976, *Second Reference Catalogue of Bright Galaxies* (Austin: University of Texas).
Einasto, J., Kaasik, A., and Saar, E. 1974, *Nature*, **250**, 309.
Holmberg, E. 1964, *Uppsala Astr. Obs. Medd*, No. 148.
———. 1975, in *Galaxies and the Universe*, ed. A. and M. Sandage and J. Kristian (Chicago: University of Chicago Press).
Hubble, E. 1926, *Ap. J.*, **64**, 321.
Krumm, N., and Salpeter, E. E. 1976, *Ap. J. (Letters)*, **208**, L7.
———. 1977, *Astr. Ap.*, **56**, 465.
Kuz'min, G. G. 1952, *Pub. Tartu Astr. Obs.*, **32**, 211.
Nilson, P. 1973, *Uppsala Obs. Ann.*, Vol. **6**.
Ostriker, J. P., Peebles, P. J. E., and Yahil, A. 1974, *Ap. J. (Letters)*, **193**, L1.
Peterson, C. J., Rubin, V. C., Ford, W. K., Jr., and Roberts, M. S. 1978, *Ap. J.*, Vol. **226**, in press (Paper III).
Roberts, M. S. 1969, *A.J.*, **74**, 859.
———. 1975, in *Galaxies and the Universe*, ed. A. and M. Sandage and J. Kristian (Chicago: University of Chicago Press).
———. 1976, *Comments Ap.*, **6**, 105.
Rubin, V. C., Ford, W. K., Jr., Strom, K. S., Strom, S. E., and Romanishin, W. 1978, *Ap. J.*, **224**, 782 (Paper II).
Rubin, V. C., Thonnard, N., and Ford, W. K., Jr. 1977, *Ap. J. (Letters)*, **217**, L1 (Paper I).
Sandage, A. 1978, in preparation.
Sandage, A., Freeman, K. C., and Stokes, N. R. 1970, *Ap. J.*, **160**, 831.
Schweizer, F. 1978, *Ap. J.*, **220**, 98.
Shostak, S. 1978, *Astr. Ap.*, submitted.
Thonnard, N., Rubin, V. C., Ford, W. K., Jr., and Roberts, M. S. 1978, *A.J.*, submitted.
Tully, R. B., and Fisher, J. R. 1977, *Astr. Ap.*, **54**, 661.
Zwicky, F., Herzog, E., Wild, P., Karpowicz, M., and Kowal, C. T. 1961–1968, *Catalogue of Galaxies and Clusters of Galaxies* (Pasadena: California Institute of Technology), Vols. **1–6**.

W. Kent Ford, Jr., Vera C. Rubin, and Norbert Thonnard: Department of Terrestrial Magnetism, Carnegie Institution of Washington, 5241 Broad Branch Road, N.W., Washington, DC 20015

Proc. Natl. Acad. Sci. USA
Vol. 90, pp. 4814–4821, June 1993
Colloquium Paper

This paper was presented at a colloquium entitled "Physical Cosmology," organized by a committee chaired by David N. Schramm, held March 27 and 28, 1992, at the National Academy of Sciences, Irvine, CA.

Galaxy dynamics and the mass density of the universe

VERA C. RUBIN

Department of Terrestrial Magnetism, Carnegie Institution of Washington, Washington, DC 20015

5

ABSTRACT Dynamical evidence accumulated over the past 20 years has convinced astronomers that luminous matter in a spiral galaxy constitutes no more than 10% of the mass of a galaxy. An additional 90% is inferred by its gravitational effect on luminous material. Here I review recent observations concerning the distribution of luminous and nonluminous matter in the Milky Way, in galaxies, and in galaxy clusters. Observations of neutral hydrogen disks, some extending in radius several times the optical disk, confirm that a massive dark halo is a major component of virtually every spiral. A recent surprise has been the discovery that stellar and gas motions in ellipticals are enormously complex. To date, only for a few spheroidal galaxies do the velocities extend far enough to probe the outer mass distribution. But the diverse kinematics of inner cores, peripheral to deducing the overall mass distribution, offer additional evidence that ellipticals have acquired gas-rich systems after initial formation. Dynamical results are consistent with a low-density universe, in which the required dark matter could be baryonic. On smallest scales of galaxies [10 kiloparsec (kpc); H_0 = 50 km·sec^{-1}·megaparsec^{-1}] the luminous matter constitutes only 1% of the closure density. On scales greater than binary galaxies (i.e., ≥100 kpc) all systems indicate a density ≈10% of the closure density, a density consistent with the low baryon density in the universe. If large-scale motions in the universe require a higher mass density, these motions would constitute the first dynamical evidence for nonbaryonic matter in a universe of higher density.

Twenty years ago, astronomers started accumulating observations that indicated that orbital velocities of stars and gas in spiral galaxies remained high out to the limits of the observable galaxy. By 1982, after a decade of initial disquiet, most astronomers reluctantly accepted the conclusion that a galaxy consists of much more than the luminous stars, gas, and dust that can be observed at various wavelengths. In fact, at least 90% of a galaxy consists of dark matter, matter that is undetectable at any wavelength and that is deduced only by its gravitational influence on the orbits of the stars and the gas that we can see. But not all astronomers were reluctant to accept the evidence. "There are reasons, increasing in number and quality, to believe that the masses of ordinary galaxies may have been underestimated by a factor of 10 or more," stated Ostriker, Peebles, and Yahill (1) in a truly prescient sentence.

This revolution in our understanding came about because theoretical arguments revealed the necessity of stabilizing the disks of galaxies with invisible halos (2), because observations with optical (3, 4) and with radio (5) telescopes showed uniformly high rotation velocities, and because an enormously influential review of the inferred distribution of matter in galaxies (6) showed the validity of a universe in which most of the matter is dark. In the decade of the 1980s, observational and theoretical investigations of the dark matter became one of the most active fields of extragalactic astronomy, and excellent reviews exist (7, 8).

For this paper, I describe a handful of recent results that I consider to be among the most significant of the last few years. They relate to the dynamics of disk and elliptical galaxies, to the distribution of nonluminous matter in galaxies and in clusters, and to estimates of the overall mass density of the universe.

Section I. The Milky Way and the Local Group Are Massive

We live in a large, massive spiral galaxy, the Milky Way, with a nearby equally large and massive companion spiral, the Andromeda galaxy. Together, these contain most of the mass within the Local Group. For the past several decades, determinations of the mass within our galaxy and the Local Group have been imprecise. This was due principally to the lack of accurate velocities and lack of knowledge of orbits for the outlying globular clusters and small satellite galaxies whose motions are used to map the gravitational potential of the Galaxy, and a lack of knowledge of the evolutionary history of the dwarf satellites. Recently, more accurate radial velocities and distances of 47 objects on the outskirts of our galaxy have become available. Based on the assumptions that (*i*) the distribution of dark and luminous matter can be modeled by a power-law function, and (*ii*) that the velocity distribution of the outer Galaxy is isotropic, Kulessa and Lynden-Bell (9) estimate a Galaxy mass of 1.3×10^{12} solar mass units ($M_\odot$) out to the distance sampled by the test objects, 230 kiloparsec (kpc; 1 pc = 3.09×10^{16} m). Based on the earlier Schmidt (10) model of the Galaxy, a mass of $\approx 1.9 \times 10^{11}$ $M_\odot$ is located within 25 kpc of the center (11). Hence, the new analysis increases the radius, the mass, and the mass-to-light ratio (M/L_B in solar units, with the luminosity measured in the blue spectral band) each by nearly an order of magnitude. At the 25-kpc limit of the optical disk, $M/L_B \approx 10$, and the dark matter constitutes 70% of the total mass (11); at 230 kpc, $M/L_B \approx 50$, and dark matter constitutes >90% of the mass (9).

A statistical analysis using the velocities of satellites orbiting a homogeneous set of spiral galaxies similar to the Milky Way (12) leads to a mass of order 2×10^{12} $M_\odot$ within 200 kpc, in agreement with that discussed above. This large mass and mass-to-luminosity ratio for our galaxy (and hence also for Galaxy M31) is also in comfortable agreement with a mass of $\approx 6.4 \times 10^{12}$ $M_\odot$ determined by Peebles (13) for the Local Group. Note that a radius of 230 kpc reaches suspiciously close to the 350-kpc half-way point between us and the Andromeda Galaxy. If the dark halos of our Galaxy and

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: M/L, mass-to-light ratio; HI disk, neutral hydrogen disk; pc, parsec; $M_\odot$, solar mass unit.

M31 reach each other, the values of M/L_B will increase to nearly 100, a value of M/L similar to that for other systems of similar radial scale. I discuss M/L values as a function of radius in *Section IV* below.

Section II. All Spiral Galaxies Need Dark Matter (Halos)

For spirals more distant than M31, velocity information is usually obtained from spectral observations of ionized hydrogen (Hα) sampled by a long-slit spectrograph aligned along the major axis and from 21-cm mapping of velocities of the neutral hydrogen gas. To derive a mass distribution from a velocity distribution, it is generally assumed that Newtonian gravitational theory is valid, that a self-gravitating luminous disk of constant M/L dominates at small radial distances, that a dark isothermal halo dominates at large radial distances, and that the internal extinction is not a function of radius. Surely the internal extinction merits a more sophisticated treatment, as does the distribution of mass in the halo, but the above simplifications suffice at present.

Outermost Regions. Virtually all observed spiral rotation curves are flat (or rising) to the limits of the observations, implying that the mass density decreases approximately as $1/r^2$. The luminosity however falls off faster, exponentially. Thus, the ratio of mass density to luminosity density increases with radial distance, requiring the presence of a halo of dark matter, which becomes increasingly dominant with increased nuclear distance. A major requirement for deriving meaningful parameters for the halo and for the disk is that rotation velocities extend from the nucleus to at least several disk scale lengths (where one scale length equals the radial distance at which the disk surface brightness has fallen by $1/e$). Velocity measures for spirals now extend this far, especially for galaxies whose neutral hydrogen disks (HI disks) are several times more extended than their optical disks. I show in Fig. 1 an optical image of the M101 Galaxy, with the superposed velocity contours of the neutral hydrogen disk; an HI disk with a radius three or four times that of the optical disk is quite impressive.

To determine the contribution to the rotational velocities from the mass of the halo, the contribution to the velocity of a constant M/L disk must first be subtracted in quadrature from the observed velocities. However, even an extended rotation curve is insufficient to constrain the mass distribution because M/L of the disk is difficult to pin down. For example, velocities in the NGC 3198 Galaxy can be well-fitted by models with a disk of a range of masses, from a disk of maximum mass (M/L = 4.4) down to a disk of zero mass (ref. 15; Fig. 2). During the mid-1980s, it was customary to analyze rotation curves by adopting the maximum mass disk consistent with the inner rise of the rotation curve (17). These models generally resulted in dark and luminous mass of about equal amounts within the optical galaxy, but it must be remembered that the halo mass so determined is only a lower limit.

A major conceptual advance came when Athanassoula and colleagues (18) pointed out that information concerning the spiral structure of the disk could be used as an additional constraint on the relative halo-to-disk mass. For galaxies with low mass halos, the disk is unstable against the formation of a nonaxisymmetric m = 1 mode. Adding mass to the halo inhibits the m = 1 mode, until the m = 2 (i.e., two-armed) mode dominates. By requiring m = 2 (but no m = 3) modes, disk masses of order one-half the maximum disk mass generally result. Even more importantly, these authors demonstrated that every spiral with a rotation curve extending several disk scale lengths requires dark matter, finally putting to rest the belief (19) that some galaxies (always only those with not very extended rotation curves) could be modeled with no dark halos.

In an observational survey of spirals with extended HI disks, Casertano and van Gorkom (20) discovered two spiral galaxies, NGC 2683 and 3521, whose extended HI disks exhibit falling velocities beyond the optical disk. Surprisingly, dark matter is required even for these spirals. For NGC

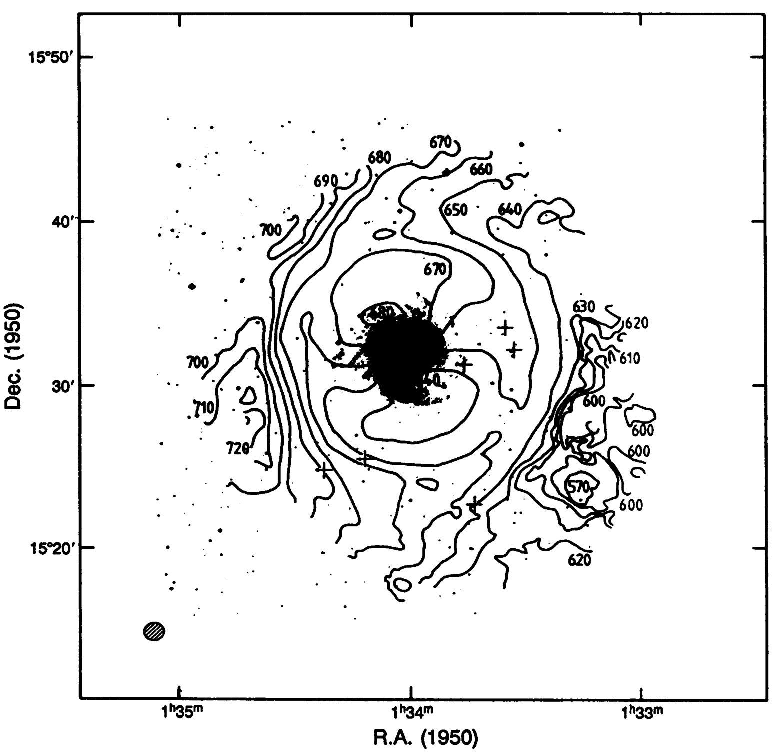


FIG. 1. An optical image of the spiral galaxy M101 with velocities (V) from 21-cm neutral hydrogen (HI) superposed (14). The hydrogen diameter is about four times the optical diameter. Angular coordinates on the sky (right ascension and declination) are the indicated coordinates.

