EXISTENCE AND NATURE OF DARK MATTER IN THE UNIVERSE

Virginia Trimble

Astronomy Program, University of Maryland, College Park, Maryland 20742, and Department of Physics, University of California, Irvine, California 92717

1. HISTORICAL INTRODUCTION AND THE SCOPE OF THE PROBLEM

The first detection of nonluminous matter from its gravitational effects occurred in 1844, when Friedrich Wilhelm Bessel announced that several decades of positional measurements of Sirius and Procyon implied that each was in orbit with an invisible companion of mass comparable to its own. The companions ceased to be invisible in 1862, when Alvan G. Clark turned his newly-ground 18½" objective toward Sirius and resolved the 10⁻⁴ of the photons from the system emitted by the white dwarf Sirius B. Studies of astrometric and single-line spectroscopic binaries are the modern descendants of Bessel’s work.

A couple of generations later, data implying nonluminous matter on two very different scales surfaced almost simultaneously. First, Oort (498, 499) analyzed numbers and velocities of stars near the Sun and concluded that visible stars fell shy by 30–50% of adding up to the amount of gravitating matter implied by the velocities. Then, in 1933, Zwicky (777) concluded that the velocity dispersions in rich clusters of galaxies required 10 to 100 times more mass to keep them bound than could be accounted for by the luminous galaxies themselves. The former result was taken much more seriously than the latter by contemporary and succeeding astronomers (being dignified by the name “the Oort limit”), which is perhaps more a statement about the personalities of Oort and Zwicky than about anything else.
The next decades were by no means devoid of relevant ideas and investigations (346, 484). The beginning of the modern era of dark-matter research can, however, be dated to 1974, when Ostriker, Yahil & Peebles (506) and Einasto, Kraasik & Saar (193) tabulated galaxy masses as a function of the radius to which they applied and found $M$ increasing linearly with $R$ out to at least 100 kpc and $10^{12} M_\odot$ for normal spirals and ellipticals.

Since then, a mainstream astronomer who seriously doubted that we are somehow not readily seeing 90% or more of the stuff in the Universe has found himself in the position of having to justify his discordant views. The low-mass torch, upheld for a time by Burbidge (109) and Woltjer (755), has recently been refueled by Valtonen (706–708).

Because dark matter has been invoked in many different objects and on many different scales, a very large fraction of astronomical research bears in some way on the issue. Necessarily, then, many aspects are given rather short shrift here.

First, nothing is said about the value of the Hubble constant, though it enters in powers from $-2$ to $+2$ into various determinations of mass and luminosity of distant objects and is arguably the largest single uncertainty in these determinations (268). Hodge’s (301) 1981 conclusion that an impartial choice of value for $H_0$ would be both difficult and unprofitable remains regrettabley correct. Besides, like Hodge, I have friends in both camps.

Next, several other relevant topics that have recently been reviewed in this series are somewhat neglected. These include evolution of galaxies in clusters (186), models of the Milky Way (36), the contribution of low-mass stars to the local mass density (415), constraints on dark matter in globular clusters (196), and properties of the Milky Way spheroid (229). In general, little is said about how calculations were done, except where methods have been substantially criticized. References cited for results generally explain how they were obtained, and, for the researcher desiring to acquire a thorough knowledge of methods, the standard starting place is Chandrasekhar’s *Stellar Dynamics* (132). For many of the cases considered, $GM/R \sim V^2$ is all the physics needed.

Finally, the literature search approached completeness only for English-language journals received by University of California, Irvine, and/or University of Maryland between May 1984 and July 1986 and catalogued under Library of Congress designation QB (astronomy and astrophysics). Some relevant work appearing earlier, in books, and in non-English or nonastronomical journals has undoubtedly been missed. Much of the pre-1979 literature can be accessed through Faber & Gallagher’s (200) fine review of masses of galaxies.
For the objects and systems discussed here, masses and luminosities will normally be given in solar units. That is, \( M/L_B = 10 \) means 10 solar masses of gravitating material for every one solar luminosity in the blue band. If no wavelength band is indicated, either bolometric luminosity is meant or the number has sufficiently large error bars that the wavelength makes no difference.

There exist a modest number of objects and systems where little or no dark matter can be present. These either have the dynamical mass equal to the luminous mass, or larger (but there exist more attractive explanations than dark matter), or seemingly smaller (so that some kind of energy input is needed to prevent collapse). The solar system, having been stable for some \( 4.5 \times 10^9 \) yr, can be used to put rather stringent limits on nearest permitted approaches of black holes, substellar mass objects, and other hypothetical forms of dark matter (299, 325), including a possible substellar companion to the Sun (746).

Other objects under this heading include high-latitude molecular clouds (seemingly unbound, but perhaps confined magnetically or by hot coronal gas pressure; 366, 433), Galactic plane molecular clouds (which require nonthermal support or continuous energy input; 621, 638), and the radio-emitting lobes of extragalactic radio sources [for which gravitational confinement by a massive black hole has been proposed (108), but pressure confinement is now more widely accepted]. One intergalactic H I cloud that looks unbound (594) may be gravitationally attached to a pair of galaxies (593) or be a genuinely transient configuration (559, 734).

Star clusters, both open (440, 441, 747) and globular (158, 159, 196, 269, 524, 535), display dynamical properties fully explicable by the mass in visible stars and predictable stellar remnants. The very large proper motions, leading to \( M = 10^8 \) \( M_\odot \) for one cluster (486), need to be confirmed. Most measurements pertain only to brighter parts of clusters and would permit very extended, massive dark halos (520), but at least in M15, the velocity dispersion declines from 15 km s\(^{-1}\) in the core to less than 1 km s\(^{-1}\) in the outskirts [see work by P. Seitzer & K. C. Freeman reported in (231)], suggesting that the light edge is also the mass edge.

2. SINGLE GALAXIES

Mass estimates applying to individual galaxies can come from velocity dispersions or rotation curves of the stars and gas (including X-ray-emitting gas) making up the galaxy itself or from positions and velocities of test particles like globular clusters and satellite galaxies. Where several methods can be applied to the same part of the same galaxy, results are frequently, but not always, in reasonable agreement. In the conventional
terminology, once a given amount of gravitating mass has been found on
one scale, it no longer counts as "dark" or "missing" on larger scales.
Thus, for instance, $M/L = 3$ is considered proper for inner disks of spirals
in clusters of galaxies—because it is what we measure for the solar neigh-
borhood—even though a third or more of the $M$ required may be in
unknown form.

2.1 Milky Way

2.1.1 Solar Neighborhood Later versions of the analysis of stellar
velocities and distributions perpendicular to the Galactic plane pioneered
by Oort (498, 499) have been quite consistent over the years. Bahcall
(35), for instance, reported $\rho(\text{dark})/\rho(\text{luminous}) = 0.5-1.5$, with the dark
matter confined to a scale height less than 0.7 kpc, rather like the old disk
population. The total disk $M/L$ is then about 3 and the local density of
dark matter 0.1 $M_\odot$ pc$^{-3}$ or 30 $M_\odot$ pc$^{-2}$ in a cylinder perpendicular to
the plane.

Most studies have, however, used rather similar tracer populations—
main-sequence F stars and K giants. For each, owing to effects of evolution
and metallicity, stellar brightnesses may have been overestimated and
distances underestimated (268), resulting in an overestimate of the local
mass density. This happens because we measure the velocity distribution
only locally, but count tracer stars far from the plane. Thus, if a given
velocity carries a star farther from the plane than you thought, the amount
of mass holding it back is less than you thought.

The dynamics of a younger population, confined more closely to the
Galactic disk, provides little evidence for dark matter (392), which could
mean that the distribution is that of a thick disk, or that there is a problem
with the analysis (37). A new, independent data set, extending to about 2
kpc from the plane, apparently requires little or no dark matter in either
a thin or thick disk, the stellar populations in these and other components
of the Galaxy providing all the necessary gravitating mass (G. Gilmore &
K. Kuijken, personal communication, 1986). The existence of a separate
dark component belonging specifically to the Milky Way disk should
probably, therefore, not be accepted without reservation (268).

2.1.2 Milky Way Rotation Curves and the Disk Mass If our Galaxy
were a spherically or even radially symmetric configuration, then the
circular orbit speeds of test particles at known distances from the center
could be turned into a curve of $M(R)$ with only modest uncertainties
arising from the difference between the gravitational potentials of spheres,
ellipsoids, and thin disks.

There are several stumbling blocks along this seemingly smooth path,
including the problem of determining accurate distances to objects whose velocities are known (or conversely) and that of converting heliocentric velocities to galactocentric ones. The former has been solved by using stars closely associated with gas clouds as the distance indicators (82, 390) or by using velocities of bright stars themselves (587). The main uncertainty in the velocity conversion comes from uncertainties in the values of our own distance from the Galactic center ($R_0$) and rotation speed ($V_c$) (81). Traditional values of 10 kpc and 250 km s$^{-1}$ are in the process of being replaced by 8.5 kpc and 220 km s$^{-1}$ (365), but the residual uncertainty is enough that masses mentioned in this section should be interpreted as containing a factor ($R_0/10$)($V_c/250$)$^2$.

The traditional $R_0$ and $V_c$ imply $1.5 \times 10^{11} M_0$ and $M/L_B = 10$ for the inner disk. Since $L$ scales as $R_0^2$ (713), the ratio $M/L$ is better known than either separately. The Galactic infrared luminosity contributes another $1-2 \times 10^{10} L_\odot$ (281), so that $M/L_{bol}$ is closer to 5 than 10.

Outside $R_0$, rotation curves determined from gas continue to rise (81, 138, 376, 390, 527), while that derived from stars remains flat at $V = V_c$ (31, 587) out to about 2 $R_0$. The implied masses are $4.3$ and $3.0 \times 10^{11} M_\odot$ for the traditional rotation constants (or 3 and $2 \times 10^{11} M_\odot$ for the new ones), and $M/L_{bol} \sim 15-20$. Some of the discrepancy probably arises from noncircular gas motions (144, 279), but there must also be some problem in at least one of the distance scales and, therefore, a real uncertainty in the mass. That $L(R)$ does not continue to rise is in no doubt, since we can see individual bright stars (or, rather, the relative deficiency of them) out to many kiloparsecs and are not dependent on deconvolutions of surface photometry as is the case for most other galaxies (Section 2.2.1).

Although these rotation curves use tracers in the disk, several lines of argument strongly suggest that much of the nonluminous mass is in a spherical or spheroidal component, especially outside $R_0$. These arguments include (a) the flaring of the Galactic H I disk (390); (b) detailed deconvolutions of both the rotation curve (36, 280) and surface photometry (384) into several Galactic components, including nucleus, bulge, thin Population I disk, thick old Population I disk, Population II spheroid, and dark halo; and (c) the need for a spheroidal or hot component (not necessarily dark) to stabilize bar modes in a thin cold disk (474, 503, 605).

Thus we have more or less talked ourselves into a relatively dark halo with $M \gtrsim M_{disk}$ out to $1-2 R_0$ without being able to say whether it is merely an unexpectedly large mass associated with the known Population II spheroid or a physically distinct component. Discussion of observations within the Galactic disk that constrain the nature (as opposed to the amount) of dark matter appears in Section 6.3.
2.1.3 The Halo of Our Galaxy. Charting the velocity field for the Galactic halo presents the same problems as for the disk in identifying a sufficient number of objects whose distances and velocities can both be measured. The uncertainties scaling with $R_0$ and $V_c^2$ persist for most techniques. An additional one arises because there is no reason to expect halo orbits to be circular, and a given position/velocity pair "belongs" to an orbit of larger semimajor axis and, therefore, of larger central mass if the orbit is nearly a circle than if it plunges straight in. The common assumption of locally isotropic velocity distributions need not be correct. The reward that compensates for these additional difficulties is that halo objects can be found out to 5–10 $R_0$ (vs. 2 $R_0$ for disk ones), enabling us to probe $M(R)$ further out. Other properties of halo objects, including ages and chemical compositions, that may be correlated with dynamics have been elegantly reviewed by Carney (124).

