SLAC – PUB – 3387
July 1984
(T/E/AS)

DARK MATTER, GALAXIES, AND LARGE SCALE STRUCTURE IN THE UNIVERSE*

JOEL R. PRIMACK

Stanford Linear Accelerator Center
Stanford University, Stanford, California, 94305

and

Santa Cruz Institute of Particle Physics,
University of California, Santa Cruz, CA 95064

Lectures presented at the
International School of Physics "Enrico Fermi"
Varenna, Italy, June 26 – July 6, 1984

* Work supported by the Department of Energy, contract DE – AC03 – 76SF00515.

ABSTRACT

These lectures aim to present an essentially self-contained introduction to current research on the nature of the cosmological dark matter and the origin of galaxies, clusters, superclusters and voids. The first lecture reviews the observational data and introduces a tentative theoretical framework within which it can be interpreted: gravitational collapse of fluctuations as the origin of structure in an expanding universe. The second lecture summarizes general relativistic cosmology, reviews the data on the basic cosmological parameters (t_o, H_o, and Ω_o), and introduces the theory of the growth and collapse of fluctuations. It also includes a brief exposition of the idea of cosmological inflation, and a briefer critique of a proposal to modify gravity as an alternative to dark matter.

The third and fourth lectures are about dark matter. Arguments that it is nonbaryonic are summarized, and the standard astrophysical classification of varieties of dark matter is introduced: *hot* (free streaming erases all but supercluster-size fluctuations), *warm* (free streaming erases fluctuations smaller than large galaxies), and *cold* (free streaming is cosmologically unimportant). The various particle physics candidates for dark matter are reviewed, together with possible tests that could constrain or eliminate them. Given a primordial spectrum of fluctuations, perhaps generated during an epoch of inflation, the subsequent evolution of this spectrum depends mainly on the free streaming length and on whether the fluctuations are adiabatic or isothermal. This evolution is discussed in some detail, both in the linear ($\delta\rho/\rho \ll 1$) and nonlinear regimes.

There appear to be several serious problems with hot (neutrino) dark matter, while the problems of accounting for cosmological observations with cold dark matter are apparently largely resolved if galaxies form only around unusually large fluctuations in the density ("biased" galaxy formation). Moreover, cold dark matter with a "Zeldovich" spectrum of primordial adiabatic fluctuations appears to lead to an attractive theory for galaxy and cluster formation.

Table of Contents

The standard theory of cosmology is the Hot Big Bang, according to which the early universe was hot, dense, very nearly homogeneous, and expanding adiabatically according to the laws of general relativity. This theory nicely accounts for the cosmic background radiation, and accurately predicts the abundances of the lightest nuclides. It is probably even true, as far as it goes; at least, I will assume so here. But as a fundamental theory of cosmology, the standard theory is seriously incomplete. One way of putting this is to say that it describes the middle of the story, but leaves us guessing about both the beginning and the end.

Galaxies and large scale structure — clusters of galaxies, superclusters and voids — are the grandest structures visible in the universe, but their origins are not yet understood. Moreover, there is compelling observational evidence that most of the mass detected gravitationally in galaxies and clusters is dark — that is, visible neither in absorption nor emission of any frequency of electromagnetic radiation.

Explaining the rich variety and correlations of galaxy and cluster morphology will require filling in much more of the history of the universe:

- *Beginnings,* in order to understand the origin of the fluctuations which eventually collapse gravitationally to form galaxies and large scale structure. This is a mystery in the standard expansionary universe, because the matter which comprises a typical galaxy, for example, first came into causal contact about a year after the Big Bang. It is very hard to see how galaxy-size fluctuations could have formed after that, but even harder to see how they could have formed earlier.

- *Denouement,* since even given appropriate initial fluctuations, we are far from understanding the evolution of clusters and galaxies, or even the origins of stars and the stellar initial mass function.

- And the mysterious *dark matter* is probably the key to unravelling the plot since it appears to be gravitationally dominant on all scales larger than the cores of galaxies. The dark matter is therefore crucial for understanding

the evolution and present structure of galaxies, clusters, superclusters and voids.

Most reviews of cosmology have until recently concentrated on explaining the Hot Big Bang, especially primordial nucleosynthesis. With the advent of grand unified theories (GUTs) in particle physics, and especially the lovely idea of cosmic inflation, it has also become possible to give an account of the very early universe which is at least coherent, if not yet very well grounded observationally.

The present lectures take a different approach, emphasizing the period *after* the first three minutes, during which the universe expands by a factor of $\sim 10^8$ to its present size, and all the observed structures form. This is now an area undergoing intense development in astrophysics, both observationally and theoretically. It is probably now ripe for major progress. It is not impossible that the present decade will see the construction at last of a fundamental theory of cosmology, with perhaps profound implications for particle physics as well.

Although I will concentrate in these lectures on the development of galaxies and large scale structure in the relatively "recent" universe, I can hardly avoid retelling some of the earlier parts of the story. Primordial nucleosynthesis will be important in this context primarily as a source of information on the amount of ordinary ("baryonic") matter in the universe; GUT baryosynthesis, for its implication that the primordial fluctuations were probably adiabatic; and inflation, for the constant curvature ("Zeldovich") spectrum of fluctuations and a plausible solution to the problem of generating these large scale fluctuations without violating causality. I will be especially concerned with evidence and arguments bearing on the astrophysical properties of the dark matter, which can also help to constrain possible particle physics candidates. The list of these now includes ~ 30 eV neutrinos, very massive right handed neutrinos, other heavy stable particles such as photinos, massive unstable neutrinos or their decay products, very light "invisible" axions, u-d-s symmetric "quark nuggets", and primordial black holes. One of these hypothetical species may be the dominant form of matter in

the universe — or perhaps it is something no one has even thought of yet!

I will begin by discussing the basic astronomical data on the distribution of matter in the universe: galaxies, clusters, superclusters and voids, and the strong evidence that all the visible matter on galaxy scales and larger is moving in the vast potential wells of the gravitationally dominant dark matter. If this is so, the inevitable question is how these enormous ghostly structures formed.

To prepare to discuss the answers that have been proposed, I will need to review the theory of gravity, not merely standard general relativistic cosmology, but also the theory of the growth and collapse of fluctuations in an expanding universe. Learning the basic theory of gravitational collapse — including "virialization" by "violent relaxation" — was a revelation to me, and it has been my experience that it is not generally appreciated outside astrophysics. Under the rubric of gravity theory, I will also discuss briefly the idea of cosmic inflation and its implications for the origin of fluctuations. And I will discuss even more briefly some recent suggestions of modified gravity, with a r^{-1} force law at large distances, as an alternative to dark matter.

Next comes the most conventional part of these lectures, describing the standard Hot Big Bang: decoupling, nucleosynthesis, recombination. This provides the essential background for the three astrophysical arguments that the dark matter is probably not baryonic: excluding various possible forms of baryonic dark matter in galaxy halos, bounding the abundance of baryonic matter using the observed deuterium abundance, and bounding the magnitude of adiabatic fluctuations at recombination from the obserational upper limits on fluctuations in the cosmic background radiation. (I will also point out explicitly the loopholes in each of these arguments.)

Finally I take up the key question: what is the dark matter that the universe is mostly made of? From the viewpoint of astrophysics, it is useful to categorize the dark matter as *hot, warm,* or *cold,* depending on its thermal velocity compared to the Hubble flow (expansion). Hot dark matter, such as ~ 30 eV neutrinos, is still

relativistic when galaxy-size masses ($\sim 10^{12} M_\odot$, where $M_\odot = 2.0 \times 10^{33}$ g is the mass of the sun) are first encompassed within the horizon. Warm dark matter is just becoming nonrelativistic then. Cold dark matter, such as axions or massive photinos, is nonrelativistic when even globular cluster masses ($\sim 10^6 M_\odot$) come within the horizon. As a consequence, fluctuations on galaxy scales are wiped out with hot dark matter but preserved with warm, and all cosmologically relevant fluctuations survive in a universe dominated by cold dark matter.

The first possibility for nonbaryonic dark matter that was examined in detail was massive neutrinos, assumed to have mass $\sim$ 30 eV — both because that mass corresponds to closure density, and because the Moscow tritium β-decay experiment continues to provide evidence that the electron neutrino has that mass. Although this picture leads to superclusters and voids of the size seen, superclusters are the first structures to collapse in this theory since smaller size fluctuations do not survive. The theory founders on this point, however, since galaxies are almost certainly older than superclusters. A related problem is that galaxy and cluster formation is sufficiently complicated in the neutrino picture that no theory of it has yet been worked out.

A currently popular possibility is that the dark matter is cold. I have been one of those who have been studying the consequences of this picture. Its virtues include an account of galaxy and cluster formation that appears — at least to me and my coworkers — to be very attractive. Its defects are less clear, perhaps at least partly because it has not yet been subjected to enough critical scrutiny. Some recent work suggests that the size of the large scale structure in a cold dark matter universe will come out right only if the density is not more than about half the critical density, but this is contrary to prejudice, the inflationary hypothesis, and the latest upper bounds on small-angle fluctuations in the microwave background radiation. Another problem with hot as well as cold dark matter is understanding the strong correlations in the locations of rich (i.e., populous) clusters of galaxies across tremendous distances, large even compared to the scale of superclusters.

These lectures end with a survey of new ideas for solving these problems, new sources of observational data which may differentiate more clearly between the various possibilities for the dark matter, and finally some possible broader implications of the picture that is emerging from particle physics and cosmology of the structure of the universe on both the smallest and largest scales.

1. Matter

1.1 SIZES

This lecture is mainly about the distribution of matter in the universe on galaxy and larger scales, and the evidence that most of the mass is dark. But I think it may be useful to provide a little orientation about sizes and distances before getting into details.

Figure 1.1 attempts to illustrate the relative distances and sizes of various objects in the universe. I also find it helpful in grasping astronomical distances to make analogies to ordinary-size objects. For example, if the sun is a grain of sand (1 mm), the orbit of the earth is 10 cm and that of Pluto is 4 m. The nearest star is 30 km away and the center of the galaxy is five times the distance to the moon.

There are $\sim 10^{11}$ stars in our galaxy, and $\sim 10^{11}$ galaxies in the visible universe — a star in the Milky Way for every grain of sand it would take to fill a large lecture room, and then a galaxy for every star. There are more stars than all the grains of sand in all the beaches of the earth.

If our galaxy is the size of a half dollar (3 cm), the nearest big galaxy is almost 1 m away, and the Virgo cluster of galaxies, located near the center of the local supercluster, is 10 m away. The most distant quasars are more than a kilometer away.

Table 1 lists the values of the most important physical constants used in these lectures. Astronomers measure distance in parsecs (pc). The sun is about 8 kpc from the center of the Milky Way galaxy, which is about halfway to the edge of the visible galaxy. As we will see, the Milky Way's dark halo extends considerably farther.

The distance to distant galaxies is deduced from their redshifts using Hubble's constant $H_o = 100h$ km s^{-1} Mpc^{-1}, the value of which remains uncertain by

about a factor of two: $\frac{1}{2} \lesssim h \lesssim 1$. Consequently, the parameter h appears in many formulas where the distance matters.

Table 1

parsec	pc	$= 3.09 \times 10^{18}$ cm $= 3.26$ light years (LY)
Newton's const.	G	$= 6.67 \times 10^{-8}$ dyne cm^2 g^{-2}
Hubble parameter	H	$= 100\,h$ km s^{-1} Mpc^{-1} , $1/2 \lesssim h \lesssim 1$
Hubble time	H^{-1}	$= h^{-1}\ 9.78 \times 10^9$ y
Hubble radius	R_H	$= cH^{-1} = 3.00\,h^{-1}$ Gpc
critical density	ρ_c	$= 3H^2/8\pi G = 1.9 \times 10^{-29} h^2$ g cm^{-3}
		$= 11\,h^2$ keV cm$^{-3} = 2.8 \times 10^{11} h^2\,M_\odot$ Mpc^{-3}
speed of light	c	$= 3.00 \times 10^{10}$ cm s$^{-1} = 306$ Mpc Gy^{-1}
solar mass	$M_\odot$	$= 2.00 \times 10^{33}$ g
solar luminosity	$L_\odot$	$= 3.83 \times 10^{33}$ erg s^{-1}
Planck's const.	$\hbar$	$= 1.06 \times 10^{-27}$ erg s $= 6.58 \times 10^{-16}$ eV s
Planck mass	M_{Pl}	$= (\hbar c/G)^{1/2} = 2.18 \times 10^{-5}$ g $= 1.22 \times 10^{19}$ GeV
proton mass	m_p	$= 1.67 \times 10^{-24}$ g $= 0.938$ GeV/c^2
Boltzmann const.	k_B	$= 1.38 \times 10^{-16}$ erg K$^{-1} = (1.16 \times 10^4)^{-1}$ eV K^{-1}
sidereal year	y	$= 3.155815 \times 10^7$ s
radian		$= 57^\circ.2958 = 3437'.75 = 206265''$

1.2 GALAXIES

The nearest large galaxy to ours is the great galaxy in the constellation Andromeda. It was first recorded on an astronomical map by Abd-al-rahman al Sufi in 964 A.D., and first drawn — as an elliptical nebula (Latin for cloud) — in an engraving of the Andromeda constellation by Bouillaud in 1667.[1] Messier included it in his catalogue of nebulas as number 31. It was not until 1923, however, that Hubble first recognized the true nature of M31.

Like our own galaxy, M31 is a typical giant spiral. It is perhaps twice as massive as the Milky Way, with a mass in stars of about $4 \times 10^{11} M_\odot$. Its radius is about 25 kpc. It is located about 0.7 Mpc from us, and its velocity along the line of sight (measured by the Doppler shift) is 270 km s^{-1} toward us.

There are about thirty other galaxies known in our local group of galaxies, but all are much smaller than these two giants. M33, the only other spiral galaxy, has perhaps a tenth the mass of the Milky Way. M32, the largest elliptical galaxy in the local group, is considerably less massive. Both M32 and M33 are fairly close to M31. The largest galaxies in the immediate vicinity of the Milky Way are two irregular galaxies, known as the Large and Small Magellanic Clouds. In addition, seven dwarf spheroidal galaxies have been found near our galaxy: Draco, Ursa Minor, Carina, and Sculptor within 100 kpc, and Fornax and Leo I and II at about twice that distance. (They are named after the constellations in which they lie.) Fornax, the most massive of them, has a mass in stars of only about $2 \times 10^7 M_\odot$. Partly because of the fact that their masses are so tiny (for galaxies), these dwarf spheroidals may give us important clues about the origin of galaxies and the composition of the dark matter, as I will discuss later on.

Figure 1.2 is the traditional Hubble "tuning-fork" diagram of galaxy types, from ellipticals, through lenticular ($S0$) and spiral galaxies (with and without central bars), to irregulars. This progression of galaxy morphologies corresponds to decreasing prominence of the spheroidal component and increasing importance of disk. Hubble thought it possible that his classification was evolutionary, and

although this is no longer believed the sequence $Sa - Sb - Sc - Sd$ is called by astronomers the progression from "early" to "late" spiral types. The disk-to-bulge luminosity ratio increases from ~ 1 for Sa to ~ 10 for Sd. Late spiral types also have more gas and young, blue stars. Roughly 10% of all bright galaxies are ellipticals, 20% are $S0$, 65% are spirals, and 5% are irregulars, with higher fractions of $S0$ and E in regions of higher galaxy number density — another important clue to galaxy origins.

Spiral Galaxies

Spiral galaxies have three visible components: the disk with its spiral arms, the nucleus or bulge, and the stellar halo or corona. In addition, spiral galaxies generally appear to possess extensive dark matter halos.

Although the spiral arms are the distinguishing feature of spiral galaxies, there is less to them than meets the eye. The spiral arms are bright because of the short-lived luminous supergiant stars and emission nebulae they contain, but the number density of long-lived stars like the sun is not much different in the arms than in the interarm regions of the disk. Following the work of C. C. Lin and Frank Shu, it is now thought that the arms are the result of density waves travelling around the galaxy: the passage of the disk matter through such a wave triggers the process of star formation. Incidentally, the spiral arms curve backward, opposite to the direction of rotation of the galaxy. Thus a spiral galaxy rotates like a pinwheel, the spiral arm density waves rotating in the same direction but slower than the stars and gas in the disk.

The disk is remarkably thin. In our galaxy, it is a few hundred parsecs thick at the radius of the sun (about 8 kpc). Perhaps 10-20% of the mass in the disk is in gas (mostly hydrogen and helium) and dust (composed of what astronomers call "metals": elements more massive than helium). The disk becomes thicker and more gaseous at large radii, and in some galaxies the outer edge of the disk is warped. Spiral galaxies are generally surrounded by a diffuse envelope of neutral atomic hydrogen (HI, observed with radio telescopes in 21cm emission),

sometimes extending to several times the optical radius.[3]

The stars of a spiral galaxy were classified by Baade in 1944 into two broad categories, Population I and II. The relatively young, metal-rich stars of the disk are Population I. The older, lower metallicity stars are Population II; these are found mainly in the nucleus and stellar halo, including the globular clusters. Globular clusters are dense spherical assemblages of stars, having typically $\sim 10^6$ stars within a radius of a few pc. There are about 200 globular clusters in the Milky Way. Thus only a tiny fraction of the $\sim 10^{11}$ stars in the galaxy are in globular clusters. The total number of stars in the diffuse stellar halo is also a tiny fraction of the total. The stellar halo and about half the globular clusters are distributed roughly spherically. The other half of the globular clusters are associated with the disk. Most of the Population II stars lie in the spheroidal bulge which occupies the center of the galaxy, with radius ~ 4 kpc and very little gas, dust, or young stars. The majority of the stars in a galaxy like ours are Population I stars in the disk.

The luminosity distribution as a function of radius in the disk component of typical S and $S0$ galaxies is of the form

$$I_D(r) = I_o \exp(-\alpha r) \quad . \tag{1.1}$$

The corresponding disk luminosity is $L_D = 2\pi I_o \alpha^{-2}$, half of which is emitted within the effective radius $r_e = 1.67\alpha^{-1}$. For example, the Milky Way has an effective radius $r_e \approx 5$ kpc, a total (disk plus bulge) luminosity $L = L_D + L_B \approx 1.6 \times 10^{10} L_\odot$, a disk-to-bulge ratio $L_D/L_B \approx 2$, and is classified as an Sb or Sc galaxy.[4]

The most important source of information about the dynamics of a galaxy is Doppler shift measurements of the line-of-sight velocities of its components. By 1979, the evidence had become overwhelming that the rotation velocity of spiral galaxies remains roughly constant from a few kpc to the largest radii at which observations are possible.[5] This is surprising, since if the mass were mainly

associated with the stars, which are centrally concentrated, then the velocity in the outer regions would fall as $v \propto r^{-1/2}$, like that of the planets in the solar system.Fig. 1.3 shows rotation curves for many spiral galaxies, obtained both from 21cm observations and from measurements of velocities of the clouds of ionized gas surrounding hot blue stars. (Because these gas clouds emit most of their light in a few spectral lines, their velocities can be measured in a fraction of the exposure time required for stellar measurements.[8]) By simple Newtonian arguments, a constant rotation velocity v_{rot} implies that the mass $M(r)$ within radius r grows linearly with radius:

$$M(r) = (v_{rot}^2/G)r. \tag{1.2}$$

Correspondingly, the mass density falls as r^{-2}. Since the luminosity falls exponentially with radius, the mass-to-light and total-to-luminous-mass ratios M/L and M/M_{lum} grow with radius. From the fact that the rotation velocity is constant to several times the effective radius, it follows that the mass associated with the dark halos of these galaxies is at least several times that of all the visible matter.

Actually, the existence of massive dark matter halos was not entirely a surprise: at least two pieces of evidence had pointed toward it. Since the mid-1930's the astronomer Fritz Zwicky had been emphasizing that there is much more mass detected dynamically in great clusters of galaxies than can be attributed to the stars in their galaxies.[9,8] And in 1973 it was pointed out[10] that a self-gravitating disk is unstable toward collapse to a rotating bar — indeed, this bar instability probably is responsible for the fact that roughly a third of spiral galaxies have central bars — but that the disk can be stabilized by a roughly spherical halo containing comparable mass at the same radius. More recent detailed studies of galactic disks have confirmed that most of the dark matter cannot be in the disk.[11] The existence of warps in the outer parts of disk galaxies is also evidence that the dark halo is roughly spherhical, since such warps would be smeared out in a nonspherical halo.

How large is the total mass associated with a typical spiral galaxy? Since according to the above equation the mass grows linearly with radius, one can equivalently ask, How large is the halo? We can set lower limits of $r_{halo} \gtrsim 70$ kpc, and correspondingly $M/M_{lum} \gtrsim 10$ and $M \gtrsim 2 \times 10^{12} M_{\odot}$ for our own galaxy from studies of its satellites; see Fig. 1.4. This mass is comparable to that suggested by studies of the dynamics of the local group of galaxies.[19] As I will discuss shortly, the evidence from studies of the dynamics of all assemblages of galaxies, from small groups to rich (i.e., very populous) clusters, is consistent with $M/M_{lum} \approx 10$. The only significant evidence to the contrary of which I am aware is a recent paper reporting the results of a new technique for measuring galaxy mass based on the distortion of the images of background galaxies by gravitational deflection of their light by foreground galaxies.[20]

Elliptical Galaxies

Elliptical galaxies are spherical or ellipsoidal stellar systems consisting almost entirely of old stars. They contain very little dust and show no evidence of spiral arms. The larger ellipticals contain many globular clusters. In all these respects, they resemble the nucleus and stellar halo components of spiral galaxies.

Elliptical galaxies are classified in several ways. One is by ellipticity, with the integer n in En designating $10(a-b)/a$ where a and b are the projected major and minor axes. Ellipticals have projected axial ratios b/a between 1.0 ($E0$) and 0.3 ($E7$). It was once widely believed that elliptical galaxies are oblate spheroids flattened by rotation, but in the past few years rotation curves and velocity dispersion data have shown that some ellipticals, especially the larger ones, rotate much too slowly to account for their flattening. There is evidence that some ellipticals, again especially the larger ones, are actually triaxial, and that their flattening is due to velocity anisotropies. It is not yet known whether these are more nearly oblate or prolate.[21]

Ellipticals vary very widely in mass, from dwarf spheroidals to supergiant galaxies. The latter are the largest known galaxies, with extensive ($\sim$ 100 kpc)

amorphous stellar envelopes and masses as much as an order of magnitude larger than that of M31. Called cD galaxies, they are usually found in the cores of rich, regular clusters of galaxies; and they are often flattened, the major axis aligned with that of the cluster. Roughly a third of all cD galaxies have multiple nuclei, which suggests that they formed through mergers. At the other end of the size scale, there are probably more dwarf ellipticals (dE) than any other type of galaxies in the universe — as is true in our local group of galaxies. Or perhaps the most populous galaxy species is dwarf irregulars.[22] In any case, dwarf galaxies represent only a small fraction of the stars and mass in the universe since they are so small.

The projected distribution of light intensity in E galaxies is well fit by the de Vaucouleurs formula

$$I(r) = I_e \exp[-7.67((r/r_e)^{1/4} - 1)], \tag{1.3}$$

where the "effective radius" r_e is the radius enclosing half of the total light. $I(r)$ falls off more slowly than r^{-2} for $r < r_e$ and more rapidly than that for $r > r_e$. The same formula fits the bulges of $S0$ and S galaxies.

The total luminosity L of an elliptical galaxy is observed to be related to its stellar velocity dispersion σ by the formula $L \approx L_*(\sigma/220 \text{ km s}^{-1})^\gamma$, where $L_* = 10^{10} M_\odot$ and $\gamma = 4 \pm 1$. This is the Faber-Jackson relation.[23] There is an analogous relation $L_H \propto v_{rot}^4$ between the total infrared luminosity and rotation velocity of spirals, called the infrared Tully-Fisher relation.[24] A similar relation has been found to hold between the total luminosity in the blue spectral band and the rotation velocity for a sample of spiral galaxies.[25] These empirical relations, displayed in Fig. 1.5, are important in providing both cosmic yardsticks and insights into the formation and dynamics of galaxies.

Luminosity Distribution

The galaxy luminosity function is defined such that $\phi(L)dL$ is the number density of galaxies having total luminosity in the interval $(L, L + dL)$. The

available data is fit by Schechter's convenient function[21]

$$\phi(L) = \frac{\phi_*}{L_*}\left(\frac{L}{L_*}\right)^{\alpha} e^{-L/L_*} \quad , \tag{1.4}$$

where[21]

$$\begin{aligned} \alpha &= -1.29 \pm 0.11 \\ \phi_* &= 1.3 \pm 0.3 \times 10^{-2} h^3 \mathrm{Mpc}^{-3} \\ L_* &= 1.1 \times 10^{10} h^{-2} L_\odot . \end{aligned} \tag{1.5}$$

This is sketched in Fig. 1.6. Actually, $\phi(L)$ must fall off more rapidly than the function (1.4) at small L, since the mean space density of galaxies corresponding to (1.4),

$$\langle n \rangle = \phi_* \Gamma(\alpha + 1) \quad , \tag{1.6}$$

diverges if $\alpha < -1$. The shape of the luminosity function for $L < 0.005 L_*$ is uncertain, but the number of small nearby galaxies is indeed less than predicted by (1.4). The shapes of the luminosity functions for the different mophological types of galaxies differ at the faint end, dwarf E and I galaxies being more numerous than dwarf S and $S0$, but the luminosity functions have similar shapes at the bright end.

The mean luminosity density corresponding to (1.4) is perfectly finite:

$$\langle \mathbf{L} \rangle = \phi_* L_* \Gamma(\alpha + 2) \approx 1.8 \times 10^8 h L_\odot \mathrm{Mpc}^{-3} \quad . \tag{1.7}$$

The majority of galaxies are faint, but most of the light comes from those that are of luminosity $\gtrsim L_*$. With (1.7) we can evaluate the mean mass-to-light ratio of the universe:

$$M/L = \Omega \rho_c / \langle \mathbf{L} \rangle \approx 1500 \Omega h (M_\odot / L_\odot) \quad , \tag{1.8}$$

where ρ_c is the critical density for closure (see Table 1 and Lecture 2) and Ω is

the average density of the universe in units of ρ_c. Typically,

$$M/L \approx 14h(M_\odot/L_\odot)$$

in the centers of galaxies;[21] thus Ω(galaxy cores) $\approx 10^{-2}$, with perhaps another factor of two including the entire visible mass in galaxies. If the total galactic mass, including that of the halo, is about ten times greater (i.e., $M/M_{lum} \approx 10$), as discussed above, then $\Omega \approx 0.2$ and the universe is open.

Interpretations

Although it is perhaps premature to sketch a theoretical framework for understanding the basic facts about galaxies, both in the context of these lectures and given the available astronomical data, I think it is nevertheless useful to do so at this point. The great advantage of keeping a tentative theory in mind as one thinks about data is that it helps in organizing and remembering the facts. If it is a good theory, it will also call attention to particularly important facts — especially those that may contradict it!

The basic picture of galaxy formation that I have in mind is that galaxies collapsed gravitationally from initially rather homogeneous mixtures of dark matter and ordinary matter (in about the ratio 10:1). As I will explain in the next lecture, the result of virialization by violent relaxation in gravitational collapse is a roughly isothermal halo, with density falling as r^{-2}, as required to produce the observed constant-velocity rotation curves. The ordinary matter continued to radiate away its kinetic energy and sink toward the center, eventually forming the visible stars. This process is called dissipational collapse. Meanwhile the dark matter retained its post-virialization velocity and density distribution, and it forms the galactic halos. We do not know what the dark matter is, but its key property, in addition to being invisible, is that it is dissipationless. Probably both properties are a consequence of its lack of significant interaction with electromagnetic radiation, perhaps because the dark matter is composed of neutral elementary particles.

In this picture, the disk in disk galaxies formed when the dissipational collapse of the baryonic matter was halted by angular momentum conservation. (The symmetrical configuration of minimum kinetic energy for given angular momentum is a disk.) Galactic spheroids resulted when dissipational collapse was halted by some other process, presumably star formation. (A collection of virialized gravitating mass points is dissipationless.) Evidently, spheroids result from matter which had either (or both) higher initial density or smaller initial angular momentum /than that which formed disks.[27,26]

It follows that all galaxies should be surrounded by massive dark matter halos. I have already discussed the strong evidence that this is true for spiral galaxies. Although relevant observations are more difficult for other galaxy types, the data available are consistent with the ubiquity of massive halos.[5,28-31]

A useful way of visualizing galaxies is sketched in Fig. 1.7, where density is plotted versus distance from the center of our galaxy, looking toward M31. In the central region of a typical large galaxy the density is high — perhaps even infinite at the very center if there is a black hole there. This is surrounded by a region of rapidly falling baryonic matter density, so that there are comparable total amounts of ordinary and dark matter enclosed within a few effective radii (r_e). The density of the dark matter halo (dashed line) declines $\propto r^{-2}$ out at least to $\sim 10^2$ kpc. If it continues to follow a r^{-2} law between the galaxies (dotted line), then the average density is approximately that required for closure, i.e. $\Omega \approx 1$. Jim Peebles calls this the "alpine model". On the other hand, if the dark matter density falls off rapidly beyond $\sim 10^2$ kpc ("crayon model"), then as I mentioned before $\Omega \approx 0.2$ and the universe is open.

The horizontal lines in Fig. 1.7 represent critical density today and at the earlier epoch when the universe had expanded only 1/10 as much; i.e., when $R = 0.1$. The expansion factor R, defined to equal unity now, is given in terms of the redshift $z \equiv \delta\lambda/\lambda$ by $R = (1+z)^{-1}$. In the next lecture I will remind you of the relationship between z or R and the time t since the Big Bang, and also

explain why the fact that the halo at 100 kpc is roughly an order of magnitude more dense than the higher of the two light horizontal lines suggests that the galaxy interior to that collapsed gravitationally before z of 10.

1.3 GROUPS AND CLUSTERS

Half or more of all galaxies are members of groups or clusters. "Groups" of galaxies are systems containing at most a few tens of bright galaxies, while "clusters" are richer (i.e., more populous) systems. They are identified as density enhancements, either in surface number density of galaxies on the sky, or in redshift-space volume density. It is thought that most of them are also gravitationally bound structures, especially those of high density.

After a particular variety of astronomical object has been discovered, it has usually proved very valuable to make a catalogue of such objects, in order to study them systematically. There are two great catalogues of clusters of galaxies, Abell's catalogue of 2712 rich clusters[33] and the more extensive Zwicky catalogue,[34] which lists and classifies poorer (i.e., less populous) clusters as well. Both catalogues are based on the Palomar Sky Survey plates, and so are limited to the northern sky.

Reliable identification of groups of galaxies requires redshift data, which has only recently become available for large numbers of galaxies. The best catalogue of groups is that recently compiled by Geller and Huchra,[35] obtained by applying a group-finding algorithm to the NB whole-sky catalogue[36] of the 1312 galaxies brighter than $m_B = 13.2$ (*), and to the Harvard-Smithsonian Center

---

$\star$ The notation m_B represents apparent magnitude in the B blue spectral band. Apparent magnitude is related to the measured flux S by $m = A - 2.5 \log_{10} S$, where the constant A depends on the spectral band; thus a galaxy of $m = 12$ appears to be 100 times brighter than one of $m = 17$. The naked eye can see to $m \approx 6.5$; a six inch (15 cm) telescope, to $m = 13$; and the Palomar 5 m telescope, to $m = 24$ (photographically). Astronomical traditions can be long lived. The magnitude scale was adopted in the 19th century to agree approximately with the brightness classification given in the catalogue of 850 stars compiled by Hipparchus in the second century B.C., whose 6th magnitude stars are about 100 times fainter than those of 1st magnitude.

for Astrophysics ("CfA") survey of the northern sky (2396 galaxies, complete to $m_B = 14.5$ for about 20% of the sky).[36] They found 92 groups in the former catalogue and 176 in the latter; about 60% of all the galaxies in the catalogues are assigned to groups.

There are several classification schemes for clusters, but a simple one that overlaps with the others is "regular" vs. "irregular".[37,38] Regular clusters have a smooth and symmetric structure, with high central galaxy density ($\gtrsim 10^3$ per Mpc^{-3}), a small fraction of spiral galaxies ($\lesssim 20\%$), high velocity dispersion (~ 1000 km s^{-1}), and a high X-ray luminosity ($> 10^{44}$ erg s^{-1}) from hot gas (of temperature $T \gtrsim 6$ keV). Examples include A85 and A2256 (the bottom two X-ray images in Fig. 1.8), A496 (upper right in Fig. 1.9), and the Coma cluster. (This cluster, designated A1656 — i.e., No. 1656 in Abell's catalogue — is the nearest rich cluster, at about 45 h^{-1} Mpc. As usual, it is named after the constellation in which it lies on the sky, Coma Berenices — Berenice's Hair.)

Only about a quarter of all rich clusters are regular. Irregular clusters have a rather lumpy structure, lower central galaxy density, a somewhat higher spiral fraction ($\gtrsim 40\%$) than regular clusters, lower velocity dispersion, lower X-ray luminosity and cooler gas (1-2 keV). Examples include A262, A1367, and the Virgo cluster. In addition, there are intermediate cases, exemplified by the middle two images in Fig. 1.8 and many of the clusters in Fig. 1.9. Many of these are elongated and have prominent subclusters.

Another distinction that is especially apparent on the X-ray images is between those clusters with a central, dominant galaxy (e.g., the three clusters on the right half of Fig. 1.8) and those without a *cD* (left half of Fig. 1.8). In their cores, *cD* galaxies look like giant ellipticals, except that some have multiple nuclei. This core is surrounded by a very extensive stellar and gaseous envelope, with optical surface brightness decreasing much more slowly than the de Vaucouleurs (eq. (1.3)) profile of a typical elliptical at large distances, and with extended, centrally peaked X-ray emission from the hot gas.

There is no sharp dividing line between "groups" and "clusters", and a substantial overlap of physical characteristics between these two categories.[41] Most groups are loose, but there are compact groups with galaxy densities comparable to those in the cores of rich clusters. Some groups even contain small cD galaxies.

Alan Dressler first demonstrated that in rich clusters there is a well-defined relationship, shown in Fig. 1.10, between the local number density of galaxies and the local fraction of each galaxy Hubble type.[42] The local density was computed using the 10 nearest (projected) neighbors of each galaxy. The fractions of E and $S0$ galaxies increase, and the fraction of $S + I$ decreases, smoothly and monotonically as the local galaxy density increases. This relation between population and density holds for individual clusters as well as, on the average, from cluster to cluster. In particular, it holds for both regular and irregular clusters. And it has recently been shown to hold for groups as well as clusters.[43]

Interpretations

As I will discuss in the next lecture, the process of dissipationless gravitational collapse produces a smooth, centrally concentrated distribution of matter. The obvious interpretation of the difference between regular and irregular clusters of galaxies is therefore that the former have undergone collapse, while the latter have not yet done so.[44] If they are indeed in virial equilibrium, regular clusters' large velocity dispersions are strong evidence for a large quantity of dark matter to provide the required gravitational binding energy. Although the mass-to-light ratio (M/L) implied for rich clusters is about a factor of 6 larger than that for galaxies (including their massive dark halos), the ratio of total to luminous mass (M/M_{lum}), which is physically more meaningful, (*) is about the same for both.

$\star$ The old red stars of the E and $S0$ galaxies in regular rich clusters are less luminous per unit mass than the younger and bluer stars of S galaxies, which are not as prevalent in rich clusters, and the X-ray observations[38] show that there is at least as much mass in the hot gas in the cores of rich clusters as there is in galaxies. M_{lum} compensates for intrinsic luminosity differences of different galaxy types and includes the mass in hot gas; that is why it is physically more meaningful than L. For more details, see Ref. 26, especially Table 1.

The analysis by Geller and Huchra[35] of groups and clusters in the CfA catalogue finds that they have approximately constant M/L. An earlier study[45] which claimed to find a trend of increasing M/L with increasing size of the cluster is now known to have been misled by a flaw in the cluster finding algorithm.

The data on M/L and M/M_{lum} are plotted in Fig. 1.11. It is apparent that the data are consistent with roughly constant M/M_{lum} across the entire range of masses from dwarf spheroidal galaxies (using the dynamical mass estimates for them) to the cores of rich clusters. The most straightforward interpretation of this constancy is that there is about an order of magnitude more dark than luminous matter in the universe.

cD galaxies are thought to form through galactic cannibalism as (gravitational) dynamical friction causes cluster galaxies to spiral into the centrally located giant, where they are disrupted by tidal forces.[38,46] The fact that many *cD* galaxies have multiple nuclei is evidently direct evidence for galactic cannibalism. Computer simulations of the evolution of groups and clusters have shown that mergers and tidal stripping are most rapid in small groups, including those that form in the early stages of the collapse of larger clusters, and that it is possible to understand the origin of *cD* galaxies in this way if cluster galaxies initially possess massive dark halos which only later become smeared out as the cluster relaxes.[47]

Finally, regarding Dressler's correlation between galaxy type and number density, the key question is whether it is caused by heredity (i.e., factors present when galaxies formed) or environment (evolutionary effects after galaxy formation, such as galaxy mergers or stripping of gas from spirals to form *S*0s). There is evidence that both heredity and environment are important.[42,48,41] I will return to all of these questions in later lectures; they are crucial to unraveling the mystery of the origin of galaxies and clusters.

1.4 SUPERCLUSTERS AND VOIDS

Thirty years ago, astronomers knew that rich clusters consist mostly of E and $S0$ galaxies, and that the majority of galaxies are spirals and lie outside these clusters in relative isolation in the "field". But they did not yet know about superclusters and voids.[49]

Gerard de Vaucouleurs was the first to define and describe the Local Supercluster, the vast aggregation of several thousand galaxies of which our own Local Group, containing the Milky Way, is an outlying member. The Local Supercluster is centered on the Virgo cluster, about $15h^{-1}$ Mpc away from us. It has recently been mapped in some detail by Tully,[50] who finds that it consists of a fairly thin disk component containing about 60% of the luminous galaxies and a halo component with 40%, and that almost all the luminous galaxies of the halo are associated with a few clusters leaving most of the volume off the disk empty.

Although there was some recognition that there are other superclusters in addition to our own on the basis of (two-dimensional) sky surveys, we have only begun to see the large scale structure of the universe clearly with the advent of large-scale redshift surveys. The limitation of these surveys is that while thousands of galaxy positions can be read off of a single photographic plate, spectral redshifts must be obtained one by one. Roughly 10^4 of them are presently available, including deep surveys of a few percent of the sky ("drilling holes in space": measuring redshifts for all galaxies brighter than a faint limiting magnitude in a small angular region) and shallower surveys covering larger angular area (the prime examples being the NB and CfA catalogues). The data is growing rapidly: the doubling time for the number of galaxy redshifts available is presently about three years. Technological advances, including image tubes and CCD (charge-coupled device) detectors that allow modern astronomers to record as much information in an exposure of a few minutes as their predecessors could in an entire night, have helped to make this possible.