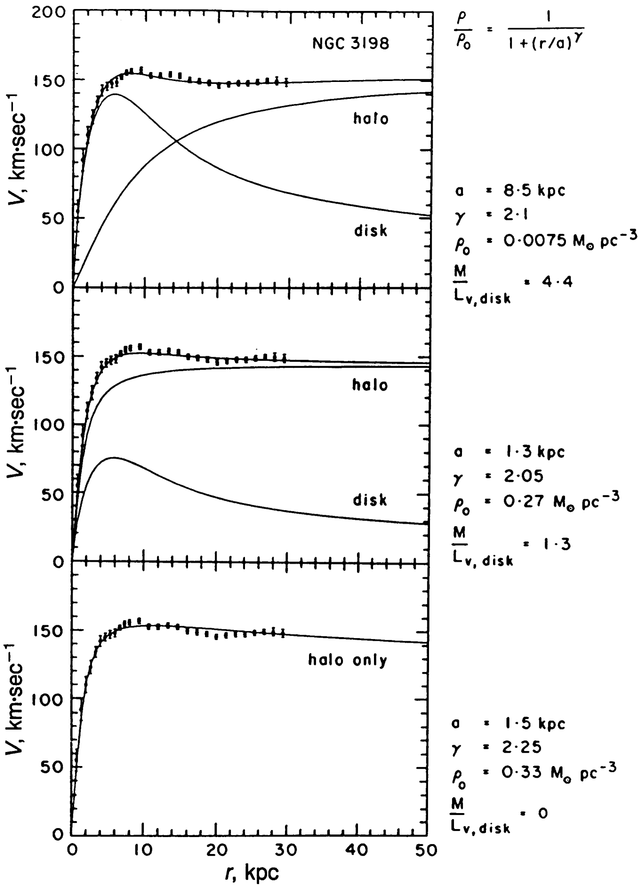


FIG. 2. Three alternative decompositions of the rotation curve of NGC 3198 (15). The observed velocities (16) are shown as points, with the model disk and halo-rotation curves plus their sum in quadrature (curves). The model parameters are given at right, for the assumed halo-density distribution ρ. Note the large ranges permitted by the models for values of central densities ρ_0 and disk M/L values.

2683, even the maximum acceptable disk (M/L = 8, set mostly by the maximum of the velocities) predicts velocities falling more rapidly than observed (Fig. 3). The authors argue that there is also a *lower* limit to the disk mass, for too much halo mass produces too flat a rotation curve. It is unclear how much this conclusion would be altered if the innermost optical velocities, discussed in *Innermost regions* below, had been available. These near-nuclear velocities are not observable at the low spatial resolution of the 21-cm observations.

From these and many other recent observations of velocities in the outermost regions of spirals, I draw the following conclusions. Virtually all spirals need dark matter in amounts significantly greater than the luminous matter in order to reproduce the observed velocities. Beyond the optical galaxy, rotation velocities in the gas disk are generally constant, implying masses that continue to rise linearly with radius, even in those regions where the optical luminosity is essentially nil. Hence, local values of M/L are enormous. Consequently, only lower limits to the mass and to M/L values are available for spiral galaxies. The only exceptions to this generalization are tidally distorted galaxies, and a few moderately peculiar spirals in dense regions near the cores of galaxy clusters (21), whose velocities decrease with increased nuclear distances.

Innermost Regions. Advances in radio astronomy techniques have permitted the detection and velocity mapping of faint outer disks and, consequently, the determination of the mass interior to large radial distances. Corresponding advances in optical instrumentation have permitted the spectral

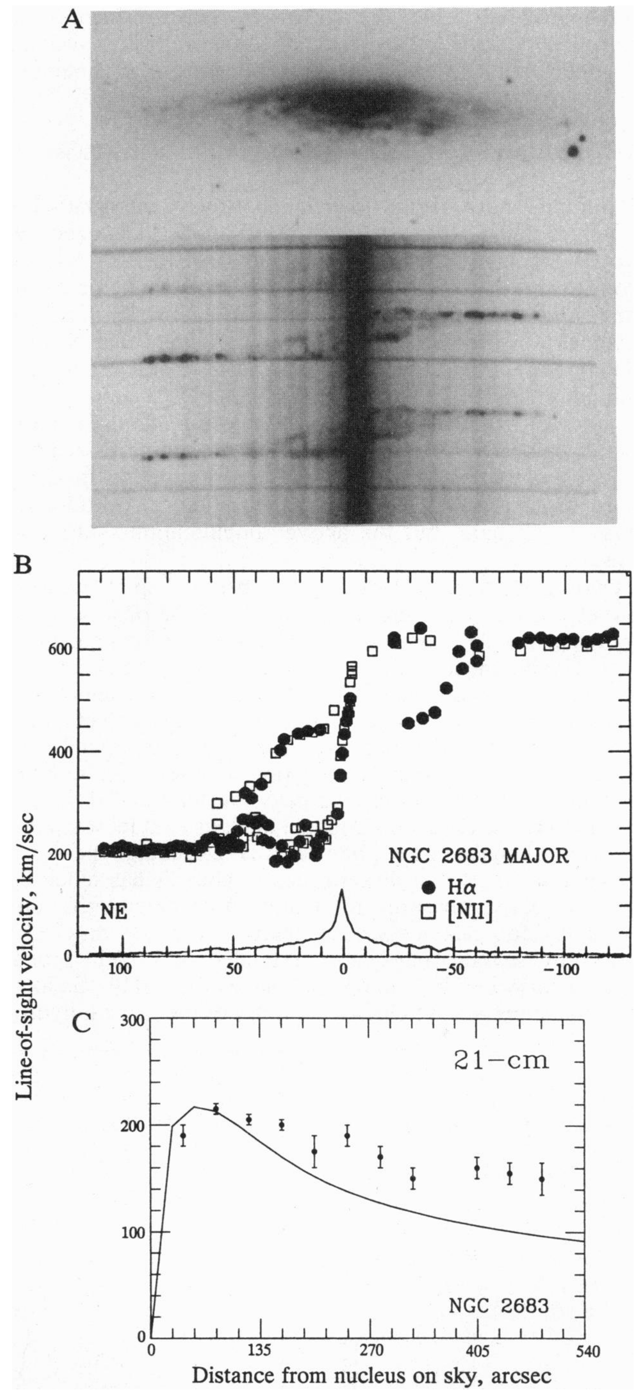


FIG. 3. (*A*) NGC 2683, an Sb spiral, from a Kitt Peak 36-inch charge-coupled device image. Below it, a portion of a Kitt Peak 4-m spectrum, showing Hα (*Upper*) and [NII] emission. Note the multivalued emission in the inner regions. (*B*) The measured optical velocities in the inner region from the spectrum in *A*, quantifying the rapid rise across the nucleus and the multivalued velocities. The surface brightness profile of the galaxy is shown by the fine line. (*C*) HI velocities in NGC 2683 (20). The solid line represents the predicted velocities based on a model with no dark matter. Thus, even the falling rotation velocities are higher than the velocities predicted by the light. Note the change in radial scale from *B* to *C*; all the velocities in *B* lie between r = 0 and the first two 21-cm points.

observations of inner-core regions at high spatial and high velocity resolution. Moreover, the use of digital detectors (charge-coupled devices) allows the smooth galaxy stellar continuum, strong in the inner regions, to be easily removed from the spectrum, thus revealing the emission of the ionized gas at enhanced signal-to-noise ratio. From these observations we now know that gas in central regions often exhibits complex velocity patterns arising from a variety of causes: warps, twists, or nonalignments of gas disks; time-dependent settling of recently acquired gas; nonaxisymmetric (triaxial) orbits that may explain why some ellipticals and early-type spirals have rotational velocities that increase too slowly for a constant M/L disk; and even holes in the innermost regions and warps in the outermost regions, so that gas along the line-of-sight at various nuclear distances and various projected velocities superposes at a single radial position.

NGC 2683, one of the two spirals with a falling 21-cm rotation curve discussed above, is a prime example exhibiting complex inner velocities. I show in Fig. 3 an image of the Galaxy, the Hα region of the spectrum in the inner 2′ with the underlying stellar continuum removed, the near-nuclear optical velocities, and the 21-cm velocity curve from Casertano and van Gorkom (20).

Within the inner 60″, velocities in NGC 2638 are multivalued, extending from the upper maximum velocity down to velocities negative with respect to the center. The steep velocity rise implies a massive nucleus, of order 5×10^8 $M_\odot$. The upper envelope represents the rotation velocities within the disk, while at least some of the lower-velocity gas is likely foreground gas warped out of the plane and intruding into the line-of-sight. Other low-velocity gas must arise from disk gas of low angular momentum, probably derived from bulge stars. None of these details are visible at the lower-resolution 21-cm observations. In fact, these optical observations all lie closer to the nucleus than the second plotted point on the 21-cm velocity curve; in velocity they rise slightly with decreased nuclear distance to the velocity $V \approx 220$ km/sec at 2″ from the nucleus. Thus, a realistic mass deconvolution into halo, disk, and bulge or central point mass (necessary to produce the steep nuclear velocity gradient) must be based on velocities for the inner regions obtained from high-resolution optical observations and for outer regions on velocities from high-sensitivity radio observations. Let us now resurrect Bosma's (5) forgotten suggestion for such composite curves.

Complex inner regions are not rare among galaxies. Rubin and Kenney (22) have discovered rapidly rotating cores with kinematic properties distinct from the normal outer rotation in ≈25% of the 70 Virgo cluster spirals and ellipticals they have observed at high-velocity resolution and high spatial scale. The galaxies are located in all regions of the cluster, span Hubble types from E through Sc, and have a wide range of central velocities and viewing angles. For a spiral viewed close to edge-on, a complex velocity field can be ascribed to an outer warp projected along the line of sight. For a galaxy viewed more nearly face-on, the inner region itself must have a warped or skew or twisted morphology. We are continuing our analysis of these interesting galaxies. Complex nuclear regions are especially prevalent in elliptical galaxies, and I discuss these in *Section VI* below.

Section III. Dwarf Spirals Contain Mostly Dark Matter

It was recognized a decade ago that low-luminosity galaxies would be important for learning about dark matter. The systematics of galaxy rotation curves as a function of luminosity suggested the query, " . . . for the smaller galaxies, . . . Is the luminous matter only a minor component of the total galaxy mass?" (4). Moreover, the low mass, and consequent low acceleration, suggests that extremely low-mass galaxies are a fertile testing ground for theories that reject the existence of dark matter but instead modify Newtonian gravitational theory.

A major discovery of the last few years is the existence of low-mass galaxies that are enormously gas rich and in which the dark mass exceeds the luminous mass by a large amount. DDO 154, one of the dwarfs richest in neutral hydrogen (23, 24), contains a neutral hydrogen gas disk with radial extent five times the optical disk. The gas disk is dynamically more important than the stellar disk; stars constitute only ≈1% of the mass. It is the very low velocity gradient beyond the nucleus, combined with the steep fall in disk luminosity, which forces the very low-mass disk. Interior to the last measured velocity, at least 95% of the mass is dark, and M/L is ≈50, almost an order of magnitude greater than a typical type Sc galaxy (within its optical radius).

The existence of dwarf spirals with such a small fraction of luminous mass is remarkable. One question that naturally follows, "Is there a galaxy that is totally dark?", has not yet been answered.

Equally intriguing is the dwarf spiral NGC 1560 (25). Gas rich, the best-fit model has the stellar and gas masses about equal, together contributing about one-half the mass of the dark matter at the last measured point. Especially remarkable is the close similarity in form with radius of the *total* mass and the *gas* mass, particularly as indicated by a dip in both the velocity and the gas density at ≈5 kpc. This similarity implies a close relation between the distribution of the HI gas and the dark matter. It also makes spiral NGC 1560 especially interesting for astronomers interested in efforts to modify Newtonian dynamics to account for flat rotation curves. Milgrom (26) has proposed modifications to Newtonian dynamics (MOND theory), such that at low accelerations the attractive force varies as $1/r$. In conventional Newtonian dynamics, "what you see is *not* what you get"; that is, the luminous galaxy is embedded in a dark halo whose mass is the dominant component in producing the observed velocities. Thus, the velocities predicted by the observed gas density would be expected to have only an insignificant relation to the total observed velocities produced by the combined gravitational field of the stars, gas, and (dominant) dark matter.

In contrast, in the MOND theory, "what you see is what you get," so the agreement of the radial variation of the observed velocities and the (Newtonian predicted) gas-plus-stars-predicted velocities is fundamental to the theory. While few astrophysicists are comfortable with a modified Newtonian theory, such theory cannot yet be ruled out on the basis of the dwarf spirals.

Section IV. Elliptical Galaxies Are Much More Complex Than We Thought

Velocity Studies from Stellar Absorption Lines. Ten years ago, discussions of galaxy dynamics were based on studies of spirals only, for little was known of the kinematics of elliptical galaxies. Ellipticals were unstudied, not because they were considered simple but because the observations were hard. Motions in spiral galaxies are relatively easy to study. Orbital motions are circular, the viewing geometry is well defined, for a given integration time the prominent gas optical emission lines are detectable with a signal-to-noise well above that of the absorption lines, and the neutral hydrogen gas disks are a prime feature for 21-cm studies. Thus, it is not surprising that during the 1970s, optical observations with image tube plus photographic plate and 21-cm observations with new interferometers advanced our knowledge of the kinematics of disk galaxies to a degree that was then not matched for elliptical galaxies. But during the last decade, knowledge of dynamics of elliptical galaxies has started to catch up, due, in part, to efficient digital recording devices and, in part, to the surprise discovery of significant emission in some ellipticals, emission that is studied in both the optical and radio spectral regions. Thus, although an uncertain viewing angle and complicated triaxial orbits keep the analysis complex, unexpectedly interesting dynamical

5

results are now available for a significant number of elliptical galaxies.