Commonly used probes include RR Lyraes and other field stars, globular clusters, and satellite galaxies. The latter two can be treated in two different ways, estimating the mass interior to their positions either from their velocities or from the assumption that their observed sizes represent tidal limits set by the Galaxy as a whole. Because stars leak gradually away from the edges of clusters and satellites, rather than vanishing instantly, tidal radii fitted to star counts or photometry are probably underestimates by factors of about 1.4 (501), and masses derived from them should be scaled up by about $(1.4)^3$. Masses assumed for the satellite dwarf galaxies (Section 2.4) also enter about linearly into calculations based on them, and the Milky Way comes out very massive only if they are.

Table 1 summarizes the results of recent investigations. Features to be noted are the upper envelope, which has $M(R)$ gradually increasing with radius to at least $10^{12} M_\odot$ at 100 kpc (implying $M/L \gtrsim 30$), and fluctuations due to different choices of velocity distribution, the argument in favor of circular orbits for the outer objects being that they are so diffuse that they would not have survived even one close passage of the Galactic center (523). The methods used in most of the velocity-based estimates are described clearly by Lynden-Bell et al. (429). Two recent results (489 and, especially, 676) are considerably lower than average despite making use of roughly the usual data base and techniques. The reasons remain to be elucidated, but if the results are correct, most of the mass of the Milky Way is actually within a few $R_0$ and the total is considerably less than $10^{12} M_\odot$.

Significant reduction of the numbers in Table 1 can be achieved in two ways. First, a nonstandard modeling of the gravitational effects of the Magellanic Clouds (707) says that cluster and satellite velocities do not currently probe the gravitational potential very well. Second, the most
Table 1 Determinations of Galactic mass as a function of galactocentric distance

<table>
<thead>
<tr>
<th>$R$ (kpc)</th>
<th>Method</th>
<th>Mass ($M_\odot$)</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>12–17</td>
<td>RR Lyr for the LMC (isotropic velocities)</td>
<td>$2.6–2.9 \times 10^{11}$</td>
<td>148</td>
</tr>
<tr>
<td>17</td>
<td>Globular cluster velocities</td>
<td>$2 \times 10^{11}$</td>
<td>106</td>
</tr>
<tr>
<td>20</td>
<td>Globular clusters (circular orbits) (isotropic orbits)</td>
<td>$0.3–0.8 \times 10^{11}$</td>
<td>487</td>
</tr>
<tr>
<td>44</td>
<td>Globular cluster and satellite galaxy tidal radii</td>
<td>$8.9 \pm 2.6 \times 10^{11}$</td>
<td>333</td>
</tr>
<tr>
<td>50</td>
<td>Halo star and cluster velocities</td>
<td>$4.4 \times 10^{11}$</td>
<td>489</td>
</tr>
<tr>
<td>65</td>
<td>Escape velocity of 3 RR Lyras</td>
<td>$10–30 \times 10^{11}$</td>
<td>282</td>
</tr>
<tr>
<td>50–100</td>
<td>Globular clusters (isotropic velocities) (no radial orbits)</td>
<td>$5 \pm 2 \times 10^{11}$</td>
<td>497</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$10 \times 10^{11}$</td>
<td></td>
</tr>
<tr>
<td>50–100</td>
<td>Globular cluster velocities</td>
<td>$2 \times 10^{11}$</td>
<td>676</td>
</tr>
<tr>
<td>100</td>
<td>Globular cluster and satellite velocities (if all bound)</td>
<td>$9 \times 10^{11}$</td>
<td>471</td>
</tr>
<tr>
<td>100</td>
<td>Globular cluster isotropic velocities</td>
<td>$10 \times 10^{11}$</td>
<td>523</td>
</tr>
<tr>
<td>118</td>
<td>One globular cluster</td>
<td>$\leq 10 \times 10^{11}$</td>
<td>654</td>
</tr>
<tr>
<td>Total</td>
<td>Escape velocities of field halo stars</td>
<td>$\geq 5 \times$ mass to $R_0$</td>
<td>125</td>
</tr>
<tr>
<td>Total</td>
<td>Globular cluster and satellite galaxy tidal radii</td>
<td>$10 \times 10^{11}$</td>
<td>138</td>
</tr>
</tbody>
</table>

Extreme members of the stellar and cluster populations may not currently be bound to the Milky Way. Some stars with high positive velocities could be runaways from disrupted binary systems, but the large negative velocities [$-161$ km s$^{-1}$ for a globular cluster at 84 kpc (523) and $-465$ km s$^{-1}$ for one RR Lyrae at 64 kpc (282)], even if the objects are in free-fall toward us for the first time, require large Milky Way masses. The only alternative is a large velocity acquired non-gravitationally, through a runaway process in some nearby galaxy, or from mutual containment in the potential of some larger, more distant mass. The ad hoc explanations will come to seem quite improbable if the samples of high-velocity stars and clusters continue to increase, and mutual containment only shoves the dark-matter problem further away without removing it.

Existing data on halo velocity distributions are rather insensitive to the difference between flat and spheroidal mass distributions, but the round shape strongly suggests round equipotentials (473).

Despite the uncertainties and reservations expressed in the preceding two sections, it seems safe to conclude that (a) within $R_0$, there is about as much mass in a spheroidal (mostly dark) halo as within the luminous...
disk; and (b) outside $R_0$, there is at least 2 and probably 3–10 times as much matter as inside. The alternative is that the Good Lord is a good deal more raffiniert, and the average astronomer considerably less so, than we would like. Because the dark matter of Section 2.1.1 is confined at least to a thick disk and that of Sections 2.1.2 and 2.1.3 necessarily roughly spherical, they are in some sense two separate components, though they could perhaps (Section 6.3) consist of physically similar objects, whether faint stars or suitably chosen black holes.

2.2 Other Spiral Galaxies

The data (liberally interpreted) base pertinent to dark matter in single spiral galaxies consists of (a) a large number of rotation curves for galaxies seen more or less edge on, determined from optical (322, 562, 564) or 21-cm (11, 93, 259) emission lines; (b) some observations bearing on the issue of whether most of the gravitating mass is in the disk or in some more spheroidal component; and (c) a grab bag of ideas about disk stability, barred and ring galaxies, gravitational lenses, and the visible spheroidal component.

In comparison with studies of the Milky Way, we have the advantage of a larger sample and of not having to assume a local rotation speed, but we lose out in angular resolution and in having the poorly known distance scale parameter $H_0$ with an uncertainty of $\pm 50\%$ (centered around 75 km s$^{-1}$ Mpc$^{-1}$) rather than $R_0$ with an uncertainty of only $\pm 20\%$. Where authors have explained which value of $H_0$ they assumed, values of $M$ and $M/L$ will sometimes include a parameter $h$, meaning ($H_0$/100).

2.2.1 Rotation Curves and Disks The rotation of M31 was first detected by Slipper (628) and followed well out into the disk by Babcock (33) as part of his PhD dissertation. His rotation curve was still rising at the last measured point and (scaling to the modern distance) implied $M = 3 \times 10^{11} M_\odot$ and $M/L = 17$ out to 18 kpc. He remarked upon the difference from the Milky Way rotation curve as then understood and on the high $M/L$ compared to solar neighborhood values. Freeman (230) was among the first to notice that such non-Keplerian rotation curves were a widespread phenomenon and to deduce that there might be considerable gravitating mass outside the observed regions. Ten years later, the entire astronomical community was dashing madly in the direction (which was, therefore, by definition, forward) of $M/L$ increasing monotonically with $R$, only to trip over Kalnajs (349), who pointed out that some of the existing rotation curves, at least the optical ones, could be equally well fit by a disk of constant $M/L$, especially if it had a sharp edge not far outside the last observed point.
There has ensued a new round of data collection and, probably, a new consensus (with a few caveats) along the following lines (11, 39, 121, 361):

Many rotation curves remain flat or even rise to large radii. Once one accepts the possibility of many components with different central densities, $M/L$'s, and exponential or power-law scale lengths, these cannot be uniquely interpreted. As far out as optical data extend, one can get away with little or no halo (provided the disk properties are chosen carefully), but most of the likely fits have at least as much mass in spheroid as in disk, independent of Hubble type (Sa to Sd), and the importance of the halo increases as we look farther out with H I rotation curves. NGC 3198 (11), for instance, has been traced to $22.5 \, h^{-1}$ kpc (11 disk scale lengths), by which point the integral $M/L_B$ has risen to $24 \, h$ and the average halo/disk ratio to at least 4. A tentative point at $37.5 \, h^{-1}$ kpc has $M/L_B = 33 \, h$, close to values for binary galaxies (Section 3.2).

Work on multiplicity of spiral arms as a function of halo/disk mass ratio provides independent support for values near one (27, 28, 674). Disk masses, analogous to the Oort limit, can be extracted from stellar and gas velocity dispersions as a function of radius in face-on spirals (386). The average old disk $M/L_B$ for six galaxies is $6 \pm 2$. Clearly, we do not know how many low-mass main-sequence stars contribute to this, but the disk colors are consistent with an initial-mass function much like the local one, so that it is at least self-consistent to conclude that disks of other spirals contain $50 \pm 20\%$ dark matter (386).

Now, about the caveats. First, one worries about noncircular gas velocities due to perturbations by companions or outlying gas that has not had time to come into dynamical equilibrium. These can be mapped and are sometimes as small as $3 \, \text{km s}^{-1}$ for very regular spirals (387), but they can be as large as $45 \, \text{km s}^{-1}$ or $30\%$ of $V_c^{\text{max}}$ for others (48, 564). It is some consolation that visible companions and crowding within clusters tend to correlate with rotation curves that remain flat or turn down rather than rising (111, 770), suggesting that this sort of noncircular velocity effect, at least, causes us to derive low masses, not high ones.

Second, there are anomalies in both directions. Some galaxies with low $M/L$ have experienced a recent burst of star formation [Mkn 348 (624a)] or may just have been placed at excessive distances (450), but for NGC 3992 the conclusion that $M/L$ from the inner rotation curve and from some very distant companion galaxies is the same (corresponding to a total mass of $2 \times 10^{11} \, M_\odot$) is independent of distance (260). Anomalously large values can occur because star formation has been depressed for a long time (96, 237), because the circular velocity rises as high as $500 \, \text{km s}^{-1}$ (251), or because the rotation curve remains flat out to $100 \, \text{kpc}$ or more (555).
Finally, if the background sky brightness around spiral galaxies has been overestimated, $L(R)$ may also continue to increase (75, 385), so that $M/L$ remains near 10 even for very massive galaxies.

2.2.2 THE SHAPE OF THE DARK MATTER DISTRIBUTION, OR HOW ROUND WAS MY HALO For external spiral galaxies as well as for the Milky Way, evidence favoring a spheroidal rather than flat dark component comes from H I flaring and warps and considerations of disk stability. Polar ring galaxies provide independent support for this picture, and a bit can be said about higher moments of the distribution.

*H I flaring* M31 (103) and other galaxies (17) share with the Milky Way H I disks that widen at large radii. Since the gas velocity dispersion perpendicular to galactic disks does not increase with radius (387), this flaring must mean that the equipotentials are becoming rounder far out, requiring an increasing fraction of the total mass to be outside the visible thin disk. Although a spheroidal component seems most likely, a thick disk of at least several kiloparsecs will also work (414, 572). The phenomenon is not universal (94).

*H I warps* Disks whose outer regions look as if their rotation axes were tilted relative to the rest of the galaxy are also shared by the Milky Way, M31 (103), and other galaxies (93). To prevent differential rotation from destroying these requires a somewhat special potential shape. Massive spherical halos are one possibility (682), but truncated disks and companion galaxies are also possibilities (635). An interesting counterargument is that, since the most remote gas in spirals may have been recently acquired through cannibalism, the fact that it finds its way into a plane quickly indicates the existence of a (dark) flat disk potential at large radii (573).

*Disk stability* The bar modes of a thin (cold) stellar disk can be damped by a comparable amount of mass in a thick (hot) halo (503, 605, 769). This halo need not be dark, and the visible spheroid contributes significantly (29). The stability consideration can at least be regarded as an argument against putting much dark matter in the inner disks of spirals, although, for a gaseous disk, a spheroid may not be a stabilizing influence at all (553).