Figure 1.12 shows an example of the results of these surveys. The top portion

shows the positions of bright galaxies in a region of the sky in the direction of the constellation Perseus. A chain of galaxies is apparent — the clearest such "filament" known. The lower portion of the figure, in which the galaxy positions are plotted in a redshift-angle "wedge" diagram, shows that these galaxies are concentrated at a particular distance, about $50h^{-1}$ Mpc; thus they really do lie in a filamentary band across the sky. Equally striking in this figure is the fact that most of the wedge diagram is empty. Such voids in the galaxy distribution are apparent on all diagrams of this sort. Galaxies are concentrated in flattened or filamentary superclusters, leaving most of the volume of the universe virtually devoid of bright galaxies.[40]

All nearby Abell clusters are now known to belong to superclusters. For example, Coma and A1367 are connected by a bridge of galaxies, including several large groups. The whole structure stretches at least 20 degrees across the sky, corresponding to a length of $\sim 30h^{-1}Mpc$; some astronomers argue that it is even larger. What is really needed now are catalogues of superclusters and voids, so that their statistical properties can be learned. Astronomers will be able to obtain enough redshift data in five to ten years for this to be possible. The largest void discovered to date is the "great void in Boötes" lying between two large superclusters, the Hercules supercluster on the near side, and the great Corona Borealis supercluster, which contains 15 Abell clusters, on the far side. The Boötes void is perhaps $60h^{-1}$ Mpc across, and the density of bright galaxies in it is probably less than a tenth, and almost certainly less than a quarter, of the average density.[52]

Any data regarding correlations of galaxy and cluster properties across the vast distances spanned by superclusters and voids is potentially important in indicating how they may have formed. Probably the most interesting data of this sort obtained thus far is Binggeli's observation that the position angles of all nearby, elongated Abell clusters lie within 45° of the direction to the nearest cluster, provided the clusters are separated by less than $\sim 15h^{-1}$ Mpc. He found a similar, though less significant, correlation on larger scales, and also a

correlation between the position angle of the brightest cluster galaxy's major axis and the direction to the nearest cluster.[53] Similar, but substantially weaker, correlations were found in a recent analysis of a larger sample of clusters.[54] In a similar vein, the analysis of local flattening of the galaxy distribution and its correlation across space may help to clarify the nature of superclustering.[55,56]

Interpretations

The cores of rich clusters and compact groups represent enhancements of 10^4 or more over the average galaxy number density. They are certainly bound and relaxed structures. On the other hand, the galaxy density enhancement represented by the Local Supercluster is much smaller, perhaps a factor of three.[58] The *peculiar velocity* (deviation from uniform Hubble flow $\vec{v} = H_o\vec{r}$) of galaxies in superclusters is typically $\lesssim 10^3$ km s^{-1}. A velocity of 10^3 km s^{-1} is equivalent to a Mpc/Gy. Thus, while galaxies in rich cluster cores have had plenty of time since the Big Bang to cross from one side to the other, probably several times, the vast majority of galaxies have hardly had time to move more than a small fraction of the distance across their local superclusters. For example, the component of the Local Group's peculiar velocity in the direction of the Virgo cluster is 200-400 km s^{-1} (measured both via the dipole anisotropy of the cosmic background radiation and with respect to an ensemble of moderately distant galaxies[57,58]), but the LG is nevertheless still expanding away from the Virgo cluster with a velocity of $\sim$ 1000 km s^{-1}. Thus the Local Supercluster has not yet had time to collapse, certainly not in its longer dimension across the disk, and it is perhaps not even gravitationally bound.

It is precisely because of their unrelaxed state that superclusters are so valuable to cosmologists: gravity has not yet had time to mix them up, so their structure may reflect in a rather simple way the nature of the primordial conditions that gave rise to them.

The big question is, Which came first, superclusters (and voids), or galaxies? One popular view, *hierarchical clustering*, has it that galaxies formed more or less

at random locations, and were subsequently gathered up into clusters of ever increasing size, culminating in the vast superclusters whose dimensions we are only now beginning to appreciate. This view has long been championed by Jim Peebles of Princeton, among others. A competing "top down" view, long advocated by the Russian astrophysicist Yakov Zeldovich and his colleagues,[59,60] among others, contends that it is the superclusters that formed first, subsequently fragmenting into smaller objects which then formed galaxies. On the face of it, there are several outstanding pieces of evidence that superclusters were primary: the very existence of superclusters and large voids is pretty direct evidence that galaxy formation could not have occurred at random locations in the early universe, and the Binggeli correlation discussed above is easy to understand as a reflection of superclustering preceeding the formation of clusters, if not also galaxies. However, there are serious problems with this view, too. As I will explain in more detail in Lecture 3, galaxies appear to be much older structures than superclusters. In addition, in the "top down" scenario it is hard to understand the observed clustering substructure[61] as well as the structure of individual galaxies.

2. Gravity

Gravity is the subject of this second lecture. In it I will try to introduce the basic ideas necessary to understand the effects of gravity both on the evolution of the entire universe and on the growth and collapse of the fluctuations that presumably formed galaxies and all larger scale structures. I will also briefly explain how the hypothesis of cosmic inflation can account for the origin of these large scale fluctuations without violating causality.

I will assume here that Einstein's general theory of relativity (GR) accurately describes gravity. Although it is important to appreciate that there is no observational confirmation of this on extragalactic scales, the tests of GR on smaller scales are becoming increasingly precise, especially with the discovery of pulsars in binary star systems.[62] There are two other reasons most cosmologists believe in GR: it's conceptually so beautifully simple that it is hard to believe it could be wrong, and anyway it has no serious theoretical competition. Nevertheless, since a straightforward interpretation of the available data in the context of this standard theory of gravity leads to the disquieting conclusion that most of the matter in the universe is dark, there have been suggestions that perhaps our theory of gravity is inadequate on large scales. I will mention them briefly at the end of this lecture.

2.1 Cosmology

The "cosmological principle" is logically independent of our theory of gravity, so it is appropriate to state it before discussing GR further. But before I can state it, some definitions are necessary:

> A *co-moving observer* is at rest and unaccelerated with respect to nearby material (in practice, with respect to the center of mass of galaxies within, say, 50 Mpc).
>
> The universe is *homogeneous* if all co-moving observers see identical properties.

The universe is *isotropic* if all co-moving observers see no preferred direction.

The *cosmological principle* asserts that the universe is homogeneous and isotropic on large scales. (It is not difficult to see that isotropy actually implies homogeneity, but the counterexample of a cylinder shows that the reverse is not true.) In reality, the matter distribution in the universe is in our common experience exceedingly inhomogeneous on small scales, and increasingly homogeneous on scales approaching the entire horizon. The cosmological principle is in practice the assumption that for cosmological purposes we can neglect this inhomogeneity. The great advantage of assuming homogeneity is that our own neighborhood becomes representative of the whole universe, and the range of cosmological models to be considered is also enormously reduced.

The cosmological principle implies the existence of a universal cosmic time, since all observers see the same sequence of events with which to synchronize their clocks. (This assumption is sometimes explicitly included in the statement of the cosmological principle; see, e.g., Ref. 63, p. 203.) In particular, they can all start their clocks with the Big Bang.

Astronomers observe that the redshift

$$z \equiv \frac{\lambda - \lambda_o}{\lambda_o} \tag{2.1}$$

of distant galaxies is proportional to their distance. We assume, for lack of any viable alternative explanation, that this redshift is a Doppler shift: the universe is expanding. The cosmological principle then implies (see, for example, Ref. 64, §4.3) that the expansion is homogeneous:

$$r = R(t) r_o, \tag{2.2}$$

which immediately implies Hubble's law:

$$v = \dot{r} = \dot{R} R^{-1} r = H r. \tag{2.3}$$

Here r_o is the present distance of some distant galaxy (the subscript "o" in cosmology denotes the present era), r is its distance as a function of time and v is its velocity, and $R(t)$ is the scale factor of the expansion (scaled to be unity at the present: $R(t_o) = 1$). The scale factor is related to the redshift by $R = (1+z)^{-1}$. Hubble's "constant" $H(t)$ (constant in space, but a function of time except in an empty universe) is

$$H(t) = \dot{R}R^{-1}. \tag{2.4}$$

Finally, it can be shown[66,63] that the most general metric satisfying the cosmological principle is the Robertson-Walker metric

$$ds^2 = c^2dt^2 - R(t)^2\left[\frac{dr^2}{1-kr^2} + r^2(\sin^2\theta d\phi^2 + d\theta^2)\right], \tag{2.5}$$

where the curvature constant k, by a suitable choice of units for r, has the value 1,0, or -1, depending on whether the universe is closed, flat, or open, respectively. For $k = 1$ the spatial universe can be regarded as the surface of a sphere of radius $R(t)$ in four-dimensional Euclidean space; and although for $k = 0$ or -1 no such simple geometric interpretation is possible, $R(t)$ still sets the scale of the geometry of space.

2.2 General Relativity

Formally, GR consists of the assumption of the Equivalence Principle (or the Principle of General Covariance[65]) together with Einstein's field equations

$$R^{\mu\nu} - \frac{1}{2}Rg^{\mu\nu} = -\frac{8\pi G}{c^4}T^{\mu\nu} - \Lambda g^{\mu\nu}. \tag{2.6}$$

The Equivalence Principle implies that spacetime is locally Minkowskian and globally (pseudo-)Riemannian, and the field equations specify precisely how spacetime responds to its contents. The essential physical idea underlying GR is that spacetime is not just an arena, but rather an active participant in the dynamics.

Fortunately, there are several excellent introductions to GR for cosmologists.[65,66,63]

It will not be necessary to discuss the details of GR here, but I think it may be useful to spend a little time on the concept of horizons, since in my experience this is one of the things that most confuse newcomers to cosmology — in particular, the apparent contradiction between Hubble's law and the speed of light as a speed limit.

I find it helpful to picture the behavior of spacetime near horizons using the somewhat artificial concept of a static point, which is fixed in space. Figure 2.1(a) shows a number of static points located at various distances from a black hole singularity. Imagine that each static point emits a pulse of light; the light circles in the figure show schematically the positions of the wavefronts a moment later. Far from the black hole, spacetime is flat and the light circle is centered on the static point. But closer to the black hole, the light is increasingly dragged toward the singularity, as if space itself were flowing into the black hole. As E. R. Harrison amusingly puts it,[67] the event horizon, located at the Schwarzschild radius, "is the country of the Red Queen where one must move as fast as possible in order to remain on the same spot." At the horizon, the light circle lies on the static point and no light can escape outward. Inside the horizon space effectively flows inward faster than light, and outward-moving light cannot even reach the horizon. It is important to understand that special relativity remains locally valid except at the singularity itself, and light always moves at the speed of light c with respect to freely falling observers.

A Hubble sphere in the expanding universe is like a Schwarzschild event horizon turned inside out. As Fig. 2.1(b) shows, the light circles are centered on their static points well inside the Hubble sphere, but dragged increasingly outward at larger radii. At the Hubble sphere, the light circle lies on the static point and no light can escape inward. And beyond the Hubble sphere, space effectively flows outward faster than the speed of light. But the galaxies in that

space are not moving at all (except for their small peculiar motions); it is the expansion of space that is carrying them away from us. The recession velocity in Hubble's law (2.3) is thus not an ordinary (local) velocity. The picture of the Hubble expansion as arising from galaxies flying apart in an underlying Euclidean space is only mildly misleading locally, but completely untenable on the scale of the Hubble radius. It is space itself that is expanding. This idea of space-time as an active participant in the dynamics of the universe is also crucial for understanding the inflationary universe.

Comoving coordinates are coordinates with respect to which comoving observers are at rest. A comoving coordinate system expands with the Hubble expansion. It is convenient to specify linear dimensions in comoving coordinates scaled to the present, as in eq. (2.2). For example, if I say that two objects were 1 Mpc apart in comoving coordinates at a redshift of $z = 9$, their actual distance then was 0.1 Mpc.

In a non-empty universe with vanishing cosmological constant, the case first studied in detail by the Russian cosmologist Alexander Friedmann in 1922-24, gravitational attraction ensures that the expansion rate is always decreasing. As a result, the Hubble radius

$$R_H(t) = cH(t)^{-1} \tag{2.7}$$

is increasing. The Hubble radius of a Friedmann universe expands even in comoving coordinates. Our backward lightcone encompasses more of the universe as time goes on.

I will conclude these preliminary reflections on horizons in the universe with Fig. 2.2. In this figure mass is plotted against linear size. The upper left portion of the graph is the region excluded by gravity: the heavy diagonal line is the Schwarzschild radius $R_S = 2GMc^{-2}$. An object of mass M having radius $\leq R_S(M)$ lies inside its horizon and has effectively no size at all. There is reason to believe that such black holes are formed in the gravitational collapse of stars,

and that massive black holes power quasars and other active galactic nuclei. There is no known way to make black holes of substellar mass except perhaps in the early universe; any lighter than 10^{15} g will already have decayed by now with the emission of Hawking radiation.

Gravity is more important, the closer an object is to the Schwarzschild line. Gravity is of course important for planets, stars, galaxies, clusters, and the universe as a whole; it is relatively unimportant for objects that are small or have low density.

The Heisenberg uncertainty principle excludes the shaded region in the lower left corner of Fig. 2.2: trying to look in smaller and smaller regions requires larger and larger amounts of energy. Combining the constraints of gravity and quantum mechanics, there is a smallest length, the Planck length $\lambda_{P\ell} = (G\hbar/c^3)^{1/2} = 2 \times 10^{-33}$ cm, and a characteristic mass of a quantum black hole, the Planck mass $M_{P\ell} = (\hbar c/G)^{1/2} = 2 \times 10^{-5}$ g (see Table 1). To understand the origins of the Big Bang before the Planck time $t_{P\ell} = \lambda_{P\ell}/c$ will require a quantum theory of gravity.

A universe of vanishing curvature $k = 0$ has critical density; the mass enclosed by the Hubble sphere lies on the heavy diagonal line in Fig. 2.2. A closed (open) universe with $k = 1$ ($k = -1$) lies above (below) this line. Presently available data indicate that the universe is actually within about an order of magnitude of critical density, as indicated by the cross in the upper left corner of the figure.

2.3 Friedmann Universes

Einstein's equations (2.6) for a homogeneous and isotropic fluid of density ρ and pressure p are

$$\frac{\dot{R}^2}{R^2} + \frac{kc^2}{R^2} = \frac{8\pi}{3}G\rho + \frac{\Lambda c^2}{3} \tag{2.8}$$

for the 00 component, and

$$\frac{2\ddot{R}}{R} + \frac{\dot{R}^2}{R^2} + \frac{kc^2}{R^2} = -\frac{8\pi}{c^2}Gp + \Lambda c^2 \tag{2.9}$$

for the ii components.[68] Multiplying (2.8) by R^3, differentiating, and comparing with (2.9) gives the equation of continuity

$$\frac{d}{dR}(\rho R^3) = -3pR^2c^{-2}. \tag{2.10}$$

Given an equation of state $p = p(R)$, this equation can be integrated to determine $\rho(R)$; then (2.8)can be integrated to determine $R(t)$.

Consider, for example, the case of vanishing pressure $p = 0$, which is presumably an excellent approximation for the present universe since the contribution of radiation and massless neutrinos (both having $p = \rho c^2/3$) to the mass-energy density is at the present epoch much less than that of nonrelativistic matter (for which p is negligible). Eq. (2.10) reduces to

$$(4\pi/3)\rho R^3 = M = \text{constant}, \tag{2.11}$$

and (2.8) yields *Friedmann's equation*

$$\dot{R}^2 = \frac{2GM}{R} + \frac{\Lambda c^2 R^2}{3} - kc^2. \tag{2.12}$$

This can be integrated in general in terms of elliptic functions, and for $\Lambda = 0$ in terms of elementary functions (see below).

Notice the analogy with Newtonian physics. Applying energy conservation to a self-gravitating sphere gives (2.12) with $k/2$ as the net energy (kinetic minus potential) per unit mass, and $\Lambda = 0$. The cosmological constant can be given a pseudo-Newtonian interpretation as a Klein-Gordon modification of the Poisson equation:[63]

$$\nabla^2\phi + \Lambda\phi = -4\pi G\rho. \tag{2.13}$$

For the time being, let us set $\Lambda = 0$. (I will discuss the case of a nonvanishing cosmological constant in Lecture 4.) Solving the Friedmann equation for k at the

present time (since k is a constant, any time will do),

$$kc^2 = R_o^2 \left(\frac{8\pi}{3} G\rho_o - H_o^2 \right). \tag{2.14}$$

Thus the universe is flat ($k = 0$) if its density equals the *critical density*

$$\rho_{c,o} \equiv \frac{3H_o^2}{8\pi G}. \tag{2.15}$$

It is convenient to specify the density in units of critical density via the *density parameter*

$$\Omega \equiv \rho/\rho_c. \tag{2.16}$$

It is also conventional to introduce the *deceleration parameter*

$$q_o \equiv -R_o \ddot{R}_o / \dot{R}_o^2. \tag{2.17}$$

It follows that if $\Lambda = 0$ and the universe is dominated today by nonrelativistic matter, $q_o = \Omega_o/2$.

The results obtained by integrating the Friedmann equation for positive, vanishing, and negative curvature universes are sketched in Fig. 2.3 and summarized below. In each case, the time since the Big Bang is given by the expression

$$t_o = H_o^{-1} f(\Omega). \tag{2.18}$$

The function $f(\Omega)$ is graphed in Fig. 2.4. It is a monotonically decreasing function, with $f(0) = 1$.

Open, $k = -1$, $\Omega_o < 1$

$$\begin{aligned} R(\eta) &= GM(\cosh\eta - 1) \\ t &= GM(\sinh\eta - \eta) \\ f(\Omega_o) &= \frac{1}{1-\Omega_o} - \frac{\Omega_o}{2(1-\Omega_o)^{3/2}} \cosh^{-1}\left(\frac{2}{\Omega_o} - 1\right). \end{aligned} \tag{2.19}$$

Flat, $k = 0$, $\Omega_o = 1$ (*Einstein-de Sitter universe*)

$$\begin{aligned} R(t) &= (9GM/2)^{1/2} t^{2/3} \\ f(1) &= 2/3 \end{aligned} \tag{2.20}$$

Closed, $k = +1$, $\Omega_o > 1$ (*Friedmann-Einstein universe*)

$$\begin{aligned} R(\eta) &= GM(1 - \cos\eta) \\ t &= GM(\eta - \sin\eta) \\ f(\Omega_o) &= \frac{\Omega_o}{2(\Omega_o - 1)^{3/2}} \cos^{-1}\left(\frac{2}{\Omega_o} - 1\right) - \frac{1}{\Omega_o - 1}. \end{aligned} \tag{2.21}$$

Figure 2.5 shows how Ω_o is related to H_o in these Friedmann models, for various values of t_o.

2.4 Comparison with Observations

Age of the Universe t_o

Observational evidence bearing on the age of the universe and other fundamental cosmological parameters was reviewed at the 1983 ESO-CERN conference by Sandage.[69] The best lower limits for t_o come from studies of the stellar populations of globular clusters (GCs). Sandage concludes that a conservative lower limit on the age of GCs is 16 ± 3 Gy, which is then a lower limit on t_o. Sandage goes on to assume (a) that the apparent cutoff in quasar redshifts at $z \sim 4$ implies that galaxy formation ended at that epoch, about 2 Gy, and (b) that the stars

in the oldest GCs studied formed at that epoch; thus he estimates $t_o \approx 18 \pm 3$ Gy. I prefer simply to conclude that $t_o > 16 \pm 3$ Gy. Fig. 2.5 shows that $t_o > 13$ Gy implies that $H_o \leq 75$ km s^{-1} Mpc^{-1} even for Ω very small, and that $H_o \leq 50$ km s^{-1} Mpc^{-1} for $\Omega = 1$. (Fig. 4.5 gives the analogous constraints for the case of a flat universe with nonvanishing cosmological constant.)

Hubble's Parameter H_o

Hubble's parameter $H_o \equiv 100h$ km s^{-1} Mpc^{-1} has in recent years been measured in two basic ways: (a) using Type I supernovae as "standard candles", and (b) using the Tully-Fisher relation between the rotation velocity and luminosity of spiral galaxies. Both methods depend on measuring the distance to nearby calibrating galaxies. Sandage has long contended that $h \approx 0.5$, and he concludes[69] that using both methods the latest data are consistent with $h = 0.50 \pm 0.07$. de Vaucouleurs has long contended that $h \approx 1$, and he has recently argued that the data still support this value.[70] Another method for determining H_o has recently been proposed which, like (a), uses Type I supernovae, but which avoids the uncertainties of the "distance ladder" by calculating the absolute luminosity of Type I supernovae from first principles (using a very plausible but as yet unproved physical model). The result obtained is that h lies between 0.38 and 0.71, with a best estimate of 0.58.[71]

Cosmological Density Parameter Ω

In the first lecture I summarized the evidence on the mass associated with galaxies from luminosity and dynamical mass measurements: Ω(luminous) $\approx 0.01 - 0.02$ and Ω(dark halos)$\approx 0.1 - 0.2$. Here I will discuss several other observations that are relevant to cosmological mass estimates: galaxy position and velocity correlation functions, the infall velocity of the Local Group toward the Virgo cluster, the dynamics of other superclusters, constraints on the density of diffuse neutral and ionized hydrogen, and attempts to measure the deceleration parameter.

Galaxy Correlation Functions

Peebles and his collaborators have analyzed the available data on the angular positions of $\sim 10^6$ galaxies in terms of low-order correlation functions.[72] More recently, redshift data from both the CfA survey[73] and a deeper redshift survey[74] have also given estimates of the relative peculiar velocity between pairs of galaxies as a function of their separation, which in turn can be used to estimate Ω.

The galaxy two-point correlation function $\xi(r)$ (also called the autocorrelation or autocovariance function) is defined by

$$\delta P = \bar{n}^2 \left[1 + \xi(r_{12})\right] \delta V_1 \delta V_2, \tag{2.22}$$

where δP is the joint probability of finding galaxies in volumes δV_1 and δV_2 separated by distance r_{12}, and $\bar{n}$ is the average number density of galaxies. Equivalently, the probability of finding a galaxy in δV at distance r, given one at the origin, is

$$\delta P(1|2) = \bar{n} \left[1 + \xi(r)\right] \delta V. \tag{2.23}$$

The three-point correlation function is defined analogously to (2.22):

$$\delta P = \bar{n}^3 \left[1 + \xi(r_{12}) + \xi(r_{23}) + \xi(r_{13}) + \zeta(r_{12}, r_{23}, r_{13})\right] \delta V_1 \delta V_2 \delta V_3. \tag{2.24}$$

The corresponding triangle geometry is sketched in Fig. 2.6.

The two-point correlation function has a remarkably simple form: it is approximately a power law over the interval $0.1h^{-1}\text{Mpc} \lesssim r \lesssim 10h^{-1}\text{Mpc}$:

$$\xi(r) = (r/r_o)^{-\gamma}, \tag{2.25}$$

with[73] index

$$\gamma = 1.77 \pm 0.04$$

and correlation length

$$r_o = 5.4 \pm 0.3h^{-1}\text{Mpc},$$

and $\xi \ll 1$ (and rather uncertain) for $r \gtrsim 10h^{-1}$ Mpc. Values of $\xi(r)$ determined from the CfA data are plotted in Fig. 2.7.The three-point function is found[76] to be well approximated by a symmetric sum of products of two-point functions:

$$\zeta(r_{12}, r_{23}, r_{31}) = Q\left[\xi(r_{12})\xi(r_{23}) + \xi(r_{23})\xi(r_{31}) + \xi(r_{31})\xi(r_{12})\right], \qquad (2.26)$$

with $Q \approx 1$.

In order to use this data to estimate the average mass density $\bar{\rho}$, it is assumed that the galaxy distribution accurately traces the mass distribution. However, it is known that rich clusters are more strongly correlated than galaxies, with $\xi_{cc}(r) \approx 10\xi_{gg}(r)$, as shown in Fig. 2.8.[77–79] Thus rich clusters and galaxies cannot both be good tracers of the mass distribution; perhaps neither is. For the time being I will blithely ignore this cautionary aside. (It will come up again.)

Let us assume then that the mass is distributed like the galaxies, with $\delta\rho/\rho \sim \xi(r)$.[21] The mass of a typical bound clump of size $r \ll r_o$ is $M \sim \bar{\rho}\xi(r)r^3$, so by the virial theorem the internal velocity should be

$$v^2 \sim GM/r \propto r^{2-\gamma}. \qquad (2.27)$$

Peebles[80] has shown more precisely that

$$\langle v_{12}^2(r)\rangle = \frac{6G\bar{\rho}}{\xi(r)}\int_r^\infty \frac{dr}{r}\int d^3z \frac{\mathbf{r}\cdot\mathbf{z}}{z^3}\zeta(\mathbf{r},\mathbf{z},|\mathbf{r}-\mathbf{z}|), \qquad (2.28)$$

which is called the *cosmic virial theorem*. Using (2.25) and (2.26), (2.28) implies

that

$$\Omega \approx Q^{-1} \left(\frac{\langle v_{12}^2 \rangle^{1/2}}{800\,\mathrm{km\,s^{-1}}} \right)^2 \left(\frac{1h^{-1}\mathrm{Mpc}}{r} \right)^{2-\gamma}. \tag{2.29}$$

The redshift survey data[73,74] give

$$\langle v_{12}^2 \rangle^{1/2} = 300 \pm 50\,\mathrm{km\,s^{-1}}, \tag{2.30}$$

for $r \sim 1h^{-1}$ Mpc, with a weak dependence on r_{12} consistent with (2.27). Applying the cosmic virialtheorem with[81,73,21] $Q = 0.7 - 1.3$ gives $\Omega \approx 0.1 - 0.2$.

Davis and Peebles[73] discuss two other methods of extracting estimates of Ω from redshift data. One is based on the Irvine-Layzer *cosmic energy equation*, which relates the single-galaxy one-dimensional velocity dispersion $\bar{v}_p$ to the potential energy stored in fluctuations. With reasonable approximations, this yields $\Omega \approx (\bar{v}_p/660\,\mathrm{km\,s^{-1}})^2 \approx 0.2$. The second method is based on the assumption that the mass clustering on scales $\lesssim 1h^{-1}$ Mpc is statistically stable, neither expanding with the Hubble expansion nor collapsing. This leads to the expression

$$\sigma^2(r) \approx 4.13 Q (H_o r)^2 \xi(r) \Omega, \tag{2.31}$$

where $\sigma(r)$ is the pair-weighted one-dimensional relative velocity dispersion, given approximately by[73]

$$\sigma(r) = \sigma_o (h r_{\mathrm{Mpc}})^{0.13 \pm 0.04} \tag{2.32}$$

with $\sigma_o = 340 \pm 40$ km s^{-1}. Then

$$\Omega = Q^{-1} (\sigma_o / 900\,\mathrm{km\,s^{-1}})^2 = 0.20(1.5^{\pm 1}). \tag{2.33}$$

All these methods give estimates of the mass on scales of order 1 Mpc. It is perhap significant that they all agree that $\Omega(\sim 1\,\mathrm{Mpc}) \approx 0.1 - 0.3$, and that

this agrees with the estimates of the amount of dark matter around galaxies and clusters discussed in Lecture 1. But all of these estimates are insensitive to a possible component of dark matter that is not clustered on small scales but instead distributed rather uniformly.

Infall toward Virgo

The best method presently available for estimating the cosmic density on scales of ~ 10 Mpc uses the infall velocity v_V of the Local Group toward the center of the Local Supercluster, which is located in the Virgo cluster at a distance of about $10h^{-1}$ Mpc. In linear perturbation theory, the infall peculiar velocity at a radius R resulting from a mass excess δM distributed spherically within that radius is given by[58]

$$v_V = \frac{2G\delta M f(\Omega)}{3H_o R^2 \Omega}, \tag{2.34}$$

where $f(\Omega) \approx \Omega^{0.6}$ [see Ref. 72, eq. (14.8)]. Assuming that the distribution of bright galaxies N traces the mass distribution on large scales so that

$$\frac{\delta M}{M} \approx \frac{\delta N}{N} \equiv \delta, \tag{2.35}$$

and using eq. (2.15), it follows that

$$\Omega^{0.6} = (3v_V/v_H\delta), \tag{2.36}$$

where v_H is the unperturbed Hubble velocity.Sandage[69,82] takes $v_V = 200 \pm 50$ km s^{-1}, $v_H = 1170$ km s^{-1}, and $\delta = 2.8 \pm 0.5$, which implies $\Omega \approx 0.06$. Davis and Peebles[56] argue that the predominance of the evidence, especially the agreement between the velocity of the Local Group with respect to the cosmic background radiation and with respect to distant galaxies[57] , suggests rather that $v_V = 400 \pm 60$ km s^{-1}, and they take $\delta = 2.2 \pm 0.3$. With the linear spherical approximation (2.36), this gives $\Omega \approx 0.2$. As I discussed in the first lecture, the Local Supercluster is not very spherical. Davis and Peebles[58] obtain

$\Omega = 0.35 \pm 0.15$ in a nonlinear spheroidal model. However, even if the mass distribution is well represented by the galaxy distribution, if it is aspherical then v_V can reflect the effect of mass outside R,[83] adding further uncertainty to the determination of $\Omega(\sim 10\,\mathrm{Mpc})$. Perhaps in a few years the galaxy flow in the neighborhood of the Local Supercluster will be better understood through the interplay of theory and observation (especially needed are reliable distance indicators independent of redshift).[70]

Dynamics of Superclusters

The dynamics of other superclusters can also be used to estimate the value of Ω, especially as better data becomes available on the peculiar velocities of the galaxies and clusters within them. As an example of this approach, Harms *et al.*[84] used a spherical model to estimate the density required to account for the observed velocity dispersion of the galaxy clusters in the supercluster 1451 + 22. The observed velocity dispersion is about half that expected from unperturbed Hubble flow. Their model gave an average density within this supercluster between 1.01 and 1.99 $\rho_{c,o}$. They estimate that the space density of galaxies is enhanced in this supercluster by a factor between about 17 and 71. Making the crucial assumption that the mass density is enhanced by the same factor, it follows that $0.014 \leq \Omega \leq 0.12$. Relaxing the assumption of spherical symmetry would allow $\Omega \lesssim 0.3$. Similar results are obtained for other superclusters.[85]

Density of Hydrogen

You may wonder how much of the dark matter could be ordinary hydrogen. Gunn and Peterson pointed out that absorption of quasar light by intervening atomic hydrogen (H_I) would cause an absorption trough on the short-wavelength side of the Lyman-α line at 1216 Å, as sketched in Fig. 2.9. The Ly α line is conveniently redshifted into the visible range for quasars with $z > 1.5$, and the absence of such an absorption trough implies that

$$\Omega(\mathrm{H_I}) < 3 \times 10^{-7} h^{-1} \tag{2.37}$$

with a similar result for molecular hydrogen

$$\Omega(H_2) < 5 \times 10^{-5} h^{-1}. \tag{2.38}$$

Although there is no absorption trough, there are many discrete absorption lines in quasar spectra caused by small "Ly α clouds" of neutral hydrogen (this interpretation is confirmed by the presence of Ly β absorption as well). These Ly α clouds are important as cosmological tracers (more on that later), but their total mass is less than that of the luminous parts of galaxies.

What about ionized hydrogen? $\Omega(H^+) \ll 1$ from nonobservation of radiation, except possibly for plasma at a temperature of $\sim 3 \times 10^8$ K. The observed X-ray background in the range 3 keV $< h\nu \lesssim$ 50 keV could be produced by nearly a closure density of ionized hydrogen at this temperature — but an enormous amount of energy would be required to heat so much gas to so high a temperature, and another explanation would still be required for the X-ray background above $\sim$ 60 keV. Moreover, as I will explain in the next lecture, the standard theory of Hot Big Bang nucleosynthesis produces the observed abundances of deuterium, ^{3}He, and ^{4}He only if the primordial baryon abundance Ω_b lies between $0.01h^{-2}$ and $0.035h^{-2} \lesssim 0.14$.[86] The upper limit is (barely) consistent with all the dark matter being baryonic, but I will disuss other arguments against this in the next lecture.

Deceleration Parameter q_o

A way of determining Ω on very large scales is to measure the deceleration parameter q_o, given by eq. (2.17). If the cosmological constant vanishes, then $q_o = 2\Omega$. Although q_o can in principle be measured by determining the deviation of very distant objects from Hubble's law, in practice it has been impossible to determinine their distances very accurately. The traditional approach, based on the assumed constant luminosity of the brightest galaxies in each rich cluster, is frought with uncertainties — in particular, the effects of evolution (time

variation in absolute luminosity, caused for example by the aging of the stellar populations) and sampling bias (near and distant samples may not be comparable). Nevertheless a recent review[87] obtains an upper limit $q_o \leq 1$ from radio galaxies observed in the near-IR having redshifts in the range ~ 0.5 to ~ 1. Alternative approaches are unfortunately also problematic. Since quasars have by far the highest observed redshifts ($z \leq 3.8$), they would provide an ideal sample for determining q_o if some feature of their spectra could be used to determine their intrinsic luminosity. A recent study, exploiting an observed correlation between the strength of the C_{IV} (triply-ionized carbon) 1550Åemission line and the luminosity of the underlying continuum in flat-radio-spectrum quasars, finds $q_o = 1^{+1.5}_{-0.7}$ assuming no evolution.[88] This result may suffer from possible selection and evolution effects,[89] however, and it is based entirely on an empirical correlation whose origin is not well understood.

To summarize, the accurate measurement of the cosmological density parameter Ω is difficult, but it probably lies in the range $0.1 \leq \Omega \leq 2$. Large Ω, such as the Einstein-de Sitter value $\Omega = 1$, is excluded unless mass density is distributed considerably more broadly than luminosity.

2.5 Growth and Collapse of Fluctuations

The continuity or energy conservation equation (2.10) can be integrated, given an equation of state $p = p(\rho)$, to determine $\rho(R)$. Then the Einstein equation (2.8) can be integrated to give $R(t)$. Consider the equation of state $p = w\rho$, where w is a constant. Integrating (2.10) gives

$$\rho \propto R^{-3(1+w)}, \tag{2.39}$$

and then integrating (2.8) in the approximation that $k = 0$, which is always valid

at early times(*), gives

$$R \propto t^{2/3(1+w)}. \tag{2.40}$$

There are two standard cases:

Radiation dominated

$$w = 1/3, \quad \rho \propto R^{-4}, \quad R \propto t^{1/2} \tag{2.41}$$

Matter dominated

$$w = 0, \quad \rho \propto R^{-3}, \quad R \propto t^{2/3}. \tag{2.42}$$

The crossover between these two regimes occurs at $R = R_{eq}$, when relativistic particles (photons and N_ν species of two-component neutrinos of negligible mass) and nonrelativistic particles (ordinary and dark matter) make equal contributions to ρ:

$$R_{eq} = \frac{4\sigma T_o^4(1+\gamma)}{\Omega\rho_{c,o}c} = \frac{4.05\times10^{-5}}{\Omega h^2}\frac{1+\gamma}{1.681}\theta^4. \tag{2.43}$$

Here the scale factor R has been normalized so that $R_o \equiv R(t_o) = 1$; γ is the ratio of neutrino to photon energy densities (discussed further in Lecture 3),

$$\gamma \equiv \frac{\rho_{\nu,o}}{\rho_{\gamma,o}} = \frac{7}{8}\left(\frac{4}{11}\right)^{4/3} N_\nu \quad (= 0.681 \text{ for } N_\nu = 3); \tag{2.44}$$

σ is the Stefan-Boltzmann constant; and $\theta \equiv (T_o/2.7K)$. The contribution of relativistic particles to the cosmological density is very small today in the standard model; for example, the contribution of photons is $\Omega_{\gamma,o} = 3.0\times10^{-5}h^{-2}\theta^4$.

* The curvature term, which is $\propto R^{-2}$, is possibly important today. But in the early universe it is always much smaller than the density term, which is $\propto R^{-3}$ (matter dominated) or $\propto R^{-4}$ (radiation dominated).

It is also possible to obtain a simple expression for $t(R)$ that is valid in both radiation- and matter-dominated eras, for the case of a flat universe (i.e., $k = 0$). Simply integrate the Einstein equation (2.8) with

$$\rho = \rho_{rel} + \rho_{nonrel} \approx \rho_{c,o}\Omega_o(R_{eq}R^{-4} + R^{-3}), \tag{2.45}$$

The result is

$$t = \frac{2}{3}H_o^{-1}\Omega_o^{-1/2}R_{eq}\left[(R - 2R_{eq})(R + R_{eq})^{1/2} + 2R_{eq}^{3/2}\right], \tag{2.46}$$

with the following limiting behaviors:

$$\begin{aligned} R \ll R_{eq}: &\qquad t \approx \frac{1}{2}H_o^{-1}\Omega_o^{-1/2}R_{eq}^{-1/2}R^2 \\ R = R_{eq}: &\qquad t_{eq} = 0.3905H_o^{-1}\Omega_o^{-1/2}R_{eq}^{3/2} \\ R \gg R_{eq}: &\qquad t \approx \frac{2}{3}H_o^{-1}\Omega_o^{-1/2}R^{3/2}. \end{aligned} \tag{2.47}$$

It is now easy to calculate the mass M_H of nonrelativistic matter encompassed by the horizon (Hubble radius) $R_H = ct(R)$ as a function of scale factor R:

$$\begin{aligned} M_H &= \frac{4}{3}\pi R_H^3 \frac{\rho_{c,o}\Omega_o}{R^3} \\ &= \frac{2.41 \times 10^{15} M_\odot}{\Omega_o^2 h^4}\left[\frac{(y-2)(y+1)^{1/2} + 2}{y}\right]^3, \end{aligned} \tag{2.48}$$

where $y \equiv R/R_{eq}$. The behavior of M_H is sketched in Fig. 2.10 (heavy solid curve).

Top Hat Model

It is now time to consider the evolution of small fluctuations in the density. In the linear regime $\delta \equiv \delta\rho/\rho \ll 1$, the growth rate is independent of shape. It is simplest to consider a spherical ("top hat") fluctuation, say a region of radius $R(1 + a)$ with uniform density $\bar{\rho}(1 + \delta)$ in a background of density $\bar{\rho}$: see Fig. 2.11.

Consider first the growth of fluctuations in a matter dominated universe. Conservation of mass implies

$$\bar{\rho}(1+\delta)R^3(1+a)^3 = \text{constant}, \tag{2.49}$$

or

$$\delta = -3a. \tag{2.50}$$

Now it is necessary to bring in gravity:

$$\ddot{R} = \frac{-4\pi G}{3}(\rho + 3p)R. \tag{2.51}$$

(This equation follows by differentiating (2.8) with respect to time and using (2.10). Alternatively, it is the 00 component of Einstein's equations in the form $R_{\mu\nu} = -(8\pi G/c^4)(T_{\mu\nu} - \frac{1}{2}g_{\mu\nu}T^{\lambda}_{\lambda})$ applied to the Robertson-Walker metric.) Applying (2.51) to the background and to the fluctuation,

$$\ddot{R}(1+a) + 2\dot{R}\dot{a} + R\ddot{a} = -(4\pi G/3)\rho R(1+a+\delta),$$

or

$$\ddot{\delta} + 2(\dot{R}/R)\dot{\delta} = 4\pi G\rho\delta. \tag{2.52}$$

Substituting $(\dot{R}/R) = \frac{2}{3}t^{-1}$, valid for a flat ($k = 0$) matter-dominated universe, and trying $\delta = t^{\alpha}$, one finds $(\alpha+1)(\alpha - \frac{2}{3}) = 0$. The general solution of (2.52) is thus

$$\delta = At^{2/3} + Bt^{-1}. \tag{2.53}$$

Notice that the amplitude of the fluctuation in the growing mode has the same rate of growth as the scale factor R in the matter-dominated universe.

An analogous calculation for a radiation-dominated universe gives

$$\delta = At + Bt^{-1}. \tag{2.54}$$

This time the growing mode for the amplitude grows as the square of the scale factor (i.e., $\delta \propto R^2$) in the radiation-dominated universe. The solution (2.54) is actually relevant only on scales larger than the horizon, since once the fluctuations come within the horizon, the radiation and baryons start to oscillate and the neutrinos freely stream away. (I will discuss this further in Lecture 3.) One must be careful in discussing behavior on scales larger than the horizon, since the freedom to choose coordinates or gauge can complicate the physical interpretation. In these lectures I am using "time-orthogonal" coordinates and the "synchronous gauge" formalism.[65,72,90] (Bardeen's gauge invariant formalism is an attractive alternative.[91]) Indeed it may seem paradoxical even to consider fluctuations larger than the horizon — but it is necessary to do so, since all cosmologically interesting fluctuations are larger than the horizon at early times. What we are doing effectively is comparing the growth rates of universes differing slightly in density. The region of slightly higher density (the fluctuation) expands slightly more slowly; consequently, the density contrast δ between it and the background grows with time. (Birkhoff's theorem[66] permits us to ignore the universe outside our spherically symmetric fluctation.)