It is now understood that many bright ellipticals are triaxial in shape, supported not by their (slow or absent) rotation but by their anisotropic velocity dispersions (27, 28). But some ellipticals are rotating, albeit slowly. In a recent study of 22 ellipticals, many of them round, Franx, Illingworth, and Heckman (29) detect rotation >20 km/sec in 21 of them. Of these, 60% of the ellipticals had their kinematical axis within 10° of the photometric minor axis. Yet for six of the galaxies, there is significant minor-axis rotation; in two of these the minor-axis rotation is more rapid than the major-axis rotation; in two of the galaxies the minor-axis rotation is approximately equal to the major-axis rotation. Minor-axis rotation in triaxial galaxies may be a clue to the formation mechanism of these galaxies but is currently incompletely understood.

In the inner regions, still more complexity exists. Two of the ellipticals exhibit skew-rotating small nuclear cores, another has a counter-rotating core, and a fourth has a core that rotates perpendicular to the rotation of the major axis. In the Franx *et al.* (29) study, almost 20% of the galaxies exhibited kinematically distinct nuclear cores, close to the fraction Rubin and Kenney observe for spirals and ellipticals in Virgo.

It is not yet possible to translate absorption line kinematical studies into statistics of M/L values for more than a few ellipticals (30, 31). Within a few scale lengths, M/L values appear generally constant, and often small, an even tighter constraint on the amount of dark matter (yet this statement has the ring of statements made early in the history of observations of spiral galaxies, when it was expected that more extended observations would show rotation velocities to fall). Important observational evidence not discussed here, such as x-ray observations of halos of ellipticals (32) and velocities of individual planetary nebula far out in ellipticals (30, 33), offers impressive evidence that elliptical galaxies do contain extended dark matter components. While absorption line studies may not be a major contributor to learning about the distribution of extended matter in ellipticals, such studies will surely teach us important details about the kinematics and evolutionary history of ellipticals.

Velocity Studies from Gas Emission Lines. Among the many recent surprises coming from the observations of elliptical galaxies has been the discovery that a significant number contain gas disks that can be used to map the gravitational potential. A collection of rotation curves for elliptical galaxies, determined from emission lines, is shown in Fig. 4. Also shown is the predicted velocity curve for a spherical galaxy with a normal surface brightness profile; the parallel form for the predicted and observed relations suggests that M/L is essentially constant over these radial ranges. A more definitive statement concerning values of M/L as a function of galaxy radius can be made for only a few special galaxies, including IC 2006.

IC 2006 (35) is an isolated spherical galaxy surrounded by a counter-rotating outer ring of HI. Motions within the galaxy imply a value of $M/L_B = 5$ for the core; velocities in the outer ring suggest that $M/L_B = 16$ overall. Just as for spirals, this increase in mass per unit of luminosity implies the existence of nonluminous matter that becomes increasingly dominant with increased radial distance.

But outer gas disks are rare. More common in elliptical galaxies are inner disks of ionized gas extending only to a small fraction of the galaxy radius. Often they are dynamically decoupled from the principal rotation of the galaxy. Such disks—e.g., NGC 5128 (36) and NGC 1316 (37)—are now understood as the product of the acquisition of a small gas-rich system. This gas and dust will require numerous orbital periods to reach dynamical equilibrium and meanwhile will be observed as a warped and twisted disk with complex internal kinematics.

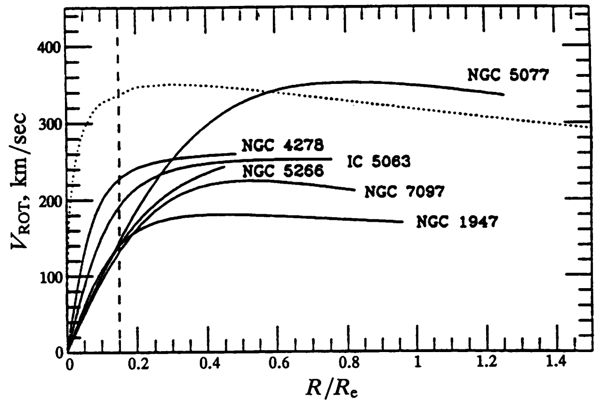


FIG. 4. Rotation curves for six elliptical galaxies plotted as a function of R_e, the half light radius (R = radius) (34). Velocities (V_{ROT}) are measured from their gas disks. The dotted curve is the rotation curve for a spherical galaxy with a normal $r^{1/4}$ surface brightness profile.

NGC 5077 is an elliptical galaxy with a disk of ionized gas along its minor axis, a not uncommon form for ellipticals. A detailed study of its gas velocities (34) shows that either triaxial or oblate mass models with constant M/L can match the velocity fields of the stars, the gas, and the central velocity dispersion. Even though the rotation velocities increase slowly from the nucleus, such velocities can arise from gas in elliptical orbits; it is not necessary to invoke an M/L increasing with radial distance. Moreover, models with the gas located in a time-evolving polar ring predict that the ring becomes enormously twisted, so that lines of sight intersect the surface two or three times, thus giving rise to the complex multivalued velocities they observe in the gas.

N-body simulations, too, produce small counter-rotating nuclear disks. Barnes and Hernquist (38, 39) show that a prograde merger of two gas-rich stellar disks produces a galaxy photometrically indistinguishable from an elliptical, which contains a kinematically distinct but small (≈800 pc) central gas core. Such cores may be skew or counter rotating, for the tidal trauma has destroyed all knowledge of the previous orientation of the gas. However, the more extended gas, settled in a disk, rotates prograde.

Many comprehensive reviews of the current state of observations and understanding of the dynamics of elliptical galaxies are available (40, 41). It is clear that the freedom afforded by triaxial orbits combined with viewing angles offers wide possibilities for models. Adding warped, twisted, precessing, and time-dependent gas disks increases the range of parameters. Arriving at the correct model may be difficult, for increasingly detailed observations may only add to the complexity, rather than constrain the range of parameters. However, the galaxy-by-galaxy studies that are continuing at an accelerating pace will ultimately make clear the evolutionary importance for ellipticals of relatively newly acquired material; with luck we will also learn about the overall distribution of matter in elliptical galaxies.

NGC 4550. Recently, I have been studying, in collaboration with Graham and Kenney (42), the kinematics of spiral and elliptical galaxies in the Virgo cluster. NGC 4550 is a featureless E7/S0 galaxy located in the core of the Virgo cluster only 1.2° from M87 (Fig. 5). We have discovered a unique pattern of stellar and gas kinematics over the inner 30% of the optical galaxy. The stellar disk contains two interpenetrating components; one system is rotating prograde and one system, cospatial with the first, is rotating retrograde. A rotating gas disk is coincident with one of the stellar disks. Thus, instead

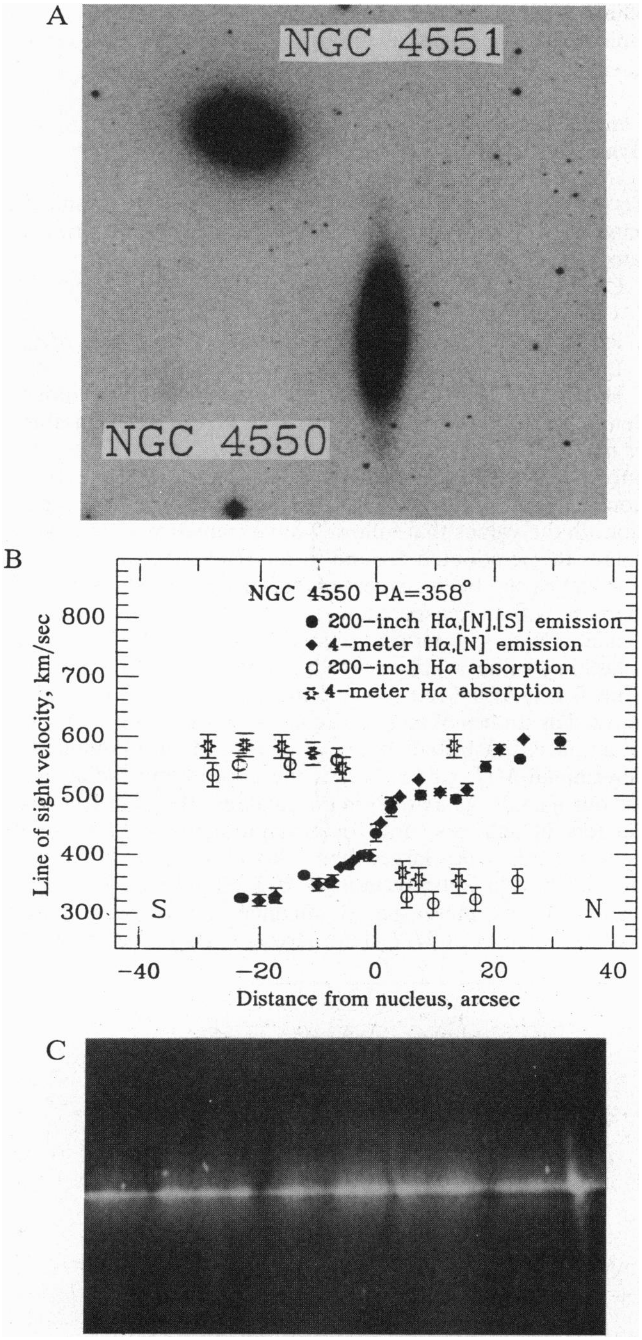


FIG. 5. (*Top*) Virgo galaxies NGC 4550 and NGC 4551 from an r-band image taken by C. Bailyn and Y.-C. Kim at the Cerro Tololo InterAmerican Observatory 0.9-m telescope. There is no evidence of morphological peculiarity in either galaxy. (*Middle*) Major-axis velocities in NGC 4550, from the stellar Hα absorption (open symbols) and from the counter-rotating emission disk (filled symbols). The counter-rotating stellar disk absorption at Hα is masked by the emission, except for that barely showing at ≈15″ N. (*Bottom*) Absorption lines from prograde and retrograde stellar disks (here shown dark) in the region blueward of [OIII]5007 (bright line at right). Note the X-shaped appearance of the absorption lines as they cross the nucleus.

of a normal spectrum composed of many single absorption lines, we observe each absorption line to be double, with the two members opposite in velocity and crossing at the nucleus. The two cospatial counter-rotating disks extend over a sizable portion of the galaxy.

A value of $V/\sigma = 1.4$, coupled with an ellipticity, $1 - b/a = 0.7$, places NGC 4550 in the Binney (43) region of rotating oblate spheroids. Thus, it is axisymmetric, flattened by rotation, with no apparent morphological or other kinematical peculiarities. Our limited material does not permit a detailed model, although absorption line intensities and measured velocities suggest that the two stellar systems have approximately similar dynamical properties, and are approximately coplanar, and absorption line strengths within the inner few kpc are only slightly stronger in the prograde disk.

Virtually all of the possibilities for the origin of the kinematics in NGC 4550 require the acquisition of secondary material after the initial stellar disk of NGC 4550 was in place. We propose the following likely scenario. A retrograde gas mass, probably small compared with the mass of NGC 4550, is captured at some angle intermediate between polar and disk, settles to the disk via dissipation caused partly by differential precession (44, 45), and forms most stars after the gas has settled to the plane. Such an uncommon system can be understood as an "equatorial ring" (46)—i.e., a disk, an alternative end point to the polar ring sequence.

The lack of morphological distortions and the presence of strong metal absorption lines suggest that the time elapsed since the acquisition is at least 10^9 yr, allowing time for the gas to come to equilibrium in the combined gravitational potential and for the merger population to form and age. An age of 10^9 yr corresponds to ≈10 orbital periods at the half-light radius (16″) and to 6 orbital periods at the limits of our velocity measurements, sufficient to destroy the morphological evidence of the event.

Cospatial stellar systems rotating both clockwise and counterclockwise are more stable than unidirectional disks (S. Tremaine, private communication) and have been discussed by Araki (47). Earlier, Lynden-Bell (48) imagined a Maxwell demon who reversed the direction (and violated conservation of angular momentum) of one-half the stars in a spherical cluster to form two interpenetrating stable counter-rotating spherical clusters. In Toomre's (49) discussion of axially symmetric flattened isothermal galaxy models, he investigated sets of models in which one-half of the stars rotate prograde, one-half rotate retrograde, noting that nothing in the analysis restricts the disk stars to orbit only prograde.

The discovery of galaxies like NGC 4550 confirms yet again the enormous variety within elliptical galaxies; Toomre's (49) models are no longer just the "elegant curiosities" he imagined. I envy observers living within NGC 4550 for the interesting stellar dynamics they are encountering.

Section V. $M/L \approx 100$ for the Perseus Cluster, a "Low" Value

Nearly 60 years ago, Zwicky (50) and Smith (51) discovered that individual galaxies in the Coma and Virgo clusters of galaxies have velocities so large with respect to the mean cluster velocity that their mutual gravitational attraction would not be sufficient to bind the clusters. Hence, either clusters are expanding and galaxies are dispersing (contrary to observation) or the total mass in the clusters is larger than that implied by the luminous material. Zwicky called this the "missing mass." For many years, this conclusion was relegated to the "things we don't understand" file and assumed to be a mysterious property of galaxy clusters. Not until we discovered the need for dark matter in individual galaxies did these early observations join the mainstream of astronomical research.

Now that we recognize that dark matter is the major constituent of clusters, we must quantify the amount of dark matter if we are to set observational limits on the mass and the ultimate fate of the universe. Is M/L averaged over a cluster as large as 500 or 1000, implying that M/L rises with the size of the region included, or is it closer to the value 100 derived for binaries and groups? During the last decade or two, attempts to answer this question were made generally via a virial analysis, an analysis that assumes galaxy membership in the cluster is known, that the distribution of mass follows the distribution of galaxies, that the cluster is relaxed, that the galaxy velocity distribution is isotropic, and that the second time derivative of the moment of inertia is zero (which would not be the case, for example, if galaxy halos were centered elsewhere than at the galaxy centers). Some years ago, from optical data for the Perseus cluster, Kent and Sargent (52) derived the large value of $M/L \approx 300$.