*Spindle or polar ring galaxies* These are a dozen or two otherwise-normal S0s, seen edge on and ringed by annuli of gas more or less perpendicular to the main galactic disks (26). Both models for the formation and, more important, measured velocities of the gas rings require a roughly spherical mass distribution (588, 599, 744).
Core radius and triaxiality  Because the inner, visible parts of spirals have much of their mass in disks, even the best existing velocity data and photometry turn out to justify only a lower limit to the halo core radius, typically in the range of 10–20 kpc (G. Lake, personal communication, 1986). This vitiates phase-space arguments that would constrain the mass of fermions contributing to dark matter if the core radii were small (516, 677). Binney (76) has compiled the arguments in favor of triaxial halos that come from numerical simulations of their formation and from detailed models of the driving and sustaining of warps, polar rings, and ripples. He mildly favors aligned spheroids for spiral galaxies and tilted ones for ellipticals. On the other hand, an antiparallel halo is an alternative cause of warps (636), and a tilted one could drive spiral arm formation where other mechanisms fail (442).

2.2.3  BARRED SPIRALS, GRAVITATIONAL LENSES, AND THE VISIBLE SPHEROID

Bars and rings  Since massive halos stabilize bar modes and make more difficult the formation of rings at the outer Lindblad resonance and of those attributed to galaxy collisions (104, 134a, 673), one might expect galaxies with these features not to show evidence for large halos. The SB NGC 3992 (261) and one barred Seyfert (58) have low $M/L$'s as expected.

Gravitational lenses  Two cases where QSOs seem to be lensed, but no galaxy is seen, require about $10^{12} M_\odot$ with $M/L_B \geq 21 h$ for 2016+112 (591) and as much as $10^{13} M_\odot$ with $M/L$ in excess of 100 for 2345+007 (696). Conversely, some cases where we see a foreground galaxy or QSO, but no lensed image of the background one produced by it (334, 695), set upper limits of $2 \times 10^{12} M_\odot$ for Q1548+114A,B and $2 \times 10^{11} M_\odot$ within a radius of 50 kpc for a number of galaxies. A proposed loosening of the latter limit (528) has probably mistaken too many of the background galaxies for dwarfs in the foreground (694) and requires rethinking. Serious worry about either the large lower limits or the small upper ones should probably be saved until gravitational lensing in general is better understood (110, 479).

The visible spheroid  Velocity dispersions in the cores of a number of S0s yield $M/L$ ratios of a few, exactly as expected from population synthesis (530) and the dominance of disk mass at small radii. Velocities of the globular clusters of M31 (278), if assumed isotropic, imply a mass of $3 \times 10^{11} M_\odot$ out to 20 kpc, very similar to the Milky Way result. This should be redone with the large sample now available (319). In a few early-type spirals, the spheroid is bright enough for photometry out to many kiloparsecs, most notably in the Sombrero galaxy (109a), for which the spheroidal $M/L_B$ is 5–6 out to 20 kpc, the disk value 10, and the halo/disk...
mass ratio about $4:1 \times 10^{11} \, M_\odot$. In such cases, the spheroid could be purely stellar (171a).

2.3 Elliptical Galaxies

The gravitational potentials of elliptical galaxies can be probed with velocities of stars, globular clusters, companion galaxies, and thermal and X-ray gas, none of which can be expected to show essentially circular motion like that in spiral disks. Most ellipticals, even highly flattened ones, are not rotationally supported (73). A few exceptions can be modeled as rotating structures within massive dark halos for which $M \propto R$ is, at any rate, a consistent solution (411). They are dynamically perhaps not very different from S0s. For the rest, the choice among isotropic, mostly radial, and mostly circular velocity distributions contributes significantly to uncertainties in most of the mass measurements addressed in the following sections.

2.3.1 Nuclei and Binary Nuclei The central black holes invoked in most models of active galactic nuclei (80) will constitute a form of dark matter once accretion ceases but make a negligible contribution to the mass inventory of the Universe. The subset of ellipticals (relatively common at the centers of rich clusters) that display two or more bright nuclei give the impression that they or their clusters must have very large $M/L$'s because the velocity differences between nuclei are often large. But this seems not to be the case (669–671). Stellar velocity profiles for these galaxies and their several nuclei are indistinguishable from those of other Es and lead to $M/L_\odot$ in the range 10–20.

2.3.2 Stellar Velocity Dispersions Normally, absorption-line profiles are deconvolved into $\sigma_v(r)$ using some standard model for $\rho(r)$ (73). The deconvolution is typically not unique, even for very good data. For instance, the central cD in Abell 2029, whose line widths increase with $R$, can be fit either by isotropic velocities and $M/L$ increasing linearly outward or by a preponderance of circular orbits and constant $M/L$ (667). Similar caveats apply to other galaxies (128). Not all conceivable distributions of orbit shapes result in self-consistent gravitational potentials within which those orbits are stable, but the difficult process of finding such self-consistent orbit distributions has only just begun (149, 598). With less than perfect line profiles, even at many radii, errors of 40% or more are easily possible (360, 551). The strongest probably safe statement is that the visible parts of most normal ellipticals have $M/L$ of 7 to 20 (34, 340).

The globular clusters have a larger velocity dispersion than the field stars at the same (two-dimensional) radius, at least in NGC 5128 (297) and M87 (475a). The difference between the two in radial distribution,
plus projection effects, can account for this (264), but additional dynamical differences are also possible.

2.3.3 THERMAL GAS Most ellipticals have very little gas, and cases showing a coherent gas velocity field are even rarer. A rotation curve for the H I in NGC 4278 (370) yielded \( M/L_B = 39 \; h \) out to two Holmberg radii. Two recent measurements with higher angular resolution suggest that \( M/L \) rises to these values from smaller central ones, as would be expected in the presence of an extended dark halo. Population synthesis to match line profiles in E cores (530) is also consistent with relatively small central \( M/L \).

The galaxies concerned are the small, gas-rich NGC 5666, for which a preliminary VLA map (G. Lake, personal communication, 1986) shows something like a normal rotation curve, with \( M/L \) increasing about a factor of two from center to outskirts, and the larger NGC 7097, which has enough optical gas to trace out to \( 1.5 \; h^{-1} \) kpc. By this point, \( M/L \) has risen to \( 7 \; h \) from \( 2 \; h \) at the center (116). Comparison of gas and stellar velocity dispersions in the latter indicates a preponderance of radial orbits. Thus, a central \( M/L \) derived assuming isotropy will be too large. Poor angular resolution in the absorption-line spectroscopy may also contribute to the discrepancy between these low central values and the larger ones of the previous section. It is important to look for color and other gradients associated with the changing \( M/L \) to distinguish population changes with radius from the effects of a hypothetical dark component.

2.3.4 SHELLS A dark halo, about as massive as that of a spiral with the same luminosity, but rather more compact, is a definite asset in modeling the shells and ripples seen in the outer regions of some normal and peculiar ellipticals (436). If the material contributing the shell light has come from a low-angular-momentum encounter with a smaller galaxy, then test particles representing it experience phase wrapping in the potential well of the elliptical and oscillate back and forth in radius. A sharp crest at the turnaround radius is seen as a shell, and multiple shells result from a spread in initial particle energies (191, 294, 776). There are other models for these shells which are not helped by a dark halo (203, 751).

2.3.5 PLANETARY NEBULAE, GLOBULAR CLUSTERS, AND COMPANION GALAXIES Emission-line velocities for the planetary nebulae in M32 yield a total mass of \( 8 \times 10^8 \; M_\odot \) and \( M/L = 3-4 \) for the visible parts of the galaxy (488). To within the statistical errors, the nebular velocities are distributed isotropically, suggesting to the authors that \( M/L \) does not increase with radius. M32 could probably not retain an independent halo in any case, given its close proximity to M31.

The most thoroughly analyzed elliptical cluster system is that of M87,
the central dominant galaxy in the Virgo cluster (316). The cluster velocity dispersion is larger than that of the stars further in, suggesting a preponderance of circular orbits. Assuming this to be correct and taking a distance of 16 Mpc, Huchra & Brodie (316) then find a mass of \(6 \times 10^{12} \, M_\odot\) and \(M/L_B = 150\) out to 18 kpc. This rises to nearly \(10^{13} \, M_\odot\) for an isotropic velocity distribution.

The velocities of companions of M87 probe much the same region and imply similarly large mass and \(M/L\), though with somewhat lower statistical significance, since fewer objects are involved (485). Because M87 is a cluster-center galaxy, one cannot be quite sure whether it is the galaxy or the cluster potential being traced out.

NGC 720, on the other hand, is a relatively isolated elliptical with six faint companions. Comparison of their velocity dispersion (353 km s\(^{-1}\)) and average radial position (23.2") with those of the main galaxy (214 km s\(^{-1}\) and 86") suggests that there is \(44^{\pm46}_{+28}\) times as much gravitating mass in the large volume as in the small one (189).

Some other central cD's seem to have a semi-bound population of companion galaxies with characteristic velocity dispersions of 250 km s\(^{-1}\) [vs. 1400 km s\(^{-1}\) for the clusters as a whole (151)], from which mass information could probably be extracted.

2.3.6 RADIO AND X-RAY SOURCES Of various models proposed for the confinement of radio jets and lobes, there is one (704) that works only with the potential of a massive halo and one (71) that works only without it.

A subset of X-ray clusters show emission strongly concentrated toward a central dominant galaxy, e.g. a core radius for Virgo of 50 kpc vs. 250 kpc for typical rich clusters. Other members of the subset are Centaurus [NGC 4696 (443)], Perseus [NGC 1275 (204)], and several poorer clusters (381, 430, 434). The strong central concentration suggests that the X-ray gas may belong mostly to the central galaxies rather than to the clusters as a whole.

The standard model of gas shed from stars, radiatively cooling and flowing toward the galactic center, and ways of using these X-ray cooling flows to measure galactic masses are comprehensively reviewed by Sarazin (578, 579). The derived mass is, unfortunately, extremely sensitive to the form of temperature gradient assumed, from which arises most of the present controversy, as well as an earlier one pertaining to M87 (77).

Given the measured Einstein temperature gradient for M87, its total halo mass amounts to \(3-10 \times 10^{12} \,(r/300 \, \text{kpc}) \, M_\odot\) at a distance of 20 Mpc (382, 652, 727). The material concerned is quite dark; beyond 100 kpc, the total \(M/L_B\) value is at least 150 and the local value greater than 500. The
circular velocity at 300 kpc corresponding to this mass is about 600 km s$^{-1}$, double the 300 km s$^{-1}$ of the central galaxy (581). For M87, then, the combination of X-ray data with that from companion galaxies, globular clusters, and stellar velocity dispersions implies $M(R)$ rising linearly with $R$ over the whole range from 0.3 to 300 kpc (579).

Results for the other central dominant galaxies mentioned have larger error bars, but are similarly suggestive of total masses near $10^{13} M_\odot$ and $M/L$'s $\gtrsim 100$. Some of this mass may, however, belong properly to the clusters rather than to the individual central galaxies.

Cooling-flow-type X-ray emission has also been seen from some non-central cluster galaxies, isolated ellipticals, and early-type spirals (227, 490, 680). These are all fainter than M87 by factors of about 100 and are at least as far away, so that even the average temperatures, let alone the gradients, are poorly known and the error bars on the masses very large. Isothermal gas at $10^7$ K (227, 659) or gas with reasonable constraints on the temperature and pressure gradients (206) leads to typical masses of $1-10 \times 10^{12} M_\odot$ and $M/L$'s of $60 \pm 20$ (for $H_0 = 50$). The X-ray gas itself is then not a major contributor toward the total mass. This is distinctly the majority view. It is supported by models of cooling flows as a function of the shape of the gravitational potential confining the gas (580), which predict X-ray surface brightnesses more centrally peaked than the observed ones unless the parent galaxies have heavy extended halos. These halos are hotter than the stellar systems they contain.