Since cosmological curvature is at most marginally important at the present epoch, it was negligible during the radiation-dominated era and at least the beginning of the matter-dominated era. But for $k = -1$, i.e. $\Omega < 1$, the growth of δ slows for $(R/R_o) \gtrsim \Omega_o$, as gravity becomes less important and the universe begins to expand freely. To discuss this case, it is convenient to introduce the variable

$$x \equiv \Omega^{-1}(t) - 1 = (\Omega_o^{-1} - 1)R(t)/R_o. \tag{2.55}$$

(Note that $\Omega(t) \longrightarrow 1$ at early times.) The general solution in the matter-

dominated era is then[92]

$$\delta = \widetilde{A} D_1(t) + \widetilde{B} D_2(t), \tag{2.56}$$

where the growing solution is

$$D_1 = 1 + \frac{3}{x} + \frac{3(1+x)^{1/2}}{x^{3/2}} \ln\left[(1+x)^{1/2} - x^{1/2}\right] \tag{2.57}$$

and the decaying solution is

$$D_2 = (1+x)^{1/2}/x^{3/2}. \tag{2.58}$$

These agree with the Einstein-de Sitter results (2.53) at early times ($t \ll t_o, x \ll 1$). For late times ($t \gg t_o, x \gg 1$) the solutions approach

$$D_1 = 1, D_2 = x^{-1}; \tag{2.59}$$

in this limit the universe is expanding freely and the amplitude of fluctuations stops growing.

Spherical Collapse

At early times, an overdense fluctuation expands with the Hubble flow. Eventually, however, it reaches a maximum radius, and then "turns around" and begins to contract, just like a small piece of a positive curvature Robertson-Walker universe. Continuing the analogy, one might suppose that it would collapse to a point — but of course it does not; "violent relaxation" rapidly brings it into virial equilibrium at a radius about half the maximum radius. Since the fluctuation is now well inside the horizon and there are no relativistic velocities, the Newtonian approximation is valid.

Figure 2.12 summarizes the collapse process with sketches of the radius, density, and density contrast as a function of scale factor R. This subsection and the next are devoted to filling in the details in this figure.

I will start by deriving an expression for the maximum radius r_m, and the time t_m at which it is reached, for a spherical "top-hat" fluctuation. As above, let the density in the fluctuation equal $\bar{\rho}(1+\delta)$, but let the radius be $r = r_i(R/R_i)+\xi$, where $\xi = 0$ at the initial time t_i. The initial time t_i is arbitrary, except that I will assume that it is in the matter-dominated era, that $\delta_i \ll 1$, and that the fluctuation is described by the growing mode $\delta \propto t^{2/3}$.

Conservation of mass ($= \rho r^3$) implies that the initial velocity at the edge of the spherical fluctuation is

$$v_i = H_i r_i + \dot{\xi}_i = H_i r_i - r_i \dot{\delta}_i/3, \tag{2.60}$$

so the corresponding kinetic energy per unit mass is

$$K_i = \frac{1}{2}H_i^2 r_i^2 - \frac{1}{3}H_i r_i^2 \dot{\delta}_i = \frac{1}{2}H_i^2 r_i^2 \left(1 - \frac{2}{3}\delta_i\right). \tag{2.61}$$

Since the potential energy per unit mass at the edge of the sphere is

$$W_i = -\frac{1}{2}H_i^2 r_i^2 \Omega_i (1+\delta_i), \tag{2.62}$$

the total energy per unit mass is

$$E = K_i + W_i = \frac{W_i}{1+\delta_i}\left[(1+\delta_i) - \frac{1}{\Omega_i}\left(1 - \frac{2}{3}\delta_i\right)\right]. \tag{2.63}$$

Maximum expansion corresponds to $K_m = 0$, so $E = W_m = (r_i/r_m)W_i$ and

$$\frac{r_m}{r_i} = \frac{W_i}{E} \approx \frac{1+\delta_i}{1 - \Omega_i^{-1} + \frac{5}{3}\delta_i}. \tag{2.64}$$

This result, derived by Blumenthal and me,[93] differs slightly from that given in Peebles[72] (§19) because I here assume a purely growing mode for δ_i and allow a

nonzero deviation of the expansion velocity from pure Hubble flow at t_i. It can be rewritten in terms of Ω_o using the fact that

$$\Omega_i = \Omega_o \frac{1+z_i}{1+\Omega_o z_i}; \tag{2.65}$$

namely,

$$\frac{r_m}{r_i} \approx \frac{(1+z_i)(1+\delta_i)}{1-\Omega_o^{-1}+\frac{5}{3}(1+z_i)\delta_i}. \tag{2.66}$$

The corresponding time can be calculated from standard Newtonian expressions. The force law

$$\ddot{r} = -GM/r^2$$

implies that

$$\dot{r}^2 = \frac{2GM}{r}\left(1-\frac{r}{r_m}\right),$$

which can be integrated giving

$$t_m = \left(\frac{\pi^2 r_m^2}{8GM}\right)^{1/2}. \tag{2.67}$$

The density in the top hat is then

$$\rho_m = \frac{3\pi}{32Gt_m^2}; \tag{2.68}$$

since the background density in the Einstein-de Sitter ($k=0$) approximation is $\bar{\rho} = (6\pi G t^2)^{-1}$,

$$\rho_m/\bar{\rho} = 9\pi^2/16 = 5.6, \tag{2.69}$$

and the density contrast is $\delta_m = 4.6$ at maximum expansion.

Violent Relaxation

Figure 2.13 shows the result of a computer "N-body simulation"[94] of the late stages of dissipationless gravitational collapse of a top-hat mass distribution: the bodies fall together (a) into a dense "crunch" (b), from which they emerge into a centrally condensed distribution (c) that remains remarkably stable thereafter. The process occurs rapidly, in a time on the order of the gravitational dynamical time

$$\tau = (G\rho)^{-1/2}. \tag{2.70}$$

It is called "violent relaxation."

The bound particles in the final configuration (c) accurately satisfy the virial theorem $\langle K \rangle = -\frac{1}{2}\langle W \rangle$. The potential energy varies inversely as the radius, $W = A/r$, so the radius after virialization r_v is given by

$$\frac{A}{r_m} = E = \frac{1}{2}\langle W \rangle = \frac{A}{2r_v}, \tag{2.71}$$

which implies that $r_v = \frac{1}{2}r_m$. As Fig. 2.13 illustrates, the radius shrinks only by roughly a factor of 2 in the collapse. (Actually, this "radius" is effectively defined by the last equality in (2.71); since the mass is redistributed in the collapse, it is somewhat arbitrary.)

Lynden-Bell[95] and Shu[96] have shown by statistical methods that the distribution resulting from dissipationless violent relaxation via chaotic changes of the collective gravitational field, with the total mass much greater than that of any component particles, is to an excellent approximation a Maxwell-Boltzmann distribution, but with components of different masses having the same velocity dispersion and not the same "temperature". In other words, the distribution is a Maxwellian in the velocities, independent of the mass. Such a distribution is nevertheless called an "isothermal sphere".[97] As I discussed in Lecture 1, constancy of the velocity implies that the total mass increases linearly with radius, or equivalently that the density falls as r^{-2}, outside the central core; this is

roughly what is found in computer simulations. Of course, this can only be true for intermediate values of r, since the total mass is finite. Another way of saying this is that high energy orbits with periods longer than the collapse time cannot be very fully populated.[21] Thus the density falls faster than r^{-2} at large r.

In any case, the simple model of a spherical top-hat initial distribution is rather unrealistic in at least two respects: it is likely that the initial density distribution is smoothly peaked rather than a step function, and moreover somewhat aspherical. The outer parts of the initial dark matter density fluctuation will collapse later, perhaps resulting in a large constant-velocity halo with density falling roughly as r^{-2} to considerable distances.[98] Asphericity is amplified in the collapse, and the most probable result is that the collapse will actually occur in one direction first: "pancake collapse".[99] This can happen even if the bulk of the matter in the fluctuation is not even bound, so that the expansion continues in the perpendicular directions; this is a popular model for the origin of superclusters.[100–102] In the case of protogalaxies, the subsequent violent relaxation of a flattened intermediate configuration produces an ellipsoidal rather than a spherical virialized distribution;[103] perhaps this is the typical shape of galactic halos.

A key feature of the dark matter is that it is dissipationless, whereas ordinary (baryonic) matter can convert its kinetic energy into radiation and thereby cool via bremsstrahlung (also called by astrophysicists "free-free scattering"), Ly α and Ly β radiation, and excitation of molecular and metallic energy levels. If the ordinary matter and dark matter are initially well mixed (which is a plausible initial condition before violent relaxation, at least in the cold DM picture, as I will discuss in Lecture 4), then dissipation during the "crunch" and afterward will cause the baryonic matter to sink to the center. The baryonic matter can radiate away energy but not angular momentum. If the dissipative collapse is halted by angular momentum, a disk will result. If it is halted by star formation (stars have negligible collision cross sections), then a spheroidal system will result. These are of course the two elements of galaxy structure.

Presumably the processes just discussed occur on a variety of scales. If, as usually assumed, smaller-mass fluctuations have higher amplitudes, then they will turn around and virialize within larger-mass fluctuations, which subsequently themselves virialize, and so on until the present. The virialization of the next larger scale of the clustering hierarchy will tend to disrupt the smaller-scale structures within it. The crucial question for galaxy formation in this gravitational collapse picture is: What sets the mass scale of galaxies? (Recall that most of the mass in galaxies is in big galaxies whose mass is within an order of magnitude of that of the Milky Way.) At least two factors must be considered: the initial fluctuation spectrum and its modification as the universe evolves, and the rate of dissipation compared to gravitational collapse on different scales.

I will return to this in Lecture 4. But first, in order to begin to discuss the fluctuation spectrum, I must ask where the fluctuations themselves came from.

2.6 Inflation and the Origin of Fluctuations

The basic idea of inflation is that before the universe entered the present adiabatically expanding Friedmann era, it underwent a period of de Sitter exponential expansion of the scale factor, termed *inflation.*[104]

The de Sitter cosmology corresponds to the solution of Friedmann's equation in an empty universe (i.e., with $\rho = 0$ or, in (2.12), $M = 0$) with vanishing curvature ($k = 0$) and positive cosmological constant ($\Lambda > 0$). The solution is

$$R = R_o e^{Ht}, \tag{2.72}$$

with constant Hubble parameter

$$H = (\Lambda/3)^{1/2}. \tag{2.73}$$

There are analgous solutions for $k = +1$ and $k = -1$ with $R \propto \cosh Ht$ and $R \propto \sinh Ht$ respectively. The scale factor expands exponentially because the positive

cosmological constant corresponds effectively to a negative pressure. de Sitter space is discussed in textbooks on general relativity (for example Refs. 63 and 105) mainly for its geometrical interest. Until recently, the chief significance of the de Sitter solution (2.72)in cosmology was that it is a kind of limit to which all indefinitely expanding models with $\Lambda > 0$ must tend, since as $R \to \infty$, the cosmological constant term ultimately dominates the right hand side of the Friedmann equation (2.12).

As Guth[104] emphasized, the de Sitter solution might also have been important in the very early universe because the vacuum energy that plays such an important role in spontaneously broken gauge theories also acts as an effective cosmological constant. A period of de Sitter inflation preceeding ordinary radiation-dominated Friedmann expansion could explain several features of the observed universe that otherwise appear to require very special initial conditions: the horizon, smoothness, flatness, rotation, and monopole problems. (A number of other people independently appreciated the power of an initial de Sitter period to generate desirable initial conditions for a subsequent Friedmann era.[106,107] A paper by Kazanas[108] is apparently the first published discussion of this in the context of grand unified theories.)

I will illustrate how inflation can help with the horizon problem. At recombination ($p^+ + e^- \to H$), which occurs at $R/R_o \approx 10^{-3}$, the mass encompassed by the horizon was $M_H \approx 10^{18} M_\odot$, compared to $M_{H,o} \approx 10^{22} M_\odot$ today. Equivalently, the angular size today of the causally connected regions at recombination is only $\Delta\theta \sim 3°$. Yet the fluctuation in temperature of the cosmic background radiation from different regions is so small that only an upper limit is presently available: $\Delta T/T < 10^{-4}$. How could regions far out of causal contact have come to temperatures which are so precisely equal? This is the "horizon problem". With inflation, it is no problem because the entire observable universe initially lay inside a single causally connected region that subsequently inflated to a gigantic scale.

This is illustrated in Fig. 2.14. The Hubble parameter $\dot{R}/R$ is constant in size during the de Sitter era; then, after reheating, the horizon size of course just grows linearly with time. A region of size r_1, initially smaller then the de Sitter horizon, inflates to a size much larger than the de Sitter horizon and is no longer causally connected (dots). After reheating, r_1 expands with the scale factor ($\propto t^{1/2}$ in the radiation-dominated Friedmann era) and eventually crosses back inside the horizon. The curve labeled r_2 shows the similar fate of a larger region. The region encompassed by the present horizon presumably all lay within a region like this that started smaller than the de Sitter horizon.

In inflationary models, the dynamics of the very early universe is typically controlled by the self-energy of the Higgs field associated with the breaking of a Grand Unified Theory (GUT) into the standard 3-2-1 model: GUT$\rightarrow SU(3)_{color} \otimes [SU(2) \otimes U(1)]_{electroweak}$. This occurs when the cosmological temperature drops to the unification scale $T_{GUT} \sim 10^{14}$ GeV at about 10^{-35} s after the Big Bang. Guth[104,109] initially considered a scheme in which inflation occurs while the universe is trapped in an unstable state (with the GUT unbroken) on the wrong side of a maximum in the Higgs potential. This turns out not to work: the transition from a de Sitter to a Friedmann universe never finishes.[110] The solution in the "new inflation" scheme[111] is for inflation to occur *after* barrier penetration (if any). It is necessary that the Higgs potential be nearly flat (i.e. decrease very slowly with increasing Higgs field) for the inflationary period to last long enough. This nearly flat part of the Higgs potential must then be followed by a very steep minimum, in order that the energy contained in the Higgs potential be rapidly shared with the other degrees of freedom ("reheating").

It turns out to be necessary to inflate by a factor $\gtrsim e^{66}$ in order to solve the flatness problem, i.e. that $\Omega_o \sim 1$. (With $H^{-1} \sim 10^{-34}$ s during the de Sitter phase, this implies that the inflationary period needs to last for only a relatively small time $\tau \gtrsim 10^{-32}$ s.) The "flatness problem" is essentially the question why the universe did not become curvature dominated long ago. Neglecting the cosmological constant on the assumption that it is unimportant after the

inflationary epoch, the Friedmann equation can be written –

$$\left(\frac{\dot{R}}{R}\right)^2 = \frac{8\pi G}{3}\frac{\pi^2}{30}g(T)T^4 - \frac{kT^2}{(RT)^2}, \tag{2.74}$$

where the first term on the right hand side is the contribution of the energy density in relativistic particles and $g(T)$ is the effective number of degrees of freedom (discussed in detail in Lecture 3). The second term on the right hand side is the curvature term. Since $RT \approx$ constant for adiabatic expansion, it is clear that as the temperature T drops, the curvature term becomes increasingly important. The quantity $K \equiv k/(RT)^2$ is a dimensionless measure of the curvature.[112] Today, $|K| = |\Omega - 1|\, H_o^2/T_o^2 \leq 2 \times 10^{-58}$. Unless the curvature exactly vanishes, the most "natural" value for K is perhaps $K \sim 1$. Since inflation increases R by a tremendous factor $e^{H\tau}$ at essentially constant T (after reheating), it increases RT by the same tremendous factor and thereby decreases the curvature by that factor squared. Setting $e^{-2H\tau} \lesssim 2 \times 10^{-58}$ gives the needed amount of inflation: $H\tau \gtrsim 66$. This much inflation turns out to be enough to take care of the other cosmological problems mentioned above as well.[113]

Of course, this is only the minimum amount of inflation needed; the actual inflation might have been much greater. Indeed it is frequently argued that since the amount of inflation is a tremendously sensitive function of the initial value of the Higgs field (for example), it is extremely likely that there was much more inflation than the minimum necessary to account for the fact that Ω_o is of order unity.[114] It then follows that the curvature constant is probably vanishingly small after inflation, which implies (in the absence of a cosmological constant today) that $\Omega_o = 1$ to a very high degree of accuracy. However, in view of our lack of knowledge of the true dynamics of the inflationary epoch (assuming that there really was one), it is at least conceivable that increasing inflation becomes increasingly less likely.[26] This could happen, for example, through the effects of a "compensating field" of the sort proposed by Abbott[115] to explain why the

cosmological constant is so small today. Then we might live in a part of the universe that happened to inflate only enough to make $\Omega_0 \approx 0.2$.

Thus far, I have sketched how inflation stretches, flattens, and smooths out the universe, thus greatly increasing the domain of initial conditions that could correspond to the universe that we observe today. But inflation also can explain the origin of the fluctuations necessary in the gravitional instability picture of galaxy and cluster formation. Recall that the very existence of these fluctuations is a problem in the standard Big Bang picture, since these fluctuations are much larger than the horizon at early times (see Fig. 2.10). How could they have arisen?

The answer in the inflationary universe scenario is that they arise from quantum fluctuations in the scalar field ϕ whose vacuum energy drives inflation. The scalar fluctuations $\delta\phi$ during the de Sitter phase are of the order of the Hawking temperature $H/2\pi$. Because of these fluctuations, there is a time spread $\Delta t \approx \delta\phi/\dot{\phi}$ during which different regions of the same size complete the transition to the Friedmann phase. The result is that the density fluctuations when a region of a particular size re-enters the horizon are equal to[116]

$$\delta_H \equiv \left(\frac{\delta\rho}{\rho}\right)_H = 2^{3/2}\Delta t. \tag{2.75}$$

The time spread Δt can be estimated from the equation of motion of ϕ (the free Klein-Gordon equation in an expanding universe)

$$\ddot{\phi} + 3H\dot{\phi} = -(\partial V/\partial\phi). \tag{2.76}$$

Neglecting the $\ddot{\phi}$ term, since the scalar potential V must be very flat in order for enough inflation to occur, $\dot{\phi}$ and hence δ_H will be essentially constant. These are fluctuations of all the contents of the universe, so they are adiabatic fluctuations.

Thus *inflation predicts the constant curvature spectrum*

$$\delta_H = \text{constant}. \tag{2.77}$$

Some time ago Harrison,[117] Zeldovich,[118] and others had emphasized that this is the only scale-invariant (i.e., power-law) fluctuation spectrum that avoids trouble at both large and small scales. If

$$\delta_H \propto M_H^{-\alpha}, \tag{2.78}$$

then if $-\alpha$ is too large the universe will be less homogeneous on large than small scales, contrary to observation; and if α is too large, fluctuations on sufficiently small scales will enter the horizon with $\delta_H \gg 1$ and collapse to black holes;[119,120] thus $\alpha \approx 0$.

Inflation predicts more: it allows the calculation of the value of the constant δ_H in terms of the properties of the scalar potential $V(\phi)$. Indeed, this has proved to be embarrassing, at least initially, since the Coleman-Weinberg potential, the first potential studied in the context of the new inflation scenario, results in $\delta_H \sim 10^2$,[116] some six orders of magnitude too large. But this does not seem to be an insurmountable difficulty. A prescription for a suitable potential has been given,[121] and particle physics models that are more or less satisfactory have been constructed.[122]

Thus inflation at present appears to be a plausible solution to the problem of providing reasonable cosmological initial conditions (although it sheds no light at all on the fundamental question why the cosmological constant is so small now). In particular, it predicts the constant curvature fluctuation spectrum $\delta_H = \text{constant}$, at the price of also predicting that the universe is essentially flat. Inflation is not the only way to get the constant curvature spectrum, however; there is also the possibility of cosmic strings.[123] Discussing cosmic strings would take us rather far afield. I just want to note here that even though they have the

same spectrum, the fluctuations generated by the motion of relativistic strings are rather different from those arising from quantum fluctuations of an essentially free field. In particular, the latter are Gaussian.[124]

2.7 Is the Gravitational Force $\propto r^{-1}$ at Large r?

In concluding this lecture on gravity and cosmology, I return to the question whether our conventional theory of gravity is trustworthy on large scales. The reason for raising this question is that interpreting modern observations within the context of the standard theory leads to the conclusion that at least 90% of the matter in the universe is dark. Moreover, there is no observational confirmation that the gravitational force falls as r^{-2} on galactic and extragalactic scales.

Tohline[125] pointed out that a modified gravitational force law, with the gravitational acceleration given by

$$a = \frac{GM_{lum}}{r^2}\left(1 + \frac{r}{d}\right), \tag{2.79}$$

could be an alternative to dark matter galactic halos as an explanation of the constant-velocity rotation curves of Fig. 1.3. (I have written the mass in (2.79) as M_{lum} to emphasize that there is not supposed to be any dark matter.) Indeed, (2.79) implies

$$v^2 = \frac{GM_{lum}}{d} = \text{constant}$$

for $r \gg d$. The trouble is that, with the distance scale d where the force shifts from r^{-2} to r^{-1} taken to be a physical constant, the same for all galaxies, this implies that $M_{lum} \propto v^2$, whereas observationally $M_{lum} \propto L \propto v^4$, as I mentioned in Lecture 1 ("Tully-Fisher law").

Milgrom[126] proposed an alternative idea, that the separation between the classical and modified regimes is determined by the value of the gravitational

acceleration a rather than the distance scale r. Specifically, Milgrom proposed that

$$\begin{aligned} a &= GM_{lum}r^{-2}, \quad a \gg a_o \\ a^2 &= GM_{lum}r^{-2}a_o, \quad a \ll a_o \end{aligned} \tag{2.80}$$

where the value of the critical acceleration $a_o \approx 8 \times 10^{-8} h^2 \,\mathrm{cm\,s^{-1}}$ (where h is the Hubble parameter) is determined for large spiral galaxies with $M_{lum} \sim 10^{11} M_\odot$. (This value for a_o happens to be numerically approximately equal to cH_o.) Eq. (2.80) implies that

$$v^4 = a_o G M_{lum} \tag{2.81}$$

for $a \ll a_o$, which is now consistent with the Tully-Fisher law $M_{lum} \propto v^4$. However, there is a problem with (2.80): data for the largest elliptical galaxies still require the existence of large amounts of dark matter.[127] For example, the cD galaxy in the Abell cluster A2029 has $M_{lum} \gtrsim 1.5 \times 10^{13} M_\odot$, which implies that the gravitational acceleration is given by the usual expression (i.e, that $a > a_o$) for $r \lesssim 100$ kpc. But the observed increasing velocity dispersion over this region implies that the mass-to-light ratio increases by a factor ~ 10 in this region. Similarly, data on M87, the giant elliptical in the Virgo cluster, imply that $a > a_o$ out to about 84 kpc, where $M/L \approx 50$. Since $M/L \approx 10$ for the nucleus, here again M/L rises dramatically with r in the supposed "classical" region. There are also problems with (2.80) in spiral galaxies; for example, eq. (2.81) predicts a universal constant in the Tully-Fisher relation, but, as Fig. 1.5 shows, the proportionality constant is different for early and late spiral types.

Thus the proposed modifications of gravity are not entirely satisfactory empirically. They are also entirely ad hoc. Indeed, it would doubtless be difficult if not impossible to fit a r^{-1} force law into the larger framework of either cosmology or theoretical physics.[128] For example, all one needs to assume in order to get the weak-field limit of general relativity is that gravitation is carried by a massless spin-two particle (the graviton); masslessness implies the standard r^{-2} force, and spin two implies coupling to the energy-momentum tensor.[129] It is not at

all clear what sort of particle physics could lead to a force law like (2.80). In the absence of an intrinsically attractive and plausible theory of gravity which leads to a r^{-1} force law at large distances, it seems to me to be preferable by far to take dark matter seriously. Moreover, dark matter is quite consistent with modern ideas in particle physics; indeed, there is an abundance of plausible particle candidates.

3. Dark Matter

3.1 The Hot Big Bang

The main subject of this lecture is the properties of the cosmological dark matter (DM). Since the various arguments regarding the properties of the DM depend in several ways on our theories regarding the evolution of the universe, I begin by reminding you about the standard theory. For more details, you may want to consult some standard references.[65,130–133] Table 2 and Fig. 3.1 summarize several major signposts in cosmic history according to the standard Hot Big Bang theory. (The numerical entries in the Table are estimates for orientation; precise values depend on cosmological parameters.)

Table 2

time	temperature	
$t_{\rm QCD} \approx 10^{-4}$ s	$\sim 10^2$ MeV	π and μ annihilation; color confinement
$t_{\nu d} \approx 1$ s	1 MeV	neutrino decoupling
$t_e \approx 4$ s	0.5 MeV	e annihilation
$t_D \approx 3$ min	0.1 MeV	D bottleneck, He synthesis
$t_{eq} \approx 3 \times 10^4$ y	2 eV	nonrelativistic matter domination
$t_{rec} \approx 4 \times 10^5$ y	0.3 eV	atomic H formation ("recombination")
$t_o \approx 15$ Gy	3×10^{-4} eV	present epoch

The hadronic era comes to an end at $T_{\rm QCD} \sim 10^2$ MeV, with quantum chromodynamic confinement of colored hadrons (quarks and gluons) into ordinary baryons and mesons, with only the slight excess of baryons over antibaryons surviving after π annihilation. There is still some uncertainty regarding the physics of this era, with such exotic possibilities as a substantial fraction of the baryonic matter ending up as u-d-s symmetric "quark nuggets" — a possible candidate for the dark matter.[134] Thereafter, the basic physical processes are thought to

be well understood, with the major uncertainties being the fluctuation spectrum and the nature of the dark matter. Shortly after νs decouple thermally (1 in Fig. 3.1), e^+e^- annihilation decreases ρ_m and heats the photon gas (2), which thereafter resumes its adiabatic decrease in temperature $\propto R^{-1}$. Since there are $\sim 10^9$ photons per baryon, there are many photons available in the high-energy tail of the distribution to photodissociate D as soon as it forms, so the temperature must decrease still further before primordial nucleosynthesis can begin. Finally the "deuterium bottleneck" is passed (3), and then most of the D is quickly bound into ^{4}He. Galaxy-size masses are first encompassed by the horizon when the scale factor $R \sim 10^7$ (see Fig. 2.10). Then fluctuations in the neutrino density are damped by "free streaming"; and fluctuations in the fluid composed of radiation and ionized hydrogen and helium do not grow — rather, they oscillate and (for adiabatic fluctuations) eventually are damped by photon diffusion ("Silk damping"). The density of nonrelativistic matter falls more slowly ($\propto R^{-3}$) than that of relativistic particles (photons and light neutrinos) ($\propto R^{-4}$). They are equal at R_{eq} (see eq. (2.43)), and nonrelativistic matter dominates thereafter in the standard model. Fluctuations in the dark matter can then begin to grow $\propto R$. Finally, at $R_{rec} \approx 10^{-3}$ (4 in Fig. 3.1), hydrogen atoms form and the universe becomes transparent. No longer tied to the radiation by Compton drag, ordinary matter can begin to form the large astronomical objects we see today: globular clusters, galaxies, clusters and superclusters.

In the next section, I will consider this story in a little more detail, and point out problems that arise if the nonrelativistic matter is only baryonic. I will begin by discussing evidence that the dark halos of galaxies are not composed of baryonic matter. If the dark matter is not baryonic, what is it? The rest of this lecture and part of the next will be concerned with a classification of dark matter candidates by their key astrophysical properties, and an outline of their consequences for the formation of structure in the universe.[135]

3.2 The Dark Matter is Probably Not Baryonic

There are three arguments that the DM is not "baryonic", that is, that it is *not* made of protons, neutrons, and electrons as all ordinary matter is. As Richard Feynman has said in other contexts, *one* argument would suffice if it were convincing. The three arguments are based on (1) excluding various possible forms of baryonic dark matter in galaxy halos; (2) bounding the abundance of baryonic matter using the observed abundance of light elements, especially deuterium; and (3) bounding the magnitude of adiabatic fluctuations at recombination from the observational upper limits on fluctuations in the cosmic background radiation. All three arguments have loopholes, and I will point them out. Nevertheless, taken together these arguments persuade me, and perhaps will persuade you, that we must take very seriously the possibility that most of the matter in the universe is not composed of atoms.

Excluding Baryonic Models

If the dark matter in galaxy halos is baryonic, then it must be gaseous, or agglomerations of atoms held together by chemical or gravitational forces. But arguments can be given against all of these possibilities:[11,136] The dark matter in galaxy halos cannot be *gas* (it would have to be hot to be pressure supported, and would radiate X-rays that are not seen); nor frozen hydrogen *"snowballs"* (they would sublimate); nor *dust grains* (their "metals", elements of atomic number $\geq$ 3, would have prevented formation of the observed low-metallicity Population II stars); nor isolated *"jupiters"* (how to make so many hydrogen balls too small to initiate nuclear burning without making a few large enough to do so?); nor *collapsed stars* (where is the matter they must have ejected in collapsing?).

The weakest argument is probably that which attempts to exclude "jupiters": arguments of the form "how could it be that way?" are rarely entirely convincing.

Deuterium Abundance

In the early universe, almost all the neutrons remaining after the deuterium

bottleneck are synthesized into ^{4}He. (Formation of nuclides with $A \geq 6$ is inhibited by the absence of any stable nuclide with $A = 5$.) The fraction remaining in D and ^{3}He is calculated[86] to be a rapidly decreasing function of η, the ratio of baryon to photon number densities, as shown in Fig. 3.2. The presently observed D abundance (compared, by number to H), is $(1-4) \times 10^{-5}$. Since D is readily consumed but not produced in stars, 10^{-5} is also a lower limit on the primordial D abundance. This, in turn, implies an *upper* limit $\eta \leq 10^{-9}$ or

$$\Omega_b = 0.0035 h^{-2} \theta^3 \eta_{10} \leq 0.035\, h^{-2} \theta^3 \tag{3.1}$$

where Ω_b is the ratio of the present average baryon density ρ_b to the critical density $\rho_{c,o}$, $\theta \equiv (T_o/2.7\mathrm{K})$, and $\eta_{10} \equiv \eta/10^{-10}$. If use is also made of the somewhat more uncertain upper limit on primordial ^{7}Li abundance[137] of (^{7}Li/H) $\lesssim 3 \times 10^{-10}$ (including an additional factor of two for good measure) then stronger upper limits are obtained: $\eta_{10} \leq 7$, and $\Omega_b \leq 0.025 h^{-2} \theta^3$. Even with the Hubble parameter at its lower limit $h = 0.5$, this corresponds to $\Omega_b \leq 0.1\, \theta^3$.

As discussed in Lectures 1 and 2, the observational limits on Ω are $0.1 \lesssim \Omega \lesssim 2$. Therefore, in a baryon-dominated universe ($\Omega \approx \Omega_b$), these bounds are consistent only with the lower limit on Ω, and then only for the Hubble parameter at its lower limit. An Einstein-de Sitter or inflationary ($\Omega = 1$) or closed ($\Omega > 1$) universe cannot be baryonic.

Galaxy and Cluster Formation

In the gravitational collapse model for the formation of large scale structure in the universe, discussed in Lecture 2, structure forms when fluctuations $\delta \equiv \delta\rho/\rho$ grow to nonlinearity ($\delta \geq 1$), when they cease to expand with the Hubble flow, and subsequently collapse and virialize. As we will see, the problem is to understand how fluctuations of galaxy and cluster size can grow to nonlinearity by the present without violating the observational bounds on small-angle fluctuations in the cosmic background radiation.

As before, consider a universe with no non-baryonic dark matter. Before recombination, ionized matter is locked to radiation by Compton scattering (this is called "Compton drag"), and it is appropriate to treat radiation plus baryonic matter as an ideal fluid. Fluctuations in this fluid of wavelength $\lambda = 2\pi R/k$ in the expanding universe obey the wave equation[138]

$$\ddot{\delta}_k + 2\frac{\dot{R}}{R}\dot{\delta}_k = 4\pi G\rho\delta_k\left[1 - \frac{\pi k_B T}{\mu\lambda^2 G\rho}\right], \tag{3.2}$$

where $\mu \approx 0.78$ is the mean mass per particle of the primordial fluid of ionized hydrogen and helium. Growth of the amplitude of fluctuations occurs only when the right hand side of (3.2) is positive, i.e. for λ satisfying the Jeans criterion

$$\lambda > \lambda_J = \left(\frac{\pi k_B T}{\mu G\rho}\right). \tag{3.3}$$

That such a condition (first derived for a non-expanding universe by James Jeans in 1902) should arise is not surprising: the effects of gravity grow with the physical size of the fluctuation, and there is a minimum size above which gravity overwhelms pressure. Equivalently, the Jeans criterion is that the sound-crossing time t_{sound} should exceed the free fall time t_{dyn}:

$$t_{sound} \approx \lambda/\sqrt{k_B T/\mu} \;>\; t_{dyn} \approx (G\rho)^{-1/2}. \tag{3.4}$$

The value of the corresponding Jeans mass

$$M_J = (4\pi/3)\rho(\lambda/2)^3 \tag{3.5}$$

was sketched in Fig. 2.10 and is also included in Fig. 3.3. In the radiation-dominated era, the Jeans mass M_J is comparable to the mass inside the horizon M_H; between matter domination and recombination, M_J levels out at $\sim 10^{17} M_\odot$; and at recombination M_J drops to about $10^6 M_\odot$ and decreases thereafter as the matter temperature drops.

Now consider what happens when a fluctuation of galaxy size comes within the horizon. (That is, consider a horizontal line across Fig. 3.3 at $M \sim 10^{11} M_\odot$.) Since the mass is below the Jeans mass, the fluctuation amplitude cannot grow. Instead, the radiation and ionized matter fluid oscillates as an acoustic wave.

I will further assume that fluctuations in baryonic matter density and in radiation density are correlated: $\delta_r = (4/3)\delta_b$. These are called adiabatic fluctuations, since the entropy per baryon is constant. These are the sort of fluctuations predicted in most grand unified models, in which $\eta \equiv n_b/n_\gamma$ is a constant determined by the microphysics.[139]

It is important to appreciate that photon diffusion damps small adiabatic fluctuations (Silk damping). The photon mean free path is $\ell_T = (\sigma_T n_e)^{-1}$, where σ_T is the Thomson cross section and n_e is the electron density. Thus the time to diffuse (random walk) a distance d is $\tau(d) = d^2/c\ell_T$. Setting this equal to the Hubble time gives the Silk damping length $d_S \approx (\ell_T ct)^{1/2}$, with the corresponding Silk mass $M_S = (4\pi/3)\rho_m d_S^3$. The Silk damping mass is sketched in Fig. 2.10 and 3.3. M_S grows until recombination, and there is strong damping of any fluctuations that lie in the hatched region below the Silk damping line.

Accurate treatment of Silk damping requires numerical calculations.[140] The result is that there is more than e^{-1} damping of fluctuations with mass smaller than[21]

$$M_S \approx 1.3 \times 10^{12} (\Omega h^2)^{-3/2} M_\odot \ . \tag{3.6}$$

Evaluating this for $\Omega = 0.1$ (consistent with the primordial nucleosynthesis bound) and $h = 1$ gives $M_S = 4 \times 10^{13}$; using $h = \frac{1}{2}$ gives $M_S = 3 \times 10^{14}$. These masses correspond to clusters of galaxies. Thus ordinary galaxies ($M_b \leq 10^{11-12} M_\odot$) can form only after the collapse of larger-scale perturbations. These would be most likely to collapse first in one dimension ("pancake" collapse), especially because of the Silk damping of smaller-scale fluctuations.[141,142]

That galaxies form in this indirect manner is a complication of this scenario,

but it is not necessarily an argument against it. The serious problem is to get large enough fluctuations. Fluctuations δ grow linearly with the scale factor

$$\delta \propto R = (1+z)^{-1} = T_o/T \tag{3.7}$$

once the universe becomes matter dominated, but fluctuations smaller than the Jeans mass ($\sim 10^{17} M_\odot$) are prevented from doing so until recombination. Moreover, growth of δ slows when an open universe goes into free expansion, when $z \lesssim \Omega^{-1}$. Thus in a baryonic universe, δ grows only between the epoch of hydrogen recombination ($z_r \approx 10^3$) and $z = \Omega^{-1} \gtrsim 10$ (when free expansion begins). It follows that there is at most a factor of $\sim 10^2$ growth — see Fig. 3.4. In order to form galaxies by the present, it is necessary that at recombination $\delta T/T = \frac{1}{3}\delta\rho/\rho \geq 3 \times 10^{-3}$ for $M \geq M_S$, which corresponds to fluctuations on observable angular scales of a few arc minutes today. Such temperature fluctuations are more than an order of magnitude larger than present observational upper limits.[143]

The main potential loophole in this argument is the assumption of adiabatic perturbations. It is true that the orthogonal mode, perturbations in baryonic density which are uncorrelated with radiation (called isothermal perturbations), do not arise naturally in most currently fashionable particle physics theories where baryon number is generated in the decay of massive grand unified theory bosons ("GUT baryosynthesis"), since in such theories $\eta \equiv n_b/n_\gamma$ is determined by the underlying particle physics and should not vary from point to point in space. Nevertheless, there may be ways of generating isothermal density fluctuations on scales much larger than the horizon during GUT baryosynthesis.[144] And galaxies originating as isothermal fluctuations do avoid both Silk damping and contradiction with present $\delta T/T$ limits.

A second loophole is the possibility that matter was reionized at some $z_i \geq 10$, by hypothetical very early sources of uv photons.[145] Then the fluctuations in $\delta T/T$ at recombination associated with baryonic proto-pancakes could be washed

out by rescattering. Suppose that the universe was entirely reionized at a redshift z_i. The optical depth in units of the Hubble radius is then

$$\tau(z_i) = (ct_H)^{-1} \int_0^{z_i} \sigma_T n_e \, d\ell = \frac{H}{c} \int_0^{z_i} \sigma_T n_{b,o} (1+z)^3 \frac{dz}{(1+z)^2(1+\Omega z)^{1/2}}$$
$$\approx \frac{2H}{3c} \sigma_T n_{b,o} [(1+z_i)^{3/2} - 1] \approx 7 \times 10^{-3} \frac{\Omega_{b,o}}{0.1} h z_i^{3/2}. \tag{3.8}$$

where

$$\sigma_T = \frac{8\pi e^4}{3m_e^2} = 0.6652 \times 10^{-24} \, \mathrm{cm}^2 \tag{3.9}$$

is the Thomson cross section. Setting $\tau(z_i)$ equal to unity, it follows that, to be effective in washing out small-angle fluctuations in the cosmic background radiation, any reionization must have occurred before

$$z_i \approx 30 \left(\frac{\Omega_{b,o} h}{0.1} \right)^{2/3}; \tag{3.10}$$

that is, before $z_i \approx 30$ for the maximal values $\Omega_b = 0.1$ and $h = 1$, and before $z_i \approx 10^2$ for $\Omega_b = 0.03$ and $h = \frac{1}{2}$. Since these redshifts are earlier than the period of galaxy formation according to most theories, especially those in which galaxies form only after the collapse of cluster-size pancakes, it is difficult to see how enough matter could have been converted to radiation to cause reionization at such early times.

Despite the loopholes in each of the three aguments against a universe with no non-baryonic dark matter, I find all the arguments together to be rather persuasive, even if not entirely compelling. If it is indeed true that the bulk of the mass in the universe is not baryonic, that is yet another blow to anthropocentricity: not only is man not the center of the universe physically (Copernicus) or biologically (Darwin), it now appears that we and all that we see are not even made of the predominant variety of matter in the universe!

3.3 Three Types of DM Particles: Hot, Warm, and Cold

If the dark matter is not baryonic, what *is* it? I will consider here the physical and astrophysical implications of three classes of elementary particle DM candidates, which are called hot, warm, and cold.[146]

Hot DM refers to particles, such as neutrinos, that were still in thermal equilibrium after the most recent phase transition in the hot early universe, the QCD confinement transition, which presumably took place at $T_{\rm QCD} \approx 10^2$ MeV. Hot DM particles have a cosmological number density roughly comparable to that of the microwave background photons, which implies an upper bound to their mass of a few tens of eV. As I shall discuss shortly, this implies that free streaming destroys any fluctuations smaller than supercluster size, $\sim 10^{15} M_\odot$.