In an important recent work, Eyles *et al.* (53) analyze x-ray images of the Perseus cluster obtained with spatial resolutions of 12′ and 3′ in the energy range 2.5–32 keV (1 eV = 1.602×10^{-19} J) during the Spacelab 2 mission. Because the x-ray-emitting gas in a cluster is expected to be in hydrostatic equilibrium, kT plays the role of mV^2 in the virial analysis and is a good tracer of the cluster gravitational potential. From the x-ray data, Eyles *et al.* (53) find that the total gravitational cluster mass is 30–65% *less* than that obtained from the application of the virial theorem, in part because the distribution of mass does not follow the distribution of galaxies. Moreover, while the dark matter is **more** centrally condensed than the galaxies and is the dominant mass component out to ≈1 Mpc, the x-ray gas is **less** so. The fraction of the total mass that is in the form of the x-ray-emitting gas increases with radius, dominating beyond 1 Mpc and reaching >55% of the total mass and still rising at the limits of the observations.

As a result of this analysis, M/L_V is of order 100 at large radii (rather than 300), with values between 55 and 145 permitted by the range of models and of statistical and systematic errors. Although based on only a single cluster, this result strengthens earlier suggestions (54) that there is not strong evidence for very large values of M/L for systems of very large size. New x-ray observations from ROSAT satellite are expected to enlarge the sample of galaxy clusters for which determinations of the mass distribution and, consequently, M/L are available.

Conclusions: $M/L \approx 100$ Will Satisfy All Dynamical Observations

To discuss the observed dynamical values of M/L and the mass density of the universe, I start with a few definitions. I adopt $H_o = 50$ km·sec^{-1}·Mpc^{-1}. Then:

Critical mass density $\rho_c = 3H_o^2/8\pi G \approx 5 \times 10^{-30}$ g·cm^{-3} ≈ 7×10^{10} $M_\odot$·Mpc^{-3},

Present mass density $\rho = 3H_o^2\Omega/8\pi G$, where $\Omega = \rho/\rho_c$,

Luminosity density $L_B = 10^8$ $L_\odot$·Mpc^{-3}.

Hence, $M/L_B = \rho/L_B \approx 750\ \Omega$, where the blue luminosity density, surely an uncertain number, is taken from Efstathiou *et al.* (55). Thus, a value of M/L_B near 750 will close the universe—i.e., Ω will be unity, and the gravitational attraction of the mass within the universe will just halt the expansion. In the values that follow, I have transformed published values to the adopted H_o and to the blue spectral band.

It had earlier been suspected that observed values of M/L increase with increased enclosed volume. However, current available data suggest that this is not so, and $M/L \approx 100$ is a limiting value required by dynamical systems on scales from 0.1 to 200 Mpc (except the highly uncertain values derived in studies of large-scale motions—e.g., ref. 56). This situation is illustrated in Fig. 6, a plot of the dynamically determined M/L values versus the appropriate radial scale for our galaxy, galaxies, binary galaxies, the Local Group, clusters of galaxies, and for gravitationally induced infall motions. For scales larger than galaxies—i.e., >100 kpc, all systems have a constant ratio of M/L. On the smaller scales (galaxies), the increasing dominance of the dark matter produces values of M/L that increase with radial scale. All

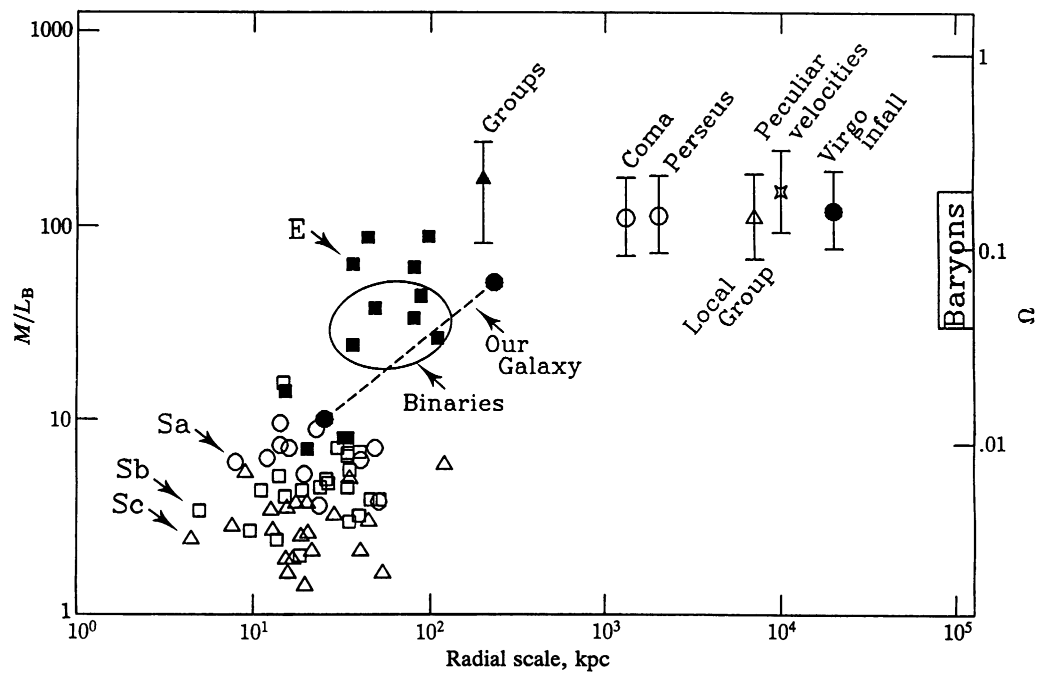


FIG. 6. Mass-to-blue light ratio as a function of the size of the system for galaxies, binaries, clusters, and larger systems, based loosely on and extended from Oemler (57). References to individual values come from Rubin (3, 4); Forman *et al.* (32); Kulessa and Lynden-Bell (9); Schweizer (58); Huchra and Geller (59); Cowie *et al.* (60); Eyles *et al.* (53); Peebles (13); van der Kruit (11); Davis and Peebles (61); Aaronson *et al.* (62); and Schramm (63); values are all transformed to $H_o = 50$ km·sec^{-1}·Mpc^{-1} and blue magnitudes. A value of $M/L = 750$ is equivalent to $\Omega = 1$. [For $H_o = 100$ km·sec^{-1}·Mpc^{-1} values of M/L increase by a factor of 2, and M/L (closure) = 1500. The form of the plot relative to Ω remains unchanged]. Error bars are generic estimates. Note that beyond ≈100 Mpc, values of M/L do not continue to rise but are constant near 100, independent of the size of the dynamical system. Dwarf spirals and irregulars, not plotted here, have M/L values ranging from ≈20–50 and radii near 10 kpc. However, their poorly known distances introduce large uncertainties in both quantities.

values fit comfortably within a universe in which the luminous mass in galaxies comprises only 1% of the closure density; the mass required by the dynamics is ten times this or 10% of the closure density, *but not higher*. Moreover, these values do not exceed the nucleosynthetic limit on the mass density of baryons in the universe (e.g., ref. 63).

Thus, observations tell us that we may inhabit a low-density universe in which the gravitational attraction of the enclosed baryonic matter is not sufficient to halt the expansion. However, theoretical cosmologists like Guth and Schramm and Turner tell us that models require that Ω is unity, that the additional 90% of the universe is still undetected, and that it is required to consist of dark, *nonbaryonic* matter. To reconcile the views of the observers with the dreams of the theorists will require heroic observational, laboratory, and theoretical efforts but will constitute a fundamental step in teaching us more about the complicated universe in which we find ourselves.

I thank my colleagues John Graham, Jeff Kenney, and Francois Schweizer for valuable help and support; Kitt Peak and Palomar Observatories for telescope time; and Drs. F. Bertola, F. Briggs, S. Casertano, and J. Kormendy for permission to reproduce their figures.

1. Ostriker, J. P., Peebles, P. J. E. & Yahill, A. (1974) *Astrophys. J. Lett.* **193**, L1–L4.
2. Ostriker, J. P. & Peebles, P. J. E. (1973) *Astrophys. J.* **186**, 467–480.
3. Rubin, V. C., Ford, W. K. & Thonnard, N. (1978) *Astrophys. J. Lett.* **225**, L107–L111.
4. Rubin, V. C., Ford, W. K. & Thonnard, N. (1980) *Astrophys. J.* **238**, 471–487.
5. Bosma, A. (1978) Dissertation (Univ. of Groningen, The Netherlands).
6. Faber, S. M. & Gallagher, J. S. (1979) *Annu. Rev. Astron. Astrophys.* **17**, 135–187.
7. Kormendy, J. & Knapp, G. R., eds. (1987) in *Dark Matter in the Universe* (Reidel, Dordrecht, The Netherlands).
8. Trimble, V. (1987) *Annu. Rev. Astron. Astrophys.* **25**, 425–472.
9. Kulessa, A. S. & Lynden-Bell, D. (1992) *Mon. Not. R. Astron. Soc.* **255**, 105–118.
10. Schmidt, M. (1985) *The Milky Way Galaxy, Int. Astron. Union Symposium 106* (Reidel, Dordrecht, The Netherlands), pp. 75–84.
11. van der Kruit, P. C. (1990) *The Milky Way as a Galaxy, Saas-Fee Advanced Course 19* (Univ. Science Books, CA), pp. 185–226.
12. Zaritsky, D., Smith, R., Frenk, C. & White, S. D. M. (1993) *Astrophys. J.*, in press.
13. Peebles, P. J. E. (1990) *Astrophys. J.* **362**, 1–13.
14. Kamphuis, J. & Briggs, F. (1992) *Astron. Astrophys.* **253**, 335.
15. Kormendy, J. (1988) in *Origin, Structure, and Evolution of Galaxies*, ed. Zhi, F. L. (World Sci., Singapore), pp. 252–280.
16. van Albada, T. S., Bahcall, J. B., Begeman, K. & Sancisi, R. (1985) *Astrophys. J.* **295**, 305–313.
17. Kent, S. M. (1988) *Astron. J.* **96**, 514–527.
18. Athanassoula, E., Bosma, A. & Papaioannou, S. (1987) *Astron. Astrophys.* **179**, 23–40.
19. Kalnajs, A. J. (1985) in *The Internal Kinematics and Dynamics of Galaxies*, ed. Athanassoula, E. (Reidel, Dordrecht, The Netherlands), pp. 87–88.
20. Casertano, S. & van Gorkom, J. H. (1991) *Astron. J.* **101**, 1231–1241.
21. Whitmore, B. C., Forbes, D. A. & Rubin, V. C. (1988) *Astrophys. J.* **333**, 542–560.
22. Rubin, V. C. & Kenney, J. D. P. (1991) *Bull. Am. Astron. Soc.* **23**, 1339 (abstr.).
23. Carignan, C. & Freeman, K. C. (1988) *Astrophys. J. Lett.* **332**, L33–L36.
24. Carignan, C. & Beaulieu, S. (1989) *Astrophys. J.* **347**, 760–770.
25. Broeils, A. H. (1992) *Astron. Astrophys.* **256**, 19–32.
26. Milgrom, M. (1983) *Astrophys. J.* **270**, 365–389.
27. Illingworth, G. (1977) *Astrophys. J. Lett.* **218**, L43–L47.
28. Davies, R. L., Efstathiou, G., Fall, S. M., Illingworth, G. & Schechter, P. L. (1983) *Astrophys. J.* **266**, 41–57.
29. Franx, M., Illingworth, G. & Heckman, T. (1990) *Astrophys. J.* **344**, 613–636.
30. Kent, S. M. (1990) in *Evolution of the Universe of Galaxies*, ed. Kron, R. G. (Brigham Young Univ., Provo, UT), pp. 109–118.
31. de Zeeuw, T. (1992) in *Morphological and Physical Classification of Galaxies*, eds. Longo, G., Capaccioli, M. & Busarello, G. (Kluwer, Dordrecht, The Netherlands), pp. 139–151.
32. Forman, W., Jones, C. & Tucker, W. (1985) *Astrophys. J.* **293**, 102–119.
33. Ciardullo, R., Jacoby, G. H. & Ford, H. C. (1989) *Astrophys. J.* **344**, 715–725.
34. Bertola, F., Bettoni, D., Danzinger, J., Sadler, E., Sparke, L. & de Zeeuw, T. (1991) *Astrophys. J.* **373**, 369–390.
35. Schweizer, F., van Gorkom, G. H. & Seitzer, P. (1989) *Astrophys. J.* **338**, 770–788.
36. Graham, J. A. (1979) *Astrophys. J.* **232**, 60–73.
37. Schweizer, F. (1980) *Astrophys. J.* **237**, 308–318.
38. Barnes, J. E. & Hernquist, L. E. (1991) *Astrophys. J.* **370**, L65–L68.
39. Hernquist, L. E. & Barnes, J. E. (1991) *Nature (London)* **354**, 210–212.
40. Wielen, E. R., ed. (1990) in *Dynamics and Interactions of Galaxies* (Springer, Heidelberg).
41. de Zeeuw, T. & Franx, M. (1991) *Annu. Rev. Astron. Astrophys.* **29**, 239–274.
42. Rubin, V. C., Graham, J. A. & Kenney, J. P. D. (1992) *Astrophys. J.* **394**, L9–L12.
43. Binney, J. (1978) *Mon. Not. R. Astron. Soc.* **183**, 501–514.
44. Sparke, L. S. (1986) *Mon. Not. R. Astron. Soc.* **219**, 657–670.
45. de Zeeuw, T. (1990) in *Dynamics and Interactions of Galaxies*, ed. Wielen, E. R. (Springer, Heidelberg), pp. 263–266.
46. Schechter, P. L. (1990) in *Dynamics and Interactions of Galaxies*, ed. Wielen, E. R. (Springer, Heidelberg), pp. 508–513.
47. Araki, S. (1987) *Astron. J.* **94**, 99–105.
48. Lynden-Bell, D. (1960) *Mon. Not. R. Astron. Soc.* **120**, 204–213.
49. Toomre, A. (1982) *Astrophys. J.* **259**, 535–543.
50. Zwicky, F. (1933) *Helv. Phys. Acta* **6**, 110.
51. Smith, S. (1936) *Astrophys. J.* **83**, 23.
52. Kent, S. M. & Sargent, W. L. W. (1983) *Astron. J.* **88**, 697–708.
53. Eyles, C. J., Watt, M. P., Bertram, D., Church, M. J., Ponman, T. J., Skinner, G. K. & Willmore, A. P. (1991) *Astrophys. J.* **376**, 23–32.
54. Davis, M. (1987) in *Dark Matter in the Universe*, eds. Kormendy, J. & Knapp, G. R. (Reidel, Dordrecht, The Netherlands), pp. 97–109.
55. Efstathiou, G., Ellis, R. S. & Peterson, B. A. (1988) *Mon. Not. R. Astron. Soc.* **232**, 431–461.
56. Guth, A. H. (1993) *Proc. Natl. Acad. Sci. USA* **90**, 4871–4877.
57. Oemler, A. (1988) in *The Minnesota Lectures on Clusters of Galaxies and Large-Scale Structure*, ed. Dickey, J. M. (Brigham Young Univ., Provo, UT), pp. 19–38.
58. Schweizer, L. (1988) *Astrophys. J. Suppl. Ser.* **64**, 427–457.
59. Huchra, J. & Geller, M. (1982) *Astrophys. J.* **257**, 423–437.
60. Cowie, L. L., Henriksen, M. & Mushotzky, R. (1987) *Astrophys. J.* **317**, 593–600.
61. Davis, M. & Peebles, P. J. E. (1983) *Astrophys. J.* **267**, 465–482.
62. Aaronson, M., Huchra, J., Mould, J., Schechter, P. L. & Tully, R. B. (1982) *Astrophys. J.* **258**, 64–76.
63. Schramm, D. N. (1993) *Proc. Natl. Acad. Sci. USA* **90**, 4782–4788.