A minority view (681) permits temperatures to rise or fall toward galaxy centers and so finds possible $M/L$ ranges that overlap values derived from stellar velocity dispersions. At least one Sa galaxy, the Sombrero, has both a rotation curve (47a) and X-ray emission measured. The lower X-ray mass fits smoothly onto an extension of the rotation curve, while the higher conventional one falls far above it (J. van Gorkom, personal communication, 1986; G. Fabbiano, personal communication, 1986). In this case also, the standard massive halo implied by isothermal X rays would be much hotter than the visible part of the galaxy.

### 2.4 Dwarf Galaxies

The dwarf galaxies are particularly interesting because they have the potential for telling us the smallest configuration that can have a dark halo and thereby constraining the minimum particle mass possible in the halos (677). The situation is somewhat different for dwarf irregulars and dwarf spheroidals, which are discussed separately without any attempt to resolve the vexing issue of whether there is an evolutionary, environmental, or other close relationship between them. It is worth remembering, however, that if you turn off star formation in the SMC for some billions of years,
it will surely become very faint and tend toward $M/L_B = 20$, and that any set of things assumed about halo masses, tidal or wind-driven stripping of gas, chemical evolution, and changes in star-formation rates and $M/L$ had better be self-consistent.

2.4.1 IRREGULAR GALAXIES The irregular galaxies (321) trail continuously downward from the latest spiral types (Sd) and clumpy irregulars with $M_B \leq -20$, through the Sm's or Magellanic types with detectable arms and rotational velocity fields, to low-surface-brightness, truly irregular Irr's and slowly rotating dIrr's (dwarf irregulars) with $M_B \geq -13$.

For representative objects brighter than $M_B = -14.5$, H I rotation curves extend far enough out to demonstrate the existence of large and rising $M/L$'s (dark matter) outside the bright optical regions (120, 121, 231, 259, 389), though as for normal spirals, the optical data alone can be fit by constant, relatively low $M/L$ (661, 662). In the range $M_B = -14.5$ to $-12$, there is still clearly rotation, but it has been traced out only to the optical edge, and the relatively small $M/L$'s pertain to the core and do not constrain dark matter one way or the other. VLA mapping of Irr's in this magnitude range (G. Lake et al., personal communication, 1986) now in progress may reveal whether they share the halo/disk ratio of about one found for brighter Irr's. Finally, below $M_B = -12$, velocity fields show little evidence of rotation, but H I line widths are still large enough to say that $M/L$ is large and the visible stars and gas are probably not self-gravitating. One caveat comes from NGC 1705 at $-15$, which has double optical emission lines split by the H I velocity width. This could plausibly be bipolar outflow rather than bound gas moving in the potential of a dark halo (231).

The blue compact dwarf galaxies are a subset dominated by bright star-forming nuclei but possessing also considerable gas and an underlying old stellar population. Existing data (660–662; T. X. Thuan, personal communication, 1986) pertain largely to the visible regions, and so the fact that the stars and gas account for the Virial mass is not surprising given the similar results for other irregular and spiral galaxy cores.

2.4.2 DWARF SPHEROIDAL GALAXIES These are the most bitterly argued of all. They have such low surface brightnesses that integrated velocity dispersions cannot be measured, and the data comprise only individual velocities for 10–20 stars per galaxy and the six globular clusters of Fornax, plus measurements of tidal radii. Because the data pertain to the central, visible parts, any necessary halos must put a good deal of mass there, so that phase-space constraints suffice to rule out neutrinos or anything else of mass $\lesssim 100$ eV (417). These galaxies are so faint that we have data only for the seven companions of our own Galaxy, but Andromeda has three
or more, and they are undoubtedly the commonest sort, at least in small groups.

The tidal $M/L$'s (201) are $\gtrsim 40$ for Carina and Ursa Minor, $\lesssim 10$ for Sculptor and Draco. Perversely, the qualitative argument for large mass coming from the ability of some dSph's to retain their own globular clusters in the presence of the parent galaxy applies only to Fornax (and the M31 companions NGC 147 and 185), whose dynamical $M/L$ proves to be low; but it is the brightest of the lot. This qualitative argument and masses derived from tidal radii both depend upon the mass assumed for the Milky Way. Thus, if our Galaxy has no massive dark halo, neither have the dwarf ellipticals from this point of view (468a).

The first velocity dispersions came from carbon stars, whose optical velocities are unreliable (344) at about the 6 km s$^{-1}$ level of the dispersions found. Much additional hard work (1–3, 22, 146, 158a, 604) has broadened the data base to include K giants and the Fornax clusters while eliminating radial velocity variables. Velocity dispersions of 6–10 km s$^{-1}$ persist for five of the companion galaxies (Leo I and II remain to be studied) and imply (158a, 231) $M/L$'s near 2 for Fornax, 5 for Sculptor, 10 for Carina (but 15–20 if it had as few young stars as the others), 40–60 for Draco, and 80–100 for Ursa Minor. The largest $M/L$ ratios belong to the least luminous galaxies, and, if they persist in larger data samples, then we have learned something important about the smallest dark configurations and what they can be made of.

3. GALAXIES IN BINARIES AND SMALL GROUPS

The existence of binary stars was proposed by Michell (463) on the statistical basis that close pairs of stars in the sky were far too common to represent accidental superpositions. The conclusion was, however, not widely accepted by natural philosophers of the time until 1803–4, when Herschel succeeded in tracing out portions of several orbits, showing them to be consistent with Newtonian gravitation. Binary galaxies were similarly discovered in a statistical way (311), and confirmation from measured orbits can be expected about the year 1,000,001,987.

In the meantime, galaxies in pairs and small groups present the problem that any particular system for which we might want to attempt a mass determination (including, unfortunately, the Local Group) could turn out to be only a brief encounter or an optical double. Even worse, the parameters used to select a sample of nominally bound systems (separations in space and velocity, $R$ and $V$) are exactly the same ones that dominate calculated masses, in the form mass proportional to $RV^2$. Inevitably, then,
the larger the range of separations admitted in a sample, the larger the masses found.

3.1 The Local Group

The mass of our own small cluster of galaxies has been estimated in three ways, using the Virial theorem (320, 470), the fact that the Milky Way and Andromeda are approaching each other in an expanding universe (293, 346, 418, 428, 574), and the effects of the Local Group on the velocity field of nearby galaxies (574). The first leads to a large, but quite uncertain, value of $M/L$, as do models of the Magellanic Stream (419).

The second consideration was pioneered by Kahn & Woltjer (346), who required that the self-gravitation of the pair suffice to have turned around their initial Hubble separation velocity by the present age of the Universe. The calculation is a pleasant application of classical mechanics, the resulting equations being given in convenient form by Lynden-Bell (428) and by Sandage (574). The chief uncertainties are the amount of time allowed for the turnaround ($6-20 \times 10^9$ yr, depending on choice of $H_0$ and $q_0$) and the present approach velocity. This is the 300 km s$^{-1}$ heliocentric speed of Andromeda (561) minus 80% (geometrical projection factor) of the local Galactic rotation speed of 184–294 km s$^{-1}$ (365, 418, 619), or 65–153 km s$^{-1}$. The most probable range of derived masses is $2-7 \times 10^{12} \, M_\odot$, and $M/L$ is about 100. Appreciable quantities of dark matter can be avoided only if the local circular speed is as improbably high as 290 km s$^{-1}$ (293). If it were, the inner Milky Way, with a flat rotation curve out to 20 kpc, would itself contain $4 \times 10^{11} \, M_\odot$ and have $M/L = 16$.

Sandage (574) has applied the third method and found that the Local Group does, indeed, perturb the nearby velocity field, but only slightly, consistent with a total mass of $4 \times 10^{11} \, M_\odot$ and $M/L$ about 5. The mass thus found is, inevitably, sensitive to choices of $H_0$, $q_0$, circular velocity, and the mass distribution within the Local Group, but is invariably less than the binding mass for the same set of assumptions. That two bright galaxies should be this close together by chance is not horribly improbable (62, 267), but the large relative velocity still has to be accounted for, either through nongravitational processes or through the gravitational effect of some other mass concentration. Random velocities of 100 km s$^{-1}$ could be contained by the potential of the Virgo supercluster at a distance of 18 Mpc if the total mass is as large as $4 \times 10^{13} \, M_\odot$. This does not eliminate the dark-matter problem but only moves it elsewhere!

A bound Local Group of relatively small mass is possible only if one follows Arp (24) in attributing an appreciable fraction of the blueshift of Andromeda to nonvelocity effects, but the absolute minimum for two test
particles falling toward each other inside a single central potential well is only $7 \times 10^{11} M_\odot$.

3.2 Binary Galaxies

The measurement of binary galaxy masses using velocity differences and projected separations was pioneered and persistently pursued by Page (511–513). He recognized most of the problems that are still with us: collecting a large enough sample that the need to average over projection angles does not produce enormous error bars; being sure that the sample includes only bound systems; and deciding the intrinsic eccentricity of the orbits. In addition, his photographically determined velocities typically had errors comparable with the velocity differences being sought, which may be the reason that his pairs of spiral galaxies yielded much smaller values of $M/L$ ($3 \pm 2$) than those in later investigations. Curiously, his results for Es and S0s ($90 \pm 40$) do not disagree.

There now exist a number of additional data sets with much more accurate optical or 21-cm velocity values (79, 354, 400, 525, 600, 688, 742). Once differences in assumed values of $H_0$, orbital eccentricity, and waveband for $L$ have been reconciled, nearly all of these are consistent with $M/L_B = 70 \pm 20 h$ within radii of 100 $h^{-1}$ kpc, assuming circular orbits (500). This drops immediately to $30 \pm 10$ if the orbital velocities are assumed to be distributed isotropically, and still lower for radial orbits (742).

The velocity separations are, on average, larger for nonisolated systems, which might reflect the effects of potentials that belong to larger groups rather than the binary pairs, but $N$-body simulations of galaxy clustering (197) suggest that this is not a problem, implying that galaxies forming in dense regions really are more massive (285). The binary galaxy $\Delta V$s are not much correlated with luminosity (742), while line widths of individual galaxies clearly are. This should not surprise us if the former probe primarily the dark halos and the latter the visible disks of galaxies.

Worrisomely, the one discordant sample (354) with $M/L_B = 7 \pm 1$ (the same as the rotation masses for the same galaxies singly) is the one that most nearly meets a strict criterion for containing only bound systems (706). One cannot, unfortunately, resolve the issue by considering only pairs that show signs of interaction. These indeed yield lower $M/L$'s (21, 706), but the members must be very close together, and so will be inside of, and not probe, most of their shared dark halo according to the conventional view.

3.3 Small Groups

Early cataloguers of groups containing three to tens of comparably bright galaxies (107, 711) typically applied the Virial theorem to their data,
finding that masses near $10^{13} \, M_\odot$ and $M/L$'s near 100 were necessary for binding, while at the same time they expressed severe reservations about both the permanence of the configurations and the appropriateness of the technique.

The most extensive published data set (238) includes 90 groups in the Center for Astrophysics redshift survey and yields a mean $M/L$ of 170 $h$ (317). A number of other observational and theoretical investigations essentially concur (51, 52, 102, 128, 197, 273, 289, 685). Once again, a dissenting analysis of triples (356) finds a mean $M/L$ of only 10, and a test using redshift asymmetries suggests that many of the systems with apparently large $M/L$'s are, in fact, unbound (706), as earlier suggested by Materne & Tammann (439a). For 36 small groups velocity-mapped in H I, the masses implied by the individual galaxy rotation curves are enough to bind the systems if the orbits are circular (594a). These groups have not been statistically tested for boundedness.

Two special cases are poor cD clusters and compact groups. The former (189) typically have velocity dispersions of the companion galaxies not much larger than those of the central galaxies, implying that much of the mass is fairly near the center. The velocity-dispersion masses (60) are sometimes rather larger than the X-ray masses (434), unlike the case of single ellipticals. For the Hickson compact groups, even the visible mass implies awkwardly short dynamical times (750), although once one allows for accidental alignments (437) and more sophisticated dynamical evolution (51, 53, 736), the number of merger products may just about agree with the number of galaxies that could have formed that way (53). The rotation curves of the individual Hickson galaxies typically turn over (i.e. no massive halos) and show other peculiarities (V. C. Rubin, personal communication, 1986).