Warm DM particles interact much more weakly than neutrinos. They decouple (i.e., their mean free path first exceeds the horizon size) at $T \gg T_{\rm QCD}$, and are not heated by the subsequent annihilation of hadronic species. Consequently their number density is roughly an order of magnitude lower, and their mass an order of magnitude higher, than hot DM particles. Fluctuations as small as large galaxy halos, $\gtrsim 10^{11} M_\odot$, could then survive free streaming. Pagels and I initially suggested that, in theories of local supersymmetry broken at $\sim 10^6$ GeV, gravitinos could be DM of the warm variety.[147] Other candidates are also possible, as I will discuss.

Cold DM consists of particles for which free streaming is of no cosmological importance. Two different sorts of cold DM consisting of elementary particles have been proposed, a cold Bose condensate such as axions, and heavy remnants of annihilation or decay such as heavy stable neutrinos. A perennial candidate, primordial black holes, is beginning to be constrained by analysis and observations.[148–150] Finally, I have already mentioned another sort of superdense objects that would behave astrophysically as cold DM: quark nuggets. As we will see, a universe dominated by cold DM looks remarkably like the one astronomers actually observe.

It is of course also possible that the dark matter is NOTA — none of the above! Maybe the dark matter is a mixture, for example "jupiters" plus neutrinos[151] or "jupiters" plus cold dark matter.[152] Some models even include unstable DM that decays into relativistic particles.[153–159]

3.4 Galaxy Formation with Hot DM

The standard hot DM candidate is massive neutrinos,[160–163] although other, more exotic, theoretical possibilities have been suggested, such as a "majoron"[164] of nonzero mass which is lighter than the lightest neutrino species, and into which all neutrinos decay.[165] For definiteness, I will discuss neutrinos.

Mass Constraints

Left-handed neutrinos of mass ≤ 1 MeV remain in thermal equilibrium until the temperature drops to $T_{\nu d}$, at which point their mean free path first exceeds the horizon size — see Fig. 3.5 — and they essentially cease interacting thereafter, except gravitationally.[65] Their mean free path is, in natural units ($\hbar = c = 1$), $\lambda_\nu \sim [\sigma_\nu n_{e^\pm}]^{-1} \sim [(G_F^2 T^2)(T^3)]^{-1}$, where $G_F \approx 10^{-5}$ GeV^{-2} is the Fermi constant that measures the strength of the weak interactions. The horizon size is $\lambda_h \sim (G\rho)^{-1/2} \sim M_{Pl} T^{-2}$, where the Planck mass $M_{Pl} \equiv G^{-1/2} = 1.22 \times 10^{19}\ GeV$. Thus $\lambda_h/\lambda_\nu \sim (T/T_{\nu d})^3$, with the neutrino decoupling temperature

$$T_{\nu d} \sim M_{Pl}^{-1/3} G_F^{-2/3} \sim 1\,\text{MeV} \ . \tag{3.11}$$

After T drops below $\frac{1}{2}$ MeV, e^+e^- annihilation ceases to be balanced by pair creation, and the entropy of the e^+e^- pairs heats the photons. Above 1 MeV, the number density $n_{\nu i}$ of each left-handed neutrino species (counting both ν_i and $\overline{\nu}_i$) is equal to that of the photons, n_γ, times the factor 3/4 from Fermi versus Bose statistics; but then e^+e^- annihilation increases the photon number

density relative to that of the neutrinos by a factor of $11/4$.(*) Thus today, for each species,

$$n_{\nu,o} = \frac{3}{4} \cdot \frac{4}{11} n_{\gamma,o} = 109\,\theta^3 \mathrm{cm}^{-3} \ , \tag{3.12}$$

where, as above, $\theta \equiv (T_o/2.7\mathrm{K})$. Since the present cosmological density is

$$\bar{\rho} = \Omega\rho_c = 11\,\Omega h^2 \mathrm{keV\,cm}^{-3} \ , \tag{3.13}$$

it follows that

$$\sum_i m_{\nu i} < \bar{\rho}/n_{\nu,o} \leq 100\,\Omega h^2 \theta^{-3} \mathrm{eV} \ , \tag{3.14}$$

where the sum runs over all neutrino species with $M_{\nu i} \leq 1$ MeV.(Heavier neutrinos will be discussed shortly.) Observational data imply that Ωh^2 is less than unity. Thus if one species of neutrino is substantially more massive than the others and dominates the cosmological mass density, as for definiteness I will assume for the rest of this section, then a reasonable estimate for its mass is $m_\nu \sim 30$ eV.

At present the experimental evidence for nonzero neutrino mass is apparently not entirely convincing. Although one group has reported[161] that 14 eV $< m_{\nu_e} <$ 40 eV and that $m_{\nu_e} > 20$ eV at the 95% confidence level[166] from tritium β-decay endpoint data, the experiment is extraordinarily difficult and according

* Since the argument giving the 11/4 factor is simple, and since the idea of warm DM is based on a generalization of it, I will sketch it here. The key ingredient is that the entropy in interacting particles in a comoving volume is conserved during ordinary Hubble expansion, even during a process such as electron-positron annihilation, so long as it occurs in equilibrium. (This fact should be intuitively obvious, since the process is reversible, and anyway it is easily derived — see e.g. Ref. 65, §15.6.) That is, $S_I = g_I(T)N_\gamma(T) =$ constant, where $N_\gamma = n_\gamma V$ is the number of photons in a given comoving volume V, and $g_I = (g_B + \frac{7}{8}g_F)_I$ is the effective number of helicity states in interacting particles (with the factor of $\frac{7}{8}$ reflecting the difference in energy density for fermions versus bosons). Just above the temperature of electron-positron annihilation, $g_I = g_\gamma + \frac{7}{8} \times g_e = 2 + \frac{7}{8} \times 4 = \frac{11}{2}$; while below it, $g_I = g_\gamma = 2$. Thus, as a result of the entropy of the electrons and positrons being dumped into the photon gas at annihilation, the photon number density is thereafter increased relative to that of the neutrinos by a factor of $\frac{11}{4}$.

to some independent authorities[167] their data are consistent with $m_{\nu_e} = 0$. Several sensitive experiments are in progress using alternative methods. The so far unsuccessful attempts to detect neutrino oscillations also give only upper limits on neutrino masses times (essentially unknown) mixing parameters.

In deriving eq. (3.14), I have been assuming that all the neutrino species are light enough to still be relativistic at decoupling, i.e. lighter than an MeV. The bound (3.14) shows that they must then be much lighter than that. In the alternative case that a neutrino species is nonrelativistic at decoupling, it has been shown[168] that its mass must then exceed several GeV, which is not true of the known neutrinos (ν_e, ν_μ, and ν_τ). (One might at first think that the Boltzmann factor would sufficiently suppress the number density of neutrinos weighing a few tensof MeV to allow compatibility with the present density of the universe. It is the fact that they "freeze out" of equilibrium well before the temperature drops to their mass that leads to the higher mass limit.) I have also been assuming that the neutrino chemical potential is negligible, i.e. that $|n_\nu - n_{\bar{\nu}}| \ll n_\gamma$. This is very plausible, since the net baryon number density $(n_b - n_{\bar{b}}) \lesssim 10^{-9} n_\gamma$, but if it is not true the consequences can be rather dramatic.[169]

Free Streaming

The most salient feature of hot DM is the erasure of small fluctuations by free streaming. Thus even collisionless particles effectively exhibit a Jeans mass. It is easy to see that the minimum mass of a surviving fluctuation is of order M_{Pl}^3/m_ν^2.[170,162]

Let us suppose that some process in the very early universe — for example, thermal fluctuations subsequently vastly inflated, in the inflationary scenario[171] — gave rise to adiabatic fluctuations on all scales. Neutrinos of nonzero mass m_ν stream relativistically from decoupling until the temperature drops to m_ν, during which time they traverse a distance $d_\nu = R_H(T = m_\nu) \sim M_{Pl}\, m_\nu^{-2}$. In order to survive this free streaming, a neutrino fluctuation must be larger in linear dimension than d_ν. Correspondingly, the minimum mass in neutrinos

of a surviving fluctuation is $M_{J,\nu} \sim d_\nu^3 m_\nu n_\nu (T = m_\nu) \sim d_\nu^3 m_\nu^4 \sim M_{Pl}^3 m_\nu^{-2}$. By analogy with Jeans' calculation of the minimum mass of an ordinary fluid perturbation for which gravity can overcome pressure, this is referred to as the (free-streaming) Jeans mass. (See Fig. 3.3.)A more careful calculation[162,172] gives

$$d_\nu = 41(m_\nu/30\,\mathrm{eV})^{-1}(1+z)^{-1}\mathrm{Mpc}\ , \tag{3.15}$$

that is, $d_\nu = 41(m_\nu/30\mathrm{eV})^{-1}$ Mpc in comoving coordinates, and correspondingly

$$M_{J,\nu} = 1.77\, M_{Pl}^3 m_\nu^{-2} = 3.2 \times 10^{15}(m_\nu/30\,\mathrm{eV})^{-2} M_\odot\ , \tag{3.16}$$

which is the mass scale of superclusters. Objects of this size are the first to form in a ν-dominated universe, and smaller scale structures such as galaxies can form only after the initial collapse of supercluster-size fluctuations.

The limits on small-angle $\delta T/T$ fluctuations are compatible with this picture. When a fluctuation of total mass $\sim 10^{15} M_\odot$ enters the horizon at $z \sim 10^4$, the density contrast δ_{RB} of the radiation plus baryons ceases growing and instead starts oscillating as an acoustic wave (as usual), while that of the massive neutrinos δ_ν continues to grow linearly with the scale factor $R = (1+z)^{-1}$ since the Compton drag that prevents growth of δ_{RB} does not affect the neutrinos. By recombination, at $z_r \sim 10^3$, $\delta_{\mathrm{RB}}/\delta_\nu \lesssim 10^{-1}$, with possible additional suppression of δ_{RB} by Silk damping. (See Fig. 3.4.) Thus the hot DM scheme with adiabatic primordial fluctuations predicts small-angle fluctuations in the microwave background radiation somewhat below current observational upper limits. Similar considerations apply in the warm and cold DM schemes.[173,174]

In numerical simulations of dissipationless gravitational clustering starting with a fluctuation spectrum appropriately peaked at $\lambda \sim d_\nu$ (reflecting damping by free streaming below that size and less time for growth of the fluctuation amplitude above it — cf. Fig. 4.2), the regions of high density form a network of filaments, with the highest densities occurring at the intersections and with voids

in between.[175,176,177] The similarity of these features to those seen in observations is cited as evidence in favor of this model.[178]

Potential Problems with ν DM

A number of potential problems with the neutrino dominated universe have emerged in recent studies, however.

1. From studies both of nonlinear[177,75] clustering ($\lambda \lesssim 10$ Mpc) and of streaming velocities[179] in the linear regime ($\lambda < 10$ Mpc), it follows that supercluster collapse must have occurred recently: $z_{sc} \leq 0.5$ is indicated,[179] and in any case $z_{sc} < 2$.[177] (See Fig. 3.6.However, the best limits on galaxy ages coming from globular clusters and other stellar populations indicate that galaxy formation took place before $z = 3$.[181] Moreover, if QSOs are associated with galaxies, as is suggested by the detection of galactic luminosity around nearby QSOs and the apparent association of more distant QSOs with galaxy clusters, the abundance of QSOs at $z > 2$ is also inconsistent with the "top-down" neutrino dominated scheme in which superclusters form first: $z_{sc} > z_{galaxies}$.

2. Numerical simulations of the nonlinear "pancake" collapse taking into account dissipation of the baryonic matter show that at least 85% of the baryons are so heated by the associated shock that they remain unable to condense, attract neutrino halos, and eventually form galaxies.[182] This could be a problem for the hot DM scheme for two reasons. With the primordial nucleosynthesis constraint $\Omega_b \lesssim 0.1$, there may be difficulty having enough baryonic matter condense to form the luminosity that we actually observe. And, where are the X-rays from the shock-heated pancakes?

3. The neutrino picture predicts[183] that there should be a factor of ~ 5 increase in M/M_b between large galaxies ($M \sim 10^{12} M_\odot$) and large clusters ($M \geq 10^{14} M_\odot$), since the larger clusters, with their higher escape velocities, are able to trap a considerably larger fraction of the neutrinos. As I discussed in Lecture 1 (see especially Fig. 1.11), although there is some

indication that M/L increases with M, the ratio of total to luminous mass M/M_{lum} is probably a better indicator of the value of M/M_b, and it is roughly the same for galaxies with large halos and for rich clusters.

4. Both theoretical arguments regarding the dwarf spheroidal (dS) satellite galaxies of the Milky Way[184] and data on Draco, Carina, and Ursa Minor[185,186] imply that dark matter dominates the gravitational potential of these dS galaxies. The phase-space constraint (*) then sets a lower limit[186] $m_\nu > 500$ eV, which is completely incompatible with the cosmological constraint eq. (3.14). (Note that even for neutrinos as the DM in large spiral galaxies, the phase space constraint implies $m_\nu > 30$ eV.)

These problems, while serious, may not be fatal for the hypothesis that neutrinos are the dark matter. It is possible that galaxy density does not closely correlate with the density of dark matter — for example, because the first generation of luminous objects heats nearby matter, thereby increasing the baryon

---

$\star$ The phase space constraint[187] follows from a theorem in classical mechanics to the effect that the maximum 6-dimensional phase space density cannot increase as a system of collisionless particles evolves. At early times, before density inhomogenitites become nonlinear, the neutrino phase space density is given by the Fermi-Dirac distribution

$$n_\nu(p) = \frac{g_\nu}{h^3}\left[1 + \exp\left(\frac{pc}{kT_\nu(z)}\right)\right]^{-1},$$

where here h is Planck's constant and $g_\nu = 2$ for each species of left-handed ν plus $\bar{\nu}$. (Since momentum and temperature both scale as redshift z as the universe expands, this distribution remains valid after neutrinos drop cut of thermal equilibrium at ~ 1 MeV, and even into the nonrelativistic regime $T_\nu < m_\nu$.) The standard version of the phase space constraint follows from demanding that the central phase space density $9[2(2\pi)^{5/2}Gr_c^2\sigma m_\nu^4]^{-1}$ of the halo, assumed to be an isothermal sphere of core radius r_c and one-dimensional velocity dispersion σ, not exceed the maximum value of the initial phase space density $n_\nu(0) = g_\nu/2h^3$. The result is

$$m_\nu > (120\,\mathrm{eV})\left(\frac{100\,\mathrm{km\,s^{-1}}}{\sigma}\right)^{1/4}\left(\frac{1\,\mathrm{kpc}}{r_c}\right)^{1/2}\left(\frac{g_\nu}{2}\right)^{-1/4}.$$

The lower limits on m_ν quoted in the text use this result. A more conservative phase space constraint has recently been obtained[188] assuming (perhaps unrealistically) that the neutrinos are in the most compact distribution possible, rather than an isothermal sphere. Assuming that for the dS galaxy Draco, $M > 10^8 M_\odot$ and $r_c < 2$ kpc, it follows that $m_\nu > 127$ eV. This is still in trouble with the cosmological upper bound on m_ν.

Jeans mass and suppressing galaxy formation. This could complicate the comparison of nonlinear N-body simulations with the data. Also, if dark matter halos of large clusters are much larger in extent than those of individual galaxies and small groups, then virial estimates would underestimate mass on large scales and the data could be consistent with M/M_{lum} increasing with M. But it is hard to avoid the constraint on z_{sc} from streaming velocities in the linear regime[179] except by assuming that the Local Group's velocity is abnormally low. And the only explanation for the high M/L of dS galaxies in a neutrino-dominated universe is the rather ad hoc assumption that the dark matter in such objects is baryons rather than neutrinos. Of course, the evidence for massive dark halos in dS galaxies is not yet solid.

In summary, the evidence against hot DM is rather impressive. At very least, it indicates that a neutrino-dominated universe must be rather more complicated than theorists have yet envisaged.

3.5 Galaxy Formation with Warm DM

Suppose the dark matter consists of an elementary particle species X that interacts much more weakly than neutrinos. The X's decouple thermally at a temperature $T_{Xd} \gg T_{\nu d}$ and their number density is not thereafter affected by particle annihilation at temperatures below T_{Xd}. With the standard assumption of conservation of entropy in comoving volumes, the X number density today $n_{X,o}$ and mass m_X can be calculated in terms of $g_I(T)$, the effective number of helicity states of interacting particles evaluated at T_{Xd}.[131] These are plotted in Fig. 3.7, assuming the "standard model" of particle physics. The simplest grand unified theories predict $g_I(T) \approx 100$ for T between 10^2 GeV and $T_{tot} \sim 10^{14}$ GeV, with possibly a factor of two increase in g_I beginning near 10^2 GeV due to $N = 1$ supersymmetry partner particles. Then for T_{Xd} in the enormous range from ~ 1 GeV to $\sim T_{GUT}$, $n_X^o \sim 5 g_X \,\mathrm{cm}^{-3}$ and correspondingly $m_X = 2\Omega h^2 g_X^{-1}$ keV,[189] where g_X is the number of X helicity states. Because of free streaming, such

"warm" DM particles of mass $m_X \sim 1$ keV will cluster on a scale $\sim M_{Pl}^3 m_X^{-2} \sim 10^{12} M_\odot$, the scale of large galaxies such as our own.

Candidates for Warm DM

What might be the identity of the warm DM particles X? Pagels and I[147] suggested that they might be the $\pm 1/2$ helicity states of the gravitino $\tilde{G}$, the spin 3/2 supersymmetric partner of the graviton G. The gravitino mass is related to the scale of supersymmetry breaking m_{SUSY} by $m_{\tilde{G}} = (4\pi/3)^{1/2} m_{SUSY}^2 M_{Pl}^{-1}$, so $m_{\tilde{G}} \sim 1$ keV corresponds to $m_{SUSY} \sim 10^6$ GeV. This now appears to be phenomenologically dubious, and $N = 1$ supersymmetry models with $m_{SUSY} \sim 10^{11}$ GeV and $m_{\tilde{G}} \sim 10^2$ GeV are currently popular.[192] In such models, the photino $\tilde{\gamma}$, the spin 1/2 supersymmetric partner of the photon, is probably the lightest particle that is odd under the supersymmetric reflection symmetry R, and hence stable. But in supersymmetric GUT models there is a relation between the mass of the gluino $\tilde{g}$, the supersymmetric partner of the gluon (the gauge boson of QCD), and that of the photino: $m_{\tilde{\gamma}} \sim \frac{1}{7} m_{\tilde{g}}$, and there is a phenomenological lower bound on the mass of the gluino $m_{\tilde{g}} \gtrsim 3$ GeV.[193] The requirement that the photinos almost all annihilate, so that they do not contribute too much mass density, also implies that $m_{\tilde{\gamma}} \geq 0.5$ GeV,[194] and they thus become a candidate for cold rather than warm dark matter.

A hypothetical right-handed neutrino ν_R could be the warm DM particle,[195] since if right-handed weak interactions exist they must be much weaker than the ordinary left-handed weak interactions, so $T_{\nu_R d} \gg T_{\nu d}$ as required. But particle physics provides no good reason why any ν_R should be light.

Thus there is at present no obvious warm DM candidate elementary particle, in contrast to the hot and cold DM cases. But our ignorance about the physics above the ordinary weak interaction scale hardly allows us to preclude the existence of very weakly interacting light particles, so we will consider the warm DM

case, mindful of Hamlet's prophetic admonition

> There are more things in heaven and earth, Horatio,
> Than are dreamt of in your philosophy.

Fluctuation Spectrum

The spectrum of fluctuations δ_ν at late times in the hot DM model is controlled mainly by free streaming; $\delta_\nu(M)$ is peaked at $\sim M_{J,\nu}$, eq. (3.16), for any reasonable primordial fluctuation spectrum. This is *not* the case for warm or cold DM.

The primordial fluctuation spectrum can be characterized by the amplitude of fluctuations just as they enter the horizon (see Lecture 2). It is expected that no mass scale is singled out, so the spectrum is just a power law

$$\delta_{DM,H} = \kappa \left(\frac{M_{DM}}{M_0} \right)^{-\alpha} . \tag{3.17}$$

Furthermore, to avoid too much power on large or small mass scales requires $\alpha \approx 0$;[117,118] and to form galaxies and large scale structure by the present epoch without violating the upper limits on both small[143] and large[196] scale (quadrupole) angular variations in the microwave background radiation requires $\kappa \sim 10^{-4}$. Equation (3.17) corresponds in Fourier transform space to $|\delta_k|^2 = k^n$ with $n = 6\alpha + 1$. The case $\alpha = 0$ $(n = 1)$ is commonly referred to as the constant curvature or (Harrison-)Zeldovich spectrum. As I discussed in Lecture 2, inflationary models predict adiabatic fluctuations with approximately the Zeldovich spectrum.

The important difference between the fluctuation spectra δ_{DM} at late times in the hot and warm DM cases is that $\delta_{\mathrm{DM,warm}}$ has power over an increased range of masses, roughly from 10^{11} to $10^{16} M_\odot$. As for the hot case, the lower limit, $M_X \sim M_{Pl}^3 m_X^{-2}$, arises from the damping of smaller-scale fluctuations

by free streaming. In the hot case, the DM particles become nonrelativistic at essentially the same time as they become gravitationally dominant, because their number density is nearly the same as that of the photons. But in the warm case, the X particles become nonrelativistic and thus essentially stop free streaming at $T \approx m_X$, well *before* they begin to dominate gravitationally at $T_{eq} \approx 6\Omega h^2$ eV. (As usual, the subscript "eq" refers to the epoch when the energy density of massless particles equals that of massive ones.) During the interval between $T \approx m_X$ and $T \approx T_{eq}$, growth of δ_{DM} is inhibited by the "stagspansion" phenomenon[155] (the generalization to adiabatic fluctuations of the "Meszaros effect"[197] effect), which we will discuss in detail in the section on cold DM. Thus the spectrum δ_{DM} is relatively flat between M_X and

$$M_{eq} = \frac{4\pi}{3}\left(\frac{ct_{eq}}{1+z_{eq}}\right)^3 \rho_{c,o}\Omega_o = 2.2 \times 10^{15}(\Omega_o h^2)^{-2} M_\odot \ . \qquad (3.18)$$

Fluctuations with masses larger than M_{eq} enter the horizon at $z < z_{eq}$, and thereafter δ_{DM} grows linearly with $R = (1+z)^{-1}$ until nonlinear gravitational effects become important when $\delta_{DM} \sim 1$. Since for $\alpha = 0$ fluctuations of all sizes enter the horizon with the same root-mean-square amplitude, and those with larger M enter the horizon later in the matter-dominated era and subsequently have less time to grow, the fluctuation spectrum at the present time falls with M for $M > M_{eq}$: $\delta_{DM,o} \propto M^{-2/3}$. For a power-law primordial spectrum of arbitrary index α,

$$\delta_{DM,o} \propto M^{-\alpha-(2/3)} = M^{-(n+3)/6} \ , \quad M > M_{eq} \ . \qquad (3.19)$$

This is true for hot, warm, or cold DM. In each case, after recombination at $z_r \approx 10^3$ the baryons "fall in" to the dominant DM fluctuations on all scales larger than the baryon Jeans mass, and by $z = 100$, $\delta_b = \delta_{DM}$.[198]

In the simplest approximation, neglecting all growth during the "stagspansion" era, the fluctuation spectrum for $M_X < M < M_{eq}$ is just $\delta_{DM,o}(M) \propto M^{-(n-1)/6} = M^{-(n_{eff}+3)/6}$ where $n_{eff} = n - 4$; i.e., the spectrum is flattened by

a factor of $M^{2/3}$ compared to the primordial spectrum. The small amount of growth that does occur during the "stagspansion" era slightly increases the fluctuation strength on smaller mass scales. Detailed calculations of these spectra are now available.[190,172] For $\alpha \leq 0$, $\delta_X(M)$ has a fairly broad peak at $M \sim M_X$. Consequently, objects of this mass — galaxies and small groups — are the first to form, and larger-scale structures — clusters and superclusters — form later as $\delta_X(M)$ grows toward unity on successively larger mass scales.

Potential Problems with Warm DM

The warm DM hypothesis is probably consistent with the observed features of typical large galaxies, whose formation would probably follow roughly the "core condensation in heavy halos" scenario.[199,200] The potentially serious problems with warm DM are on scales both larger and smaller than M_X. On large scales, the question is whether the model can account for the observed network of filamentary superclusters enclosing large voids. The most productive approach to this question has employed sophisticated N-body simulations with $N \sim 3 \times 10^4$ in order to model the large mass range that is relevant.[201] The N-body results suggest that warm and cold DM can reproduce the observed large scale structure, although to get good agreement it may be necessary to assume that galaxies do not accurately trace the DM distribution. I will discuss this further in the next Lecture.

On small scales, the preliminary indications that dwarf spheriodal galaxies have large DM halos[184,185] pose problems nearly as serious for warm as for hot DM. Unlike hot DM, warm DM is (barely) consistent with the phase space constraint.[185,186] But since free streaming of warm DM washes out fluctuations δ_X for $M \lesssim 10^{11} M_\odot$, dwarf galaxies with $M \sim 10^7 M_\odot$ can form in this picture only via fragmentation following the collapse of structures of mass $\sim M_X$, much as ordinary galaxies form from supercluster fragmentation in the hot DM picture. The problem here is that dS galaxies, with their small escape velocities $\lesssim 10\,\mathrm{km\,s^{-1}}$, would not be expected to bind more than a small fraction of the

X particles, whose typical velocity must be $\sim 10^2 km\,s^{-1}$ ($\approx$ rotation velocity of spirals). Thus we expect M/M_{lum} for dS galaxies to be much smaller than for large galaxies — but the indications are that they are comparable.[184,185,186] Understanding dwarf galaxies may well be crucial for unravelling the mystery of the identity of the DM.[200]

4. Cold Dark Matter

To summarize the terminology introduced in the previous lecture, the dark matter (DM) that appears to be gravitationally dominant on all astronomical scales larger than the cores of galaxies[5] can be classified, on the basis of its characteristic free-streaming damping mass M_D, as hot ($M_D \sim 10^{15} M_\odot$), warm ($M_D \sim 10^{11} M_\odot$), or cold ($M_D < 10^8 M_\odot$). For the case of cold DM, the main subject of this lecture, the shape of the DM fluctuation spectrum is determined by (a) the primordial spectrum (on scales larger than the horizon), which is usually assumed to have a power spectrum of the form $|\delta_k|^2 \propto k^n$ (inflationary models predict the "Zeldovich spectrum" $n = 1$); and (b) "stagspansion",[135] the stagnation of the growth of DM fluctuations that enter the horizon while the universe is still radiation-dominated, which flattens the fluctuation spectrum for $M \lesssim 10^{15} M_\odot$.[202,203,204]

An attractive feature of the cold dark matter hypothesis is its considerable predictive power: the post-recombination fluctuation spectrum is calculable, and it in turn governs the formation of galaxies and clusters. As I will discuss in this lecture, good agreement with the galaxy and cluster data is obtained in the cold DM model for a Zeldovich spectrum of primordial fluctuations, and the model also appears to be reasonably consistent with the observed large-scale clustering, including superclusters and voids.[26]

4.1 Cold DM Candidates

Besides the evidence summarized in Lecture 3 against hot and warm DM, a further reason to consider cold DM is the existence of several plausible physical candidates, including axions of mass $\sim 10^{-5}$ eV;[205,206] a heavy, weakly interacting, stable particle, such as the photino, with a mass $\gtrsim \frac{1}{2}$ GeV;[194] and primordial black holes[207] with $10^{17} g \lesssim m_{PBH} \lesssim M_\odot$.[26] Still another exotic cold DM candidate has recently been proposed by Witten: "nuggets" of $u - d - s$ symmetric quark matter.[134] There is thus no shortage of cold DM candidate particles —

although there is admittedly no direct evidence that any of them actually exists. In the remainder of this section, I will briefly discuss the rationale for each of these cold DM candidates.

First, the axion. Quantum chromodynamics (QCD) with quarks of nonzero mass violates CP and T due to nonperturbative instanton effects. This leads to a neutron electric dipole moment that is many orders of magnitude larger than the experimental upper limit, unless an otherwise undetermined complex phase $\theta_{\rm QCD}$ is arbitrarily chosen to be extremely small. Peccei and Quinn[208] have proposed the simplest and probably the most appealing way to avoid this problem, by postulating an otherwise unsuspected symmetry that is spontaneously broken when an associated pseudoscalar field — the *axion*[209] — gets a nonzero vacuum expectation value $\langle\phi_a\rangle \sim f_a e^{i\theta}$. This occurs when $T \sim f_a$. Later, when the QCD interactions become strong at $T \sim \Lambda_{\rm QCD} \sim 10^2$ MeV, instanton effects generate a mass for the axion $m_a = m_\pi f_\pi / f_a = 10^{-5}$ eV (10^{12} GeV$/f_a$). Thereafter, the axion contribution to the energy density is[210] $\rho_a = 3 m_a T^3 f_a^2 (M_{Pl}\Lambda_{\rm QCD})^{-1}$. (A coherent state of axions behaves cosmologically like pressureless dust, despite the fact that $m_a \ll T_{\rm QCD}$.[211]) The requirement $\rho_a^0 < \rho_0 \Omega$ implies that $f_a \leq 10^{12}$ GeV, and $m_a \geq 10^{-5}$ eV.[210] The longevity of helium-burning stars implies[212] that $f_a > 10^9$ GeV, $m_a < 10^{-2}$ eV. Thus if the hypothetical axion exists, it is probably important cosmologically, and for $m_a \sim 10^{-5}$ eV gravitationally dominant. (The mass range 10^{9-12} GeV, in which f_a must lie, is also currently popular with particle theorists as the scale of supersymmetry[192] or family symmetry breaking, the latter possibility connected with the axion.[213]) If axions comprise the dark halo of our galaxy, laboratory experiments have recently been proposed that could detect them.[206]

A quite different sort of cold DM elementary particle arises naturally in supersymmetry. The lightest supersymmetric partner (LSP) particle is stable because of R-parity, a reflection symmetry equivalent to[214]

$$R = (-1)^{3(B-L)}(-1)^F, \tag{4.1}$$

where B, L, and F are baryon, lepton, and fermion numbers, respectively. R-parity is therefore an exact symmetry in any theory in which $(B-L)$ and F are conserved. Under R-parity, all ordinary particle fields are unchanged and all superpartner fields change sign.

In many supersymmetric theories, the LSP is a photino of mass $m_{\tilde{\gamma}} \gtrsim \frac{1}{2}$ GeV, this lower limit corresponding to cosmological critical density and dependent on the theoretical parameters controlling the $\tilde{\gamma}$ mass and interactions.[194] (Cf. also §3.6 above.) The $\tilde{\gamma}$s almost all annihilate at high temperatures, leaving behind a small remnant that, because $m_{\tilde{\gamma}}$ is large, can contribute a critical density today. Remarkably, there is actually one piece of observational evidence that the Milky Way's DM halo is composed of photinos: the calculated[215] annihilation rate $\tilde{\gamma}\tilde{\gamma} \longrightarrow p\bar{p}$ for $m_{\tilde{\gamma}} \approx 3 - 10$ GeV leads to a flux of low-energy (0.6-1.2 GeV) anti-protons comparable to the observed primary cosmic ray flux, and no plausible alternative explanation is known for $\bar{p}$s below the ~ 2 GeV kinematic threshold for secondary production. This theory can be checked by additional observations of the spectrum and anisotropy of cosmic ray anti-protons, positrons, and gamma rays.

Another possibility for the LSP is a sneutrino $\tilde{\nu}$ (a scalar partner of a neutrino). The main difference from the photino as LSP is that the $\tilde{\nu}$ annihilation rate can be much larger (semiweak rather than weak because it is mediated by $\tilde{Z}$, a fermion, with amplitude proportional to $m_{\tilde{Z}}$ rather than m_Z^2), so that there is no lower limit on the $\tilde{\nu}$ mass from the cosmological density.[216] However, the assumption that $\tilde{\nu}$ is the LSP implies that $m_{\tilde{\nu}} \lesssim 2$ GeV if $\rho_{\tilde{\nu},o} \sim \rho c, o$.

The next cold DM candidate to be considered may seem rather contrived: a particle, such as a ν_R, that decouples while still relativistic but whose number density relative to the photons is subsequently diluted by entropy generated in a first-order phase transition such as the Weinberg-Salam $SU(2) \otimes U(1) \longrightarrow U(1)$ symmetry breaking.[189] (Recall that the m_X bound in Fig. 3.7, corresponding to warm DM, assumes no generation of entropy.) More than a factor $\sim 10^3$ entropy

increase would overdilute $\eta = n_b/n_\gamma$, if we assume η was initially generated by GUT baryosynthesis; correspondingly, $m_X \leq 1$ MeV, and $M_X \geq 10^5 M_\odot$.

Finally, there is the possibility that the DM consists of objects more massive than any stable elementary particle. Two cases that have been investigated are "quark nuggets" and black holes.

A "quark nugget" is a hypothetical chunk of quark matter with roughly equal numbers of u, d, and s quarks. Because the Fermi energy is shared among three species, $u-d-s$ symmetric quark matter is more stable than $u-d$ quark matter, and perhaps absolutely stable, as long as the strange quark mass $m_s \lesssim 200$ MeV.[134,217] Note that even if an ordinary nucleus of atomic weight A is more massive than an amount of $u-d-s$ quark matter with equal baryon number, the decay rate of the ordinary nucleus is extremely small since it is proportional to $\sim A$ powers of the weak interaction constant G_F. At the time of the QCD confinement phase transition in the early universe, however, there were almost as many s quarks as u and d quarks, since m_s is comparable to $T_{\rm QCD}$. So if the bubbles of the high temperature (quark plasma) phase can radiate the annihilation energy of the excess $q\bar{q}$ pairs primarily by bulk neutrino emission, then these bubbles can become $u-d-s$ quark nuggets. But as Witten[134] admits, "There is a rich element of wishful thinking here, since this picture assumes neutrino losses are the main way for the high temperature phase to lose energy, while in fact neutrino losses and surface evaporation appear comparable."

How large might quark nuggets be? Witten argues that when bubbles of the high temperature phase begin to shrink, their radius is of order $R_1 \approx 1 - 10^4$ cm. The lower limit arises under the assumption that the bubbles grow by coalescence of smaller bubbles of initial size $R_0 \ll R_1$. In this case, the effective bubble size is independent of R_0 and is determined by surface tension, with the result that $R_1 \approx M_{Pl}^{2/3}/T_{\rm QCD}^{5/3} \approx 1 - 10$ cm. The upper limit, $R_1 \approx 10^4$ cm, applies if R_0 is as large as it could possibly be (a little smaller than the horizon length $M_{Pl}/T_{\rm QCD}$ at the QCD transition). The linear dimensions of the bubbles then shrink by

a factor of $\eta^{1/3} \approx 10^{-3}$, where as usual η is the baryon to photon ratio. The result is that a quark nugget is expected to have a radius of $10^{-3} - 10$ cm, and correspondingly a mass $m_{qn} \sim 10^6 - 10^{18}$ g.

It follows from the discussion in Lecture 1 that the density of dark matter in our part of the galaxy is

$$\rho_{DM} \sim \frac{v^2}{4\pi G r^2} \approx 10^{-24}\,\mathrm{g\,cm^{-3}}, \tag{4.2}$$

where $v \approx 220$ km s^{-1} is the circular velocity and $r \approx 8$ kpc the galactic radius of the sun. The flux intercepted by the earth is thus $\sim 10^9$ g per year. Given the rather large uncertainty in m_{qn}, this flux corresponds to anywhere from 10^3 light quark nuggets striking the earth per year, to one heavy one in the earth's lifetime. Such dense projectiles would pass right through the earth. De Rujula and Glashow[218] call any aggregate of stable nuclear matter intermediate in size between nuclei and neutron stars that happen to hit the earth, a "nuclearite". If they exist in the mass range discussed, they can be detected by looking for linear "astroblems" (the debris of their passage through rock) or linear seismic sources.

Primordial black holes of typical mass $M_{BH} \lesssim M_\odot$ could have formed if there were fluctuations of large amplitude when that mass scale first entered the horizon, or black holes of mass $M_{BH} \gtrsim M_\odot$ could have formed from collapse of ordinary matter after recombination. A rather wide mass range is allowed,

$$10^{-16} M_\odot \lesssim M_{BH} \lesssim 10^6 M_\odot, \tag{4.3}$$

the lower limit implied by the non-observation of γ rays from black hole decay by Hawking radiation, and the upper limit required to avoid disruption of galactic disks and star clusters.[207,219] Stronger but less certain upper limits are $M_{BH} \lesssim 10^3 M_\odot$ from non-observation of accretion onto BHs as they plunge through the galactic disk,[220,28] $M_{BH} \lesssim 10^2 M_\odot$ from dwarf spheroidal halos,[186] and $M_{BH} \lesssim 10^{-2} M_\odot$ from the non-observation of the focusing of quasar cores.[221]

You will doubtless have noticed that, while there is no direct evidence against any of these candidates for cold DM, there are no convincing arguments in their favor either. Actually, it is not clear that we have a good basis to judge the plausibility of any DM candidates, since in no case except possibly "quark nuggets" is there a fundamental explanation — or, even better, a prediction — for the ratio $\omega \equiv \rho_{DM,o}/\rho_{lum,o}$, which is thought to lie in the range $10 \leq \omega \leq 10^2$, the lower limit corresponding to an open universe and the upper limit to $\Omega \approx 1$. Two fundamental questions about the universe which the fruitful marriage of particle physics and cosmology has yet to address successfully are the value of ω and of the cosmological constant Λ.

4.2 Galaxy and Cluster Formation with Cold DM

"Stagspansion"

I will follow the current conventional wisdom and assume that the primordial fluctuations were adiabatic. In the standard formalism, fluctuations $\delta \equiv \delta\rho/\rho$ grow as $\delta \sim R^2$ on scales larger than the horizon (cf. §2.5), where $R = (1+z)^{-1}$ is the scale factor normalized to 1 at the present. When a fluctuation enters the horizon in the radiation-dominated era, the photons (together with the charged particles) oscillate as an acoustic wave, and the non-interacting neutrinos freely stream away (they are still relativistic, since in the cold DM case their masses are $\ll 30$ eV). As a result, the main driving terms for the growth of δ_{DM} disappear and the growth accordingly stagnates as the universe continues to expand ("stagspansion") until matter dominates; see Fig. 4.1. Matter domination first occurs at $z = z_{eq}$, where (cf. eq. (2.43))

$$\begin{aligned} z_{eq} &= 4.2 \times 10^4 h^2\ \Omega\ (1 + 0.68 N_\nu)^{-1} \\ &= 2.5 \times 10^4 h^2\ \Omega \quad \text{for } N_\nu = 3 \end{aligned} \tag{4.4}$$

The first study of the growth of cold DM fluctuations was the numerical calculations of Peebles,[202] who for simplicity ignored neutrinos: $N_\nu = 0$ in the

above equation. Subsequent numerical calculations have included the effects of the known neutrino species ($N_\nu = 3$, $m_\nu \approx 0$) both outside and inside the horizon.[185,203,204,173,174,222] Numerically, the largest effect of including neutrinos is the change in z_{eq}.