Physics Today

The Dynamical Evidence for Dark Matter

Scott Tremaine

Citation: Physics Today **45**(2), 28 (1992); doi: 10.1063/1.881329
View online: http://dx.doi.org/10.1063/1.881329
View Table of Contents: http://scitation.aip.org/content/aip/magazine/physicstoday/45/2?ver=pdfcov
Published by the AIP Publishing

THE DYNAMICAL EVIDENCE FOR DARK MATTER

OIL ON CANVAS, 29×36 1/4". COLLECTION, THE MUSEUM OF MODERN ART, NEW YORK. ACQUIRED THROUGH THE LILLIE P. BLISS BEQUEST

'The Starry Night,' by Vincent van Gogh. The 1889 oil painting suggests how the night sky might look if all of the mass in the universe were luminous. Observations of galaxy dynamics and modern theories of galaxy formation imply that the visible components of galaxies, composed mostly of stars, lie at the centers of vast halos of dark matter that may be 30 or more times larger than the visible galaxy. In most models of galaxy formation, the halos are comparable in size to the distance between galaxies. The halos form as a result of the gravitational instability of small density fluctuations in the early universe; the star-forming gas collects at the minima of the halo potential wells. Infall of outlying material into existing halos and mergers of small halos with larger ones continue at the present time. If the halos were visible to the naked eye, there would be well over 1000 nearby galaxies with halo diameters larger than the full Moon. **Figure 1**

Studies of the dynamics of galaxies show that at least 90% of the mass in the universe is in some invisible, unknown form.

Scott Tremaine

Almost all of our information about the universe beyond Earth comes from photons—visible photons from stars, x-ray photons from hot plasmas, radio photons from the 21-cm hyperfine transition in hydrogen, microwave photons from the cosmic background radiation and so forth.

It would be folly to assume that all the matter in the universe emits detectable photons. Thus we should not be surprised if the mass of a galaxy or other astronomical system, as measured by its gravitational field, exceeds the sum of the masses of those of its components that shine brightly enough to be detected in our telescopes. The difference between this "luminous mass" and the total mass is ascribed to "dark matter"—matter whose existence is inferred solely from its gravitation.[1]

An early example of this reasoning was the prediction in 1846 of the existence and location of Neptune from unexplained residuals in the motion of Uranus. Another example from the solar system was the anomalous precession of Mercury's perihelion. A hypothetical planet ("Vulcan"), or else a ring of material, inside Mercury's orbit was invoked to explain this anomaly, but Einstein showed in 1916 that it was a consequence of general relativity rather than of dark matter. This is a cautionary reminder that dark matter may sometimes be explained away by revisions to the accepted laws of physics.

At present there is no significant dynamical evidence for dark matter in the solar system. On larger scales, however, the story is quite different. There is convincing evidence not just that dark matter is present but that most of the mass in galaxies is dark. The visible parts of galaxies, composed mainly of stars, are surrounded by extended halos of dark matter that may be a factor of 30 or more larger in both mass and size. Van Gogh's famous painting "The Starry Night" (figure 1) provides a surprisingly accurate view of what the dark halos might look like if they were visible. The average mass density of the dark matter could exceed the critical value needed to close the universe.

An equally remarkable conclusion, based on nucleosynthesis arguments, is that most of the dark matter—and hence most of the mass in the universe—is not composed of protons or neutrons. Thus the material that makes up the stars that we see and the everyday world that we know is only a minor pollutant in a sea of invisible material of unknown nature.

The solar neighborhood

The first natural place beyond the solar system to look for dark matter is the solar neighborhood—an imaginary volume centered on the Sun that is large enough to contain plenty of stars for statistical analyses but small enough compared with the size of the Galaxy that the bulk properties of the stellar distribution are constant within it.

The distance of the nearest star to the Sun is 1.3 parsecs (1 pc is 3.086×10^{13} km). About the smallest volume containing a statistically useful sample of stars is a Sun-centered sphere of radius 10 pc, in which there are 300 known stars.[2] An instructive exercise is to divide these into an inner sample of 61 stars within 5 pc and an outer sample of 239 stars between 5 and 10 pc from the Sun. The corresponding densities are 0.12 and 0.065 stars per cubic parsec. Since the density ought to be constant over such small distances, the drop in density by a factor of two from the inner to the outer sample implies that the outer sample is seriously incomplete. Thus even at the smallest interstellar distances, many of the stars are so faint that they have yet to be discovered, which is a hint that substantial dark mass might lurk in faint stars.

Most stars are in a state of thermal equilibrium, in which energy generated by hydrogen fusion is balanced by heat lost through thermal radiation. However, below a transition mass M_c of 0.08 times the mass $M_\odot$ of the Sun, stars cannot fuse hydrogen, as their electrons become degenerate before they are dense and hot enough for fusion to proceed. The luminosity of stars with $M<M_c$, usually called "brown dwarfs," is supplied by slow gravitational contraction rather than by fusion[3] and hence is much smaller than the luminosity of hydrogen-burning stars. (The luminosity drops by more than two orders of magnitude between $0.10M_\odot$ and $0.07M_\odot$.) Brown dwarfs are so faint that they would be very difficult to detect even if they were much more numerous than the brighter, hydrogen-burning stars. Hence they are a natural candidate for dark mass.

We can attempt to estimate the number density of brown dwarfs by extrapolating the number density of brighter stars. The plausible assumption in this extrapolation is that the transition mass M_c plays no special role in the physics of star formation: Since the transition mass involves nuclear physics and star formation probably does not, the rates of formation of stars above and below M_c should be similar. The luminosity $L(M)$ of a star of mass M can be computed from stellar structure theory and checked by observations of binary star orbits. Let $\phi(L)\,\mathrm{d}L$ be the number density of stars with luminosity in the range $[L, L+\mathrm{d}L]$, as determined from star catalogs. Then

Scott Tremaine is director of the Canadian Institute for Theoretical Astrophysics and is a professor of physics and astronomy at the University of Toronto.

the number density $n(M)\,dM$ of stars with mass in the range $[M, M + dM]$ is given by

$$n(M) = \phi[L(M)]\,\frac{dL(M)}{dM}$$

Figure 2 shows the number density derived in this way. The density becomes quite uncertain as we approach the transition mass M_c, both because dL/dM becomes very large and because measurement of $\phi(L)$ becomes harder and harder at low luminosities. The figure suggests that $n(M)$ is relatively flat for masses below M_c, which would imply that brown dwarfs contain a negligible fraction of the mass in the solar neighborhood. A sharp upturn in $n(M)$ below about $0.15M_\odot$ is not excluded by the data, although there is no reason to suppose that it is present.

Fortunately there is a dynamical method of measuring the total mass density in this region. The stars in the solar neighborhood belong to the Galactic disk, which has a radius of about 10 kpc but a thickness of only a few hundred parsecs. Because the disk is so thin, it can be approximated as an infinite slab. The gravitational potential of the slab is $U(z)$, where z is the distance perpendicular to the slab's midplane. The phase-space density $f(\mathbf{x},\mathbf{v},t)$ of stars of a given type obeys the collisionless Boltzmann equation, which expresses the conservation of phase-space density along a trajectory:[4]

$$\frac{\partial f}{\partial t} + \mathbf{v}\cdot\nabla_x f - \nabla_x U\cdot\nabla_v f = 0 \qquad (1)$$

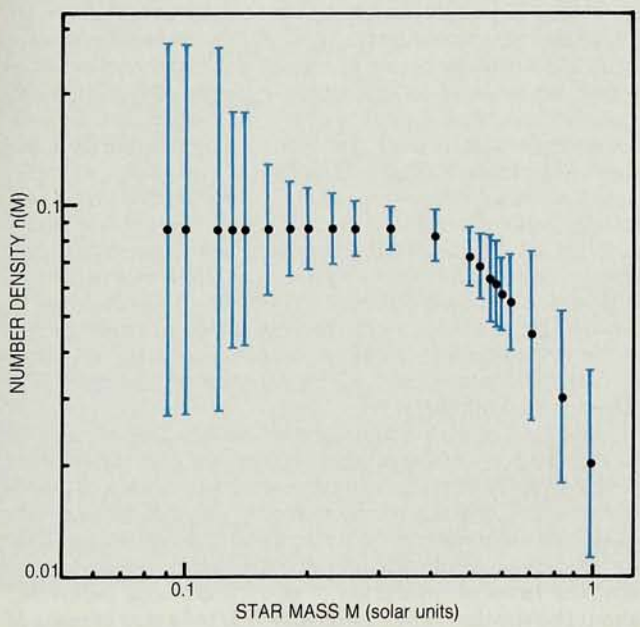
Here ∇_x and ∇_v denote gradients with respect to position and velocity. Assuming slab symmetry and a stationary distribution (that is, no dependence on x, y or t), we multiply the equation by v_z and integrate over velocity to get

$$\frac{d}{dz}\nu\sigma_z^2 = -\nu\frac{dU}{dz} \qquad (2)$$

where $\nu(z) = \int f(z,\mathbf{v})d\mathbf{v}$ is the number density of stars and $\sigma_z{}^2(z) = \int v_z{}^2 f(z,\mathbf{v})d\mathbf{v}/\nu(z)$ is their mean-square velocity in the z direction. Thus measurements of the number density and velocity dispersion of any given type of star as a function of height above the Galactic midplane determine the potential $U(z)$ through equation 2 and the mass density $\rho(z)$ through Poisson's equation. The method is difficult to apply in practice, mostly because statistical uncertainties in ν and $\sigma_z{}^2$ are amplified by the two differentiations needed to derive the density.

This argument was first used in 1922 by Jacobus C. Kapteyn, who deduced that the total density in the solar neighborhood was no more than a factor of 2 or so larger than the density in visible stars. Modern estimates have not substantially changed this conclusion: Two recent studies found[5] that the ratio of the total density to the density in known objects (stars and gas) was 1.0 ± 0.3 and $2.6^{+1.9}_{-1.2}$ ($1-\sigma$ limits). Thus neither the extrapolation of the number density of stars per unit mass $n(M)$ nor the dynamical estimates strongly suggest that there is substantial dark matter in the solar neighborhood, although a dark matter density similar to the density in known objects is not excluded.

A simple way to parameterize the ratio of dark to luminous matter is the mass-to-light ratio Υ, which is usually measured in solar units: $\Upsilon_\odot = 1$ solar mass/1 solar luminosity, or about 0.5 in cgs units. Here "light" usually means light in the visible part of the spectrum. This is the natural wavelength range to use, since stars are the largest known contributor to the mass in galaxies and they are most easily studied in visible light. In the likely case that there is no dark matter in the solar neighborhood, its mass-to-light ratio is about $1.5\Upsilon_\odot$—close to unity since the Sun is an average star—but the larger of the dynamical estimates above allows a value up to $7\Upsilon_\odot$.