The situation for all the systems mentioned in the preceding three sections seems to be that dark matter, exceeding that implied by rotation curves and other local measurements by about a factor of 10, is needed to guarantee their stability over a Hubble time, but that there is in every case some alternative possible interpretation of the data. Unfortunately, the only discriminants so far suggested are likely to be believed only by those who already believed a particular answer anyway. A signature for bound but noninteracting systems is badly needed!

4. RICH CLUSTERS, SUPERCLUSTERS, AND GLOBAL CONSIDERATIONS

4.1 Cluster Cores

The clustering of the spiral nebulae was recognized long before their extragalactic nature was universally accepted (296, 754), and a dynamical
study of Virgo (777) was among the first indicators of dark matter in the Universe. Nearly all modern studies concur with this pioneering one in finding cluster mass-to-light ratios considerably larger than those of individual galaxies, even ones with extensive rotation curves. Readers who agree with Fasenko (212) and Zabrenowski (766) that clusters and larger entities merely reflect the effects of nonuniform absorption in our own Galaxy can skip the rest of this section.

Published $M/L$ values range from 30 to 1000, but this narrows considerably when normalized to a single value of $H_0$, face-on galaxies, and a unique waveband for $L$. By way of reminder, $M/L(H = 50) = 1/2M/L(H = 100)$, $M/L_\nu \approx 0.7 M/L_B$, and, probably, $M/L_{\text{bol}} \approx 1/2M/L_\nu$. There is still another factor of two between the Virial mass ($2T + U = 0$) and the just-bound mass ($T + U = 0$). And, finally, $M/L$ drops by yet another factor two when corrected for absorption in our Galaxy and normalized to face-on external galaxies (217, 218; J. E. Felten, personal communication, 1986). It is not always very easy to determine which set of parameters was used in a particular study, but virtually all the very large values found in the literature drop to 100–200 for face-on galaxies, the correct choice of solar $L_\nu$, and $H = 50$ km s$^{-1}$ Mpc$^{-1}$ (J. E. Felten, personal communication, 1986).

Virial $M/L$'s of 100 or more (in the sense of the previous paragraph) have been derived for Virgo (253, 318, 686), Coma (178, 362, 469, 657), Perseus (363), and a number of other rich clusters (95, 239, 507), including one at a redshift of 0.39 (592). Estimators other than the Virial theorem yield concordant, but typically slightly larger, masses (289, 631). Some uncertainty attaches to these numbers owing to faint galaxies in the clusters which cannot be counted directly and are included using an average luminosity function (469, 558). Light from between obvious galaxies or from their extreme outskirts is also rather poorly measured, but almost certainly less than that seen in the galaxies (664). $M/L$ could be changed by a factor of 2–4.

More uncertainty comes from cluster dynamics. If, for instance, $M/L$ is not constant throughout Coma, acceptable fits to measured velocities can be achieved with masses from $6 \times 10^{14} M_\odot$ and circular orbits in the outskirts to $5 \times 10^{15} M_\odot$ and radial orbits far out (461, 658), corresponding to $M/L_B = 40$ to 300 ($H_0 = 50$). If outer parts of clusters are not relaxed, Virial theorem masses could be too large by factors of 3–5 (130, 481). An extreme version of this idea (708) associates dark matter only with central dominant galaxies, the others acting like unrelaxed test particles. Finally, some clusters are certain to be contaminated by foreground and background galaxies that will push up the velocity dispersion and so $M/L$ (219, 239, 522) or to contain substructures that can have the same effect (426).
Even Virgo may be a victim of this latter phenomenon (657, 710). All these considerations serve primarily to widen the error bars on $M/L$, though perhaps more in the direction of low values than high ones.

Potential wells of rich clusters can also be probed via X-ray emission of hot gas in them. The standard assumptions (578) of hydrostatic, isothermal gas lead to $M/L$ values of 100 or more (208, 698), with the usual noise coming from different choices of distance, luminosity scale, etc. Once this noise is eliminated, results from analyses that assume different gas distributions and temperature structures (152) are not terribly discordant, with $M/L \sim 200 \, h$, although the X-ray gas itself contributes as much as 30% of the mass in this case.

One very large value, $M/L = 10^{3.3}$ for the cluster that lenses the QSO 0857+561 (262, 554), is subject to the same caveat as single galaxy lenses— that the process is not yet very well understood (110, 479).

How is this dark matter distributed? Clearly, 100-kpc halos must have blended in cluster cores where the intergalactic distance is small. In addition, halos attached to the individual galaxies would make them hefty enough to segregate, the most massive ones falling to the center. While morphological segregation in clusters is conspicuous (252), luminosity segregation is much less so (118, 236, 306, 758), leading to a majority conclusion that $M/L \leq 30$ for individual cluster galaxies (557, 571, 738). A similar limit comes from the requirement that gravitational drag heating of X-ray gas not be so large as to prevent cooling flows where they are seen (468). The amount of dark matter in rich clusters could, however, easily be the same as would have been associated with the individual galaxies in a less dense environment (227). The likelihood of this depends mostly on how you think galaxies formed, but, in any case, most of the dark matter must now belong communally to the cluster, not individually to the galaxies.

4.2 Superclusters

These larger scale structures, catalogued by Abell (6) and others (46, 609, 614, 709), provide another opportunity for measuring mass-to-luminosity ratios. Results on this and larger scales are, however, usually given in terms of the ratio $\Omega$ of density in a particular component to the density needed to close the Universe, $\rho_c = 3H^2/8\pi G = 2 \times 10^{-29} \, h^2 \, g \, cm^{-3}$. So defined, $\Omega$ coming out of most dynamical analyses is independent of $H$. Since the average luminosity density of the Universe is about $2.4 \times 10^8 \, h \, L_\odot \, Mpc^{-3}$ (217), a component with a particular value of $(M/L)h$ will contribute $\Omega = (M/L)h/1000h$ toward closure.

Necessarily, the values of $\Omega$ found from these large-scale structures can pertain only to volumes where there is at least some luminous material! In
addition, many, though not all, of the cluster and supercluster analyses assume a constant ratio of dark to luminous mass throughout the volume considered. The effect of both these limitations is probably to make derived values of $\Omega$ lower limits to the real value, though the latter could, in principle, go either way.

Abell (7) first noted that binding typical superclusters would require $10^{16-17} M_\odot$ in a 50-Mpc radius. More recent studies of our own Virgo supercluster are countably infinite in number (4, 187, 416, 656, 687, 715, etc.). The measurements are often stated in terms of a "Virgo-centric infall." This does not mean that we are actually moving toward the Virgo cluster, but only that we are moving away less rapidly than in a Hubble velocity field unperturbed by the large mass concentration. Most determinations fall in the range $250 \pm 50$ km s$^{-1}$, implying $\Omega = 0.2 \pm 0.1$ if all superclusters have the same $M/L$.

Other superclusters, if bound structures, also tell us that $\Omega = 0.2 \pm 0.1$ out to scales of 30–50 Mpc (172, 226, 276, 521). Smaller values have been claimed for a few structures (263, 656), and some things that look like superclusters may not be bound (47), but the main remaining controversy seems to be whether we can rule out $\Omega = 1$ only at the 1-$\sigma$ level (113, 456) or with considerably more confidence (304, 410, 521).

Structure on scales of clusters to superclusters can, as an alternative to looking at particular objects, be probed via the correlation function

$$\xi(r) = \xi(r/r_0)^{-n},$$

which expresses the excess probability of finding an object a distance $r$ away from an index object relative to the probability in a random distribution. The method was pioneered by Peebles (517), and the probability of its use shows a large excess at small distances from Princeton. Calculations of $\Omega$ from the correlation function (208a) are sometimes dignified by the name "Cosmic Virial Theorem" and sometimes undignified by the phrase "finding $\Omega$ written on the sky" (mostly by people who say they didn’t).

There is reasonable agreement that $n = 1.8$ for galaxies and about 2.0 for clusters, that $r_0$ is about 10 Mpc for galaxies and considerably larger for clusters, and that there are real differences among kinds of galaxies and among the catalogues that have been analyzed (44, 142, 167, 194, 403, 531, 568, 617). Some disagreement persists about the shape of the function at large distances (353, 534, 560, 651) and the amount of clustering of clusters and superclusters (42, 44, 348, 597). These differences and disagreements should be kept in mind when considering a particular model of galaxy formation that claims to match or not to match "the" observed $\xi(r)$.
Determinations of $\Omega$ from the form of $\xi(r)$ may not be particularly meaningful, in the sense that a number of different initial conditions can relax to what we see by the present time (582). Where the attempt has been made, however, the best fit is about 0.2 (168, 519), the same as found from the superclusters considered as separate entities. In so far as $\Omega$ is written in the sky, it is written small.

4.3 Very Large Scale Structures and Streaming (VLSS)

Structures on scales still larger than superclusters provide an opportunity for probing a larger fraction of the total mass density, and the difficulty of forming them without simultaneously introducing detectable lumpiness into the 3-K microwave background radiation puts very severe constraints on the properties of dark matter participating in galaxy formation.

The rich clusters themselves apparently extend out to $30 \, h^{-1} \, \text{Mpc}$ and so include about a quarter of all galaxies (43, 249). Abell superclusters are clumped on scales up to $150 \, h^{-1} \, \text{Mpc}$ (57). And a number of recent investigations (224, 287, 313, 459) have pointed out bridges, filaments, and connections among well-known superclusters, including Virgo, Perseus-Pisces, Hydra-Centaurus, and Abell 569, 2634, and 2666. The current record is held by Tully's (684) joining of Virgo to Coma to Hydra-Centaurus and beyond, for a structure at least $279 \, h^{-1} \, \text{Mpc}$ across. The most famous hole in the distribution of galaxies, the great void in Bootes (368), is at least $60 \, h^{-1} \, \text{Mpc}$ across (491).

The density contrast between voids (or the largest lumps) and the Universe as a whole is an important boundary condition for formation processes. The Bootes void definitely contains a few low-luminosity, emission-line galaxies (491) and perhaps some gas clouds (105). Other voids in the KOSS (368) and other surveys (377, 404) approach the size of the Bootes one and are not completely empty either. The density contrast could be as low as $\delta \rho / \rho = 2-3$ (easily produced in models of biased galaxy formation) but is more probably about 10, presenting severe difficulties (171).

The topology of the largest structures has been described as chains (342), cells (341), pancakes (772), filaments (254), disks (683), and bubbles (404). The critical question is whether it is the high- or the low-density regions that form a connected network, and one recent statistical investigation (258) finds that both do, so that the topology is most like that of a sponge.

Velocity deviations from uniform Hubble flow on scales of about $100 \, h^{-1} \, \text{Mpc}$ and with amplitudes of $400 \, \text{km s}^{-1}$ or more showed up more than a decade ago (563) and were initially greeted with some scepticism, partly because the implied direction of motion for the Milky Way was roughly orthogonal to that given by the dipole anisotropy of the 3-K
radiation, though the magnitude was about the same (749). Additional work on both spiral (5, 147) and elliptical (112) galaxies has reproduced roughly the magnitude and direction of the early result, but the way of describing the motion has changed. Rather than as deviant motion of the Local Group, it is now perceived (112) as coherent streaming at 600–700 km s\(^{-1}\) of the galaxies in a 100-Mpc diameter region relative to the microwave background rest-frame. The vectorial sum of this streaming and our motion relative to the sample galaxies (= Local Group velocity + motion of Milky Way in LG + rotation of MW at position of Sun) is then the solar motion measured relative to the 3-K radiation.

It is not clear that these large-scale motions have much to do with dark matter, except to suggest the physical reality of very large systems and, therefore, perhaps, of density enhancements 100 Mpc or more across. They are, however, very nearly impossible to account for within the currently population-biased cold dark-matter models of galaxy formation (726a, 739) and could, if confirmed, force the rejection of such models and, therefore, of cold dark matter in general.