It is instructive to make the further approximation of setting $\delta_{\gamma+b} = \delta_\nu = 0$ once a fluctuation is inside the horizon. Then one can analytically match the solution for $R > R_{\rm horizon}$

$$\begin{aligned} \delta_{DM}(R) &= A_1 D_1(R) + A_2 D_2(R) \ , \\ D_1 &= 1 + 1.5y \quad \text{where} \quad y = R/R_{eq} \ , \\ D_2 &= D_1 \ln\left[\frac{(1+y)^{1/2}+1}{(1+y)^{1/2}-1}\right] - 3(1+y)^{1/2} \ . \end{aligned} \tag{4.5}$$

to the growing mode $\delta_{DM} \sim R^2$ for $R < R_{\rm horizon}$. Matching the derivatives requires $A_2 D_2$ comparable to $A_1 D_1$ but opposite in sign. For $R \gg R_{\rm horizon}$ only the growing solution D_1 survives, which explains the moderate growth in δ_{DM} between horizon crossing and matter dominance. In the limit of large k, one finds[203] $|\delta_k| \propto k^{n/2-2} \ln k$, where as usual $n = 1$ corresponds to a primordial Zeldovich spectrum. Correspondingly, for $M \ll M_{eq} \approx 10^{16} M_\odot$, the rms fluctuation in the mass within a random sphere containing average mass M is $\delta M/M \propto |\ln M|^{3/2}$. Some authors have considered only the Meszaros solution D_1 and erroneously inferred that the fluctuation spectrum would be essentially flat for $M < M_{eq}$ for a Zeldovich primordial spectrum, which would then be inconsistent with observations.[223] One can match the boundary conditions without D_2 only for isothermal primordial fluctuations, for which the amplitide does not grow at very early times since the fluctuating component is subdominant. (This was the case considered by Meszaros.[197])

Figure 4.2(a) is a sketch of $\delta\rho/\rho(M)$ for hot, warm and cold DM. Figure 4.2(b) shows numerical results[204] for $\delta M/M$, again assuming a Zeldovich ($n = 1$) spectrum and normalized so that $\delta M/M = 1$ at $R = 8h^{-1}Mpc$.[202] Notice that

$\delta M/M$ is relatively flat for $M < 10^9 M_\odot$, and steepens to $\delta M/M = M^{-2/3}$ (that is, $n = 1$, reflecting the primordial spectrum) for $M \gg M_{eq}$.[26,204]

Galaxy and Cluster Formation

The key features of galaxy formation in the cold DM picture are these: after recombination (at $z_{rec} \approx 10^3$) the amplitude of the baryonic fluctuations rapidly grows to match that of the DM fluctuations; smaller-mass fluctuations grow to nonlinearity and virialize, and then are hierarchically clustered within successively larger bound systems; and finally the ordinary matter in bound systems of total mass $\sim 10^{8-12} M_\odot$ cools rapidly enough within their DM halos to form galaxies, while larger mass fluctuations form clusters.

At any mass scale M, when the fluctuation $\delta M/M$ approaches unity, nonlinear gravitational effects become important. The fluctuation then separates from the Hubble expansion, reaches a maximum radius, and begins to contract. Spherically symmetric fluctuations, for example, contract to about half their maximum radii (see Lecture 2). During this contraction, violent relaxation[224] due to the rapidly varying gravitational field converts enough potential energy into kinetic energy for the virial theorem, $\langle PE \rangle = -2 \langle KE \rangle$, to be satisfied. After virialization, the mean density within a fluctuation is roughly eight times the density corresponding to the maximum radius of expansion.[225]

Since the cold-DM fluctuation spectrum $\delta M/M$ is a decreasing function of M, smaller mass fluctuations will, on the average, become nonlinear and begin to collapse at earlier times than larger mass fluctuations. Small mass bound systems are subsequently clustered within larger mass systems, which go nonlinear at a later time. This hierarchical clustering of smaller systems into larger and yet larger gravitationally bound systems begins at the baryon Jeans mass ($M_{J,b} \sim 10^5 M_\odot$ at recombination) and continues until the present time.

Although small-mass fluctuations are the first to go nonlinear in the cold DM picture, pressure inhibits baryons from falling into such fluctuations if $M < M_{J,b}$.

More importantly, even for $M > M_{J,b}$, the baryons are not able to contract further unless they can cool by emitting radiation. Without such mass segregation between baryons and DM, the resulting structures will be disrupted by virialization as fluctuations that contain them go nonlinear.[226] Moreover, successively larger fluctuations will collapse relatively soon after one another if they have masses in the flattest part of the $\delta M/M$ spectrum, i.e., (total) mass $\leq 10^9 M_\odot$.

Gas of primordial composition (about 75% atomic hydrogen and 25% helium, by mass) cannot cool significantly unless it is first heated to $\gtrsim 10^4$ K, when it begins to ionize.[26] Assuming a primordial Zeldovich spectrum normalized so that at the present time $\delta M/M = 1$ at $R = 8h^{-1}$ Mpc, as recommended by Peebles,[202] the smallest protogalaxies for which the gas is sufficiently heated by virialization to radiate rapidly and contact have total mass $M = 10^9 M_\odot$. One can also deduce an *upper* bound on galaxy masses by requiring that the cooling time be shorter than the dynamical time[227] ; this upper bound is $M \lesssim 10^{12} M_\odot$. These limits are illustrated in Fig. 4.3, which shows the density of ordinary (baryonic) matter versus internal kinetic energy (temperature) of typical fluctuations of various sizes, just after virialization, calculated from $\delta M/M$ for $\Omega = h = 1$. This is superimposed upon the Rees-Ostriker cooling curves (for which cooling time equals gravitational free fall time) and data on galaxies (with kinetic energy determined from rotation velocity for spirals and velocity dispersion for ellipticals).*

Only in protogalaxies for which the cooling time is short compared to the dynamical time can the baryons dissipate and contract. This dissipation leads to higher baryonic densities and somewhat higher temperatures.

The collapse of fluctuations having mass $> 10^{13} M_\odot$ leads to clusters of galaxies in this picture. In clusters, only the outer parts of member galaxy halos are stripped off; the inner baryonic cores continue to contract, presumably until star

* See Ref. 26 for a considerably more elaborate version of Fig. 4.3, with virialization curves for several multiples of the rms fluctuation spectrum for an open as well as an Einstein-deSitter universe, much more detailed galaxy and cluster data, and discussion of molecular and Compton cooling.

formation halts dissipation.

Fluctuations that start with greater amplitude than average will turn around earlier, at higher density, and thus lie below the virialization curve on Fig. 4.3. As the baryons in a virialized fluctuation dissipate, their density will initially increase at constant T within the surrounding isothermal halo of dissipationless material (DM), and then T will increase as well when the baryon density exceeds the DM density, as suggested by the dashed line in the figure. The Zeldovich primordial spectrum is more consistent with the data in Fig. 4.3 than an $n = 2$ (or $n = 0$) primordial spectrum, which lies too low (too high) on the figure compared to the galaxies. With the Zeldovich spectrum, the important conclusion is that one should observe dissipated systems with large halos having total mass $10^8 M_\odot \lesssim M \lesssim 10^{12} M_\odot$. This is essentially the range of observed galaxy masses.

While the $n_b - T$ diagram (Fig. 4.3) is useful for comparing data and predictions with the cooling curves, it is also useful to consider total mass M versus T, as in Fig. 4.4. This avoids having to take into account the differing amounts of baryonic dissipation suffered by various galaxies. The heavy solid and dashed curves again correspond to the $n = 1$ cold DM spectrum, for $(\Omega = 1, h = 0.5)$ and $(\Omega = 0.2, h = 1)$ respectively. It is striking that the galaxies in the $M - T$ diagram lie along lines of roughly the same slope as these curves. This occurs because the effective slope of the $n = 1$ cold DM fluctuation spectrum in the galaxy mass range is $n_{eff} \approx -2$, which corresponds to the empirical Tully-Fisher and Faber-Jackson laws: $M \propto v^4$. The light dashed lines in Fig. 4.4 are the post-virialization curves for primordial fluctuation spectra with $n = 0$ (white noise) and $n = 2$. Again, the $n = 1$ (Zeldovich) spectrum is evidently the one that is most consistent with the data.

The points in Fig. 4.4 represent essentially all of the clusters identified by Geller and Huchra[228] in the CfA catalog within 5000 km s^{-1}. The cluster data lie about where they should on the diagram, and even the statistics of the distribution seem roughly to correspond to the expectations represented by the 0.5, 1, 2,

and 3σ curves.

Notice that spiral galaxies lie roughly along the 1σ curve while elliptical galaxies lie along the 2σ curve. Although this displacement is not large compared to the uncertainties, it is consistent with the fact that more than half of all bright galaxies are spirals, while only about 10 % are ellipticals. In hierarchical clustering scenarios, it seems likely that the higher σ fluctuations will develop rather smaller angular momenta, as measured by the dimensionless parameter λ ($\equiv JE^{\frac{1}{2}}G^{-1}M^{-\frac{5}{2}}$). There are two reasons for this: high-overdensity fluctuations collapse earlier than average fluctuations, and are thus typically surrounded by a relatively homogeneous matter distribution;[229] also, higher amplitude fluctuations are typically rounder[230] and consequently have lower quadrupole moments. Both effects result in less torque. This difference appears to exist with either white noise or a flatter spectrum, but to be somewhat larger in the latter case. If high σ fluctuations have little angular momentum, their baryons can collapse by a large factor in radius, forming high-density ellipticals and spheroidal bulges, as shown in Fig. 4.3. Since, with a flat spectrum, higher σ fluctuations occur preferentially in denser regions destined to become rich clusters, one expects[231] to find more ellipicals there — as is observed.[42,48,43] Note that the rich clusters lie along the same 2 and 3σ curves in Fig. 4.4 as do the elliptical galaxies.

Note also that while the galaxy data lies below the rms virialization curve, the data on groups and clusters of galaxies lies more or less evenly around it. This suggests that galaxy formation may be an inefficient process, with lower-amplitude fluctuations of galaxy mass not giving rise to visible galaxies.[26]

Presumably the collapse of the low-λ protoellipical galaxies is halted by star formation well before a flattened disk can form, yielding a stellar system of spheroidal shape. The mechanism governing the onset of star formation in these systems is unfortunately not yet understood, but may involve a threshold effect which sets in when the baryon density exceeds the DM halo density by a sufficient factor.[226,181] Disks (spirals and irregulars) form from average, higher-λ

protogalaxies, which, for a given mass, are larger and more diffuse than their protoelliptical counterparts. The collapse of disks thus occurs via relatively slow infall of baryons from $\sim 10^2$ kpc, halted by angular momentum. Infall from such distances is consistent both with the extent of dark halos inferred from observations and with the high angular momenta of present-day disks ($\lambda \sim 0.4$).[232] The location of the galaxies in Fig. 4.3 is consistent with these ideas if the baryons in all galaxies collapsed by roughly the same factor, about an order of magnitude, but somewhat less for late-type spirals and irregulars and somewhat more for early-type E's and spheroidal bulges.

It has been theorized that the Hubble sequence originates in the distribution of either the initial angular momenta or else the initial densities[234] of protogalaxies. However, if overdensity and angular momentum are linked, with the high-σ fluctuations having lower λ, then these two apparently competitive theories become the opposite sides of the same coin.

It is interesting to ask whether the cold DM picture can account for the wide range of morphologies displayed by clusters of galaxies in X-ray and optical-band observations, ranging from regular, apparently relaxed configurations to complex, multicomponent structures. (Cf. Figs. 1.8 and 1.9.) Preliminary results are encouraging. In particular, simulations show that large central condensations form quickly and can grow by subsequent mergers to form cD galaxies if most of the DM is initially in halos around the baryonic substructures, as expected for cold DM, but not if the DM is initially distributed rather diffusely throughout the whole cluster, as expected for hot DM.[47]

Consider finally the difference in Fig. 4.4 between the solid and dashed lines. The dashed lines, representing a lower-density universe ($\Omega = 0.2$), curve backward at the largest masses and lie far away from the circle representing the cores of the richest clusters, Abell classes 2 and 3. Since these regions of very high galaxy density contain at least several percent of the mass in the universe, the circle should lie between the 2 and 3σ lines (assuming Gaussian statistics).

It does so for the solid ($\Omega = 1$) lines, but not for the dashed lines. At face value, this is evidence favoring an Einstein-de Sitter universe for cold DM. However, there are at least two reasons why this argument should probably not be taken too seriously. First, the velocity dispersions represented by the Abell cluster circle in Fig. 4.4 correspond to the cluster cores. The model curves on the other hand refer to the entire virialized cluster, over which the velocity dispersion is considerably lower (as indicated by the arrow attached to the circle in Fig. 4.4). Second, the assumption of spherical symmetry used in obtaining both sets of curves in the figure is only an approximation. The initial collapse is probably often quite anisotropic — more like a Zeldovich pancake than a sphere. It is therefore preferable to compare these data with N-body simulations rather than with the simple model represented by the curves in Fig. 4.4. This will require N-body simulations of large dynamical range, which can perhaps be achieved by putting many mass points into one cell of the P^3M–type simulations. Until this becomes possible the data in the figure do not allow a clear-cut discrimination between the $\Omega = 0.2$ and $\Omega = 1$ cases, especially if the Hubble parameter h is allowed to vary simultaneously within the observationally allowed range, as has been assumed.

It is important to appreciate that some means of suppressing galaxy formation in regions of lower than average density is required in the cold DM model both for $\Omega = 1$ and for $\Omega \approx 0.2$.[26] In the former case, this is needed to hide most of the mass. In a low-Ω universe, on the other hand, large regions of much lower density than average cannot form by gravitation alone.[235] The amplitude of linear fluctuations stops growing in an open universe when it goes into free expansion (i.e. for $z \lesssim \Omega_o^{-1}$), and large voids would stop expanding in comoving coordinates. Thus formation of large galaxies must be inhibited somehow in regions of moderately low density if the number density of of galaxies in voids is less than one-quarter of the average, the quoted upper limit for the Boötes void.[236] Since it is possible to imagine various physical phenomena which might contribute to suppression of galaxy formation in underdense regions,[26] this is

not necessarily a problem for the cold DM picture. But it does imply that the distribution of luminous galaxies probably cannot be an accurate tracer of the mass distribution on large scales in any version of the cold DM universe.

4.3 N-BODY SIMULATIONS OF LARGE SCALE STRUCTURE

Elaborate N-body simulations of the evolution of large scale structure in a universe dominated by cold DM have recently been carried out by Davis, Efstathiou, Frenk, and White (DEFW)[201] using the P^3M (particle-particle/particle-mesh) scheme,[237] in which the equations of motion are integrated directly for nearby particles while the gravitational effects of more distant masses are calculated by applying fast Fourier transform methods to Poisson's equation, with periodic boundary conditions at the edges of the computational volume. A wide range of structures arise in such simulations, including filaments, superclusters, and large regions of fairly low density. As Fig. 3.6 shows, the distribution of galaxies is qualitatively similar to observations. A more quantitative comparison reveals various problems, however.

Comparison with Obserations

Ensembles of simulations were run for $\Omega = 1$ and also for $\Omega = 0.2$. In the $\Omega = 1$ simulations, there is no time when the mass autocorrelation function is a good match in both slope and amplitude to that of the galaxies. Moreover, the rms peculiar velocities of pairs of particles are several times larger than those actually observed for galaxies (discussed in §2.4 above). For $\Omega = 0.2$, the mass autocorrelation function matches the galaxy autocorrelation function fairly well for $h \approx 1.1$, and the peculiar velocities are in better agreement with observation at separations of $\sim 5h^{-1}$ Mpc although still too large on smaller scales. The parameter Q which relates the galaxy three-point and two-point correlation functions (see eq. 2.26) is observed to be about unity and weakly (if at all) dependent on the size of the three-point triangle; Q determined from the simulations is too large by almost a factor of 2 on small scales for $\Omega = 1$ and

even a little larger for $\Omega = 0.2$, though in better argeement with $Q \approx 1$ on larger scales. An independent set of N-body simulations[238] found that in neutrino models Q varies strongly with scale in a manner completely inconsistent with observations; while the weaker dependence of Q in cold DM models is more in accord with observations, the residual variation still exceeds any seen in the data.

Thus, while the cold DM simulations are not in such gross disagreement with observations as the ν simulations, neither the Einstein-de Sitter nor the open universe cold DM simulations agree quantitatively with observations of galaxy spatial and velocity distributions. Although the open universe simulations are perhaps in better agreement, they are in potential conflict with the latest data[143] on small-angle fluctuations in the cosmic background radiation, which imply[174,173] that $\Omega \geq 0.2h^{-4/3}$ unless there is significant reheating of the intergalactic medium after recombination (cf. eq. 3.8). There is also a serious problem with the age of the universe (see Figs. 2.4 and 2.5): $t_o = 8.3h^{-1}$ Gy for $\Omega = 0.2$ (assuming $\Lambda = 0$), which is in severe conflict with globular cluster age estimates and even in conflict with nuclear cosmochronometry for $h \gtrsim 1$. Finally, $\Omega = 0.2$ is in conflict with the expectations of the inflationary universe hypothesis.

Flat Universe with Positive Cosmological Constant

The age problem can be relaxed and inflationary models allowed if the cosmological constant is non-zero.[69,239] If, as is usually supposed, there is much more than enough inflation to solve the flatness and other cosmological problems, then the curvature k becomes vanishingly small. Einstein's equation (2.8) evaluated at the present is then

$$1 = \Omega + \frac{\Lambda}{3H^2}. \tag{4.6}$$

If $\Lambda = 0$, this of course implies $\Omega = 1$, but a positive cosmological constant can compensate for $\Omega < 1$ in an inflationary universe.

In a universe characterized by eq. (4.6), the time dependence of the scale factor in the epoch of nonrelativistic matter domination ($\rho \propto R^{-3}$) is

$$R(t) \propto \sinh[(3\Lambda)^{1/2}t/2]^{2/3}, \tag{4.7}$$

and the age of the universe is

$$t_o = \frac{2}{3H_o(1-\Omega_o)^{1/2}} \sinh^{-1}(\Omega_o^{-1}-1)^{1/2} \tag{4.8}$$

for positive Λ. Figure 4.5 shows the corresponding values of cosmological density parameter Ω_o and Hubble parameter h for various values of t_o. Comparison with the analogous plot for the case $\Lambda = 0$, Fig. 2.5, shows that it is now somewhat easier to bring the predicted age t_o into consistency with globular cluster age estimates. Moreover, since fluctuations grow considerably more after $z \lesssim \Omega_o^{-1}$ in the present case compared to the case $\Lambda = 0$,[239] there is now no conflict with the observational constraints[143] on $\Delta T/T$ for $\Omega \approx 0.2$.

DEFW ran one N-body simulation of the case (4.6) with $\Omega_o = 0.2$, and found that the results closely resembled those found for $\Lambda = 0$ and the same value of Ω_o. Matching the galaxy autocorrelation function again leads to $h \approx 1.1$, and although this now corresponds to $t_o = 10$ Gy rather than 7.5 Gy, it still leads to a universe which is embarrassingly young.

"Biased" Galaxy Formation

Thus far, I have discussed DEFW's attempts to fit their cold dark matter N-body simulations to the observed universe under the assumption that the distribution of bright galaxies closely approximates the distribution of dark matter. The results are better than for the neutrino simulations, but not really very impressive. DEFW found much better agreement with the data, however, when they tried assuming that bright galaxies form only at relatively high peaks of the linear density distribution. The visible galaxies thus are pictured as being like the unsubmerged peaks of a vast flooded mountain range.

DEFW's biasing procedure depends on two parameters: the threshold ν and the smoothing size r_o. After smoothing, they identified all peaks of amplitude at least ν times the rms density fluctuation and tagged the nearest particle to each as a "galaxy", whose motion is subsequently followed over the course of the simulation. With a biasing threshold of $\nu = 2.5$ for their Einstein-de Sitter ($\Omega = 1$) simulations, DEFW found that the correlation length of the galaxies exceeds that of the dark matter by a factor of about 2.4, and the observed two- and three-point correlations are fit quite well for $h = 0.44$ (corresponding to $t_o = 15$ Gy). The simulations also fit the observed peculiar velocity distribution rather well (about as well as the unbiased $\Omega = 0.2$ simulations did). Since the dark matter is mostly rather smoothly distributed in the biased model, the peculiar velocities generated are much smaller than in the unbiased model with the same density ($\Omega = 1$).

Since both star and galaxy formation are very poorly understood, it is going to be hard to put biased galaxy formation on a solid theoretical footing. But the basic idea that bright galaxies should form only at the highest peaks of the linear density distribution is a rather plausible one, and a number of possible physical mechanisms have been proposed.[240] Moreover, as BFPR already pointed out in connection with Fig. 4.4, the greater range of velocity dispersions observed in groups and clusters of a given mass as compared to galaxies is evidence for biasing: Groups of relatively low velocity dispersion, which presumably arose from lower amplitude fluctuations, are visible because of the galaxies that they contain. The absence of bright galaxies of correspondingly low rotation velocity (for spirals) or velocity dispersion (for ellipticals) suggests that only the fluctuations of larger amplitude form visible galaxies.

Very Large Scale Structure

How well can the cold DM picture account for the largest structures observed in the universe: superclusters and voids? To answer quantitatively, it is necessary to compare predictions with observations statistically. The challenge is to devise

appropriate statistical tests which not only are amenable both to calculation and observation, but also actually discriminate between alternative theories. I will here discuss several statistical tests: cluster autocorrelations, the size distribution of voids, percolation, and correlations between the orientation of rich clusters and the directions to nearby ones (Binggeli's statistic).

The autocorrelation function of rich clusters appears to have the same slope as that of galaxies, but with a correlation length r_o about five times greater.[172,79] (See Fig. 2.8.) As Kaiser[241] has emphasized, statistical effects in the formation of rich clusters can lead to just such an effect: $\xi_{clusters}(r) \approx A\xi_{density}(r)$, with the amplification factor A increasing with the richness of the clusters, in qualitative agreement with observation. Three-point and higher order correlations are also enhanced in a calculable way,[242] and it may soon be possible to check whether these statistical predictions are in accord with observations. If this is indeed the right explanation for the enhanced clustering of rich clusters, then if Abell clusters are actually positively correlated at $r \approx 100h^{-1}$ Mpc, the underlying mass density must also be. DEFW point out that the correlation function corresponding to the cold DM linear density distribution is negative for $r > 18(\Omega h^2)^{-1}$ Mpc, and thus argue that for this picture to make sense, $\Omega h < 0.18$. But this is in trouble with $\Delta T/T$ unless $\Lambda > 0$, and with globular cluster ages. Biased galaxy formation does not help. Of course, it is possible that the strong correlation of rich clusters is at least partly caused by something entirely different from the effect that Kaiser discussed, in which case this line of argument is irrelevant.Indeed, N-body simulations of the clustering of clusters[243,66] suggest that some additional physical effect beyond the Gaussian statistics of rare events will be necessary to account for the large amplitude A that is observed.

The several thousand acurate galaxy redshifts presently available are not yet enough to tell us very much about the statistics of voids, but DEFW made some preliminary comparisons with their $\Omega = 1$ N-body simulations. They did this by distributing the particles in redshift space, in which two of the coordinates are positions and the third is the corresponding velocity component, and then

constructing what amounts to a volume-limited redshift catalogue. They find that 8-10% of the cubes of size 32 h^{-1} Mpc have density less than half the mean, while only 0.6% of such cubes have density less than 30% of the mean. Large, low-density regions are even less common in open universe simulations, and in the biased galaxy versions of $\Omega = 1$ simulations. For comparison, an analysis of galaxy counts in randomly placed spheres of radius 20 h^{-1} Mpc (i.e. with the same volume as a 32 h^{-1} Mpc cube) suggests that 20% of such spheres have density less than half the mean, about twice the number found in the simulations.[201] The 30 h^{-1} Mpc radius Boötes void appears to have a density of bright galaxies less than 25% of the mean.[236] There is nothing like it in the simulations.

The percolation statistic was advocated by Shandarin[244] and his colleagues[178,245] for comparing models with the observed galaxy distribution. The basic idea is to draw a sphere of radius r around each galaxy, and determine how the length of the longest chain of overlapping spheres depends on r. Unfortunately, percolation turns out not to be a very useful statistic: it is rather sensitive to irrelevant features such as sampling parameters and the depth of the galaxy redshift survey, but not very sensitive to differences between neutrino and cold DM simulations.[246,247]

Although there is no statistically significant alignment of ordinary galaxies with larger-scale structures,[56] the brightest galaxies in rich clusters are often aligned with their parent clusters, and Binggeli[53] found that nearby rich clusters tend to point toward each other. More precisely, Binggeli found a correlation between the position angle of the major axis of a cluster and the position angles of the lines connecting its center to those of neighboring clusters. This correlation was found to be strongest for clusters separated by less than 20 h^{-1} Mpc. Dekel and collaborators[248,56] have looked for similar effects in N-body simulations. They found that there are none in Poisson simulations (i.e., simulations that start with randomly distributed mass points), and that in neutrino simulations the correlations are comparable to but a little stronger than those Binggeli found. Correlations similar to Binggeli's were obtained in Dekel's "AI"

simulations (which have Poisson noise superimposed on a neutrino-type power spectrum), and in DEFW's cold DM simulations only for $\Omega h < 0.2$.

Very recently, Struble and Peebles[54] have repeated and extended Binggeli's observations with a sample of 237 clusters, more than five times the number (44) in Binggeli's sample. They find that there is at most marginal alignment, and conclude that Binggeli most probably hit a statistical fluke. They also remark that in the light of their data, the success of Dekel et al.[248] in reproducing the alignments in Binggeli's data sample with neutrino-type simulations is evidence against such models.

Binggeli's statistic does discriminate between different theoretical models, and allows ready comparison with data. Related statistics are being developed by Dekel, Vishniac, and others (personal communications). It will be very interesting to see the results of further analysis of models and comparison with data.

4.4 Summary and Prospect

I have a hard time keeping all this information straight, so I imagine that you might also. In order to help, I have prepared a "Consumer's Report on Dark Matter", Fig. 4.6 (following the notation of the *Consumer Reports* magazine, which is known for its authoritative reviews of American consumer goods). I have tried to be reasonably objective in rating how well each of the three models discussed here — hot, warm, and cold DM — accords with each category of data considered.

The first column of Fig. 4.6 lists the three sorts of dark matter, and candidates for each ($\widetilde{g}$ = gravitino, $\widetilde{\gamma}$ = photino, PBH = primordial black holes). The next two columns summarize theoretical expectations regarding the fluctuation spectrum (solid line), the dependence of the ratio of total to luminous mass on size, and the contents of voids. The evidence on M/M_{lum} summarized in Fig. 1.11 is evidently most consistent with cold DM. Because free streaming damps

out all but supercluster-size structures in the hot DM model and galaxies can form only after the pancake collapse of these protosuperclusters, the voids (i.e., the regions between the pancakes) are expected to contain essentially no galaxies. But since galaxies form before large scale structure in the warm and cold DM models, voids are expected to contain at least some small galaxies. Preliminary, indirect evidence favoring the latter expectation has recently been published:[249] absorption lines corresponding to triply-ionized silicon and carbon have been detected at the redshifts of the Perseus-Pisces and Boötes voids in the ultraviolet spectra of background quasars. In the hot DM picture, any gas in the voids should be of primordial composition.

According to the "Consumers Report", the two most severe difficulties of the hot DM model are the phase space constraint indicating that the heavy halos that are apparently associated with dwarf spheroidal galaxies cannot be made of light neutrinos, and the mismatch between the recent ($z < 2$) formation of superclusters indicated by density and velocity data and the early formation of galaxies and quasars. I have just discussed the difficulty for hot DM posed by the new data of Struble and Peebles on Bingelli's statistic; and I mentioned earlier in this lecture the indications that it is hard to form cD galaxies by mergers in small clusters unless the protogalaxies are assumed to have individual heavy halos, as expected for cold or warm but not hot DM. Peculiar velocities are too large in all $\Omega = 1$ models unless galaxy formation is "biased"; and although peculiar velocities are smaller in open cold models their distribution does not agree in detail with observations. Yet another difficulty with hot DM is that the dense clumps that form in this model cannot be correlated with any known population of objects.[250] Even if galaxy formation were for some reason suppressed in the densest regions, they would provide a highly visible population of X-ray sources.

The cold dark matter model can be credited with several successes in explaining galaxy and cluster formation.[26] It predicts roughly the observed mass range of galaxies, the dissipational nature of galaxy collapse, and dissipationless galactic halos and clusters with the observed Faber-Jackson and Tully-Fisher relations

for galaxies. In addition, it may also provide natural explanations for galaxy-environment correlations and for the differences in angular momenta between ellipticals and spiral galaxies.

If the cold dark matter is distributed like bright galaxies, then $\Omega \approx 0.2$. However, this version of the cold DM model is in serious conflict with the observational limit on small-angle $\Delta T/T$ fluctuations, and also with globular cluster age estimates since $h \approx 1$ for $\Delta T/T$ fitting the galaxy autocorrelation function. The $\Delta T/T$ problem is removed and the age problem is ameliorated, but only a little, in the zero-curvature version of this model, with cosmological constant $\Lambda = 3H^2(1 - \Omega)$. Perhaps the best overall fit to the data is obtained in the $\Omega = 1$ cold DM model with biased galaxy formation. This version may have more trouble than the open universe ones in accounting for the distribution of voids, however.

The cold DM models' possible difficulty in producing voids like that in Boötes and in explaining the enhanced clustering of rich clusters both suggest that there is still an important physical ingredient missing in the cold DM scheme — maybe a feature in the fluctuation spectrum on a "superpancake" scale of order 100-150 h^{-1} Mpc, as Dekel has suggested.[251,56] Perhaps the least ugly way of obtaining this is from the added growth of fluctuations exceeding the photon-baryon Jeans mass in a universe with adiabatic primordial fluctuations and roughly equal amounts of cold dark matter (to preserve the galaxy-size fluctuations, as usual) and baryonic matter (mostly dark now — perhaps "jupiters"): $\Omega_b \approx \Omega_{CDM} \approx 0.1$.

As always, there is work for theorists. But the special excitement of the present era in astrophysics is that there is also plenty of relevant observational data, with the prospect of lots more coming soon. Fig. 4.7 summarizes the relative sensitivities of a number of existing and planned astronomical instruments. Just to give one example to illustrate the sorts of new data that will soon be available, consider the observation of ultraviolet Lyman α absorption lines with

the Space Telescope, now scheduled to be launched in 1986. As I explained in connection with Fig. 2.9, some of the Ly α absorption lines seen in quasar spectra are apparently caused by clouds of neutral hydrogen distributed along the line of sight from the quasar to us. Dekel[252] has argued that the absence of autocorrelation in the Ly α absorption line "forest" of individual quasars,[253] and of cross-correlation in Ly α forests of two quasars separated by $\sim 1\,h^{-1}$ Mpc perpendicular to the line of sight,[254] indicate that the formation of superclusters did not begin until $z \lesssim 2$. (This conclusion also follows from other lines of argument, as I have discussed.) Oort[255] argues on the contrary that the observation[256] of two pairs of quasars at $z = 2.83$ and 2.85 and $z = 3.14$ and 3.16 suggests that superclustering has existed since $z \approx 3$, and that the lack of cross-correlation in the Ly α forests of the neighboring quasars can be explained by a filamentary distribution of hydrogen clouds in superclusters. If Dekel is right, then ultraviolet observations of the Ly α forest at redshifts $z \lesssim 1.5$ will show the onset of superclustering. (It is of course also possible that the Ly α clouds are not associated with superclusters, but instead occur only in voids.) Similar observations can also determine the composition of the gas in voids.[249] Such data would obviously be an important constraint on cosmological theories for galaxy formation and large scale structure.

Remarkably enough, such data may also shed important light on the interactions of elementary particles on very small scales. Fig. 4.8 is redrawn from a sketch by Shelley Glashow which was reproduced in the *New York Times Magazine*.[257] Glashow uses the snake eating its tail — the uroboros, an ancient symbol associated with creation myths[258] — to represent the idea that gravity may determine the structure of the universe on both the largest and smallest scales. But there is another fascinating aspect to this figure. There are left-right connections across it: medium small to medium large, even smaller to even larger, and so on. Not only does electromagnetism determine structure from atoms to mountains,[259] and the strong and weak interactions control the properties and compositions of stars and solar systems. The dark matter, which may hold the

key to understanding the origin of galaxies, clusters, and superclusters, may itself reflect fundamental physics below the weak interaction scale. And if cosmic inflation is to be believed, cosmological structure on scales even larger than the present horizon arose from interactions on the grand unification scale.

Acknowledgments

These lectures are based on research done in collaboration with George Blumenthal, Sandra Faber, and Martin Rees — all of whom I would like to thank for their efforts to teach me astrophysics, and also for allowing me to use material from our joint publications, especially Refs. 26, 135, 203, and 204. In preparing these lectures, I have also benefitted from George Blumenthal's unpublished notes for his course on Cosmology at the University of California, Santa Cruz. I thank Howard Haber for his careful reading of this manuscript and for several helpful suggestions. I am very grateful to Sidney Drell for the hospitality of SLAC during 1984, and to Nicola Cabibbo for inviting me to give these lectures at Varenna — surely one of the lovliest places in the civilized world. Finally, I acknowledge partial support from the National Science Foundation under grant PHY-81-15541.

REFERENCES

1. J. Heidmann: in *Cosmology, History, and Theology*, edited by W. Yourgrau and A. D. Breck (Plenum, New York, 1977), p. 39.
2. P. W. Hodge: *Galaxies and Cosmology* (McGraw-Hill,New York, 1966).
3. R. Sancisi: in *The Structure and Evolution of Normal Galaxies*, edited by S. M. Fall and D. Lynden-Bell (Cambridge University Press, 1981), p. 149.
4. S. M. Fall: "General Properties of Galaxies," in *The Structure and Evolution of Normal Galaxies*, edited by S. M. Fall and D. Lynden-Bell (Cambridge, 1981), p. 149.
5. S. M. Faber and J. S. Gallagher: *Ann. Rev. Astron. Astrophys.*, **17**, 135 (1979).
6. A. Bosma: *Astron. J.*, **86**, 1825 (1981).
7. V. C. Rubin, W. K. Ford, and N. Thonnard: *Astrophys. J.*, **238**, 471 (1980).
8. V. C. Rubin: *Science*, **220**, 1339 (1983).
9. F. Zwicky: *Helv. Phys. Acta*, **6**, 110 (1933).
10. J. P. Ostriker and P. J. E. Peebles: *Astrophys. J.*, **186**, 467 (1973). J. P. Ostriker, P. J. E. Peebles, and A. Yahil: *Astrophys. J. Lett.*, **193**, L1 (1974).
11. L. Blitz, M. Fich, S. Kulkarni: *Science*, **220**, 1233 (1983), and references therein.
12. C. Frenk and S. D. M. White: *Mon. Not. R. Astron. Soc.*, **193**, 295 (1980).
13. K. A. Innanen, W. E. Harris, and R. F. Webbink: *Astron. J.*, **88**, 338 (1983).
14. F. D. A. Hartwick and W. L. W. Sargent: *Astrophys. J.*, **221**, 512 (1978).
15. T. Murai and M. Fujimoto: *Pub. Astron. Soc. Japan*, **32**, 581 (1980).

16. D. N. C. Lin and D. Lynden-Bell: *Mon. Not. R. Astron. Soc.*, **198**, 707 (1982).

17. M. R. S. Hawkins: *Nature*, **303**, 406 (1983).

18. V. C. Rubin: *Sci. Am.*, **248**(6), 96 (June 1983).

19. J. Einasto, A. Kaasik, and E. Saar: *Nature*, **250**, 309 (1974); D. Lynden-Bell: in *Astrophysical Cosmology: Proceedings of the Study Week on Cosmology and Fundamental Physics*, edited by H. A. Brück, G. V. Coyne, and M. S. Longair (Vatican, 1982), p. 85.

20. J. A. Tyson, F. Valdes, J. F. Jarvis, and A. P. Mills, Jr.: *Astrophys. J. Lett.*, **281**, L59 (1984).

21. G. Efstathiou and J. Silk: *Fund. Cosmic Phys.*, **9**, 1 (1983).

22. J. S. Gallagher and D. A. Hunter: *Ann. Rev. Astron. Astrophys.*, **22**, 37 (1984).

23. S. M. Faber and R. E. Jackson: *Astrophys. J.*, **204**, 668 (1976).

24. M. Aaronson, J. Huchra, and J. Mould: *Astrophys. J.*, **229**, 1 (1979).

25. D. Burstein, V. C. Rubin, N. Thonnard, and W. K. Ford: *Astrophys J.*, **253**, 539 (1982).

26. G. R. Blumenthal, S. M. Faber, J. R. Primack, and M. J. Rees: *Nature*, **311**, 517 (1984). [BFPR]

27. S. Faber: in *Astrophysical Cosmology*, edited by H. A. Brück, G. V. Coyne, and M. S. Longair (Vatican, 1982).

28. J. E. Gunn: *Phil. Trans. R. Soc. Lond.*, **A296**, 313 (1980).

29. G. Illingworth: in *The Structure and Evolution of Normal Galaxies*, edited by S. M. Fall and D. Lynden-Bell (Cambridge, 1981), p. 27.

30. W. Forman, C. Jones, and W. Tucker: CfA preprint 2096 (1985).

31. G. R. Blumenthal, S. M. Faber and I are presently preparing a paper for *Arr. Rev. Astron. Astrophys.* on dark matter.

32. J. R. Primack and G. R. Blumenthal: in *Fourth Workshop on Grand Unification*, edited by H. A. Weldon, P. Langacker, and P. Steinhardt (Boston, Birkhäuser, 1983), p. 256.

33. G. O. Abell: *Astrophys. J. Suppl.*, **3**, 211 (1958).

34. F. Zwicky, E. Herzog, P. Wild, M. Karpowicz, and C. T. Kowal: *Catalogue of Galaxies and Clusters of Galaxies*, 6 vols. (Pasadena, Caltech, 1961-68).

35. M. J. Geller and J. P. Huchra: *Astrophys. J. Suppl.*, **52**, 61 (1983).

36. J. P. Huchra, M. Davis, D. W. Latham, and J. Tonry: *Astrophys. J. Suppl.*, **52**, 89 (1983).

37. N. Bahcall: *Ann. Rev. Astron. Astrophys.*, **15**, 505 (1977).

38. W. Forman and C. Jones: *Ann. Rev. Astron. Astrophys.*, **20**, 547 (1982).

39. M. J. Geller and T. C. Beers: *Pub. Astron. Soc. Pacific*, **94**, 421 (1982).

40. J. H. Oort: *Ann. Rev. Astron. Astrophys.*, **21**, 373 (1983).

41. M. J. Geller: in *Clusters and Groups of Galaxies*, edited by F. Mardirossian, G. Giuricin, and M. Mezzetti (Dordrecht, Holland, Reidel, 1984).

42. A. Dressler: *Astrophys. J.*, **236**, 351 (1980).

43. M. Postman and M. J. Geller: *Astrophys. J.*, **281**, 95 (1984).

44. H. R. Butcher and A. Oemler: *Astrophys. J.*, **226**, 559 (1978).

45. W. H. Press and M. Davis, *Astrophys. J.*, **259**, 449 (1982).

46. S. D. M White: in *Morphology and Dynamics of Galaxies*, edited by L. Martinet and M. Mayor (Geneva Observatory, 1982).

47. T. X. Thuan and W. Romanishin: *Astrophys. J.*, **248**, 439 (1981). A. Cavaliere, P. Santangelo, G. Tarquini, and N. Vittorio: in *Clusters and Groups*

of Galaxies, edited by F. Mardirossian, G. Giuricin, and M. Mezzetti (Dordrecht, Holland, Rei del, 1984), p. 499.

48. A. Dressler: *Ann. Rev. Astron. Astrophys.*, **22**, 185 (1984).

49. S. A. Gregory and L. A. Thompson: *Sci. Am.*, **246**(3), 106 (March 1982).

50. R. B. Tully: *Astrophys. J.*, **257**, 389 (1982).

51. S. A. Gregory, L. A. Thompson, and W. G. Tifft: *Astrophys. J.*, **243**, 411 (1981).

52. R. F. Kirshner, A. Oemler, P. L. Schechter, and S. A. Shectman: *Astrophys. J. Lett.*, **248**, L57 (1981) and in *Evolution of the Universe and Its Present Structure, IAU Symposium No. 104*, edited by G. O. Abell and G. Chincarini (Dordrecht, Ho lland, Reidel, 1983), p. 197.

53. B. Binggeli: *Astron. Astrophys.*, **107**, 338 (1982).

54. M. F. Struble and P. J. E. Peebles: *Astron. J.* (submitted).

55. A. Dekel, M. J. West, and S. J. Aarseth: *Astrophys. J.*, **279**, 1 (1984).

56. A. Dekel: *Eighth Johns Hopkins Workshop on Current Problems in Particle Theory: Particles and Gravity*, edited by G. Domokos and S. Kovesi-Domokos (Singapore, World Scientific Publishing, in press).