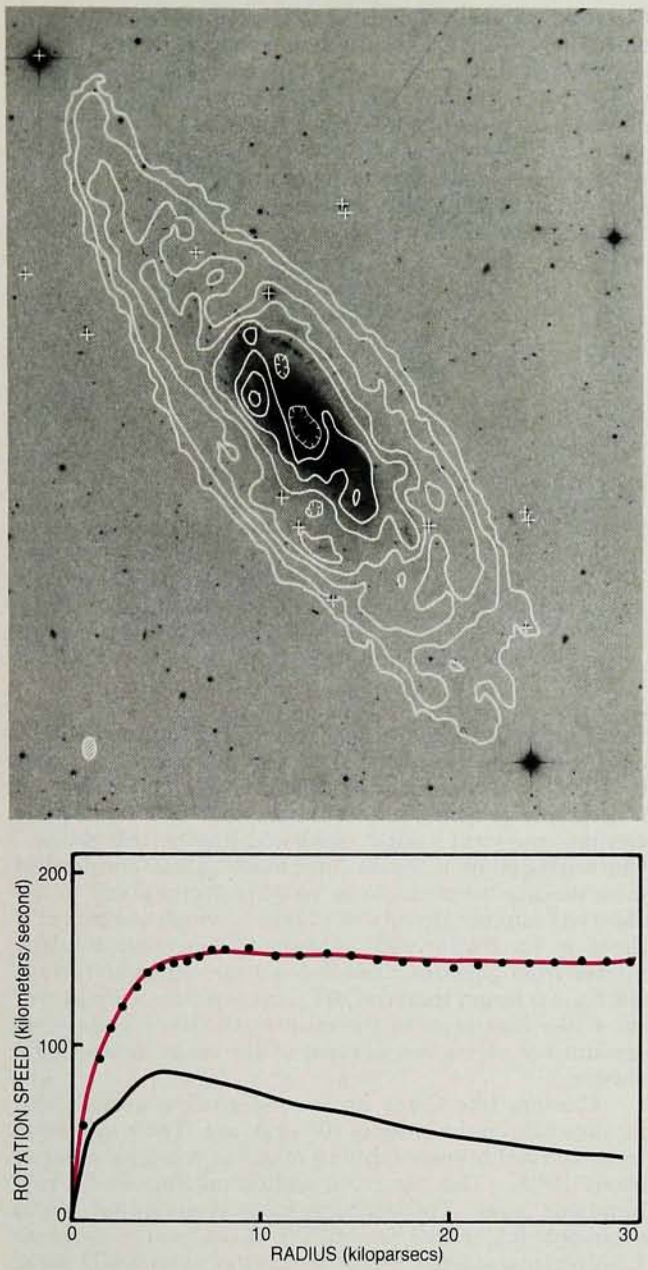


Number density of stars in the solar neighborhood as a function of mass. The black points denote a smooth fit to the data, but other curves lying within the error bars are also consistent with the data. Units are stars per cubic parsec per solar mass. (Adapted from P. Kroupa, C. A. Tout, G. Gilmore, Mon. Not. R. Astron. Soc. **244**, 76, 1990.) **Figure 2**

Galaxy rotation curves

The stars in most galaxies lie mainly in a thin disk and travel on nearly circular orbits around the galactic center. The circular speed v_c at a given radius R can be determined from the Doppler shift of spectral lines in either the integrated starlight or the interstellar gas that rotates with the stars.

To look for dark matter in a galaxy, we compare the observed centripetal acceleration $v_c{}^2/R$ with the calculated gravitational acceleration due to the luminous mass,



Disk galaxy and its rotation curve. Top: An optical image of the spiral galaxy NGC 3198 with a superimposed contour map of the column density of hydrogen gas. The hydrogen is detectable out to much larger radii than the stars, and so provides a better probe of the mass distribution at large radii. The shaded ellipse at lower left indicates the resolution of the hydrogen observations. Bottom: The rotation speed of the gas as a function of radius (black points), together with the circular speed derived from the assumption that all of the mass is in the visible stars and gas (black curve). The red curve shows the best-fit circular speed derived by assuming that the galaxy also contains a spherical dark halo with a density that follows the empirical law $\rho \propto 1/(1 + r^2/r_c^{\,2})$, where r_c is the core radius. (Adapted from T. S. van Albada *et al.*, Astrophys. J. **295**, 305, 1985, and S. M. Kent, Astron. J. **93**, 816, 1987.) **Figure 3**

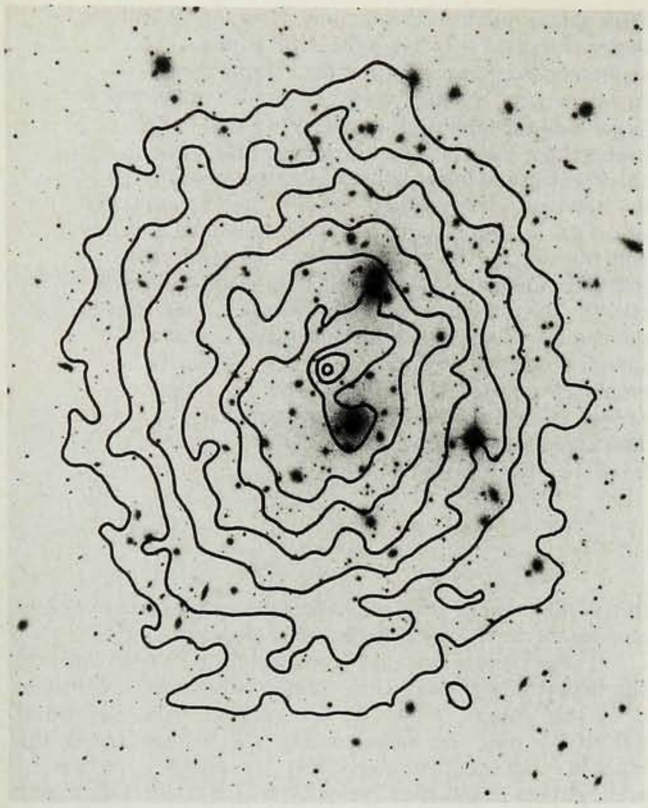
which is mostly in stars. Figure 3 shows the observed rotation curve $v_c(R)$ in the disk galaxy NGC 3198, along with the circular speed derived from the assumptions that the disk surface brightness is proportional to the surface density and that there is no dark mass. To obtain this particular curve, the mass-to-light ratio of the disk was chosen to be as large as possible. (With any larger value, the predicted speed would exceed the observed speed in the inner parts.) Even with this extreme assumption, the predicted speed is more than a factor of 3 lower than the observed speed at the outermost measured point. (At larger radii the density of interstellar gas is too low to permit measurement of the velocity.) This implies that the calculated gravitational field from the disk is too small by a factor of 10 to account for the observed rotation.

We conclude that stars and other luminous mass make up less than 10% of the total mass in that galaxy. The remaining 90% or more is dark matter. Most of the dark matter must be located at radii larger than that of the stars; otherwise the rotation speed would exhibit Keplerian behavior—that is, $v_c(R) = (GM/R)^{1/2} \propto R^{-1/2}$— in the outer parts, whereas in fact the speed is more or less constant over the outer two-thirds of the galaxy.

These results can also be considered a lower limit on the mass-to-light ratio. The limit depends on the distance d to the galaxy, which is determined from its radial velocity v and the relation $v = H_0 d$, where H_0 is the Hubble constant. (More precisely, $v = H_0 d + v_p$, where v_p is the galaxy's "peculiar" velocity, which is typically under 500 km/sec in magnitude; a galaxy's peculiar velocity is defined in this equation as the difference between its actual velocity and the Hubble velocity.) We shall use $H_0 = 75$ km sec^{-1} Mpc^{-1}, although respectable estimates[6] of H_0 range from below 50 to above 100. With this value for the Hubble constant, the distance of NGC 3198 is 9.2 Mpc and its mass-to-light ratio Υ is at least $40\Upsilon_\odot$, which is about a factor of 20 larger than the mass-to-light ratio in the solar neighborhood. Similar flat rotation curves and mass-to-light ratios are found in most disk galaxies, including our own.[7]

The shapes of rotation curves suggest that the dark matter is distributed in extended halos that surround the visible stars. A simple empirical model for the halo density is a spherical distribution $\rho(r) = \rho_0/(1 + r^2/r_c^{\,2})$, where the core radius r_c and central density ρ_0 are fitting parameters. The rotation curve fits produced by this model are quite good, as figure 3 shows. The corresponding dark mass within radius r is proportional to r for $r \gg r_c$; evidently this growth must stop at some sufficiently large radius r_{max}, since otherwise the mass of the galaxy would be infinite, but the rotation curves imply only that r_{max} must lie near or beyond the last measured points on the rotation curve. Less accurate than rotation-curve analysis, such methods as measurement of the relative velocities of galaxy pairs or the kinematics of satellite galaxies[8] suggest that r_{max} is 100 kpc or even larger. Thus we reach two remarkable conclusions: The total mass and extent of ordinary galaxies are almost completely unknown, and between 90% and 99% of the mass in galaxies is dark.

Before about 1970, measurements of rotation curves were restricted to the inner parts of galaxies. It was natural for observers to extrapolate the rotation curves assuming Keplerian behavior beyond the last measured point, since most of the light from the galaxy was contained well within that point. This extrapolation gave a direct—but spurious—estimate of the total mass of the galaxy. In retrospect, it is remarkable that the dangers of this extrapolation were not more clearly recognized. By the early 1970s, high-resolution observations of interstel-



Cluster of galaxies. Superimposed on the optical image of the Coma cluster is a contour map of the x-ray surface brightness measured by the Einstein satellite. The distance between the two most prominent galaxies is 7 minutes of arc, or 190 kiloparsecs. (Courtesy of William Forman and Christine Jones, Harvard–Smithsonian Center for Astrophysics, Cambridge, Mass.) **Figure 4**

lar gas began to provide flat rotation curves that clearly showed that the mass was not contained solely in the visible disk stars.[7] By now, rotation curves of dozens of galaxies provide convincing evidence that most of the mass in disk galaxies is dark.

Clusters of galaxies

Galaxies are not distributed uniformly throughout the universe but instead have a rich hierarchy of structure ranging from binary galaxies through groups containing a few galaxies to clusters containing thousands of galaxies. One of the largest nearby clusters is the Coma cluster, shown in figure 4.

The central regions of clusters of galaxies are the largest equilibrium structures in the universe and hence are natural sites to prospect for dark mass. The phase-space density $f(\mathbf{x},\mathbf{v},t)$ of cluster galaxies obeys the collisionless Boltzmann equation (equation 1), which can be analyzed by assuming spherical symmetry (the substantial ellipticity of many clusters does not strongly affect the results), no time dependence and an isotropic velocity distribution. Then multiplying equation 1 by the radial velocity v_r and integrating over velocity space yields a result reminiscent of the hydrostatic equilibrium equation,

$$\frac{\mathrm{d}(\nu\sigma^2)}{\mathrm{d}r} = -\nu\frac{\mathrm{d}U}{\mathrm{d}r} \tag{3}$$

Here $\nu(r)$, defined as $\int f(r,v)\mathrm{d}\mathbf{v}$, is the number density of galaxies at radius r; $\sigma^2(r)$, defined as $\frac{1}{3}\int v^2 f(r,v)\mathrm{d}\mathbf{v}/\nu(r)$, is their mean-square velocity in one dimension; and $U(r)$ is the gravitational potential.

Observations yield the projected number-density distribution and the velocity dispersion along the line of sight, and these determine $\nu(r)$ and $\sigma(r)$. The potential $U(r)$ then follows from equation 3, and we can estimate the mass $M(r)$ contained within radius r through $GM(r)/r^2 = \mathrm{d}U/\mathrm{d}r$.

This analysis shows that the mass contained within $r = 1.3$ Mpc of the center of the Coma cluster is about $8\times10^{14}M_\odot$. Our assumption that the velocity distribution is isotropic is somewhat arbitrary, but other plausible assumptions yield similar results, at least at this radius.[9] One selling point of the isotropic model is that the derived mass density turns out to be roughly proportional to the observed number density of galaxies, which is natural if there is no cluster-wide process that segregates dark matter from galaxies. The derived mass-to-light ratio is $300\Upsilon_\odot$, far larger than the $10\Upsilon_\odot$ expected for a mixture of stars like that seen in the cluster galaxies. Thus stars account for only a few percent of the mass in the Coma cluster.

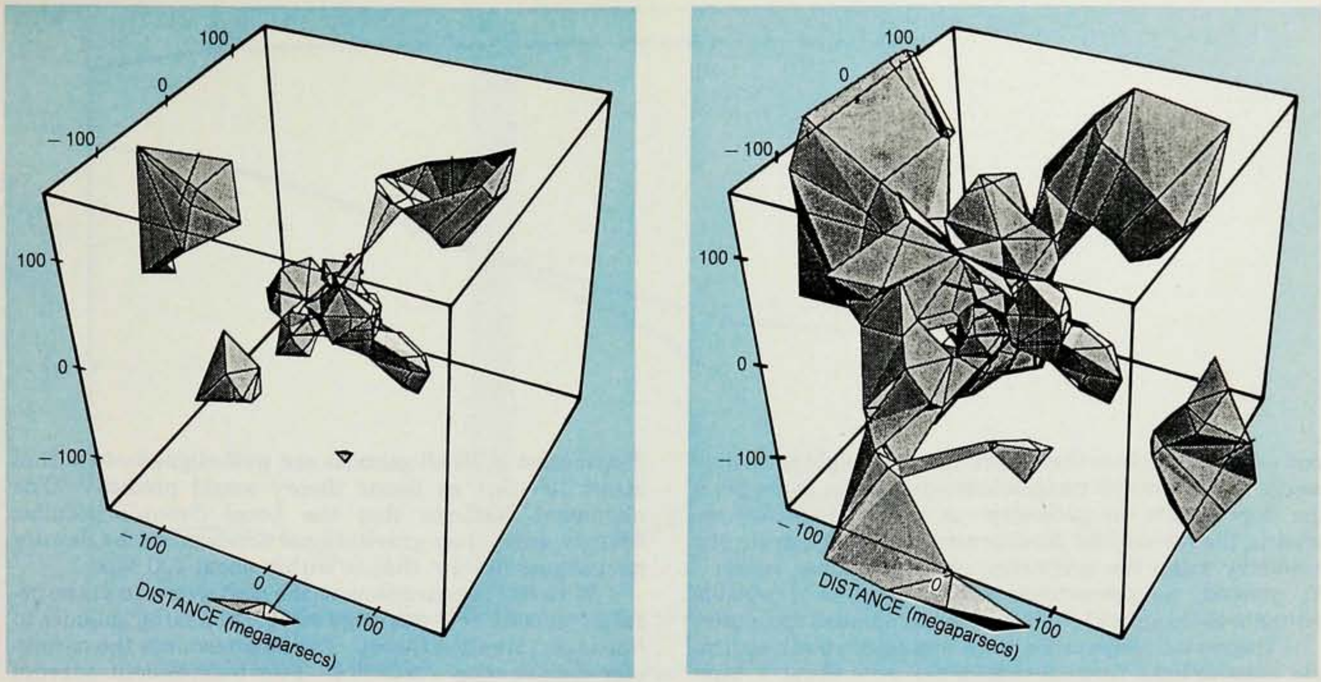
Clusters like Coma are strong x-ray sources,[9] with luminosities on the order of 10^{44} ergs/sec. The x rays arise from thermal bremsstrahlung in gas at a temperature of about 10^8 K. The gas is an additional source of x-ray-luminous mass. The total gas mass is somewhat model dependent but cannot exceed 20% of the total mass inside 1.3 Mpc; thus at least 80% of the cluster mass is still dark.