In addition, several attempts have been made to calculate \(\Omega\) directly from the largest scale clustering. Lahav (397) finds 0.3 from 15,000 catalogued optical galaxies at an average distance of 50 \(h^{-1}\) Mpc. This differs little (but in the “right” direction) from the 0.2 given by supercluster scales. Finally, two analyses of the distribution of IRAS galaxies (449, 762) yielded \(\Omega = 0.5\) and \(0.85 \pm 0.15\), respectively. The main difference between the IRAS and optical samples is that the former has an average, baseline number density of bright galaxies twice that of the latter (760), so that a given fractional density enhancement contains more total mass. Which is “right” cannot be certainly determined until more of the IRAS galaxies have measured distances (166), but, in the meantime, it is just possible that something very close to the closure density has already been detected, though its form and distribution remain largely unknown.

### 4.4 The Global Value of \(\Omega\)

Given that the known mass-energy in the Universe comes within a factor of about five of the closure density, and that redshifts and apparent magnitudes of distant galaxies limit the effective \(\Omega\) to at most four (257), one might well guess \(\Omega = 1\) exactly. The motivation for this is strengthened when it is recognized (180) that a value within a factor 10 of one now can be achieved only if \(\Omega\) fell within about 1 part in \(10^{15}\) of unity during the epoch of nucleosynthesis \((T \approx 1\) MeV) and 1 part in \(10^{49}\) at the time of the GUT phase transition \((T \approx 10^{14}\) GeV).

The case for \(\Omega = 1\) is enhanced by a model for the evolution of the very early Universe called inflation (100, 271). The general idea is that an early
period of exponential expansion, triggered by a high-temperature phase transition, guarantees that widely separated parts of the Universe were once in communication (and so can reasonably have the same density and temperature now) and $\Omega = 1$ to very high precision [though there are alternative versions of inflation that have most of the standard virtues but permit $\Omega < 1$ (256, 424)]. The whole subject of the relationship between particle physics and cosmology and its implications for values of $H_0$, $\Omega$, and the cosmological constant $\Lambda$ is an exceedingly active one, with archival journal papers often lagging current research by three or four ideas even when they first appear. The interested reader is advised to start with recent conference proceedings (375, 607) and then to haunt the preprint department of his local library.

The other main argument for a critical-density universe arises from the difficulties of forming galaxies and larger structures if $\Omega < 1$ without simultaneously introducing into the 3-K radiation larger inhomogeneities than the present upper limit, $\Delta T / T \lesssim \text{few} \times 10^{-5}$, on the angular scales that correspond to galaxies at the epoch ($Z \sim 10^3$) when matter and radiation decouple (701).

Qualitatively, to make $\Delta \rho / \rho \approx 1$ now requires $\Delta \rho / \rho$ of about $10^{-3}$ at decoupling. (Lumps grow at most linearly with $Z$ in an expanding universe.) This, in turn, imposes fluctuations of $\Delta T / T = (1/3) \Delta \rho / \rho$, which we will still see if the Universe has remained optically thin. These are too big even if $\Omega = 1$. The situation is worse if $\Omega < 1$, so that the galaxies, etc., we see correspond to $\Delta \rho / \rho \gg 1$. On the other hand, $\Delta T / T$ can be greatly reduced if there is a background of nonbaryonic dark matter that does not interact with radiation at $T \lesssim 1 \text{ GeV}$. Then lumps in it can start growing early and reach sizable amplitudes without perturbing the radiation. The baryons begin to follow the dark matter only after decoupling, eventually (because they are dissipative) becoming more clumped than the dark background.

Quantitative versions of this argument are very much more complex, but lead to the same conclusion that galaxy formation is enormously simplified if the Universe is closed with at most weakly interacting dark matter (83, 89, 169, 427, 723, 775), though the very large scale streaming velocities are still a problem (726a, 739). This subject also is in a phase of rapid development and most easily followed through conferences and preprints.

5. INTERMISSION

Most of the authors cited have been convinced of the existence of significant amounts of nonluminous mass. This is not entirely a bandwagon
effect; in addition, if you work hard on something, you want it to be important. Table 2 summarizes the current majority view, but the dark matter needed on many scales can be reduced or possibly eliminated if you are willing to accept some of the following alternative hypotheses.

2. Rotation curves of Milky Way and other spirals: outer gas in non-circular and probably impermanent orbits owing to effects of recent arrival, companions, etc.; luminosity at large radius underestimated because sky background brightness overestimated.
3. Velocities of globular clusters, companion galaxies, and outlying stars of Milky Way and other galaxies: outer high-speed objects not in permanent bound orbits.
4. X-ray emission from elliptical galaxies: gas temperature distribution that declines steeply toward galaxy centers.
5. Dwarf spheroidals: not tidally distorted by Milky Way because Galaxy mass small; stellar velocity dispersion due partly to binary and pulsating stars.
6. Velocity dispersion of stars and globular clusters in galaxies and galaxies in clusters: preponderance of circular orbits at large radius.
7. Binary galaxies: preponderance of radial orbits or isotropic distribution.
8. Small groups: many unbound or bound only as part of larger structures (including Local Group).

There is no way of avoiding sizable quantities of nonluminous mass if the advantages of inflation or some other argument incline you to favor

<table>
<thead>
<tr>
<th>Table 2</th>
<th>Amounts of mass on various scales implied by &quot;mainstream&quot; investigations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scale</td>
<td>$\langle M/L \rangle$</td>
</tr>
<tr>
<td>Visible stars and clusters</td>
<td>1</td>
</tr>
<tr>
<td>Visible parts of galaxies</td>
<td>10</td>
</tr>
<tr>
<td>Binary galaxies and groups</td>
<td>10–100</td>
</tr>
<tr>
<td>Rich clusters and superclusters</td>
<td>100–300</td>
</tr>
<tr>
<td>Largest scale coherent structures</td>
<td>$700 \pm 150? !$</td>
</tr>
<tr>
<td>Inflationary scenario</td>
<td>1000 $h$</td>
</tr>
</tbody>
</table>
\[ \Omega = 1. \] The final section is devoted to a brief discussion of possible kinds
of dark matter and observational constraints on them.

6. OBSERVATIONAL CONSTRAINTS ON THE
NATURE OF DARK MATTER

If one accepts the evidence for nonluminous mass on many scales, the next
obvious question is, what is it made of? A single candidate would be the
most elegant solution, but it is asking a great deal of any one sort of thing
that it be "cold" enough to settle down into the Milky Way disk and
simultaneously "hot" enough to remain much less clustered than the
superclusters over the age of the Universe.

Curiously ordinary, baryonic matter may come closest to meeting this
stringent condition, since it is known to exist in both dense stars and very
diffuse X-ray-emitting gas. The objection to \( \Omega = 1 \) in baryons derived from
nucleosynthetic considerations and possible ways around it are addressed
in Section 6.2. Other candidates currently under consideration include
black holes, neutrinos, and a whole zoo of (mostly hypothetical) particles
predicted by various branches of theoretical high-energy physics. Section
6.4 discusses these.

6.1 The Nonstarters

A few dark-matter candidates (stars and gas) would be so conspicuous
that they can be ruled out easily, and a few others (gravitational radiation
and primordial black holes) would be so inconspicuous that little can be
said about them.

True stars, which derive most of their energy from nuclear reactions,
extend down to \( 0.085 M_\odot \) and \( \lesssim 10^{-3} L_\odot \) (345). Dim though they are, they
could be seen individually in our own Galaxy (161, 248) and collectively in
other galaxies (626) if they were responsible even for the \( M/L = 5-10 \)
imbied by spiral rotation curves.

Gas is similarly well inventoried in all possible forms in the Milky Way
and other galaxies. On intergalactic scales, a closure density of cold or
warm gas would produce conspicuous emission and/or absorption lines
which we do not see (515). Very hot gas would radiate X rays, and indeed
we see a highly isotropic X-ray background well matched by \( 40 \pm 5 \) keV
thermal bremsstrahlung (439), which if produced by hot intergalactic gas
would require something close to the closure density (266). There are two
objections. First, heating the gas requires more than 10% of all available
nuclear energy in the Universe, and there are no very obvious sources (220,
266). Second, after removing the contributions of galaxies, quasars, and

© Annual Reviews Inc. • Provided by the NASA Astrophysics Data System
clusters that are known X-ray sources, the remaining spectrum can no longer be fit by thermal radiation (244).

Gravitational radiation of many but not all possible wavelengths could close the Universe without having yet been detected (19, 556). Clearly it cannot be clustered or clumped, but it can mimic dark matter in clusters in some ways (543). Primordial black holes could also dominate $\Omega$ (126, 480), and the extent to which they might cluster depends upon their (unknown) velocities. Because PBHs must form very early, if at all, they do not count as baryonic matter in the context of nucleosynthesis and are not conspicuous in any other way either.

6.2 *Baryons: The Nucleosynthesis Problem*

The ability of the conventional hot big bang to account for about 25% helium in the matter expanding and cooling from it (284) is generally regarded as one of its great triumphs. Simultaneous minor reactions can also yield small amounts of $H_2$, $He^3$, and $Li^7$, consistent with the quantities observed (728). But the initial conditions must be chosen very carefully. If the baryon density is too high, or the number of neutrino species larger than three (or several other if's), then too much He and not enough $H_2$ come out.

The largest $\Omega$ (baryon) consistent with deuterium production depends on the present abundance [which varies at least from $D/H = 3 \times 10^{-6}$ to $2 \times 10^{-5}$ (717)], on how much has been destroyed by passage through stars [perhaps as much as two thirds of the original supply (143)], and on a number of details of nuclear physics and cosmology (86). Given conventional physics, a standard hot big bang, and generous error bars, $0.015 \leq \Omega h^2 \leq 0.15$ (86, 170).

This nucleosynthesis limit overlaps the $\Omega$ determined from clusters and superclusters, which could, therefore, consist entirely of baryonic matter. It can also be stretched toward one in a variety of ways, most simply by lowering $H_0$ to $\lesssim 25$ km s$^{-1}$ Mpc$^{-1}$ (610), in which case clusters are bound by their X-ray gas. A second, more drastic modification is nonzero lepton number, which changes the equilibrium n/p ratio and can therefore either raise or lower $H_2$ and He production at a given baryon density (228, 703). Third, various inhomogeneities in density, temperature (569a), and n/p ratio can also push production up or down, possibly enough to permit $\Omega = 1$ in baryons. Variations in n/p arising because neutrons can drift out of dense regions in the presence of a magnetic field, but protons cannot, look particularly promising (20, 307). Finally, one might abandon big-bang nucleosynthesis completely and produce deuterium, helium, and perhaps the photon background elsewhere. Although early efforts at this were not terribly successful (549), considerably recent effort has focused
on pregalactic stars, especially supermassive ones of $10^6 \pm 2 \, M_\odot$ that have long since become black holes (90, 126, 359, 398, 540, 616).

In light of these possibilities, it may be premature to rule out baryonic dark matter, even at the level of $\Omega = 1$. Conversely, the idea that dark matter at least up to the scale of galaxies must be baryonic to account for the near-constancy of $\rho(\text{dark})/\rho(\text{luminous})$ (39) is subject to the reservation that formation of the luminous disks might have been directly controlled by the dark halos, or the halos by the disks (202, 209).

6.3 Baryonic Dark Matter: The Candidates (Brown Dwarfs, White Dwarfs, Black Holes)

It is not absolutely certain that we need baryonic dark matter even in the Galactic disk. The process of disk formation could draw inward enough of the hypothetical nondissipative matter of the halo to account for the Oort limit (54, 140). But baryons are clearly a possibility, at least up to $\Omega \sim 0.1$–0.2. How might they be assembled? Normal stars and gas were among the nonstarters in Section 6.1. In addition, even the Oort limit in neutron stars and stellar-remnant black holes can probably be ruled out. First, ongoing accretion of interstellar gas would result in the radiation of more X rays than we see. Second, the stellar evolution and demise that formed them would have scattered into the interstellar gas far more heavy elements than are present (288, 447). The additional upper limit of $2 \, M_\odot$ on disk dark objects based on the long-term survival of wide binary systems (40) is probably a bit less stringent than the two already mentioned (733).