57. L. Hart and R. D. Davies: *Nature*, **297**, 191 (1982).

58. M. Davis and P. J. E. Peebles: *Ann. Rev. Astron. Astrophys.*, **21**, 109 (1983).

59. Ya. B. Zeldovich, J. Einasto, and S. F. Shandarin: *Nature*, **300**, 407 (1982).

60. S. F. Shandarin, A. G. Doroshkevich, and Ya. B. Zeldovich: *Sov. Phys. Usp.*, **26**, 46 (1983).

61. P. J. E. Peebles: *Science*, **224**, 1385 (1984).

62. C. M. Will: *Theory and Experiment in Gravitational Physics*, (Cambridge University Press, 1981).

63. W. Rindler: *Essential Relativity: Special, General, and Cosmological*, Revised Second Ed. (New York, Springer-Verlag, 1977).

64. M. Rowan-Robinson: *Cosmology*, Second Ed. (Oxford, 1981).

65. S. Weinberg: *Gravitation and Cosmology: Principles and Applications of the General Theory of Relativity* (New York, Wiley, 1972).

66. C. W. Misner, K. S. Thorne, and J. A. Wheeler: *Gravitation* (San Francisco, Freeman, 1973).

67. E. R. Harrison: *Cosmology: the Science of the Universe* (Cambridge, 1981).

68. See, e.g., Ref. 63, §9.9.

69. A. Sandage and G. A. Tammann: in *First ESO-CERN Symposium: Large-Scale Structure of the Universe, Cosmology and Fundamental Physics*(1984).

70. G. de Vaucouleurs: in *Tenth Texas Symposium on Relativistic Astrophysics*, edited by R. Ramaty and F. C. Jones, *Ann. N. Y. Acad. Sci.*, **375**, 90 (1981); and in *Clusters and Groups of Galaxies*, edited by F. Mardirossian, G. Giuricin, and M. Mezzetti (Dordrecht, Holland, Reidel, 1984), p. 29.

71. W. D. Arnett, D. Branch, and J. C. Wheeler: *Nature*, in press.

72. P. J. E. Peebles: *The Large-Scale Structure of the Universe* (Princeton University Press, 1980).

73. M. Davis and P. J. E. Peebles: *Astrophys. J.*, **267**, 465 (1983).

74. A. J. Bean, G. Efstathiou, R. S. Ellis, B. A. Peterson, and T. Shanks: *Mon. Not. Roy. Astron. Soc.*, **205**, 605 (1983).

75. A. Dekel and S. J. Aarseth: *Astrophys. J.*, **283**, 1 (1984).

76. P. J. E. Peebles and E. J. Groth: *Astrophys. J.*, **196**, 1 (1975).

77. M. G. Hauser and P. J. E. Peebles: *Astrophys. J.*, **185**, 757 (1973).

78. N. A. Bahcall and R. M. Soneira: *Astrophys. J.*, **270**, 20 (1983).

79. A. A. Klypin and A. I. Kopylov: *Sov. Astron. Lett.*, **9**, 41 (1984).

80. Ref. 72, eq. (75.10).

81. P. J. E. Peebles: in *Proceedings of the Tenth Texas Symposium on Relativistic Astrophysics, Ann. NY Acad. Sci.*, **375**, 157 (1981).

82. Cf. also A. Dressler: *Astrophys. J.*, **281**, 512 (1984).

83. G. L. Hoffman and E. E. Salpeter: *Astrophys. J.*, **263**, 485 (1982).

84. R. J. Harms, H. C. Ford, R. Ciardullo, and F. Bartko: in *Tenth Texas Symposium on Relativistic Astrophysics*, edited by R. Ramaty and F. C. Jones, *Ann. N. Y. Acad. Sci.*, **375**, 178 (1981).

85. R. J. Harms, H. C. Ford, and R. Ciardullo: in *Early Evolution of the Universe and Its Present Structrure*, I. A. U. Symp. No. 104, edited by G. O. Abell and G. Chincarini (Reidel, Dordrecht, 1983), p. 285.

86. J. Yang, M. S. Turner, G. Steigman, D. N. Schramm, and K. A. Olive: *Astrophys. J.*, **281**, 493 (1984).

87. R. G. Kron: Yerkes Observatory preprint (1983).

88. E. J. Wampler, C. M. Gaskell, W. L. Burke, and J. A. Baldwin: *Astrophys. J.*, **276**, 403 (1984).

89. H. S. Murdoch: *Mon. Not. Roy. Astr. Soc.*, **202**, 987 (1983).

90. W. H. Press and E. T. Vishniac: *Astrophys. J.*, **239**, 1 (1980).

91. J. M. Bardeen: *Phys. Rev. D* **22**, 1882 (1980).

92. Ref. 72, §11.

93. G. R. Blumenthal and J. R. Primack: preprint SLAC-PUB-3388 (1984).

94. P. J. E. Peebles: *Astrophys. J.*, **75**, 13 (1970).

95. D. Lynden-Bell: *Mon. Not. Roy. Astr. Soc.*, **136**, 101 (1967).

96. F. H. Shu: *Astrophys. J.*, **225**, 83 (1978).

97. S. Chandrasekhar: *Principles of Stellar Dynamics* (New York, Dover, 1960).

98. J. E. Gunn: in *Astrophysical Cosmology: Proceedings of the Study Week on Cosmology and Fundamental Physics*, edited by H. A. Brück, G. V. Coyne, and M. S. Longair (Vatican, 1982), p. 233.

99. See, e.g., Peebles[72] §20-21 and the references cited there.

100. Ya. B. Zeldovich: *Astron. and Astrophys.*, **5**, 84 (1970).

101. Ya. B. Zeldovich: in *The Large Scale Structure of the Universe*, IAU Symposium No. 78, edited by M. S. Longair and J. Einasto (Dordrecht, Reidel, 1978), p. 409.

102. A. Dekel: *Astrophys. J.*, **264**, 373 (1983).

103. S. J. Aarseth and J. Binney: *Mon. Not. Roy. Astr. Soc.*, **185**, 227 (1978).

104. A. Guth: *Phys. Rev. D* **23**, 347 (1981).

105. S. W. Hawking and G. F. R. Ellis: *The Large Scale Structure of Space-Time* (Cambridge University Press, 1973).

106. E. B. Gliner and I. G. Dymnikova: *Sov. Astron. Lett.*, **1**, 93 (1975).

107. R. Brout, F. Englert, and E. Gunzig: *Gen. Rel. Grav.*, **10**, 1 (1979); R. Brout, F. Englert, and P. Spindel: *Phys. Rev. Lett.*, **43**, 417 and (E) 890 (1979); R. Brout et al.: *Nucl. Phys.*, **B170**, 228 (1980). This work is reviewe d by F. Englert: in *Physical Cosmology*, Les Houches Session XXXII (North Holland Pub. Co., Amsterdam, 1980), and in the 1981 Cargese Lectures, edited by M. Levy.

108. D. Kazanas: *Astrophys. J. Lett.*, **241**, L59 (1980).

109. A. H. Guth and P. J. Steinhardt: *Sci. Am.* **250**(5), 116 (May 1984).

110. A. H. Guth and E. Weinberg: *Phys. Rev. D***23**, 826 (1981); *Nucl. Phys.*, **B212**, 321 (1983).

111. A. D. Linde: *Phys. Lett.*, **108B**, 389 (1982). A. Albrecht and P. J. Steinhardt: *Phys. Rev. Lett.*, **48**, 1220 (1982).

112. K. A. Olive: in *Grand Unification With and Without Supersymmetry, and Cosmological Implications*, International School for Advanced Studies Lecture Series No. 2 (Singapore, World Scientific Publishing, 1984).

113. Recent reviews include K. A. Olive, Ref. 112; M. S. Turner: in NATO Advanced Study Institute on Quarks and Leptons, Max Planck Institute, Munich, September 1983 (Fermilab preprint Conf-84/60-A, to be published); P. J. Steinhardt: in *Fourth Worksh op on Grand Unification*, edited by H. A. Weldon, P. Langacker, and P. Steinhardt (Birkhäuser, Boston, 1983); and R. Brandenberger: *Phys. Rep.*, (in press).

114. A. H. Guth: in *The Very Early Universe*, edited by G. W. Gibbons, S. Hawking, and S. Siklos (Cambridge University Press, 1983).

115. L. Abbott: Brandeis preprint (1984).

116. A. H. Guth and S.-Y. Pi: *Phys. Rev. Lett.*, **49**, 1110 (1982). Similar calculations were done by S. Hawking: *Phys. Lett.*, **115B**, 295 (1982); A. Starobinskii: *Phys. Lett.*, **117B**, 175 (1982); and J. Bardeen, P. J. Steinhardt, a nd M. S. Turner: *Phys. Rev. D* **28**, 679 (1983).

117. E. R. Harrison: *Phys. Rev. D***1**, 2726 (1970).

118. Ya. B. Zeldovich: *Mon. Not. Roy. Astr. Soc.*, **160**, 1P (1972).

119. B. J. Carr: *Astrophys. J.*, **201**, 1 (1975); J. D. Barrow and B. J. Carr: Mon. Not. Roy. Astr. Soc., **182**, 537 (1978).

120. B. J. Carr: *Comm. Astr. and Space Phys.*, **7**, 161 (1978).

121. P. J. Steinhardt and M. S. Turner: *Phys. Rev. D* **29**, 2162 (1984).

122. For example, Q. Shafi and A. Vilenkin: *Phys. Rev. Lett.*, **52**, 691 (1984).

123. Ya. B. Zeldovich: *Mon. Not. Roy. Astr. Soc.*, **192**, 663 (1980). A. Vilenkin: *Phys. Rev. D***24**, 2082 (1981). The recent work on cosmic strings has been reviewed by T. W. B. Kibble: in *Proceedings of the International*

Symposium on Ph ase Transitions in the Very Early Universe (Bielefeld, June 1984).

124. L. F. Abbott and M. B. Wise: *Astrophys. J. Lett.*, **282**, L47 (1984).

125. J. Tohline: in *Eleventh Texas Conference on Relativistic Astrophysics*, edited by D. E. Evans, *Ann. NY Acad. Sci.*, **422**, 390 (1984), and in *Internal Kinematics and Dynamics of Galaxies*, I.A.U. Symp. 100, edited by E. Athanassoula (Rei del, Dordrecht, 1983).

126. M. Milgrom: *Astrophys. J.*, **270**, 365, 371, 384 (1983).

127. A. Dressler and M. Lecar: Astrophys. J. Lett. (submitted).

128. J. Beckenstein and M. Milgrom: *Astroiphys. J.*, **286**, 7 (1984). J. E. Felton: *Astrophys. J.*, **286**, 3 (1984).

129. S. Weinberg: Ref. 65 and *Phys. Rev.*, **138**, 988 (1965).

130. S. Weinberg: *The First Three Minutes* (Basic Books, New York, 1977).

131. G. Steigman: *Ann. Rev. Astron. Astrophys.*, **14**, 339 (1976).

132. A. D. Dolgov and Ya. B. Zeldovich: *Rev. Mod. Phys.*, **53**, 1 (1981).

133. Ya. B. Zeldovich and I. D. Novikov: *The Structure and Evolution of the Universe* (University of Chicago Press, 1983).

134. E. Witten: *Phys. Rev. D30*, 272 (1984).

135. J. R. Primack and G. R. Blumenthal: in *Formation and Evolution of Galaxies and Large Structures in the Universe*, edited by J. Audouze and J. Tran Thanh Van (Reidel, Dordrecht, 1983), and in *Fourth Workshop on Grand Unification*, edited by H. A. Weldon, P. Langacker, and P. Steinhardt (Birkhäuser, Boston, 1983), p. 256.

136. D. J. Hegyi and K. A. Olive: *Phys. Lett.* **126B**, 28 (1983).

137. M. Spite and F. Spite: *Astron Astrophys.*, **115**, 357 (1982).

138. See. e.g., Ref. 72, §16.

139. E. W. Kolb and M. S. Turner: *Ann. Rev. Nucl. Part. Sci.*, **33**, 645 (1983).

140. P. J. E. Peebles and J. T. Yu: *Astrophys. J.*, **162**, 815 (1970); W. H. Press and E. Vishniac: *Astrophys. J.*, **236**, 323 (1981); M. L. Wilson and J. Silk: *Astrophys. J.*, **243**, 14 (1981); P. J. E. Peebles: *Astrophys. J.*, **248**, 885 (1981).

141. Ya. B. Zeldovich: in *The Large Scale Structure of the Universe*, I. A. U. Symposium No. 78, edited by M. S. Longair and J. Einasto (Reidel, Dordrecht, 1978), p.409. For reconsiderations of the adiabatic scheme with neutrinos, see the following referen ce.

142. A. G. Doroshkevich, M. Yu. Khlopov, R. A. Sunyaev, A. S. Szalay, and Ya. B. Zeldovich: *Tenth Texas Symposium on Relativistic Astrophysics*, edited by R. Ramaty and F. C. Jones, *Ann. N.Y. Acad. Sci.*, **375**, 32 (1981); H. Sato: *ibid* 43; S. F. Shandarin, A. G. Doroshkevich, and Ya. B. Zeldovich: *Sov. Phys. Usp.*, **26**, 46 (1983).

143. R. B. Partridge: *Astrophys. J.*, **235**, 681 (1980). J. M. Uson and D. T. Wilkinson: *Astrophys. J. Lett.*, **277**, L1 (1984), and *Astrophys. J.*, **283**, 471 (1984).

144. J. D. Barrow and M. S. Turner: *Nature*, **291**, 469 (1981). J. R. Bond, E. W. Kolb, and J. Silk: *Astrophys. J.*, **255**, 341 (1982).

145. M. J. Rees: in *The Very Early Universe*, edited by G. Gibbons, S. Hawking, and S. Siklov (Cambridge University Press, 1983), and references therein.

146. I thank Dick Bond for suggesting this terminology to me at the 1983 Moriond Conference, where I used it in my talk (Ref. 135). George Blumenthal and I had thought of this classification independently, but not the names.

147. H. R. Pagels and J. R. Primack: *Phys. Rev. Lett.*, **48**, 223 (1982).

148. B. J. Carr: *Comments on Astrophys.*, **7**, 161 (1978).

149. R. Canizares: *Astrophys. J.* **263**, 508 (1982).

150. C. Lacey: in *Formation and Evolution of Galaxies and Large Structures in the Universe*, edited by J. Audouze and J. Tran Thanh Van (Reidel, Dordrecht, 1983).

151. D. N. Schramm and G. Steigman: *Astrophys. J.*, **241**, 1 (1981).

152. G. Blumenthal, A. Dekel, and J. Primack, in preparation.

153. M. Davis, M. Lecar, C. Pryor, and E. Witten: *Astrophys. J.*, **250**, 423 (1981).

154. P. Hut and S. White: *Nature*, **310**, 637 (1984).

155. S. Bludman and Y. Hoffman: *Phys. Rev. Lett.*, **52**, 2087 (1984).

156. A. G. Doroshkevish and M. Yu. Khlopov: *Mon. Not. Roy. Astr. Soc.* (in press).

157. M. Turner, G.Steigman, and L. M. Krauss: *Phys. Rev. Lett.*, **52**, 2090 (1984); M. S. Turner: Fermilab-Pub.-84/89-A (1984); K. A. Olive, D. Seckel, and E. Vishniac: Fermilab-Pub.-84/86-A (1984).

158. M. Fukugita and T. Yanagida: RIFP preprint 561 (1984).

159. G. Gelmini, D. N. Schramm, and J. W. F. Valle: preprint CERN-TH-3865 (1984).

160. S. S. Gershtein and Ya. B. Zeldovich: *JETP Lett.*, **4**, 174 (1966). R. Cowsik and J. McClelland: *Phys. Rev. Lett.*, **29**, 669 (1972). G. Marx and A. S. Szalay: in *Neutrino '72* (Technoinform, Budapest, 1972), vol. 1, p. 123. A. S. Sz alay and G. Marx: *Astron. Astrophys.*, *49*, 437 (1976).

161. V. A. Lyubimov, E. G. Novikov, V. Z. Nozik, E. F. Tretyakov, and V. S. Kosik: *Phys. Lett.*, **94B**, 266 (1980).

162. J. R. Bond, G. Efstathiou, and J. Silk: *Phys. Rev. Lett.*, *45*, 1980 (1980).

163. F. R. Klinkhamer and C. A. Norman: *Astrophys. J. Lett.*, **243**, L1 (1981).

164. Y. Chikashige, R. N. Mohapatra, and R. D. Peccei: *Phys. Rev. Lett.*, *45*, 1926 (1980); *Phys. Lett.*, *98B*, 265 (1981).

165. G. B. Gelmini, S. Nussinov and M. Roncadelli: *Nucl. Phys.*, **B209**, 157 (1982).

166. S. Boris et al.: in *Proc. HEP 83*, edited by J. Guy and C. Costain (Oxford, Rutherford Lab., 1983), p. 386.

167. For example, F. Boehm: in *Fourth Workshop on Grand Unification*, edited by H. A. Weldon et al. (Birkhäuser, Boston, 1983). Cf. also F. Boehm and P. Vogel: *Ann. Rev. Nucl. Part. Sci.*, **34**, 125 (1984).

168. B. W. Lee and S. Weinberg: *Phys. Rev. Lett.*, **39**, 165 (1977); M. I. Vysotsky, A. D. Dolgov, and Ya. B. Zeldovich: *JETP Lett.*, **4**, 120 (1977); P. Hut: *Phys. Lett.*, **B69**, 85 (1977); K. Sato and H. Kobayashi: *Prog. Theor. Phys.*, **58**, 1775 (1977). J. E. Gunn, B. W. Lee, I. Lerche, D. N. Schramm, and G. Steigman: *Astrophys. J.*, **223**, 1015 (1978).

169. P. Langacker, G. Segrè, and S. Soni: *Phys. Rev. D* **26**, 3425 (1982). K. Freese, E. W. Kolb, and M. S. Turner: *Phys. Rev. D***27**, 1689 (1983).

170. G. S. Bisnovatyi-Kogan and I. D. Novikov, *Sov. Astron.* **24**, 516 (1980).

171. See Lecture 2, above, and *The Very Early Universe*, edited by G. Gibbons, S. Hawking, and S. Siklov (Cambridge University Press, 1983) and references therein.

172. J. R. Bond and A. S. Szalay: *Astrophys. J.*, **274**, 443 (1984).

173. N. Vittorio and J. Silk: *Astrophys. J. Lett*, **285**, L39 (1984).

174. J. R. Bond and G. Efstathiou: *Astrophys. J. Lett.*, **285**, L45 (1984).

175. A. Melott: *Mon. Not. Roy. Astr. Soc.* **202**, 595 (1983); J. Centrella and A. Melott: *Nature*, **305**, 196 (1983); and references therein.

176. A. A. Klypin and S. F. Shandarin: *Mon. Not. Roy. Astr. Soc.*, **204**, 891 (1983).

177. C. Frenk, S.D.M. White and M. Davis: *Astrophys. J.*, **271**, 417 (1983).

178. Ya. B. Zeldovich, J. Einasto and S. F. Shandarin: *Nature* **300**, 407 (1982), and references therein.

179. N. Kaiser: *Astrophys. J. Lett.*, *273*, L17 (1983).

180. S. D. M. White: in *Inner Space/Outer Space*, edited by E. W. Kolb, et al. (Univ. of Chicago Press, in press).

181. S. Faber: in *Large-Scale Structure of the Universe, Cosmology and Fundamental Physics*, First ESO-CERN Symposium, edited by G. Setti and L. Van Hove (CERN, Geneva, 1984).

182. J. R. Bond: in *Formation and Evolution of Galaxies and Large Structures in the Universe*, edited by J. Audouze and J. Tran Thanh Van (Reidel, Dordrecht, 1983). P. R. Shapiro, C. Struck-Marcell and A. L. Melott: *Astrophys. J.*, **275**, 413 (198 3).

183. J. R. Bond, A. S. Szalay and S.D.M. White, *Nature* **301**, 584 (1983).

184. S. M. Faber and D.N.C. Lin: *Astrophys. J. Lett.*, **266**, L17 (1983).

185. M. Aaronson: *Astrophys. J. Lett.*, **266**, L11 (1983); M. Aaronson and K. Cook: *Bull. Am. Astr. Soc.*, **15**, 907 (1983); K. Cook, P. Schechter, and M. Aaronson: *ibid.*

186. D.N.C. Lin and S. M. Faber: *Astrophys. J. Lett.*, **266**, L21 (1983).

187. S. D. Tremain and J. E. Gunn: *Phys. Rev. Lett.*, **42**, 407 (1979).

188. J. Madsen and R. I. Epstein: *Astrophys. J.*, **282**, 11 (1984).

189. J. R. Primack: in *Particles and Fields 2* (Plenum, 1983), p. 607.

190. J. R. Bond, A. S. Szalay and M. S. Turner, *Phys. Rev. Lett.* **48**, 1636 (1982).

191. G. R. Blumenthal, H. Pagels and J. R. Primack, *Nature* **299**, 37 (1982).

192. H. Haber and G. Kane: *Phys. Rep.* (in press).

193. Cf. the preceeding reference and S. Dawson, E. Eichten, and C. Quigg: *Phys. Rev. D* (in press).

194. H. Goldberg, *Phys. Rev. Lett.* **50**, 1419 (1983). J. Ellis, J. S. Hagelin, D. V. Nanopoulos, K. Olive, and M. Srednicki: *Nucl. Phys.*, **B238**, 453 (1984).

195. K. A. Olive and M. S. Turner: *Phys. Rev.* **D25**, 214 (1982).

196. P. M. Lubin, G. L. Epstein and G. F. Smoot: *Phys. Rev. Lett.* **50**, 616 (1983); D. J. Fixin, E. S. Cheng and D. T. Wilkinson: *Phys. Rev. Lett.* **50**, 620 (1983).

197. M. Guyot and Ya. B. Zeldovich: *Astron. Astrophys.*, **9**, 227 (1970); P. Meszaros: *Astron. Astrophys.*, **37**, 225 (1974).

198. A. G. Doroshkevich, Ya. B. Zeldovich, R. A. Sunyaev and M. Yu Khlopov, *Sov. Astron. Lett.* **6**, 252 (1980); A. D. Chernin, *Sov. Astron.* **25**, 14 (1981).

199. S.D.M. White and M. Rees: *Mon Not. R. Astr. Soc.* **183**, 341 (1978).

200. J. Silk: *Nature* **301**, 574 (1983).

201. M. Davis, G. Efstathiou, C. Frenk and S.D.M. White: *Astrophys. J.* (in press). [DEFW]

202. P. J. E. Peebles: *Astrophys. J. Lett.*, **263**, L1 (1982); *Astrophys. J.*, **258**, 415 (1982).

203. J. R. Primack and G. R. Blumenthal: in *Clusters and Groups of Galaxies*, edited by F. Mardirossian, G. Giuricin, and M. Mezzetti (Reidel, Dordrecht, Holland, 1984), p. 435.

204. G. R. Blumenthal and J. R. Primack: SLAC-PUB-3388 (1984).

205. J. Ipser and P. Sikivie: *Phys. Rev. Lett.*, **50**, 925 (1983).

206. P. Sikivie: *Phys. Rev. Lett.*, **51**, 1415 (1983); *erratum* **52**, 695 (1984).

207. B. J. Carr: *Comments on Astrophys.*, **7**, 161 (1978). F. W. Stecker and Q. Shafi: *Phys. Rev. Lett.*, **50**, 928 (1983). K. Freese, R. Price, and D. N.Schramm: *Astrophys. J.*, **275**, 405 (1983).

208. R. Peccei and H. Quinn: *Phys. Rev. Lett.*, **38**, 140 (1977).

209. S. Weinberg: *Phys. Rev. Lett.*, **40**, 223 (1978); F. Wilczek: *Phys. Rev. Lett.*, **40**, 279 (1978).

210. L.Abbott and P. Sikivie: *Phys. Lett.*, **120B**, 133 (1983); M. Dine and W. Fischler: *Phys. Lett.*, **120B**, 137 (1983); J. Preskill, M. Wise and F Wilczek: *Phys. Lett.*, **120B**, 127 (1983); J. Ipser and P. Sikivie: *Phys. Rev. Lett.*, **50**, 925 (1983).

211. R. Brandenberger: preprints NSF-ITP-84-93 and 144 (1984).

212. D. Dicus, E. Kolb, V. Teplitz and R. Wagoner: *Phys. Rev.*, **D18**, 1829 (1978); M. Fukugita, S. Watamura and M. Yoshimura: *Phys. Rev. Lett.*, **48**, 1522 (1982).

213. F. Wilczek: *Phys. Rev. Lett.*, **49**, 1549 (1982).

214. G. R. Farrar and P. Fayet: *Phys. Lett.*, **76B**, 575 (1978); **79B**, 442 (1978). G. R. Farrar and S. Weinberg: *Phys. Rev. D27*, 2732 (1983).

215. J. Silk and M. Srednicki: *Phys. Rev. Lett.*, **53**, 624 (1984)

216. J. S. Hagelin, G. Kane, and S. Raby: *Nucl. Phys.*, **B241**, 638 (1984). L. E. Ibanez: FTUAM preprint 83-28 (1984). S. Raby: "Supersymmetry and Cosmology", Lectures at the XV GIFT Seminar, June 1984, Los Alamos preprint LA-UR-84-2693.

217. E. Farhi and R. L. Jaffe: *Phys. Rev. D*, **30**, 2379 (1984).

218. A. De Rujula and S. Glashow: *Nature*, **312**, 734 (1984).

219. C. G. Lacey: in *Formation and Evolution of Galaxies and Large Scale Structures in the Universe*, edited by J. Audouze and J. Tran Thanh Van (Dordrecht, Reidel, 1983), p. 351.

220. J. Ipser and R. Price: *Astrophys. J.*, **216**, 578 (1977).

221. C. Canizares: *Astrophys. J.*, **263**, 508 (1982).

222. Bardeen, J. (in preparation).

223. M. S. Turner, F. Wilczek, and A. Zee: *Phys. Lett.*, **125B**, 35-40 and (E) 519 (1983). T. Hara: *Prog. Theor. Phys.*, **6**, 1556-1568 (1983).

224. Lynden-Bell, D. *Mon. Not. R. astr. Soc.* **136**, 101 (1967). Shu, F. H. *Astrophys. J.* **225**, 83 (1978).

225. Peebles, P. J. E. *The Large Scale Structure of the Universe (Princeton University Press, 1980). Efstathiou, G. and Silk, J. Fundamentals of Cosmic Phys.* **9**, 1-138 (1983).

226. S. D. M. White and M. J. Rees: *Mon. Not. R. astr. Soc.*, **183**, 341-358 (1978).

227. M. J. Rees and J. P. Ostriker: *Mon. Not. R. Astr. Soc.* , **179**, 541 (1977).

228. Geller, M. and Huchra, J. *Astrophys. J. Suppl.* **52**, 61-87 (1983).

229. S. M. Faber, G. R. Blumenthal, and J. R. Primack: UCSC preprint (1984).

230. A. G. Doroshkevich: *Astrophysics* **6**, 320-330 (1973).

231. The statistics of such correlations is discussed by G. R. Blumenthal, S. M. Faber, and J. R. Primack: (in prep., 1985).

232. G. Efstathiou and B. J. T. Jones: *Mon. Not. R. Astron. Soc.* **186**, 133 (1979). S. M. Fall and G. Efstathiou: *Mon. Not. R. astr. Soc.* **193**, 189 (1980).

233. A. Sandage, K. C Freeman, and N. R. Stokes: *Astrophys. J.* **160**, 831 (1970). G. Efstathiou and J. Barnes: in *Formation and Evolution of Galaxies and Large Scale Structures in the Universe*, edited by J. Audouze and J. Tran Thanh Van (Reide l, Dordrecht, 1984), p. 361.

234. J. R. Gott and T. X. Thuan: *Astrophys. J.* **204**, 649 (1976).

235. G. L. Hoffman, E. E. Salpeter, and I. Wasserman: *Astrophys. J.*, **268**, 527 (1983).

236. R. F. Kirshner, A. Oemler, P. L. Schechter, and S. A. Schectman: *Astrophys. J. Lett.*, **248**, L57 (1981) and in *Early Evolution of the Universe and Its Present Structure*, IAU Symp. No. 104, edited by G. O. Abell and G. Chincarini (Dordrecht, Reidel, 1983), p. 197.

237. G. Efstathiou, M. Davis, C. S. Frenk, and S. D. M. White: *Astrophys. J. Suppl.* (in press).

238. J. N. Fry and A. L. Melott: preprint NSF-ITP-84-143.

239. P. J. E. Peebles: *Astrophys. J.*, **284**, 439 (1984).

240. J. Bardeen: in *Inner Space/Outer Space*, edited by E. W. Kolb et al. (University of Chicago Press, in press). N. Kaiser: *ibid.* J. Silk, and R. Schaeffer and J. Silk, Berkeley preprints.

241. N. Kaiser: *Astrophys. J. Lett.*, **284**, L9 (1984).

242. H. D. Politzer and M. B. Wise: Caltech preprint.

243. J. Barnes, A. Dekel, G. Efstathiou, and C. Frenk: NSF-ITP preprint.

244. S. F. Shandarin: *Sov. Astron. Lett.*, **9**, 104 (1983).

245. A. Melott, J. Einasto, E. Saar, I. Suisalu, A. A. Klypin, and S. F. Shandarin: *Phys. Rev. Lett.*, **51**, 935 (1983).

246. J. D. Barrow and S. P. Bhavsar: *Mon. Not. Roy. Astron. Soc.*, **205**, 66 (1983).

247. A. Dekel and M. J. West: *Astrophys. J.* (in press).

248. A. Dekel, M. J. West, and S. J. Aarseth: *Astrophys. J.*, **279**, 1 (1984).

249. N. Brosch and P. M. Gondhalekar: *Astron. Astrophys.*, **140**, L43 (1984). I thank Adrian Melott for bringing this article to my attention.

250. S. D. M. White, M. Davis, and C. Frenk: *Mon. Not. Roy. Astron. Soc.*, **209**, 15P (1984).

251. A. Dekel: *Astrophys. J.*, **284**, 445 (1984).

252. A. Dekel: *Astrophys. J. Lett.*, **261**, L13 (1982).

253. W. L. W. Sargent, P. J. Young, A. Boksenberg, and D. Tytler: *Astrophys. J. Suppl.*, **42**, 41 (1980).

254. W. L. W. Sargent, P. Young, and D. P. Schneider: *Astrophys. J.*, **256**, 374 (1982).

255. J. H. Oort: *Astron. Astrophys.* (submitted 1984).

256. J. H. Oort, H. Arp, and H. de Ruiter: *Astron. Astrophys.*, **95**, 7 (1981).

257. T. Ferris: *New York Times Magazine*, Sept. 26, 1982, p. 38.

258. E. Neumann: *Origins and History of Consciousness* (Princeton Univ. Press, 1954).

259. V. F. Weisskopf: *Science*, **187**, 605 (1975). W. H. Press and A. P. Lightman: *Phil. Trans. R. Soc. Lond. A* (submitted).

Figure Captions

Fig. 1.1. Sizes and distances.

Fig. 1.2. Hubble's classification of galaxy types.[2] Elliptical (E) galaxies have a spheroidal appearance with no visible disk, while lenticular ($S0$) and spiral (S) galaxies have a disk in addition to a spheroidal central bulge or nucleus. $S0$'s have no spiral arms, and relatively little dust and ionized hydrogen in the disk; these are more prominent in spiral galaxies. A large minority of S galaxies have "bars" across their nuclei, with the spiral arms beginning from the ends of the bar rather than winding out directly from the nucleus. Some irregular (Irr) galaxies resemble disk galaxies but with less symmetry; others are even more irregular than that.

Fig. 1.3. Rotation curves for many spiral galaxies, obtained both (a) from 21cm observations and (b) from optical measurements (see text). (Sources: (a) Ref. 6, reprinted in Ref. 5; (b) Ref. 7. Reproduced by permission.)

Fig. 1.4. Mass of the Milky Way Galaxy interior to radius R, deduced from the dynamics of its components and satellites: carbon monoxide clouds, globular clusters,[12,13,83] the Magellanic Clouds[15,16], a distant RR Lyrae star (assumed to be bound),[17] and satellite galaxies. (Adapted from Ref. 18.)

Fig. 1.5. Total mass (assumed to be $10 \times M_{lum}$) vs. velocity (rotation velocity for disks, velocity dispersion for spheroids) for various galaxy types. Slopes are roughly consistent with $L \propto v^4$. (From Fig. 4 of Ref. 26; reprinted here as Fig. 4.4.)

Fig. 1.6. Schechter luminosity function. Axes are $\log_{10} \phi(L)$ and $L\phi(L)$ vs. $\log_{10} L/L_*$.

Fig. 1.7. Schematic plot of mass density vs. distance from center of the Milky Way. (See text for explanation. From Ref. 32.)

Fig. 1.8. X-ray isodensity contours for six rich clusters of galaxies. The clusters on the left represent those having large X-ray core radii and no central,

dominant galaxy. Those on the right have smaller X-ray core radii and contain cD galaxies. The clusters at the top of the figure are less dynamically evolved than those at the bottom. (From 0.5-3.0 keV imaging proportional counter images from the Einstein satellite. Reproduced, with permission, from Ref. 217; ©1982 by Annual Reviews, Inc.)

Fig. 1.9. Contours of galaxy density in rich clusters. (Reproduced, with permission, from Ref. 39, as reproduced in Ref. 40.)

Fig. 1.10. Fractions of elliptical, lenticular, and spiral plus irregular galaxies as a function of the log of the projected density, in galaxies per Mpc^2. Also shown is an estimate of the space density, in galaxies per Mpc^3. The upper histogram shows the number distribution of the galaxies in the sample ($\sim$ 6000 galaxies in 55 rich clusters) over the bins of projected density. (Reproduced, with permission, from Ref. 42.)

Fig. 1.11. Mass-to-light ratio, M/L_B, and total-to-luminous mass, M/M_{lum}, for structures of various sizes in the universe. Although M/L_B increases systematically with mass, the more physically meaningful ratio M/M_{lum} appears to be constant on all scales within the errors. If the velocity dispersion data for the dwarf spheroidal galaxies are interpreted to imply heavy halos, the upper estimates in the figure result. The lower estimates follow from assuming that all the mass is visible. The upper estimates are probably more realistic. (Reproduced from Ref. 26. The data come from Table 1 of this Ref.)

Fig. 1.12. Perseus supercluster. (Top) Map of the surveyed region of the sky (coordinates are declination vs. right ascension), showing galaxies with $3700 < V_o < 8200$ km s^{-1}. (Bottom) Wedge diagram of all galaxies in the survey with $V_o < 10{,}000$ km s^{-1}. The position variable is each galaxy's projected position on the majop axis of the Perseus supercluster filament with arbitrarily chosen zero point. The foreground galaxies are clustered, although one galaxy is found in an otherwise void region. At redshifts higher than the supercluster's, very few galaxies are found. (Reproduced, with permission, from Ref. 51.)

Fig. 2.1. (Left) Schwarzschild and (Right) Hubble spheres. (Left figure adapted from Ref. 67, Fig. 9.7.)

Fig. 2.2. Mass and linear size.

Fig. 2.3. Sketch of solutions of Einstein's equations for the scale factor $R(t)$ in open ($k < 0$), flat ($k = 0$), and closed ($k > 0$) Friedmann universe models. Hubble's constant H_o is the slope $\dot{R}$, and deceleration due to gravity acting since the Big Bang (at $t = 0$ in the figure) implies that the actual age of the universe t_o is less than the Hubble time H_o^{-1}; the precise relationship is given by eq. (2.18).

Fig. 2.4. The function $f(\Omega) = t_o H_o$ for a Friedmann universe ($\Lambda = 0$).

Fig. 2.5. Relationship of cosmological density parameter Ω_o and Hubble's parameter $H_o \equiv 100h$ km s^{-1} Mpc^{-1} for various values of t_o (in Gy).

Fig. 2.6. Geometry for the three-point correlation function.

Fig. 2.7. Two-point galaxy correlation function $\xi(r)$ calculated from the CfA data,[73] as reploted and fitted by Dekel and Aarseth. Dots ($\bullet$) are ξ and pluses ($+$) are $1 + \xi$. (Reproduced, with permission, from Ref. 75, which discusses a possible physical interpretation of the fits and the corresponding length scales r_c, r_o, r_b, r_z.)

Fig. 2.8. The cluster-cluster correlation function has the same slope as the galaxy-galaxy correlation function, but is roughly an order of magnitude larger.

Fig. 2.9. Schematic sketches of quasar spectra, showing (Left) the unattenuated redshifted Ly α emission feature and the absorption trough that would be seen if there were significant absorption by intervening hydrogen, and (Right) the actual sort of spectra seen, with discrete absorption lines from small intervening "Ly α clouds" of neutral hydrogen. (From unpublished maunscript by George Blumenthal.)

Fig. 2.10. Mass M_H (heavy solid line) of nonrelativistic matter encompassed by the horizon as a function of redshift. The Jeans mass M_J is approximately

equal to M_H until the era (z_{eq}) when nonrelativistic matter begins to dominate gravitationally; at recombination (z_r), M_J drops sharply (light solid line). Photon diffusion (Silk damping) erases adiabatic fluctuations in radiation plus ordinary matter that cross into the region below the dashed line. Fluctuations in a universe dominated by hot or warm dark matter are damped by free streaming until the temperature drops to the mass of the dark matter particles and they become nonrelativitic; the corresponding (free-streaming) Jeans masses are indicated. (See Lecture 3.)

Fig. 2.11. "Top-hat" fluctuation: constant density $\rho_m(1+\delta)$ inside a spherical volume of radius $R(1+a)$ in a universe of background density $\rho_m = \bar{\rho}$.

Fig. 2.12. Schematic sketches of radius, density, and density contrast of an overdense fluctuation. It initially expands with the Hubble expansion, reaches a maximum radius (solid vertical line), and undergoes violent relaxation during collapse (dashed vertical line), which results in the dissipationless matter forming a stable halo. Meanwhile the ordinary matter (ρ_b) continues to dissipate kinetic energy and contract, thereby becoming more tightly bound, until dissipation is halted by star or disk formation.

Fig. 2.13. N-body simulation of gravitational collapse. (Reproduced, with permission, from Ref. 94.)

Fig. 2.14. During the inflationary (de Sitter) expansion, regions once causally connected expand much faster than light (dotted lines). After reheating, the universe becomes dominated by matter rather than the effective cosmological constant associated with the Higgs self-energy; and as the horizon expands faster than the Hubble expansion ($R \propto t^{1/2}$ in a radiation-dominated universe), regions again become causally connected.

- Figure 3.1. Radiation temperature (T_γ) and density (u_γ) and nonrelativistic matter density (ρ_m) from neutrino decoupling (1) to the present, according to the standard Hot Big Bang picture. (See text.)

Figure 3.2. Calculated primordial abundances, by number relative to hydrogen, of light isotopes D, ^{3}He, and ^{7}Li, and by mass of ^{4}He (Y_p) for N_ν=2, 3, and 4 light ν species (the error bar shows the effect of $\Delta\tau = \pm 0.2$ min uncertainty in the neutron lifetime). The primordial abundances inferred from observations are $Y_p = 0.23 - 0.25$, $(\mathrm{D/H}) \geq 1 \times 10^{-5}$, $(\mathrm{D}+^3\mathrm{He})/\mathrm{H} \leq 10^{-4}$, $(^7\mathrm{Li/H}) = (1.1 \pm 0.4) \times 10^{-10}$. These are consistent with the predictions for $\eta = (4 - 7) \times 10^{-10}$ and $N_\nu \leq 4$. (Reproduced, with permission, from Ref. 86.)

Fig. 3.3. Jeans mass vs. scale factor for a universe dominated by baryons (solid line), neutrinos ($m = 30$ eV), and warm dark matter ($m = 1$ keV). (From Ref. 191.)