The x-ray observations can also be used to check our estimate of the cluster mass. The gas in the central parts of the cluster is in hydrostatic equilibrium, which implies that

$$\frac{\mathrm{d}p}{\mathrm{d}r} = -\rho_g\frac{\mathrm{d}U}{\mathrm{d}r} \tag{4}$$

where $\rho_g(r)$ and $p(r)$ are the gas density and pressure. Imaging observations by the Einstein satellite (figure 4) and spectral observations by the Tenma and Exosat satellites can be fit to models for the temperature and density distribution of the gas. These can be combined with equation 4 and the ideal-gas equation to yield the potential gradient $\mathrm{d}U/\mathrm{d}r$ and hence the mass distribution. The mass contained within 1.3 Mpc is found to be 5–$10\times10^{14}M_\odot$, close to the value obtained from the galaxy kinematics.[9]

The dark mass in clusters cannot be attached to extended galaxy halos, as these will be shorn off by tidal forces to form a smooth dark matter distribution spread throughout the cluster, through which the truncated galaxies swim.



Density of galaxies around the Sun. The shaded volumes denote regions in which the fractional enhancement in the number density of galaxies, $\delta_n = [n(\mathbf{x})/\bar{n}] - 1$, exceeds $+0.5$ (left panel) and $+0.3$ (right panel). All distances are in megaparsecs. (Adapted from ref. 20.) **Figure 5**

Fritz Zwicky first pointed out the presence of dark matter in clusters of galaxies in 1933. His original paper was based on only seven galaxy radial velocities and a distance to Coma that was too small by at least a factor of 5. Fortunately, distance errors scale out of his calculation, and his conclusion remains unchanged and by now seems inescapable: Almost all of the mass in the Coma cluster is dark. Studies of other galaxy groups and clusters yield similar results, as do measurements of the distortion and splitting of images of distant galaxies by the gravitational fields of clusters.[10]

Dynamics on larger scales

In the standard Friedmann–Robertson–Walker cosmological model, the universe is homogeneous and isotropic on sufficiently large scales.[11] Of course there are small-scale irregularities such as stars, galaxies and clusters, but if the FRW approximation is correct there must be some distance $r_\star$ such that the mean density $\bar{\rho}$ and all other local properties are approximately the same in every cube of side $r_\star$, wherever it may be in the universe. Surveys suggest that $r_\star$ is about 50 Mpc, or only about 1% of the size of the visible universe, which is of order $c/H_0 = 4000$ Mpc, and so the FRW approximation is reasonable.

A useful measure of the cosmological significance of dark matter in FRW models is the density parameter Ω, defined as the ratio of the mean density of the universe to the critical density: $\Omega = \bar{\rho}/\rho_c$, where $\rho_c = 3H_0^{\,2}/8\pi G = 1.06\times 10^{-29}\,\mathrm{g/cm^3}$. The density parameter is the ratio of the kinetic energy of the Hubble expansion to the absolute value of the gravitational potential energy; thus an expanding universe with $\Omega<1$ will expand forever, whereas one with $\Omega>1$ must eventually collapse. (I assume the cosmological constant[12] is zero.) The geometry of the universe is also determined by Ω: If $\Omega>1$ the universe is closed and the geometry is spherical, whereas if $\Omega<1$ the universe is open and the geometry is hyperboloidal.

The mean mass density $\bar{\rho}$ is the product of the mean luminosity density $\bar{\epsilon}$ and the mean mass-to-light ratio $\bar{\Upsilon}$. Galaxy surveys[13] show $\bar{\epsilon} = 1.3\times 10^8\, L_\odot/\mathrm{Mpc}^3$ to within a factor of 2, which implies

$$\Omega \equiv \bar{\Upsilon}/\Upsilon_c \tag{5}$$

where $\Upsilon_c = 1200\,\Upsilon_\odot$. Thus, for example, if $\bar{\Upsilon}$ is equal to the mass-to-light ratio of the Coma cluster—$300\Upsilon_\odot$—then Ω is 0.25 and the universe is open.

Many cosmologists believe that $\Omega = 1$ to high accuracy. Reasons for this belief include the inflation hypothesis,[14] which resolves several traditional problems with FRW cosmology and predicts $\Omega = 1$, and the temporal Copernican principle. (To understand that principle, suppose, for example, that $\Omega \approx 0.25$ and the universe is open; then in an FRW model Ω is very near unity at early times and very near zero at late times. There is no obvious reason why we should be living during the special epoch at which Ω first peels away from unity.) A further advantage of $\Omega = 1$ is that infall of distant bound material surrounding a protogalaxy provides a natural explanation for the origin of extended dark halos.

There is also a worrisome argument against $\Omega = 1$, based on a comparison of the age of an FRW model with stellar ages. Stellar evolution models show that the oldest stars are $t_\star = 15\pm 3$ billion years old.[15] The age of an FRW universe is $t_0 = f(\Omega)/H_0$, where $f(\Omega) = 1$ for $\Omega = 0$ and $\frac{2}{3}$ for $\Omega = 1$. If $\Omega = 1$, then for consistency ($t_\star < t_0$) we must have $H_0 < 45 \pm 9\ \mathrm{km\,sec^{-1}\,Mpc^{-1}}$, far smaller than our preferred value of 75 and near the lowest values obtained by any of the methods of determining the Hubble constant.[6] The significance of this argument is difficult to assess, however, without more secure limits on the Hubble constant and without a compelling alternative to the standard FRW model.

Because the geometry of the universe depends on Ω, the mean density $\bar{\rho}$ can in principle be determined from geometrical measurements, such as the dependence on distance of the brightness or the number density of galaxies. However, the uncertain effects of galaxy evolution—galaxies fade as their stars age, and brighten as they merge with nearby companions—generally overwhelm the dependence on geometry. A more promising approach, though still far from practical, is to constrain the geometry using the properties of gravitational lenses.[10] At present, no geometrical approach gives a reliable estimate of Ω, and so I shall focus on dynamical estimates.

The measurements we have discussed so far sample the mass-to-light ratio on scales not over about 1 Mpc, whereas equation 5 requires the average mass-to-light ratio on scales exceeding $r_\star$, which is about 50 Mpc. The two may be different: If, for example, stars and galaxies form preferentially in high-density regions such as clusters of galaxies, then the average mass-to-light ratio may be substantially larger than the mass-to-light ratio in the central parts of clusters like Coma. In one simple empirical model for such "biased" galaxy formation,[16] fluctuations in the number density $n(\mathbf{x})$ of galaxies are proportional to fluctuations in the mass density $\rho(\mathbf{x})$:

$$\delta_n(\mathbf{x}) = b\delta_\rho(\mathbf{x}) \qquad (6)$$

where $\delta_n = n(\mathbf{x})/\bar{n} - 1$, $\delta_\rho = \rho(\mathbf{x})/\bar{\rho} - 1$ and the constant $b \gtrsim 1$ is called the bias factor.

The determination of mass-to-light ratios on scales of order $r_\star$ or greater requires methods different from those applicable to galaxies or clusters. Structures of this size are still just starting the process of gravitational collapse, and their evolution is described by linear perturbations to an FRW model. One simple prediction of linear theory is that the peculiar velocity $\mathbf{v}_p$ is directly proportional to the peculiar gravitational acceleration $\mathbf{g}$ arising from the density fluctuations $\delta\rho(\mathbf{x})$, which in turn, if equation 6 applies, can be determined from the fluctuations $\delta n(\mathbf{x})$ in the number density of galaxies.[17]

This prediction can be checked against the peculiar velocity of our own Galaxy, or more properly the peculiar velocity of the center of mass of the Local Group of galaxies. The Local Group includes our own Galaxy plus its bound companion M31 at a distance of 0.7 Mpc. The peculiar velocity of the Local Group, determined from the motion of the Sun relative to the cosmic background radiation,[17] is 600 ± 27 km/sec toward Galactic longitude 268° and latitude 27°.

The best available sample for determining the number-density fluctuations $\delta n(\mathbf{x})$ is the set of galaxies detected by the Infrared Astronomical Satellite. The IRAS survey is unaffected by dust obscuration, covers almost the whole sky, has well-calibrated flux limits and samples a sufficiently large volume of the universe—to distances beyond 100 Mpc. Figure 5 shows an estimate of $\delta n(\mathbf{x})$ from the IRAS survey.

It turns out that the Local Group's peculiar velocity $\mathbf{v}_p$ and the peculiar acceleration $\mathbf{g}$ determined from the distribution of IRAS galaxies are well aligned—to within about 10°—just as linear theory would predict.[18] The alignment confirms that the Local Group's peculiar velocity arose from gravitational acceleration by density fluctuations nearby, that is, within about 100 Mpc.

With this encouragement, the next step is to fit the peculiar-velocity field of a large sample of nearby galaxies to linear perturbation theory. The fit determines the parameter combination $\lambda \equiv \Omega^{0.6}/b$. Two independent surveys yielded $\lambda = 0.95 \pm 0.20$ and 0.89 ± 0.16, consistent with $\Omega = 1$ if the bias factor b is near unity on large scales.[19,20]

Thus analysis of the phase-space distribution of galaxies out to about 100 Mpc supports the density parameter $\Omega = 1$ that is favored on theoretical grounds, so long as the galaxy density traces the mass density on large scales.

What is the dark matter?

Although my main goal in this article is to describe the evidence for dark matter, I will briefly summarize what little we know about its nature.

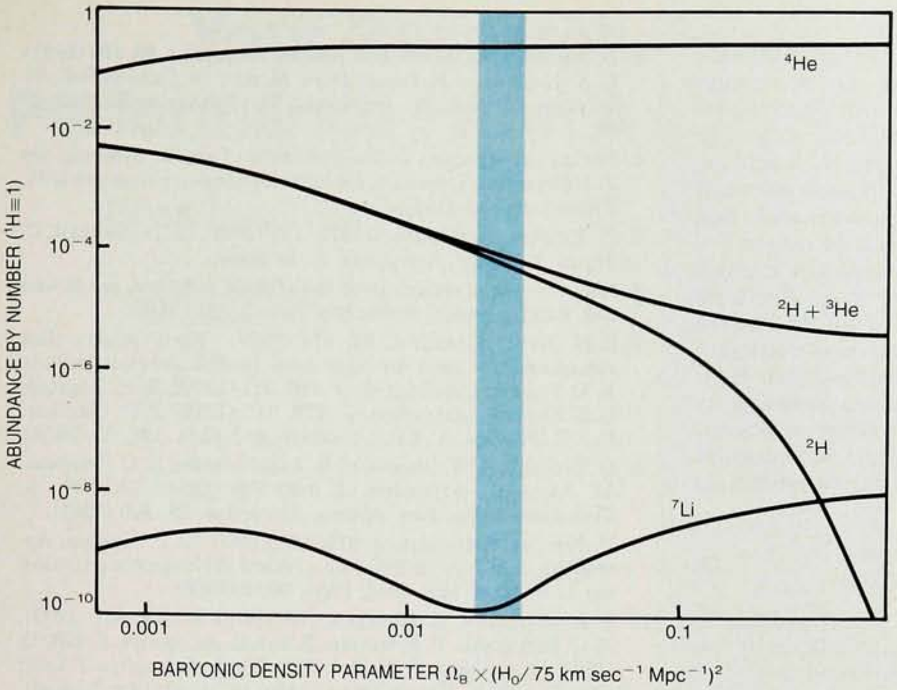
A strong constraint on the total density of baryons comes from the abundances of the light elements. In the standard FRW cosmology, ^{2}H, ^{3}He, ^{4}He and ^{7}Li are all formed in the first 10^3 seconds after the Big Bang, when the temperature exceeds 5×10^8 K. The abundance predictions of the standard model depend on a single parameter, which may be taken to be the present mean density of baryons, $\bar{\rho}_B$. The observed abundances are consistent with the predictions of the standard model[21] if and only if the baryonic density parameter Ω_B, which is defined as $\bar{\rho}_B/\rho_c$, satisfies

$$0.02 \leqslant \Omega_B (H_0/75 \text{ km sec}^{-1} \text{ Mpc}^{-1})^2 \leqslant 0.03 \qquad (7)$$

As figure 6 indicates, if Ω_B exceeds the upper limit, cosmological production of ^{2}H is too small. Nuclear reactions in stars do not provide a loophole, since they only destroy ^{2}H. Below the lower limit the abundance of ^{2}H plus ^{3}He is too large. (The combination is used because most of the ^{3}He is produced by burning ^{2}H.) Variations in either direction tend to produce too much ^{7}Li.