Brown dwarfs are substellar objects whose only energy source is contraction. It is not certain that any (apart from the Jovian planets) have ever been detected (627). A number of searches have identified no candidates (345). On the other hand, at least one survey (283) found interesting numbers of very red (presumably cool, faint) images, and at least one set of models for the evolution of low-mass stars and substars (163) implies that the known faint stars are part of a population whose number is still rising as $M^{-1.68}$ at $0.1 \, M_\odot$. In addition, a separate population of objects $\lesssim 0.08 \, M_\odot$ as old as the Galaxy could not yet have been seen in any way, and might be expected to arise from star formation in gas under pressure (205) or in pregalactic objects of $10^6 - 8 \, M_\odot$ (359). The upturn of star numbers in M13 for star masses below $0.5 \, M_\odot$ (427a) is interesting in connection with the possibility of a separate, low-mass population.

For other galaxies, we can say very little about the initial mass function of small stars (171a, 585), and it could continue to rise smoothly, putting large amounts of material into faint stars and substars. Brown dwarfs in sufficient numbers to be dynamically important might show up in deep exposures taken with the Wide Field/Planetary Camera of the Hubble
Space Telescope (640) or in the infrared flux from galaxies at moderate redshift (639).

White dwarfs are the normal remnants of 0.5–8 \( \pm 2 \) \( M_\odot \) stars and can fade below detectability in less than the age of the Galaxy (163). Even so, the ones remaining from stars produced at constant initial mass function and formation rate over the history of the Galaxy contribute at most 0.007 \( M_\odot \) \( \text{pc}^{-3} \) locally (163, 221), only about 5% of the Oort limit. One model of galactic chemical evolution does, however, posit an early generation of intermediate-mass stars whose white-dwarf remnants might account for the dynamical mass in the Milky Way disk and, probably, the halo as well (406, 493). This early generation would have had to be even more prominent to contribute significantly on the scale of clusters and superclusters. Since these very old degenerate dwarfs could be as cool as the hypothetical brown dwarfs and will surely be much smaller, they will be exceedingly difficult to detect or rule out.

Black holes of stellar masses have already been noted as nonstarters, and very small primordial ones are essentially unconstraining. This leaves massive ones that might have formed in the early Universe (126, 127, 398). If these make up the halo dark matter, then, in addition to having contributed (perhaps) helium and photons in the past, they will now be stirring up the disk stars and contributing to the known increase of stellar velocity dispersion with age that is generally blamed on giant molecular clouds (393) or spiral arms (122). Black holes that are too massive will be too efficient at this, but a \( 10^{12} \) \( M_\odot \) halo consisting of \( 10^6 \) \( M_\odot \) objects is just right (336, 350, 394). If similar black holes make up the dark matter in dwarf ellipticals, then they must be clustered towards the center, predicting an outward decrease of stellar velocity dispersion as a test of the model (394). In addition, \( 10^6 \) \( M_\odot \) black holes in halos should reveal themselves by gravitationally lensing radiation from objects behind them (504, 509). Resolution of radio components may be possible with (VLBI), but optical observations will have to wait for the space optical interferometer.

By way of summary, baryonic matter definitely cannot be ruled out and even has some advantages at least up to the \( \Omega \sim 0.15 \) level, consistent with conventional nucleosynthesis. If one takes very seriously the nucleosynthetic lower limit on \( \Omega_b \) or the need for dissipative material in galactic disks, then at least some of the nonluminous mass must be baryonic.

### 6.4 Nonbaryonic Dark Matter

One’s first impression is that this category includes a countably infinite number of indistinguishable, hypothetical entities. There is some truth in this. But a count of words mentioned in three recent reviews of particle
physics and cosmology (492, 532, 690) uncovered only 36 names, not quite all of which in fact designate physically distinct entities. These can be classified in terms of their predicted masses and the theories that predict them and by their astrophysical contribution as dark-matter mimics, hot dark matter (relativistic when galaxies form and so promoting large-scale structure), or cold dark matter [nonrelativistic when galaxies form and so promoting small-scale structure (89)].

Table 3 attempts such a classification. For most of the entities, the requirement that $\Omega(\text{baryon})/\Omega(\text{dark}) \sim 0.1$ constitutes a new form of fine tuning, replacing that required to get $\Omega$ close to but not exactly one. The ratio may arise naturally for quark nuggets (753), but it must otherwise require a particular value of the energy scale of a symmetry breaking or some other process (690).

6.4.1 THE DARK MATTER MIMICS—$G(R)$ AND $\Lambda$

A gravitational coupling constant that increases monotonically with separation or varies suitably with acceleration would mean that the amount of matter present in luminous galaxies and parts of galaxies could act like a larger amount of dynamical mass and suffice to account for flat rotation curves, large velocity dispersions, etc. (466, 576, 666). There are theoretical (245) and experimental (482) indications that $G$ may indeed vary, though probably on distance scales much smaller than those required to simulate dark matter. Although many details of such models remain to be worked out (216, 577), preliminary results are that they do not match observations as well as dark-matter models (188, 295). Their main virtue is that of producing the required effect over a range of distance scales.

A nonzero cosmological constant $\Lambda$, on the other hand, acts like the zero-point energy of a quantum field theory (771) and contributes homogeneously. It also permits the curvature of space $k$ and the deceleration parameter $q_0$ to become independent parameters, so that most of the standard cosmological tests (apparent magnitude or angular diameter vs. redshift, source counts, etc.) do not really tell us $\Omega$, but only some combination of $\Omega$ and $\Lambda$ or $k$ or $q_0$, even when authors (422) hope they are measuring $\Omega$. To make $q_0 = 1/2$ or $k = 0$ with $\rho < \rho_c$ takes $\Lambda$ of order $\pm 10^{-35}$ s$^{-2}$. This is also a sort of fine-tuning problem, at least in an inflationary universe, where $\Lambda$ is briefly enormous (100). Astrophysicists have debated the likelihood of nonzero $\Lambda$ for years (339, 446, 665) without reaching any definite conclusion. It shares with most "real" dark matter the virtue of making galaxy formation a bit easier (724), but cannot solve the dark-matter problem on all scales, being unable to cluster. This objection need not apply to the zero-point energy of a real field theory (J. Weber, personal communication, 1986).
Table 3  Summary of nonbaryonic dark matter candidates

<table>
<thead>
<tr>
<th>Candidate/particle</th>
<th>Approximate mass</th>
<th>Predicted by</th>
<th>Astrophysical effects</th>
</tr>
</thead>
<tbody>
<tr>
<td>$G(R)$</td>
<td>—</td>
<td>Non-Newtonian gravitation</td>
<td>Mimics DM on large scales</td>
</tr>
<tr>
<td>$\Lambda$ (cosmological constant)</td>
<td>—</td>
<td>General relativity</td>
<td>Provides $\Omega = 1$ without DM</td>
</tr>
<tr>
<td>Axion, majoron, goldstone boson</td>
<td>$10^{-5}$ eV</td>
<td>QCD; PQ symmetry breaking</td>
<td>Cold DM</td>
</tr>
<tr>
<td>Ordinary neutrino</td>
<td>10–100 eV</td>
<td>GUTs</td>
<td>Hot DM</td>
</tr>
<tr>
<td>Light higgsino, photino, gravitino, axino, sneutrino</td>
<td>10–100 eV</td>
<td>SUSY/SUGR</td>
<td>Hot DM</td>
</tr>
<tr>
<td>Para-photon</td>
<td>20–400 eV</td>
<td>Modified QED</td>
<td>Hot/warm DM</td>
</tr>
<tr>
<td>Right-handed neutrino</td>
<td>500 eV</td>
<td>Superweak interaction</td>
<td>Warm DM</td>
</tr>
<tr>
<td>Gravitino, etc.</td>
<td>500 eV</td>
<td>SUSY/SUGR</td>
<td>Warm DM</td>
</tr>
<tr>
<td>Photino, gravitino, axino, mirror particle, simpson</td>
<td>keV</td>
<td>SUSY/SUGR</td>
<td>Warm/cold DM</td>
</tr>
<tr>
<td>neutrino b</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Photino, sneutrino, higgsino, gluino, heavy neutrino</td>
<td>MeV</td>
<td>SUSY/SUGR</td>
<td>Cold DM</td>
</tr>
<tr>
<td>Shadow matter</td>
<td>MeV</td>
<td>SUSY/SUGR</td>
<td>Hot/cold (like baryons)</td>
</tr>
<tr>
<td>Preon</td>
<td>20–200 TeV</td>
<td>Composite models</td>
<td>Cold DM</td>
</tr>
<tr>
<td>Monopoles</td>
<td>$10^{16}$ GeV</td>
<td>GUTs</td>
<td>Cold DM</td>
</tr>
<tr>
<td>Pyron, maximon, perly pole, newtorites,</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Schwarzschild</td>
<td>$10^{19}$ GeV</td>
<td>Higher-dimension theories</td>
<td>Cold DM</td>
</tr>
<tr>
<td>Supersymmetric strings</td>
<td>$10^{19}$ GeV</td>
<td>SUSY/SUGR</td>
<td>Cold DM</td>
</tr>
<tr>
<td>Quark nuggets, nucleartes</td>
<td>$10^{15}$ g</td>
<td>QCD, GUTs</td>
<td>Cold DM</td>
</tr>
<tr>
<td>Primordial black holes</td>
<td>$10^{15–30}$ g</td>
<td>General relativity</td>
<td>Cold DM</td>
</tr>
<tr>
<td>Cosmic strings, domain walls</td>
<td>$10^{8–10}$ $M_\odot$</td>
<td>GUTs</td>
<td>Promote galaxy formation, but cannot contribute much to $\Omega$</td>
</tr>
</tbody>
</table>

*Abbreviations: DM, dark matter; QCD, quantum chromodynamics; PQ, Peccei & Quinn; GUTs, grand unified theories; SUSY, supersymmetric theories; SUGR, supergravity; QED, quantum electrodynamics.

*bOf these various supersymmetric particles predicted by assorted versions of supersymmetric theories and supergravity, only one, the lightest, can be stable and contribute to $\Omega$, but the theories do not at present tell us which one it will be or the mass to be expected.
6.4.2 HOT DARK MATTER AND GALAXY FORMATION All constraints derived from models of galaxy formation should be accepted with some caution, because a number of additional, unrelated parameters and processes undoubtedly contribute, and their effects are not always separable. Among these are (a) the spectrum of the initial perturbations; (b) the possibility that galaxies do not trace mass but form only at 2–3 \( \sigma \) peaks [biasing (50, 546, 623, 757)]; (c) gravitational clustering and relaxation after formation (155, 655), some kinds of which can mimic biasing (305); and (d) nongravitational mechanisms for galaxy formation including the selective shadowing of radiation pressure by dust (309) and gas-dynamical effects of exploding stars, supermassive objects, or active nuclei that pile up gas (330). All of these complicate the extrapolation back from observed galaxies to the nature of the underlying dark matter.

Nevertheless, the hot/cold distinction is an important one because in an expanding, cooling universe dominated by particles of mass \( m_\text{p} \), the Jeans mass is

\[
M_J = 3 \times 10^{18} M_\odot / m_\text{p} \text{ (eV)}
\]

(92, 773). Thus, with a dominant particle of 10–100 eV (HDM), supercluster-sized structures will acquire their identity first and later fragment into galaxies, while masses of MeV, GeV, or more will lead to galaxies or smaller structures forming first and larger things being built up by gravitational clustering (CDM). Observations of amount of clustering as a function of redshift ought eventually to be able to tell us which happened (343) but have not yet done so.

Hot dark matter was the first to be considered and has several virtues. First, one sort is actually known to exist—the neutrinos and antineutrinos of electron, muon, and (presumably) tau flavors—which in most modern pictures of symmetry breaking should have some rest mass (690). Second, small extrapolations back in time of the hot big-bang conditions known from nucleosynthesis permit a fairly precise calculation of their number density at about 100 cm\(^{-3} \) for each species. Thus, the rest masses must be 10–100 eV if they are to add up to \( \Omega = 1 \) (154, 243). Other (hypothetical) particles in the same mass range do not share these virtues, but will behave in the same way during galaxy formation.