Fig. 3.4. Overdensity $\delta \equiv \delta\rho/\rho$ of a fluctuation vs. redshift. The light solid line represents the growth of a baryonic fluctuation in a universe with no dark matter. With the constraint $\Omega_b \lesssim 0.1$ from big bang nucleosysthesis, growth of the fluctuation amplitude δ slows when $z \lesssim 0.1$ and there is only a factor of $\sim 10^2$ growth from recombination (z_r) until the present. Even without this constraint (dashed line), the factor of 10^3 growth of δ since recombination requires too large an amplitude of adiabatic fluctuations at recombination to be consistent with $\Delta T/T$ constraints. This problem is avoided with dark matter (DM): δ_{DM} (heavy solid line) grows $\propto (1 + z)^{-1}$ after dark matter becomes gravitationally dominant at z_{eq}; but Compton drag prevents baryonic fluctuations (dotted curve) from growing until after recombination, when they rapidly grow to match δ_{DM}.

Fig. 3.5. Decoupling of neutrinos occurs when their mean free time t_f exceeds the Hubble time t_H.

Fig. 3.6. (a) CfA galaxies shown on equal area plot of the northern sky (outer circle corresponds to Galactic latitude +40°, lower blank region below declination 0° is blocked by our galaxy's disk). "Galaxy" locations in N-body simulations, plotted as in (a), for universe dominated by (b) neutrinos and (c) cold dark matter. The neutrino simulation assumes $\Omega = 1$, $h = 0.54$, and that galaxy formation began at $z = 2.5$; the triangles represent galaxies while the

dots are points in whose neighborhood the matter has not yet collapsed. The cold dark matter simulation assumes $\Omega = 0.2$ and $h = 1.1$. (Reproduced, with permission, from Ref. 180.)

Fig. 3.7. Mass m_X and present number density of warm dark matter particles X, assuming the standard particle physics model with no entropy generation after T_{GUT}. The mass (left axis) scales as Ωh^2.

Fig. 4.1. Numerical results for the growth of $\delta = k^{3/2}\delta_k$ versus scale factor a for fluctuations of various masses $M = \frac{4}{3}\pi^4 k^{-3}\rho_c$. The curves are drawn for $n = 1$, $\Omega = h = 1$, and a baryonic to total mass ratio of 0.1. The vertical line represents the value of a when the universe becomes matter dominated, and the dashed line shows the (constant for $n = 1$) value of δ when each mass scale crosses the horizon. These curves illustrate the stagnation of perturbation growth after small mass scales cross the horizon and show why at late times $\delta(k)$ is nearly flat for large k (small M). (From Ref. 203.)

Fig. 4.2. Density fluctuations as a function of mass. a) $k^{3/2}\,|\delta_k| = \delta\rho/\rho(M)$, where $M = 4\pi^4\rho_o/3k^3$, for isothermal white noise ($n = 0$), and adiabatic Zeldovich ($n = 1$) neutrino[190] and cold dark matter spectra. b) Root-mean-square mass fluctuation within a randomly placed sphere containing average mass M in a cold dark matter universe, for ($\Omega = 1$, $h = 0.5$), ($\Omega = 0.2$, $h = 1$). The curves are normalized[202] at 8 Mpc and assume a primordial Zeldovich ($n = 1$) fluctuation spectrum.

Fig. 4.3. The baryonic density versus temperature as root-mean-square perturbations having total mass M become nonlinear and virialize. The numbers on the tick marks are the logarithm of M in units of $M_\odot$. This curve assumes $n = 1$, $\Omega = h = 1$, and a baryonic to total mass ratio of 0.07. The region where baryons can cool within a dynamical time lies below the cooling curves. Also shown are the positions of observed galaxies, groups, and clusters of galaxies. The dashed line represents a possible evolutionary path for dissipating baryons. (From Ref. 32.)

Fig. 4.4. Total mass M versus temperature T. The quantity T is $\mu V^2/3k$, where μ is mean molecular weight (≈ 0.6 for ionized, primordial H + He) and k is Boltzmann's constant. M for groups and clusters is total dynamical mass. For galaxies, M is assumed to be 10 M_{lum} (corresponding to Fig. 1.11). If dwarf spheroidals actually have $M/L_B = 30$, they may have suffered baryon stripping[186], in which case M is a lower limit (arrows). Details of the region occupied by massive galaxies are shown in the inset in upper left. Model curves represent the equilibria of structures that collapse dissipationlessly from the cold dark matter initial fluctuation spectra with $n = 1$. The curves labeled 1σ refer to fluctuations with $\delta M/M$ equal to the rms value. Curves labeled 0.5σ, 2σ, and 3σ refer to fluctuations having 0.5, 2, and 3 times the rms value. Heavy curves: $\Omega = 1$, $h = 0.5$; dashed curves: $\Omega = 0.2$, $h = 1$; these cases were chosen to span the astrophysically interesting range. In addition to the n = 1 curves, two 1σ curves for $n = 0$ and $n = 2$ are also shown (light dashes).

Major conclusions from the figure: 1) Either set of curves for $n = 1$ (Zeldovich spectrum) provides a good fit to the observations over 9 orders of magnitude in mass. Curves with $n = 0$ and $n = 2$ do not fit as well. 2) The apparent gap between galaxies and groups and clusters in Fig. 4 (which stems from baryonic dissipation) vanishes in this figure, and the clustering hierarchy is smooth and unbroken from the smallest structures to the largest ones. 3) The Fisher-Tully and Faber-Jackson laws for galaxies ($M \propto V^4$ or T^2) arise naturally as a consequence of the slope of the cold DM fluctuation spectrum in the mass region of galaxies. 4) Groups and clusters are distributed around the $n = 1$ loci about as expected. The apparent upward trend among the groups is not physically meaningful but arises from their selection as minimum-density enhancements (see constant-density arrow). 5) The exact locations of galaxies are somewhat uncertain. In particular, the temperatures of E's and S0's may be overestimated owing to the use of nuclear rather than global velocity dispersions. Taken at face value, however, the data suggest that early-type galaxies (E's and S0's) arise from high-$\delta M/M$ fluctuations, whereas late-type galaxies (Sc's and Irr's) arise

from low-$\delta M/M$ fluctuations. 6) Groups and clusters appear to fill a wider band than galaxies. If real, this difference may indicate that very weak, low-$\delta M/M$ fluctuations on the mass scale of galaxies once existed but did not give rise to visible galaxies. This suggests further that galaxy formation, at least in some regions of the universe, may not have been fully complete and that galaxies are therefore not a reliable tracer of total mass. 7) There seems to be a real trend along the Hubble sequence to increasing mass among early-type galaxies. Neither this trend nor the rather sharp demarcation between galaxies and groups and clusters is fully understood. (This figure is from Ref. 26.)

Fig. 4.5. Cosmological density Ω vs. Hubble parameter for various values of t_o in a flat universe with cosmological constant given by eq. (4.6).

Fig. 4.6. Consumers' Report on Dark Matter.

Fig. 4.7. Relative sensitivities of astronomical instruments. (Figure courtesy of the Space Telescope Science Institute.

Fig. 4.8. Cosmological uroboros. (Adapted from Ref. 32.)

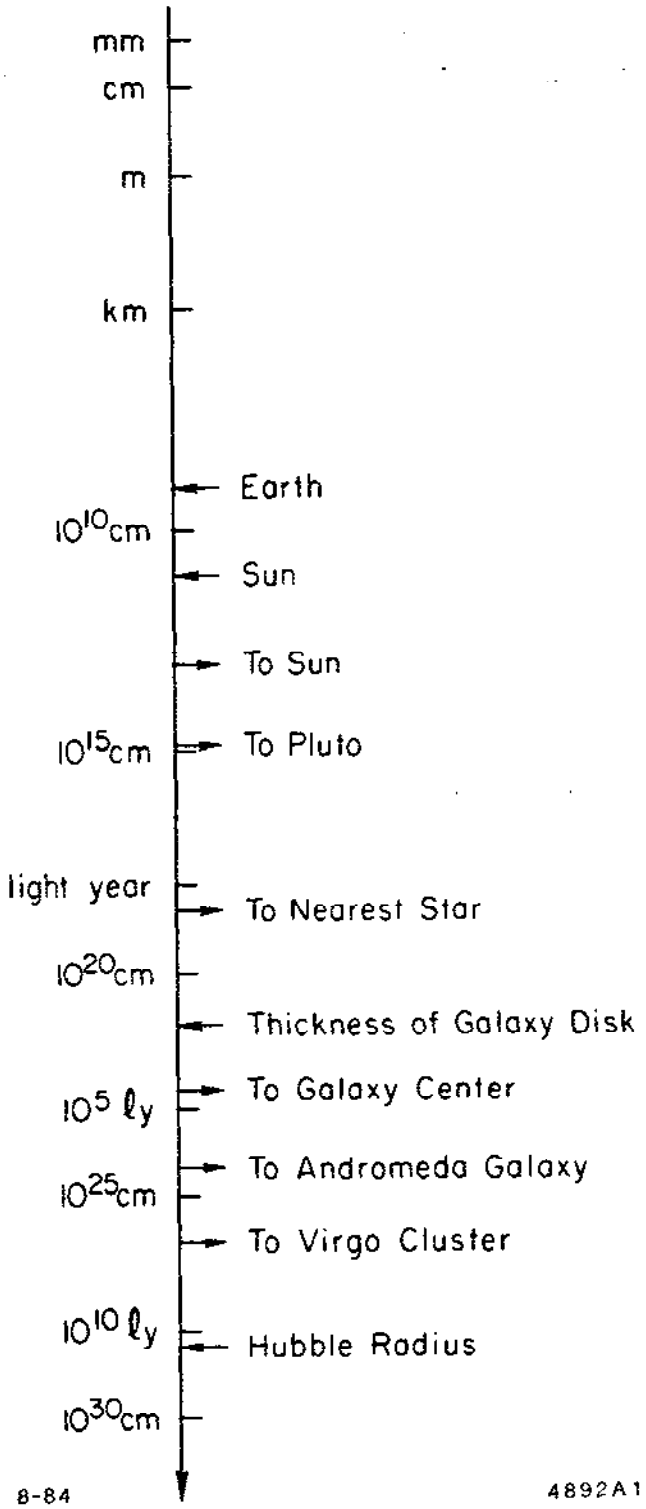

mm
cm
m
km
Earth
10^{10}cm
Sun
To Sun
10^{15}cm
To Pluto
light year
To Nearest Star
10^{20}cm
Thickness of Galaxy Disk
10^5 ℓy
To Galaxy Center
10^{25}cm
To Andromeda Galaxy
To Virgo Cluster
10^{10} ℓy
Hubble Radius
10^{30}cm
8-84
4892A1

Fig. 1.1

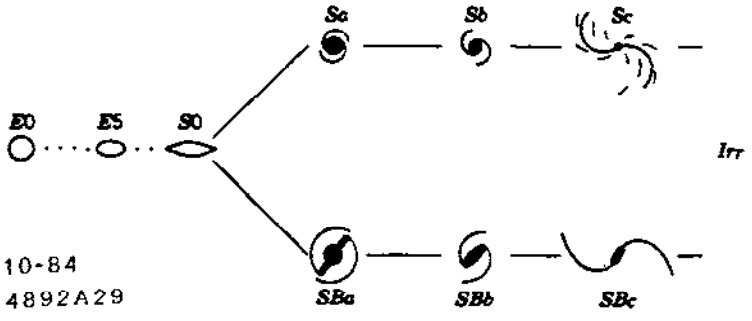


Fig. 1.2

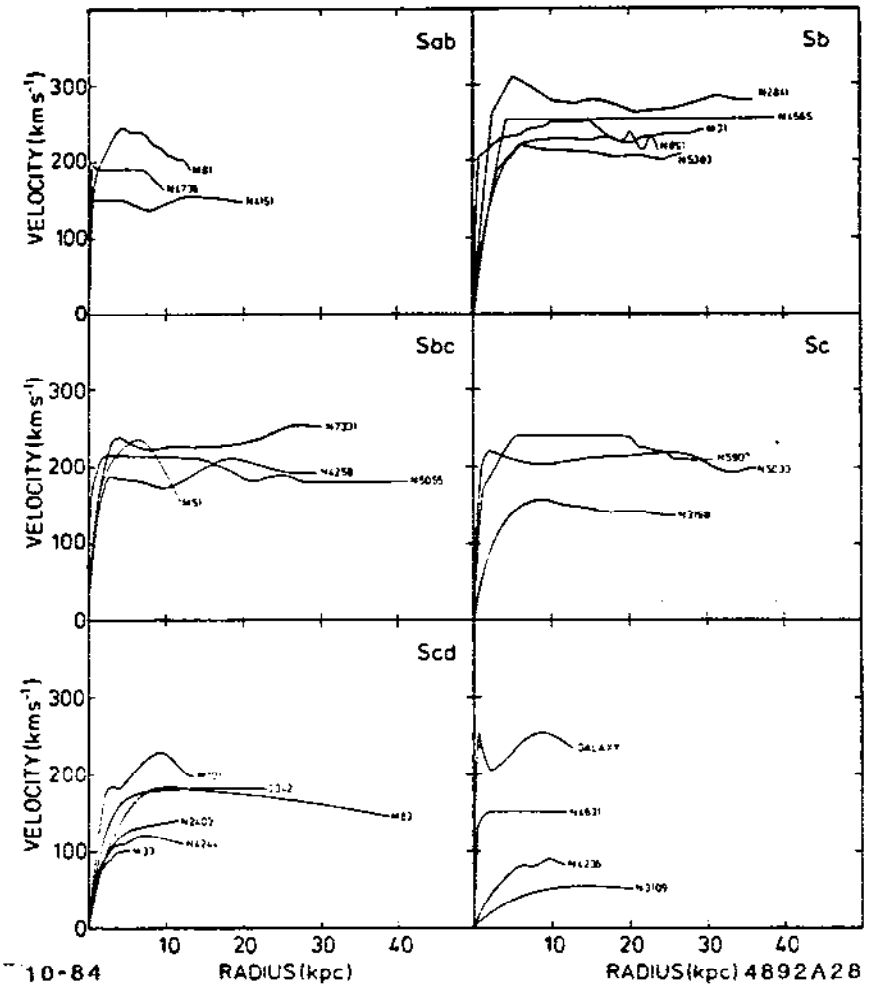
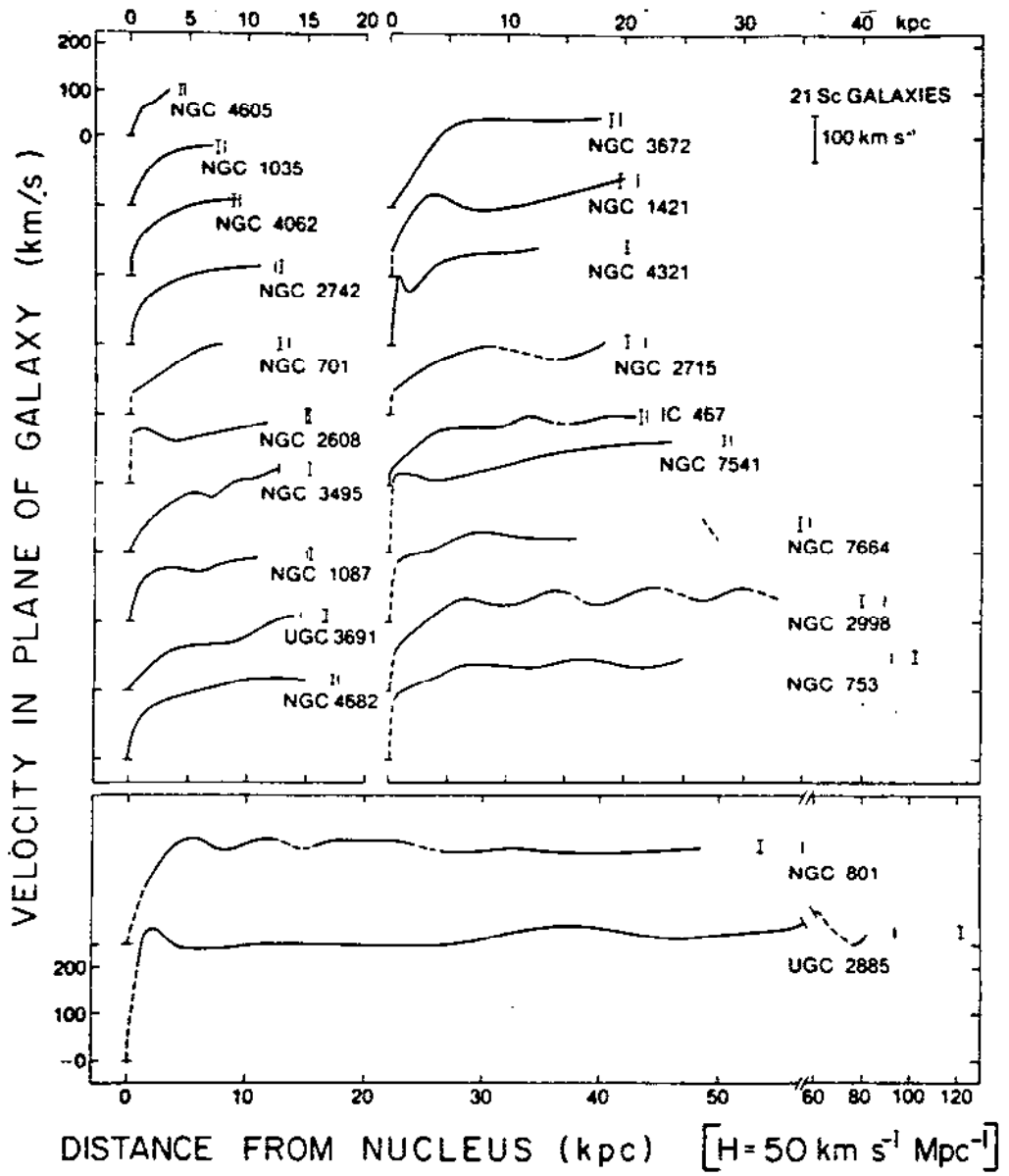

Sab
Sb
Sbc
Sc
Scd
VELOCITY (kms⁻¹)
300
200
100
0
10
20
30
40
RADIUS (kpc)
10-84
4892A28

Fig. 1.3(a)



10-84

4892A27

Fig. 1.3(b)

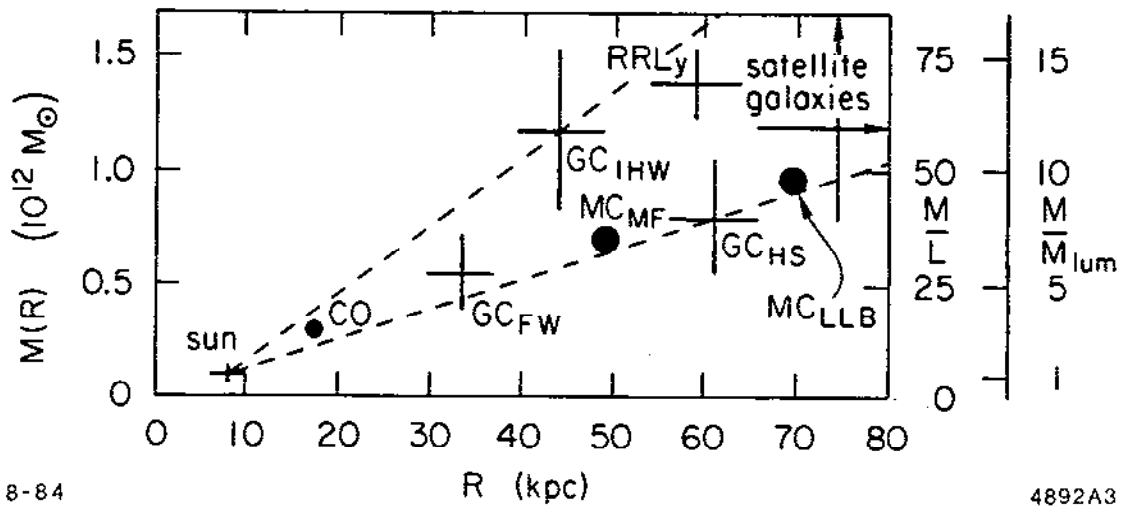

M(R) (10^12 M⊙)
1.5
1.0
0.5
0
RRL_y
satellite galaxies
GC_IHW
MC_MF
GC_HS
MC_LLB
GC_FW
CO
sun
75
50
25
0
M/L
15
10
5
1
M/M_lum
0 10 20 30 40 50 60 70 80
R (kpc)
8-84
4892A3

Fig. 1.4

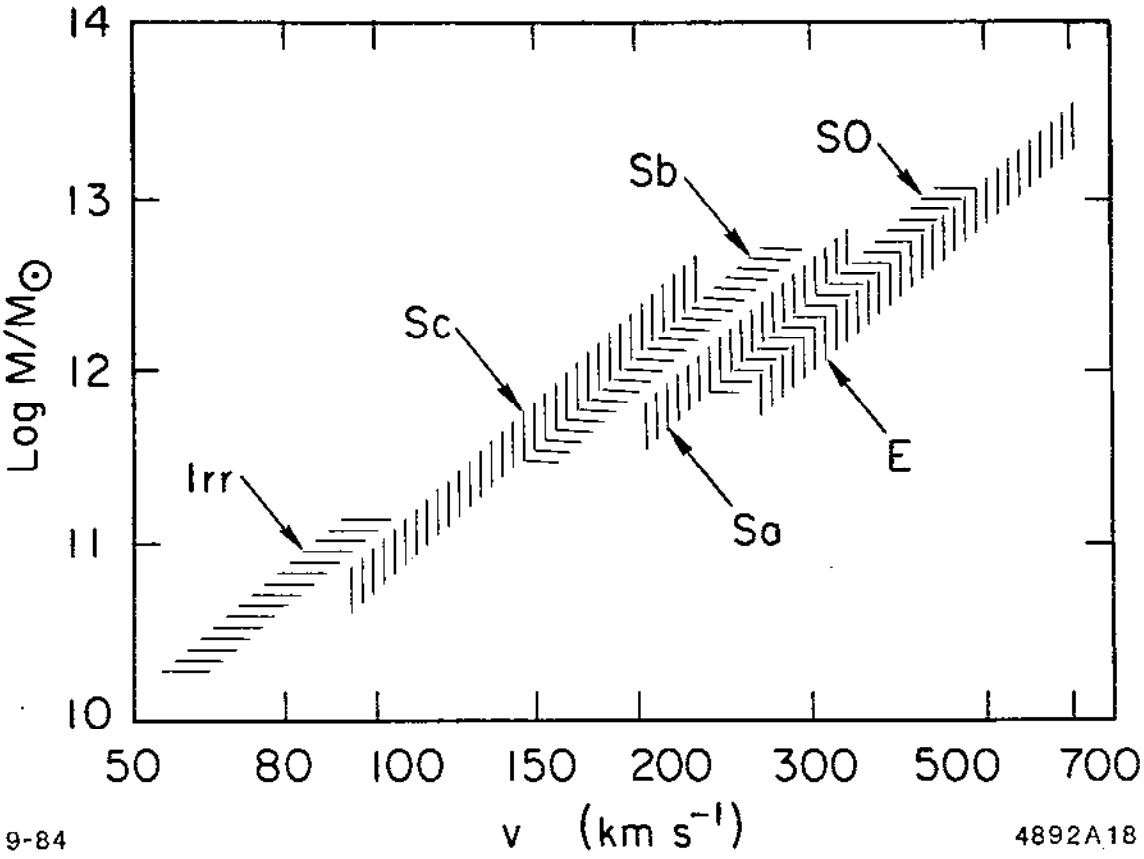

14
13
12
11
10
Log M/M⊙
50
80
100
150
200
300
500
700
v (km s⁻¹)
Irr
Sc
Sb
S0
Sa
E
9-84
4892A18

Fig. 1.5

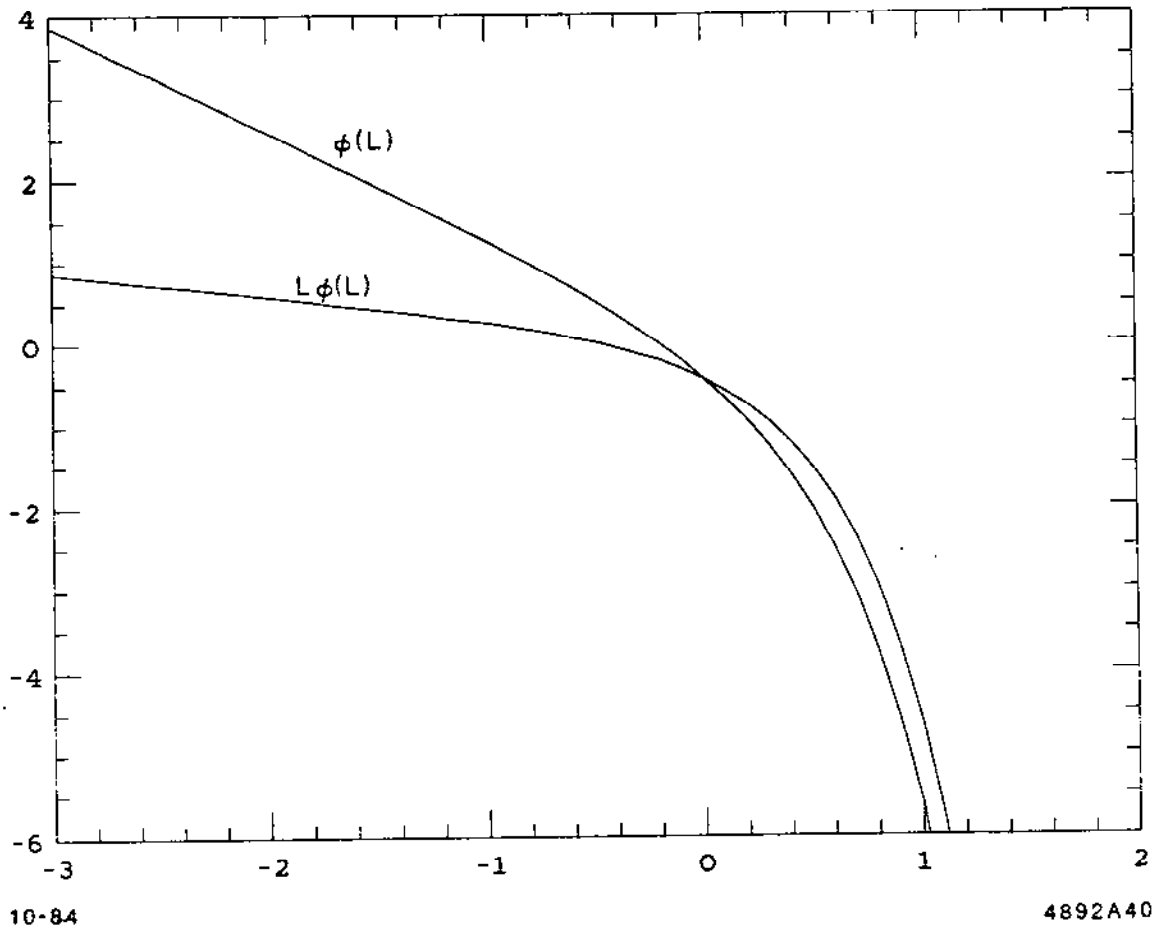


Fig. 1.6

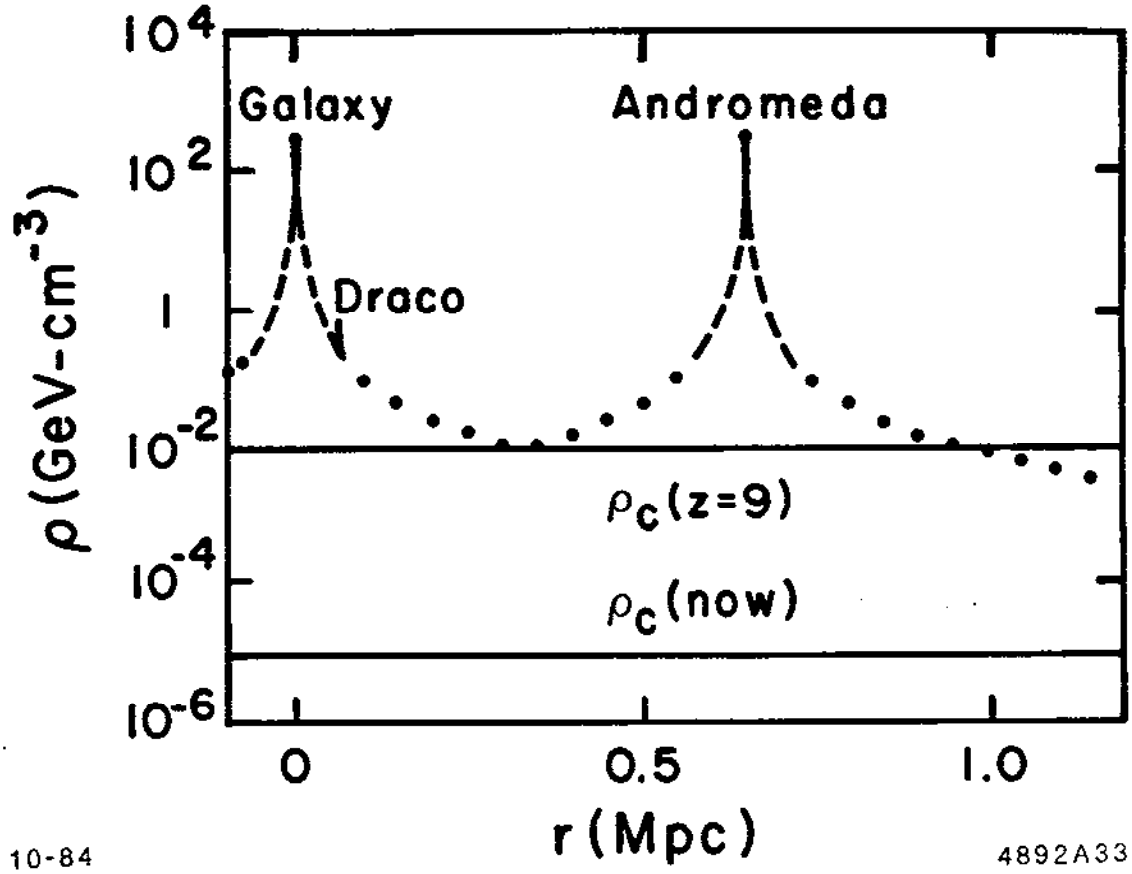
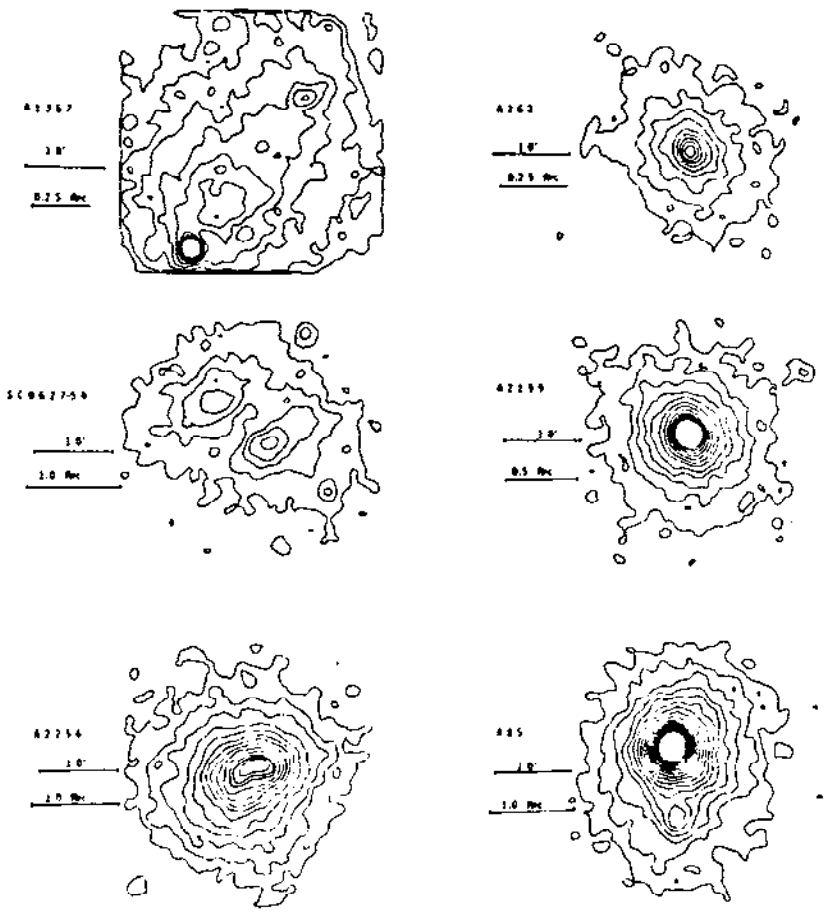

10^4
10^2
1
10^{-2}
10^{-4}
10^{-6}
ρ(GeV-cm^{-3})
Galaxy
Andromeda
Draco
ρ_c(z=9)
ρ_c(now)
0
0.5
1.0
r(Mpc)
10-84
4892A33

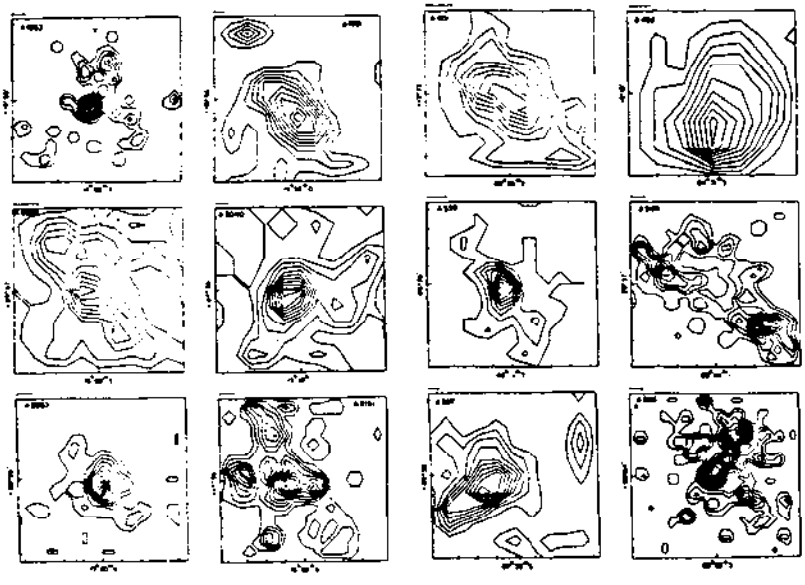
Fig. 1.7



10-84

4892A25

Fig. 1.8



10-84

4892A22

Fig. 1.9

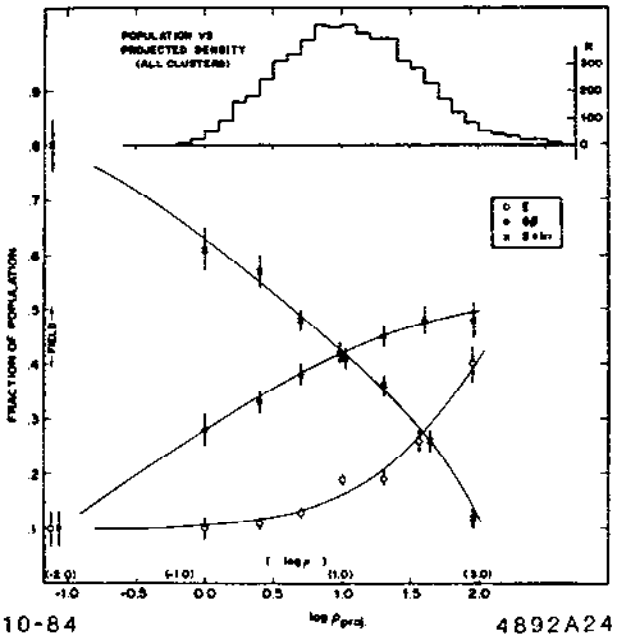

POPULATION VS
PROJECTED DENSITY
(ALL CLUSTERS)
N
300
200
100
0
FRACTION OF POPULATION
FIELD
(-2.0)
(-1.0)
(log ρ)
(1.0)
(2.0)
-1.0
-0.5
0.0
0.5
1.0
1.5
2.0
log ρ proj.
10-84
4892A24

Fig. 1.10

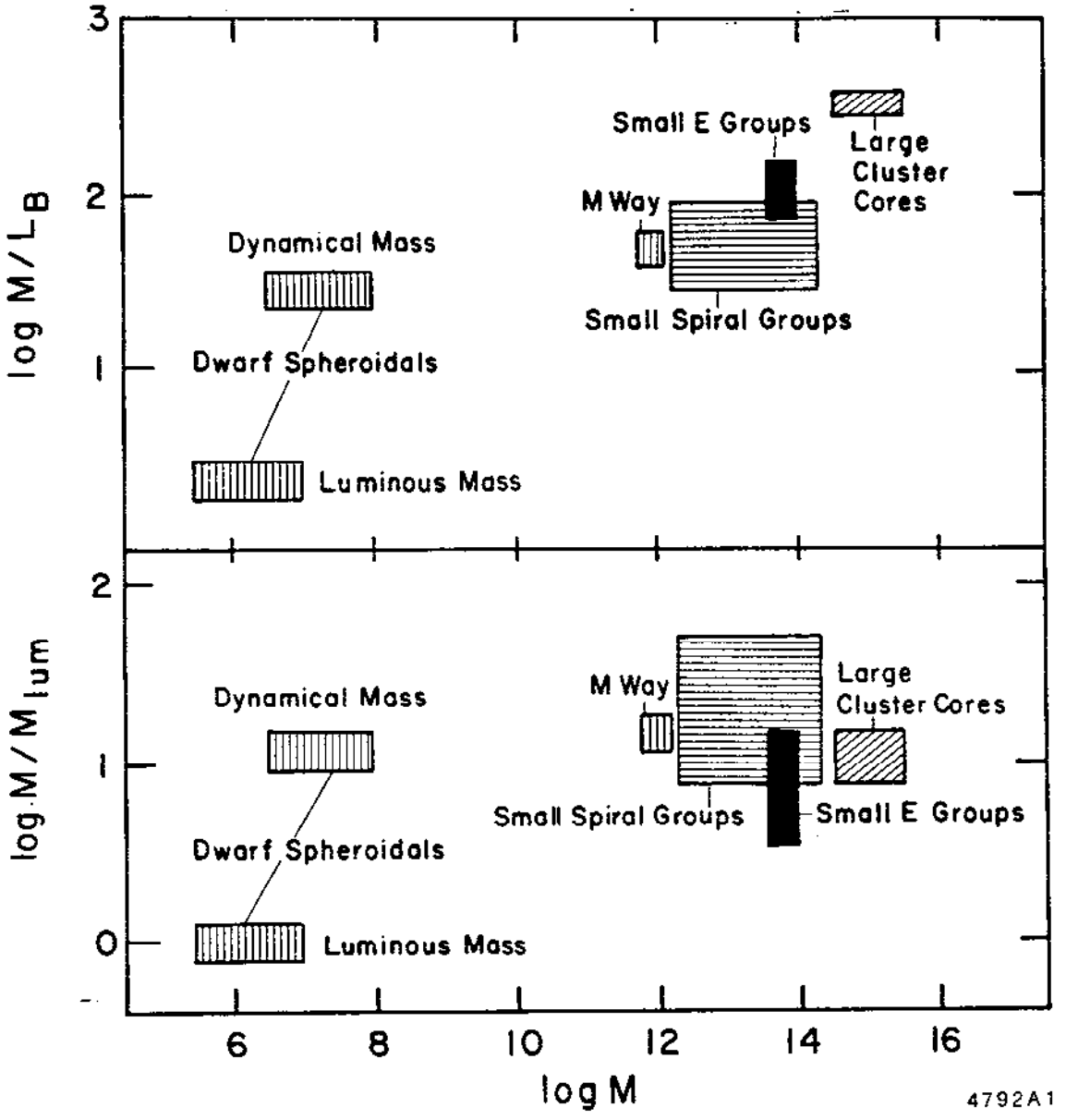
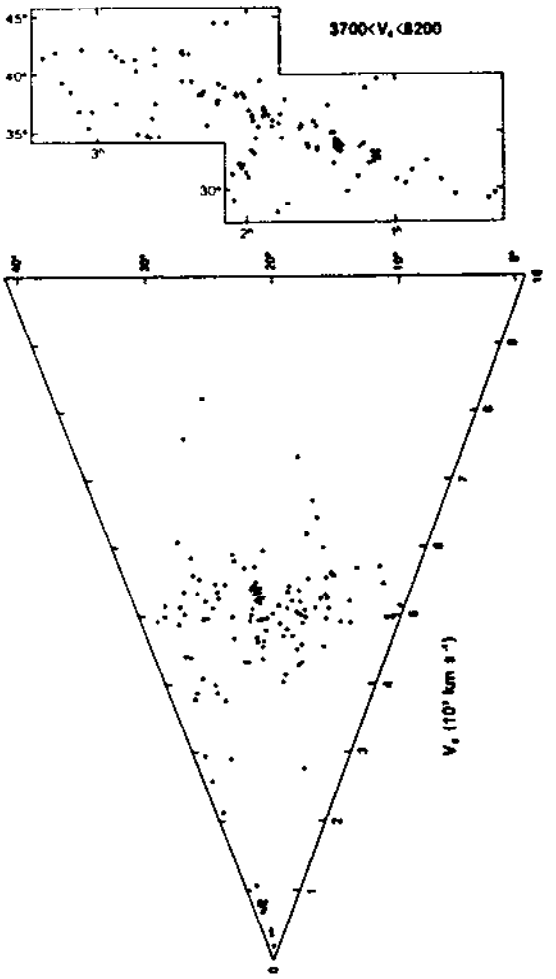