The principal uncertainty in Ω_B arises from the uncertain value of the Hubble constant H_0. Even if H_0 is as small as 50 km sec^{-1} Mpc^{-1}, the constraint in equation 7 implies that $\Omega_B \lesssim 0.07$. This density is far too small to explain the Ω value inferred from the mass of the Coma cluster ($\Omega \simeq 0.25$) or from large-scale density fluctuations and peculiar velocities ($\Omega \simeq 0.75$–1.15 if the bias factor for IRAS galaxies is near unity) or to provide the density required by inflation ($\Omega = 1$). Thus most of the mass of the universe cannot be in baryons.

An appealing possibility is that the nonbaryonic dark matter consists of weakly interacting, massive, stable

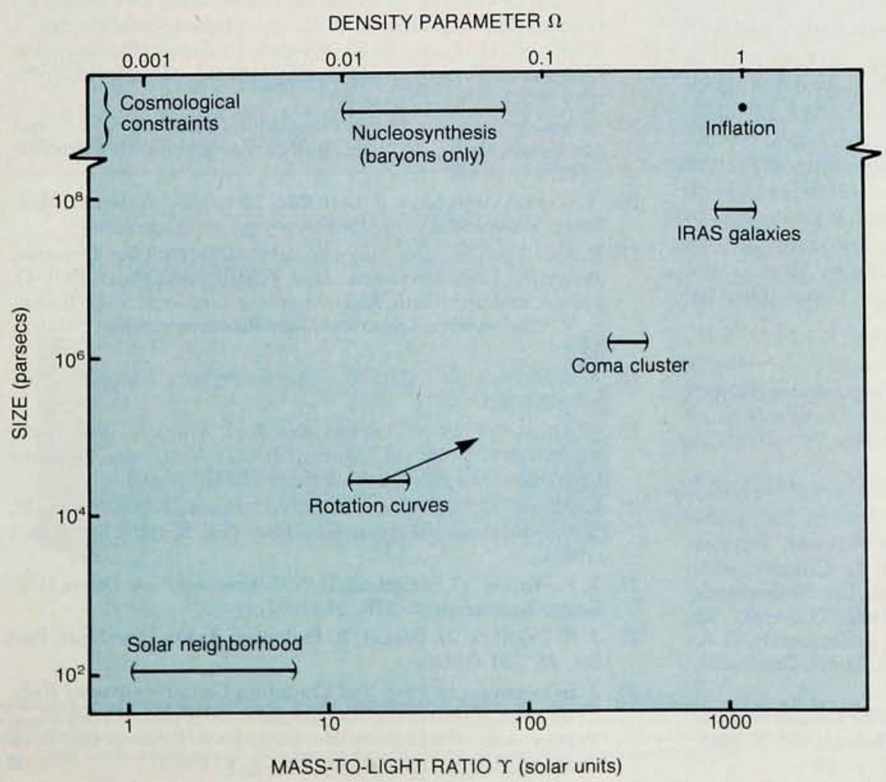


Predicted abundances from cosmological nucleosynthesis as a function of Ω_B (the ratio of the baryon density to the critical density) and the Hubble constant H_0. The shaded band marks the range consistent with observations. (Adapted from ref. 21.) **Figure 6**

elementary particles—"WIMPs"—formed in the hot, dense early universe. There are several plausible but hypothetical candidates in nonstandard models of particle physics, including massive neutrinos, neutralinos and axions, some of which may be detectable in laboratory experiments.[22] (See Leo Stodolsky's article in PHYSICS TODAY, August 1991, page 24.)

The constraint in equation 7 also implies that $\Omega_B > 0.01$, even if H_0 is as large as 100 km $\mathrm{sec^{-1}\ Mpc^{-1}}$. The mean mass-to-light ratio of baryonic material thus exceeds $12\Upsilon_\odot$, from equation 5. This is substantially larger than the mass-to-light ratio of the stars and gas in the solar neighborhood, which is typical for disk galaxies. Thus there must be a substantial component of baryonic dark matter as well, most likely in the form of brown dwarfs or compact stellar remnants—white dwarfs, neutron stars or black holes.

Finally, it is possible that the apparent evidence for dark matter arises from inadequacies in the conventional laws of gravity or dynamics. There is little evidence that Newtonian gravity is accurate on scales much larger than 0.1 pc, the size of the solar system comet cloud. Thus, for ex-



Constraints on the mass-to-light ratio Υ and density parameter Ω as a function of scale, with determinations based on the standard FRW cosmological model plotted at the top, beyond the size scale. The plot assumes that the Hubble constant H_0 is 75 km $\mathrm{sec^{-1}\ Mpc^{-1}}$, except that the error bar for the nucleosynthesis constraint includes the uncertainty in the Hubble constant. The constraint from galaxies detected by IRAS, the Infrared Astronomical Satellite, assumes that the bias factor b is 1. The arrow on the rotation curve determination is a reminder that we can measure only a lower limit on the halo mass and size from rotation curves. The relation between Υ and Ω is given by equation 5. **Figure 7**

ample, we might consider modifying the gravitational acceleration from a point mass M from the Newtonian expression $a = GM/R^2$ to $a = (GM/R^2) + (GM/R_0R)$, where R_0 is some new fundamental length. Then the circular speed around a mass M at distances $R \gg R_0$ would be $v_c = (GM/R_0)^{1/2}$, consistent with the flat rotation curves of disk galaxies. One difficulty (among several) with this proposal is that if the mass-to-light ratio is constant—which it should be, if there is no dark mass—the circular speed should scale as $v_c \propto L^{1/2}$, where L is the total luminosity of the galaxy. This contradicts the observation that for disk galaxies $v_c \propto L^{0.25}$ over more than two orders of magnitude in luminosity (the infrared Tully–Fisher law).

A much more interesting modification, proposed by Mordehai Milgrom of the Weizmann Institute of Science in Rehovot, Israel, is to introduce a new fundamental acceleration a_0, so that the acceleration from a point mass M is

$$a = \begin{cases} GM/R^2, & \text{for } a \gg a_0 \\ (GMa_0)^{1/2}/R, & \text{for } a \ll a_0 \end{cases} \tag{8}$$

In this case the circular speed at large distances is $v_c = (GMa_0)^{1/4}$, and for constant mass-to-light ratio we have $v_c \propto L^{1/4}$, consistent with the Tully–Fisher law.

The modified acceleration specified by equation 8 is surprisingly successful at explaining most of the dynamical evidence for dark matter, with the constant $a_0 \simeq 1 \times 10^{-8}$ cm/sec^2. The modified acceleration can be derived from a nonrelativistic Lagrangian, but so far there is no fully satisfactory replacement for general relativity that yields equation 8 in its weak-field limit.[23]

Figure 7 summarizes the evidence for dark matter described in this article. The following general trends seen in the figure are confirmed by many other dynamical arguments, of varying accuracy and rigor, that are not described here:

▷ The dark mass exceeds the luminous mass in virtually all systems of galaxy size or larger.

▷ The ratio of dark to luminous mass generally increases with scale.

▷ On scales greater than about 10 Mpc the ratio of dark to luminous mass is independent of scale and of order 10^3 in solar units, large enough that the total density could equal or exceed the critical value needed to close the universe.

▷ At least 90% of the mass in the universe is not baryonic.

At present we must admit with some embarrassment that we do not know what most of the universe is made of. A more positive view is that the light that has been shed on dark matter over the past two decades is the first stage of a revolution against "barycentric" cosmology that is the direct descendant of the revolution that Copernicus led against geocentric cosmology.

* * *

I am grateful to Stephen Kent, Nicholas Kaiser, Konrad Kuijken, Michael Merrifield, Anthony Tyson and Don VandenBerg for advice and discussions, and to Jeremiah Ostriker for introducing me to this interesting subject.

References

1. For discussions of dark matter, see M. Turner, Physica Scripta **T36**, 167 (1991); D. Lynden-Bell, G. Gilmore, eds., *Baryonic Dark Matter*, Kluwer, Dordrecht, The Netherlands (1990); E. W. Kolb, M. S. Turner, *The Early Universe*, Addison–Wesley, Redwood City, Calif. (1990); J. Kormendy, G. R. Knapp, eds., *Dark Matter in the Universe*, Reidel, Dordrecht, The Netherlands (1987).
2. W. Gliese, H. Jahreiss, A. R. Upgren, in *The Galaxy and the Solar System*, R. Smoluchowski, J. N. Bahcall, M. S. Matthews, eds., U. Arizona P., Tucson (1986), p. 13.
3. D. J. Stevenson, Annu. Rev. Astron. Astrophys. **29**, 163 (1991). L. A. Nelson, in *Baryonic Dark Matter*, D. Lynden-Bell, G. Gilmore, eds., Kluwer, Dordrecht, The Netherlands (1990), p. 67.
4. For an introduction to the dynamics of stellar systems, see J. J. Binney, S. Tremaine, *Galactic Dynamics*, Princeton U.P., Princeton, N. J. (1987).
5. K. Kuijken, Astrophys. J. **372**, 125 (1991). J. N. Bahcall, C. Flynn, A. Gould, Astrophys. J., in press.
6. For a review of estimates of the Hubble constant, see S. van den Bergh, Astron. Astrophys. Rev. **1**, 111 (1989).
7. S. M. Kent, Astron. J. **93**, 816 (1987). Early papers that recognized the need for dark mass in disk galaxies include K. C. Freeman, Astrophys. J. **160**, 811 (1970); D. H. Rogstad, G. S. Shostak, Astrophys. J. **176**, 315 (1972); J. P. Ostriker, P. J. E. Peebles, A. Yahil, Astrophys. J. Lett. **193**, L1 (1974).
8. D. Zaritsky, E. W. Olszewski, R. A. Schommer, R. C. Peterson, M. Aaronson, Astrophys. J. **345**, 759 (1989). M. Fich, S. Tremaine, Annu. Rev. Astron. Astrophys. **29**, 409 (1991).
9. D. Merritt, Astrophys. J. **313**, 121 (1987). J. P. Hughes, Astrophys. J. **337**, 21 (1989). For a review of clusters of galaxies, see C. Sarazin, Rev. Mod. Phys. **58**, 1 (1986).
10. S. A. Grossman, R. Narayan, Astrophys. J. **344**, 637 (1989). A. G. Bergmann, V. Petrosian, R. Lynds, Astrophys. J. **350**, 23 (1990). J. A. Tyson, F. Valdes, R. A. Wenk, Astrophys. J. Lett. **349**, L1 (1990). For an introduction to gravitational lenses, see R. D. Blandford, C. S. Kochanek, in *Dark Matter in the Universe*, J. Bahcall, T. Piran, S. Weinberg, eds., World Scientific, Singapore (1987), p. 133; R. D. Blandford, R. Narayan, Annu. Rev. Astron. Astrophys. **30** (1992), in press.
11. For discussions of FRW cosmology, see R. M. Wald, *General Relativity*, U. Chicago P., Chicago (1984); S. Weinberg, *Gravitation and Cosmology*, Wiley, New York (1972).
12. S. M. Carroll, W. H. Press, E. L. Turner, Annu. Rev. Astron. Astrophys. **30** (1992), in press.
13. G. Efstathiou, R. S. Ellis, B. A. Peterson, Mon. Not. R. Astron. Soc. **232**, 431 (1988). J. Loveday, B. A. Peterson, G. Efstathiou, S. J. Maddox, to appear in Astrophys. J. (1992).
14. A. H. Guth, Phys. Rev. D **23**, 347 (1981). For reviews of inflation, see R. H. Brandenberger, Rev. Mod. Phys. **57**, 1 (1985); A. Guth, in *Inner Space/Outer Space*, E. W. Kolb, M. S. Turner, D. Lindley, K. Olive, D. Seckel, eds., U. Chicago P., Chicago (1986), p. 287; Ya. B. Zeldovich, Sov. Sci. Rev. E **5**, 1 (1986); W. H. Press, D. N. Spergel, in *Dark Matter in the Universe*, J. Bahcall, T. Piran, S. Weinberg, eds., World Scientific, Singapore (1987), p. 197.
15. D. VandenBerg, in *The Extragalactic Distance Scale*, S. van den Bergh, C. Pritchet, eds., Astron. Soc. Pacific, San Francisco (1988), p. 187.
16. N. Kaiser, Astrophys. J. Lett. **284**, L9 (1984). A. Dekel, M. J. Rees, Nature **326**, 455 (1987).
17. P. J. E. Peebles, *The Large-Scale Structure of the Universe*, Princeton U.P., Princeton, N. J. (1980). D. Lynden-Bell, O. Lahav, in *Large-Scale Motions in the Universe*, V. C. Rubin, G. V. Coyne, eds., Princeton U.P., Princeton, N. J. (1988), p. 199.
18. R. T. Harmon, O. Lahav, E. J. A. Meurs, Mon. Not. R. Astron. Soc. **228**, 5p (1987).
19. A. Yahil, in *Particle Astrophysics*, J.-M. Alimi, A. Blanchard, A. Bouquet, F. M. de Volnay, J. T. T. Van, eds., Editions Frontières, Gif-sur-Yvette, France (1990), p. 483.
20. N. Kaiser, G. Efstathiou, R. Ellis, C. Frenk, A. Lawrence, M. Rowan-Robinson, W. Saunders, Mon. Not. R. Astr. Soc. **252**, 1 (1991).
21. T. P. Walker, G. Steigman, D. N. Schramm, K. A. Olive, H.-S. Kang, Astrophys. J. **376**, 51 (1991).
22. J. R. Primack, D. Seckel, B. Sadoulet, Annu. Rev. Nucl. Part. Sci. **38**, 751 (1988).
23. J. Bekenstein, in *Proc. 2nd Canadian Conf. on General Relativity and Relativistic Astrophysics*, A. Coley, C. Dyer, T. Tupper, eds., World Scientific, Singapore (1988), p. 68. R. H. Sanders, Astron. Astrophys. Rev. **2**, 1 (1990).