Three different experimental groups have reported evidence for neutrino rest masses in the cosmologically interesting range (423, 550, 625). Each has in turn been firmly doubted (16, 64, 85, 164, 421), and the case must currently be regarded as not proven. One of the (many) suggested resolutions of the solar neutrino problem (70) invokes a rest mass that might also fall in the interesting range.

Simulations of galaxy formation in a universe dominated by HDM have
been carried out by many groups, first enthusiastically (596), next with the realization that they make galaxies too late \(Z \lesssim 3\) and with velocity dispersions in the small-scale structure much larger than we see (135, 233, 740), and then in a spirit of “save the HDM” (455, 612). The current situation can probably be summarized by saying that hot dark matter is exceedingly useful in producing large-scale structure, including perhaps even the very large-scale streaming (153, 183, 234, 457, 700), and may have some part in scenarios with decaying dark matter or two “tooth fairies” (see Sections 6.4.4 and 6.4.5), but that it cannot be the only nonbaryonic component present at the time of galaxy formation.

6.4.3 COLD DARK MATTER A large fraction of the candidates listed in Table 3 come under this heading and are collectively called “ino’s,” WIMPs (weakly interacting massive particles), or just cold dark matter. Axions! are formed cold (690) and so belong here despite their small masses. It was recognized from the beginning that CDM scenarios would require something like biasing to provide correlations over large scales (270). If galaxies form only at several-sigma density peaks, then they will be more clustered than the underlying dark matter, because most of these peaks will be small fluctuations sitting atop larger scale, smaller amplitude ones. The biased cold dark matter program has now been explored in considerable detail and has achieved a number of successes in accounting for the observed properties of galaxies and clusters (50, 83, 84, 123, 169, 328, 539, 548, 570, 586, 634, 743, 765). There are minor disagreements about how best to do the calculations and interpret them (629), but the chief problem is in accounting for the largest scale voids and superclusters and, especially, for the very large-scale streaming motions (725, 726a, 739).

Because HDM gets into trouble with small-scale phenomena and CDM with large-scale ones, a natural thought is to try to combine their virtues in some way. This has been attempted through the assumption of massive particles that decay to relativistic ones (Section 6.4.4) and through the positing of two or more kinds of dark matter (Section 6.4.5).

6.4.4 DECAYING DARK MATTER The general idea here is that the Universe came out of its inflationary epoch with \(\Omega = 1\) in some unstable WIMP that helped to make galaxies and then decayed away, leaving \(\Omega = 1\) total in the decay products but only 0.2 in bound structures. There must be no photons in the decay products, thus the WIMPs cannot be photinos (495), but gravitinos and heavy neutrinos are possibilities. Calculations of this type are numerous (185, 240, 303, 326, 496, 583, 647, 691, 703a) and reasonably successful in making galaxies without disturbing the 3-K background. They require, however, fine tuning of the decay epoch, make the look-back age of the Universe uncomfortably short (since \(R \propto t^{1/2}\) in a
radiation-dominated universe), and must leave at least half the initial dark mass of a spiral galaxy in the halo in order to reproduce flat rotation curves (223). The advantages, therefore, probably do not outweigh the disadvantages.

6.4.5 SCENARIOS WITH TWO TOOTH FAIRIES  By way of explanation, when an American child loses a milk (baby) tooth, he is supposed to put it under his pillow at bedtime. The tooth fairy comes during the night, takes away the tooth, and leaves a suitable coin in its place. The hope in these scenarios (595) is that we can put dark matter under our computers at quitting time and find that a good fairy has left galaxies by morning, except that in most of the schemes currently under investigation, it really takes two of them.

Biased CDM belongs to this category, unless the biasing comes automatically out of the way the perturbations grow (548). Percolation of explosive galaxy formation plus HDM is another possible combination (133). More popular, however, are the combination of (a) one hot and one cold entity, which together with the baryons add up to $\Omega = 1$ (9, 195, 211, 608, 699, 716), (b) one dark matter candidate plus nonzero cosmological constant (532), or (c) a dark matter background in which the seeds for galaxy formation are quark nuggets, primordial black holes, or cosmic strings (774). The strings, at least, automatically introduce large-scale correlations or biasing and are the 1986 "best-buy" model (12, 68, 255, 308, 547, 548, 643, 692, 718). It is, however, a safe bet that, by the time you read this in 1987 or later, some other combination will seem at least as likely to leave realistic galaxies under the computer.

6.4.6 DETECTION OF DARK MATTER CANDIDATES—ASTROPHYSICAL METHODS  Existing or proposed observations can constrain, or possibly provide evidence for, the presence of several possible kinds of DM. Clearly the two are not entirely distinct; an effect attributable, for instance, to photino decay either proves the existence of photinos or sets an upper limit to them, depending on the likelihood of other causes of the same effect.

The excluded candidates are those with combinations of mass, lifetime, and cross section that would more than close the Universe, spoil nucleosynthesis, or produce a larger background of photons than we see in some waveband (86, 91, 268, 374, 590, 624, 649, 726). No whole class can be eliminated in this way.

Prospects for future detection include baryons (gas, faint galaxies, supermassive black holes, etc.) in voids whose presence would favor biasing and whose absence would favor strings as the mechanism for producing large-scale coherent structure. Gamma rays from WIMP decays or annihilations in the Milky Way halo should also eventually be measurable (565, 637).

Of the observations that could indicate detection of DM candidates has
already occurred, some gravitational lenses and a feature near the center of our own Galaxy have been blamed on strings (141, 510), a possible 1667-Å feature in the ultraviolet radiation background has been attributed to light neutrino decay (644), and photons at higher energies attributed to decay or annihilation of more massive WIMPs (601–603). These could also be responsible for the unexplained high flux of low-energy cosmic-ray antiprotons (646); and sneutrinos might cool white dwarfs faster than conventional energy-transport mechanisms (472).

The most far-reaching of these "we may already have seen them" proposals is the simultaneous use of WIMPs to close the Universe, reduce the solar neutrino flux to the observed value, and adjust the frequencies of solar normal modes to match data (215, 241, 246, 532, 650). The particles must necessarily be trapped in cores of other stars as well and influence their structure and evolution, in ways that remain to be worked out but could be either good (J. Faulkner, personal communication, 1986) or bad (A. Renzini, personal communication, 1986) from the point of view of bringing theory and observation into accord.

6.4.7 DETECTION OF DARK MATTER CANDIDATES—LABORATORY METHODS
As in the astrophysical case, some volumes of parameter space can be ruled out on the basis of existing experiments (10, 177, 236a). All positive results so far reported—neutrino rest mass (550), axions (756), and monopoles (119)—have been questioned and need further work for their confirmation.

Future prospects include both the use of existing proton-decay apparatuses to look for high-energy neutrinos from WIMP decay (379) and the construction of new, dedicated apparatus. Some of these, like a proposed large, high-Q microwave cavity to detect axions through their propensity to convert to photons in the presence of a strong magnetic field (620), carry price tags in the modern high-energy physics range. Two others, germanium or silicon spectrometers (32, 732) and bolometers (115, 190, 232), that detect WIMPs through their propensity to interact with nuclei and deposit energy in superconducting grains, could register roughly one count per day with a 10-kg detector and a cost in the individual PI grant range. These are, of course, enormously exciting possibilities, although we confess that we have not yet started urging our friends and relations to build suitable detectors.

7. L’ENVOI
Is there a dark-matter problem? Not necessarily, though there clearly are a number of astrophysical problems to which different kinds of dark matter
are among the possible solutions. There are also astrophysical problems [e.g. the Great Red Spot on Jupiter (19a)] to which dark matter is not a solution.

No one kind of dark matter, with the remotely possible exception of ordinary baryons, can solve all the problems at once. On the other hand, to invoke five or six different kinds to match the properties of spiral disks and halos, ellipticals, dwarf galaxies, clusters, superclusters, and galaxy formation is probably also the wrong strategy. How are we betting? At least even on baryons on scales up to superclusters, and maybe one chance in four on baryons all the way (either closing the Universe or managing to exist in an open one). The remaining probability spreads rather uniformly over a very large number of candidates at the moment. Several different events (confirmation of neutrino oscillations, laboratory detection of 10-GeV photinos, or whatever) could collapse the wave function considerably.

Finally, there are well-defined, finite, observational, theoretical, and experimental programs that can be expected to improve our understanding of the amount, distribution, and nature of nonluminous mass. These range from studying the brightness of K giants as a function of metal abundance and looking for baryonic matter in cosmic voids to pursuing higher dimensional theories of particle physics and building superconducting microwave cavities. There is surely at least one task suitable for every scientist who is interested in the problem. Failing this, feel free to keep in mind this thought from James Russell Lowell ("A Fable for Critics"): "Nature fits all of her children with something to do. He who would write, but can't write, can surely review."

ACKNOWLEDGMENTS


Literature Cited

42. Barnes, J. 1984. MNRAS 208: 873
76. Binney, J. 1986. See Ref. 369, p. 303
80. Blandford, R. D. 1986. See Ref. 654a, p. 359
81. Blitz, L. 1983. See Ref. 25, p. 43
110. Burke, B. 1986. See Ref. 654a, p. 517
125. Carney, B. W., Latham, D. 1986. See Ref. 369, p. 39
166. Davis, M. 1986. See Ref. 369, p. 97
218. Felten, J. 1986. See Ref. 369, p. 111

© Annual Reviews Inc. • Provided by the NASA Astrophysics Data System
1987 ARA&A 25 425

265. Gunn, J. E. 1986. See Ref. 369, p. 537

© Annual Reviews Inc. • Provided by the NASA Astrophysics Data System
385. Kruij, P. van der 1986. See Ref. 369, p. 415
426. Lukash, V. N. 1986. See Ref. 369, p. 379
448. Melott, A. 1986. Preprint (Some Like It Hot)
455. Millington, S. J. C., Peach, J. V. 1986. MNRAS 221: 15
474. Olive, K. A. 1986. See Ref. 375
481. Ostriker, J. P. 1986. See Ref. 369, p. 85
Ap. J. 243: 1
Nature 314: 718
Astron. J. 88: 909
MNRAS 198: 1P
Phys. Lett. 121B: 119
Phys. Lett. 137B: 169
Astron. J. 90: 1796
MNRAS 217: 127
Phys. Rev. Lett. 53: 1292
Astron. J. 59: 285
Vistas Astron. 28: 595
Ap. J. Suppl. 57: 205
612. Shapley, H. 1933. 
Astrophys. Space Sci. 123: 103
Ap. J. Suppl. 57: 77
MNRAS 194: 851
Phys. Rev. Lett. 51: 1415
Ap. J. 297: 1
Science 235: 1367
Phys. Rev. Lett. 54: 1891
623. Slipher, V. M. 1914. 
Lowell Obs. Bull. 2: 65
Sov. Astron. AJ 29: 367
MNRAS 211: 911
Phys. Rev. Lett. 56: 263
Ap. J. 308: 697
Astron. Astrophys. 98: 140
Ap. J. Lett. 303: L21
Phys. Rev. Lett. 45: 1460
Nucl. Phys. B 252: 73
Astron. J. 83: 1050
MNRAS 213: 953
Ap. J. 278: 536
Astron. J. 90: 1481
Quasars, IAU Symp. No. 119. Dordrecht: Reidel
Nature 314: 718
Astron. J. 92: 1248
MNRAS 220: 949
Ap. J. 205: 696
Ann. NY Acad. Sci. 302: 423
653. Tohill, J. E. 1983. See Ref. 25, p. 205
Ap. J. 266: 58
Ap. J. 279: 13
Astron. J. 90: 2431
Ap. J. 291: 45
690. Turner, M. S. 1986. See Ref. 369, p. 445
751. Williams, R. E., Christiansen, W. A. 1984. ESO Preprint 322
760. Yahil, A. 1986. See Ref. 432