3
2
1
log M/L_B
Small E Groups
Large
Cluster
Cores
M Way
Small Spiral Groups
Dynamical Mass
Dwarf Spheroidals
Luminous Mass
2
1
0
log M/M_lum
M Way
Large
Cluster Cores
Small Spiral Groups
Small E Groups
Dynamical Mass
Dwarf Spheroidals
Luminous Mass
6
8
10
12
14
16
log M
4792A1

Fig. 1.11



10-84 4892A23

Fig. 1.12

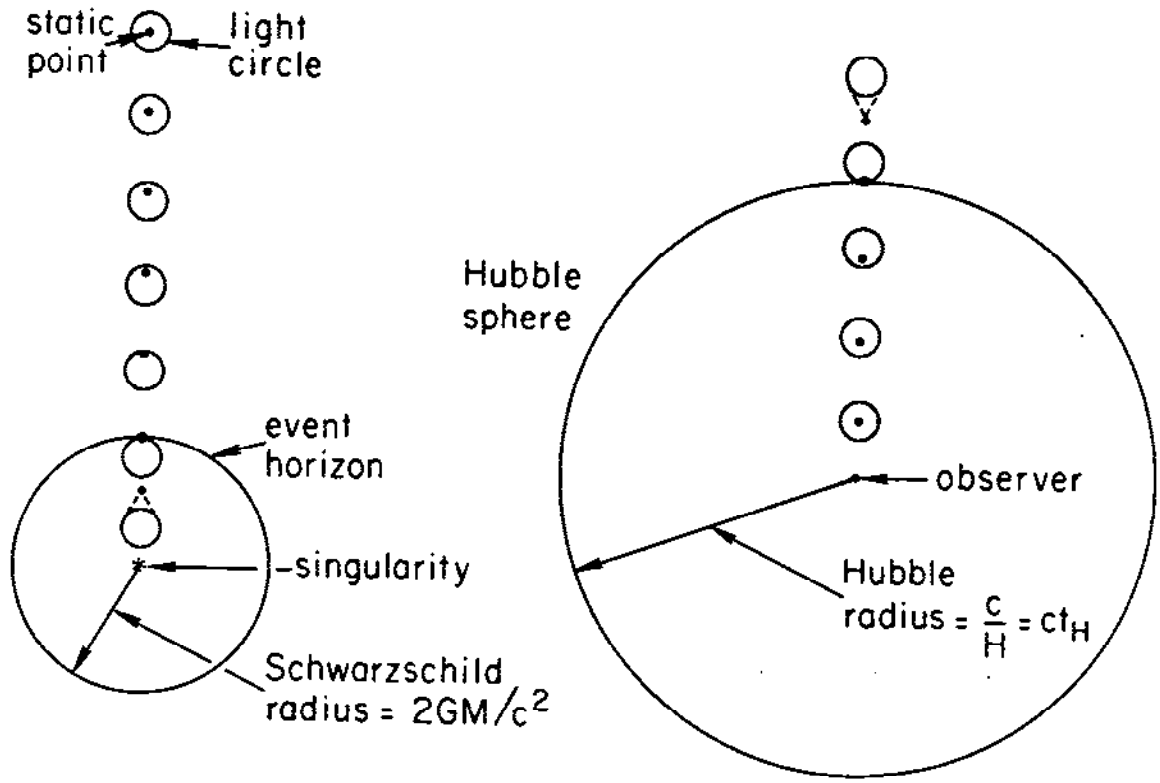

static point
light circle
event horizon
singularity
Schwarzschild radius = $2GM/c^2$
black hole
8-84
Hubble sphere
observer
Hubble radius = $\frac{c}{H} = ct_H$
expanding universe
4892A2

Fig. 2.1

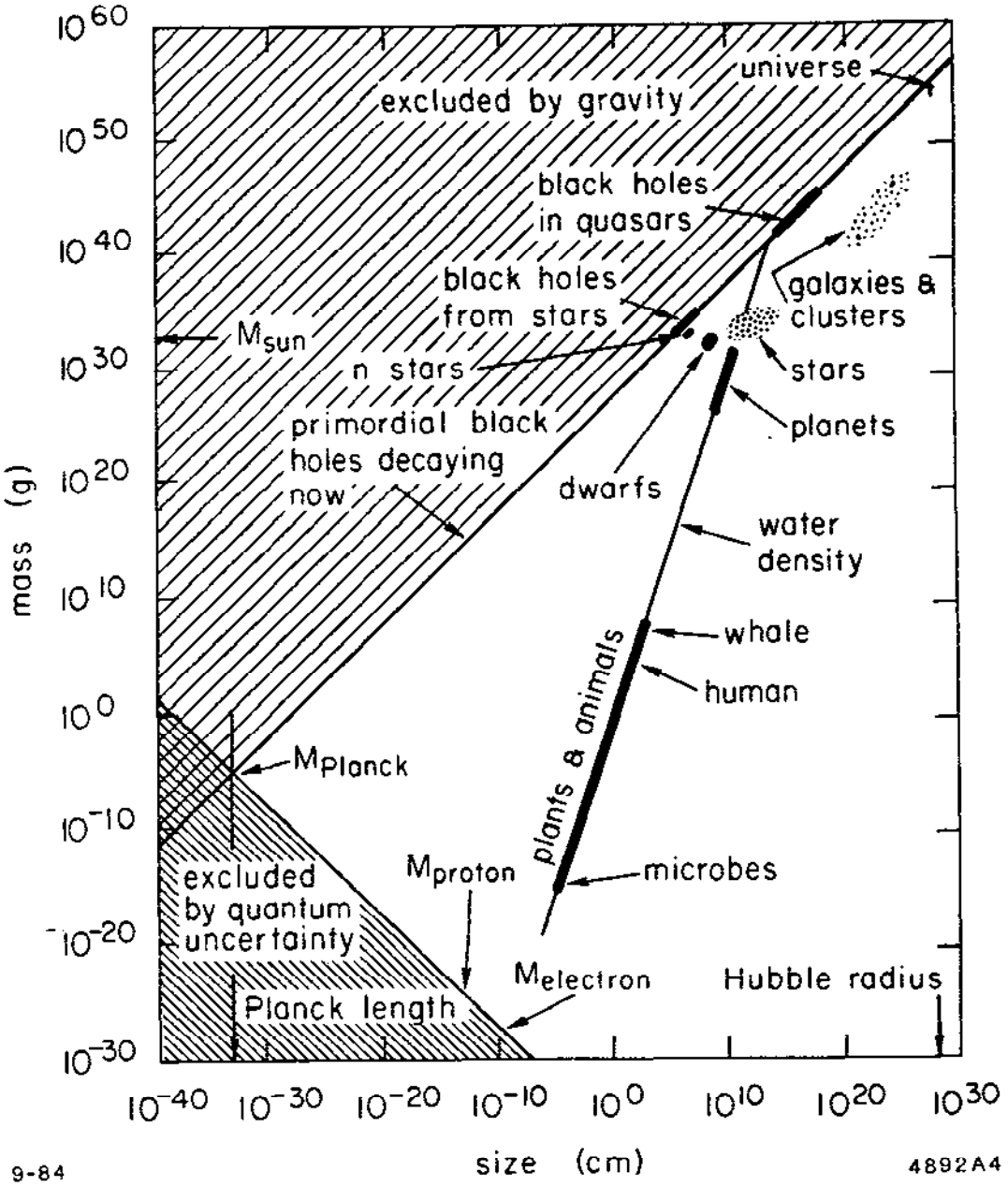

10^{60}
10^{50}
10^{40}
10^{30}
10^{20}
10^{10}
10^{0}
10^{-10}
10^{-20}
10^{-30}
mass (g)
excluded by gravity
universe
black holes in quasars
black holes from stars
galaxies & clusters
M_{sun}
n stars
stars
planets
primordial black holes decaying now
dwarfs
water density
whale
human
plants & animals
microbes
M_{Planck}
excluded by quantum uncertainty
M_{proton}
$M_{electron}$
Hubble radius
Planck length
10^{-40} 10^{-30} 10^{-20} 10^{-10} 10^{0} 10^{10} 10^{20} 10^{30}
size (cm)
9-84
4892A4

Fig. 2.2

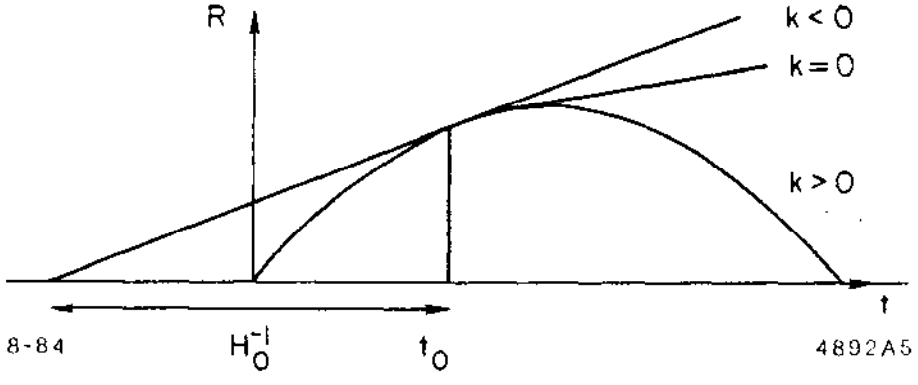


Fig. 2.3

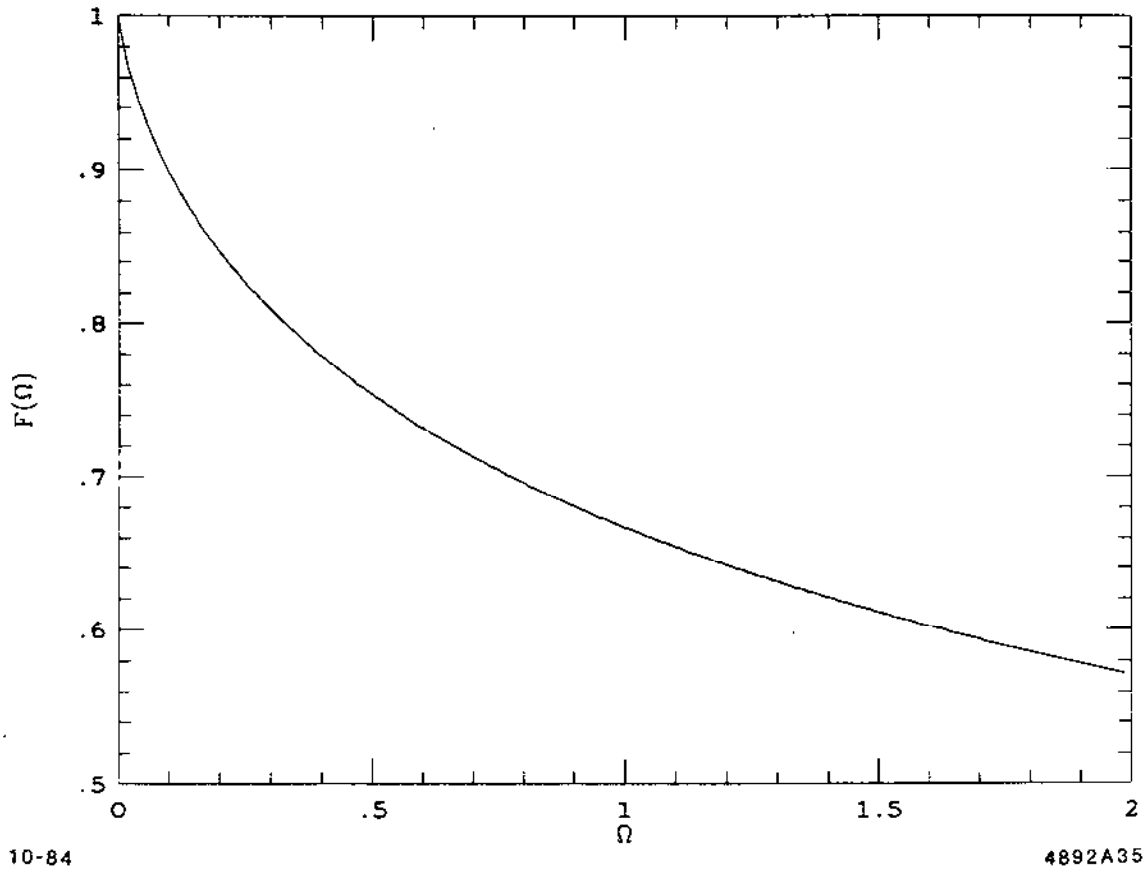


Fig. 2.4

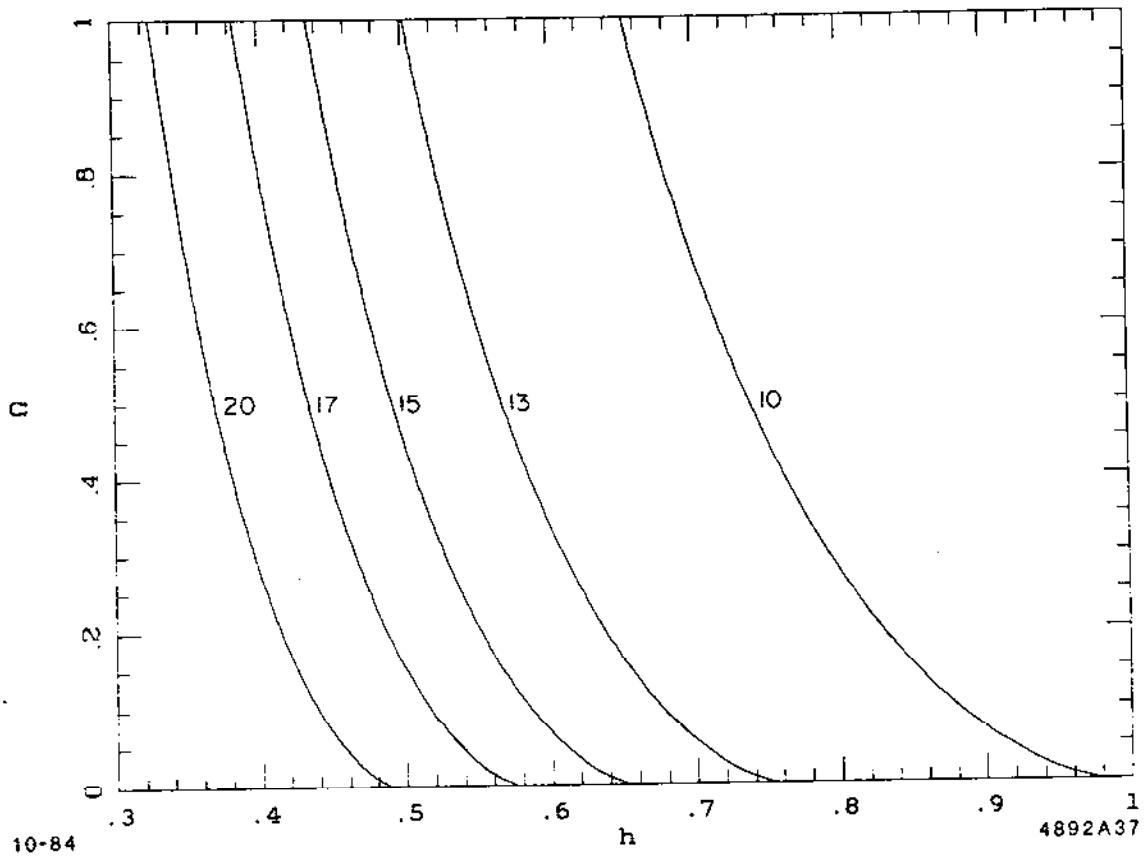


Fig. 2.5

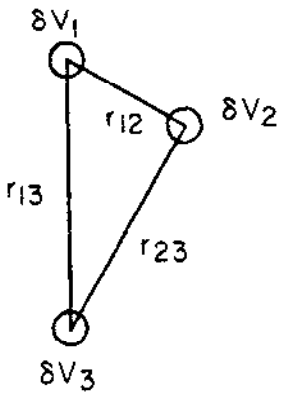


Fig. 2.6

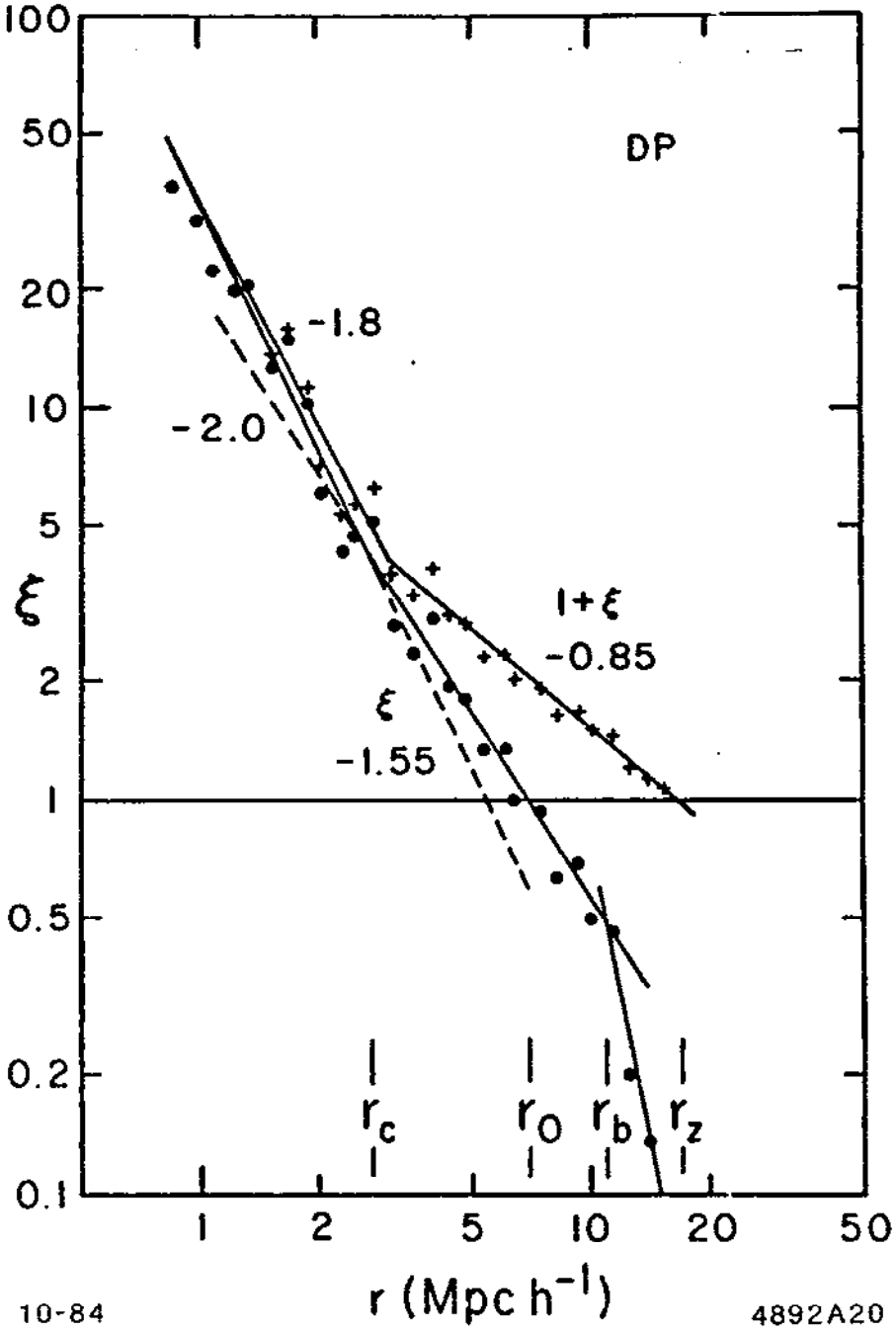

100
50
20
10
5
2
1
0.5
0.2
0.1
ξ
DP
-1.8
-2.0
1+ξ
-0.85
ξ
-1.55
r_c
r_0
r_b
r_z
1
2
5
10
20
50
r (Mpc h^-1)
10-84
4892A20

Fig. 2.7

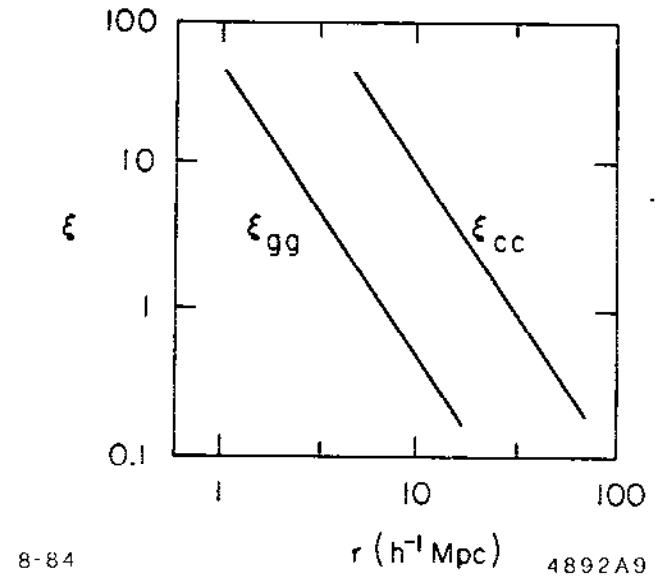


Fig. 2.8

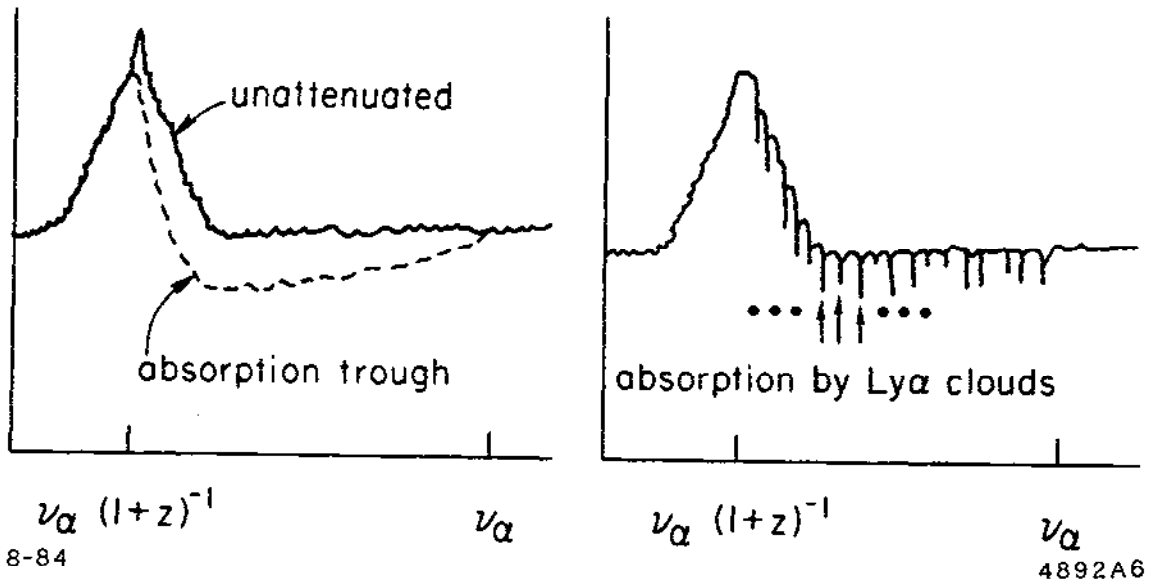

unattenuated
absorption trough
absorption by Lyα clouds
$\nu_\alpha\ (1+z)^{-1}$
ν_α
$\nu_\alpha\ (1+z)^{-1}$
ν_α
8-84
4892A6

Fig. 2.9

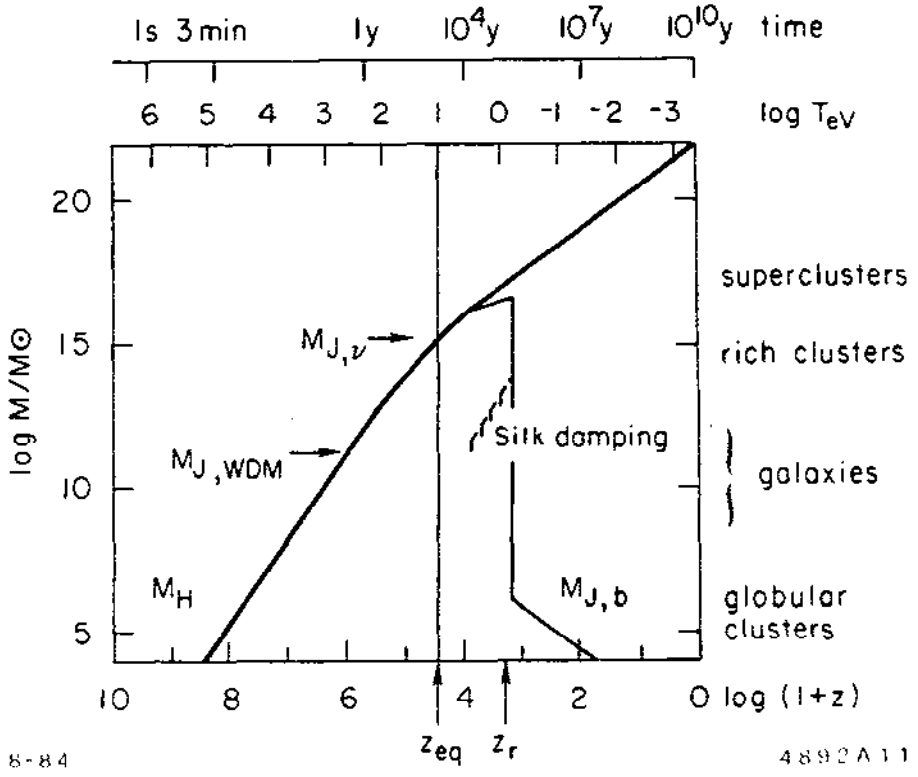

1s 3 min
1y
10^4y
10^7y
10^{10}y time
6 5 4 3 2 1 0 -1 -2 -3
log T_{eV}
20
15
10
5
log M/M⊙
superclusters
rich clusters
galaxies
globular clusters
$M_{J,\nu}$
$M_{J,WDM}$
M_H
Silk damping
$M_{J,b}$
10 8 6 4 2 0 log (1+z)
z_{eq}
z_r
8-84
4892A11

Fig. 2.10

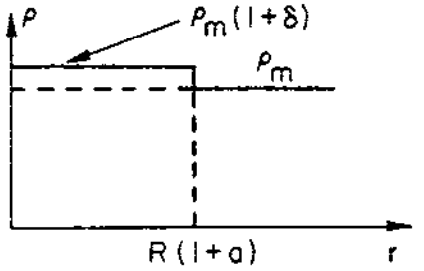


Fig. 2.11

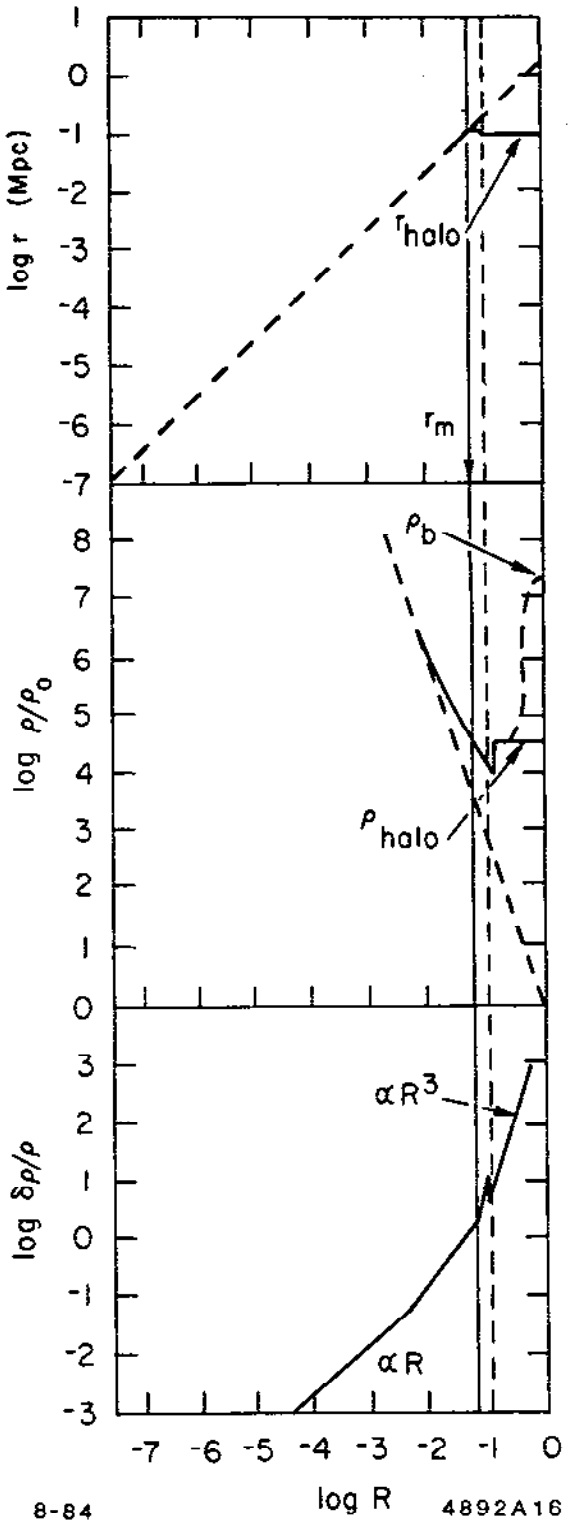

log r (Mpc)
r_{halo}
r_m
log ρ/ρ_o
ρ_b
ρ_{halo}
log $\delta\rho/\rho$
$\propto R^3$
$\propto R$
log R
8-84
4892A16

Fig. 2.12

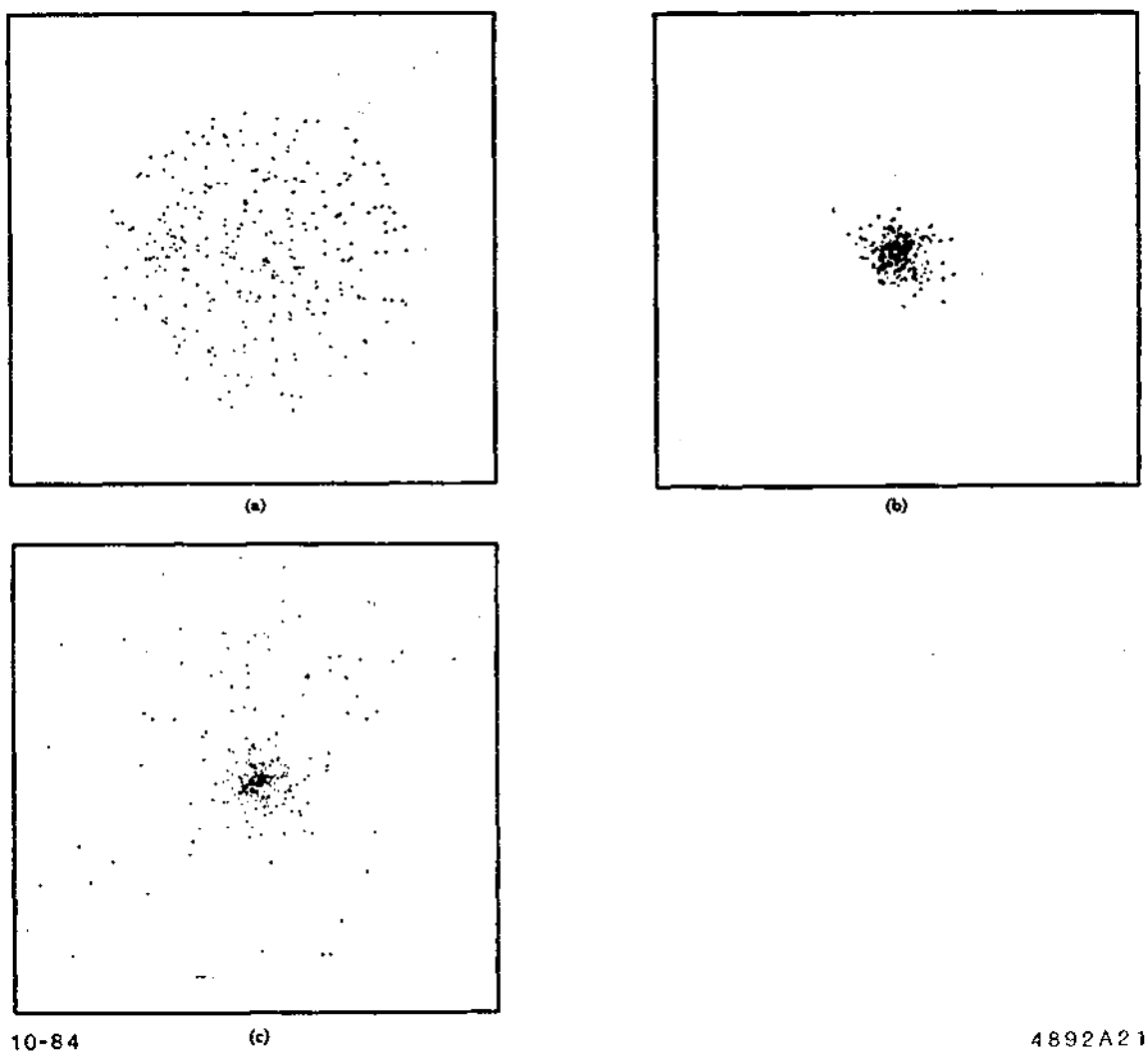


Fig. 2.13

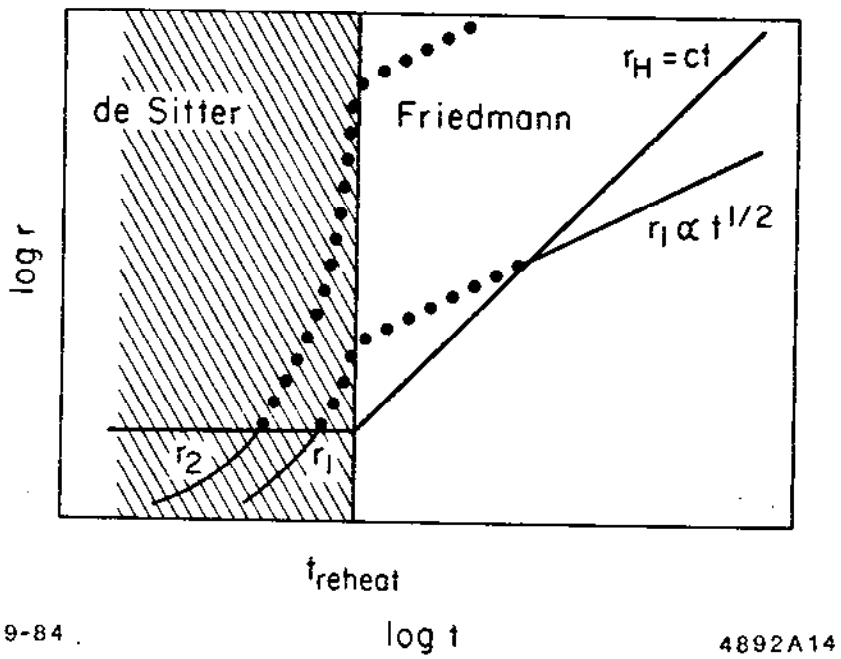

de Sitter
Friedmann
$r_H = ct$
$r_1 \propto t^{1/2}$
r_2
r_1
log r
t_{reheat}
log t
9-84
4892A14

Fig. 2.14

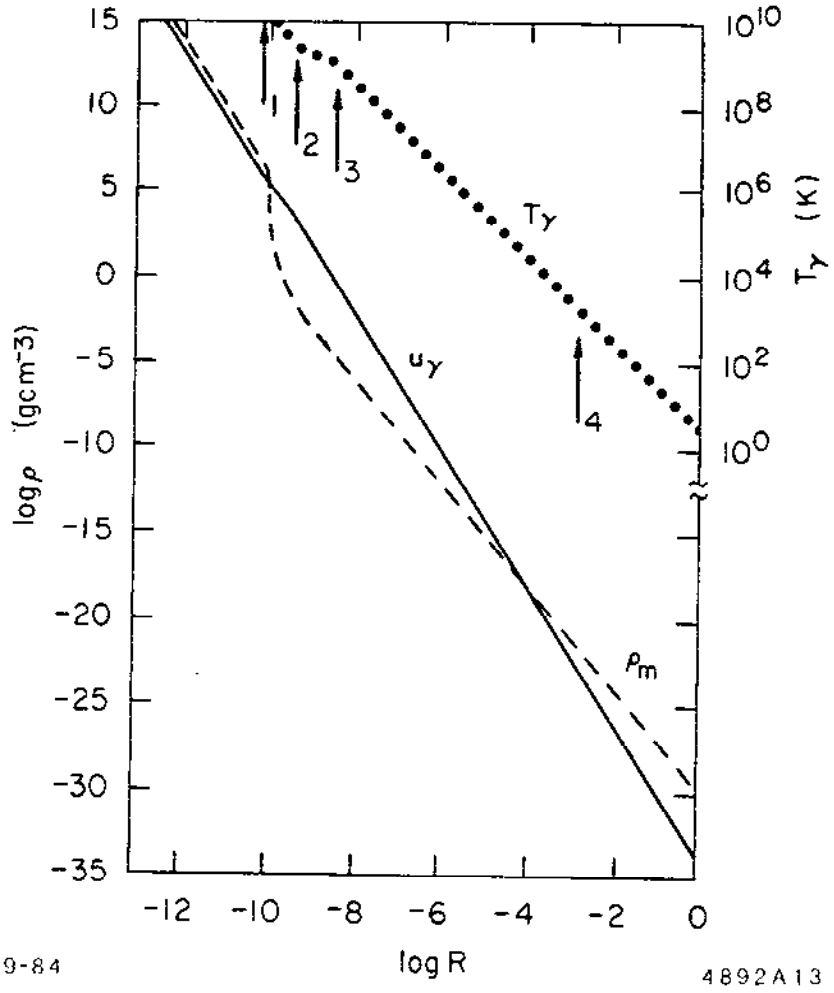

15
10
5
0
-5
-10
-15
-20
-25
-30
-35
log ρ (gcm⁻³)
10¹⁰
10⁸
10⁶
10⁴
10²
10⁰
T_γ (K)
-12
-10
-8
-6
-4
-2
0
log R
1
2
3
4
T_γ
u_γ
ρ_m
9-84
4892A13

Fig. 3.1

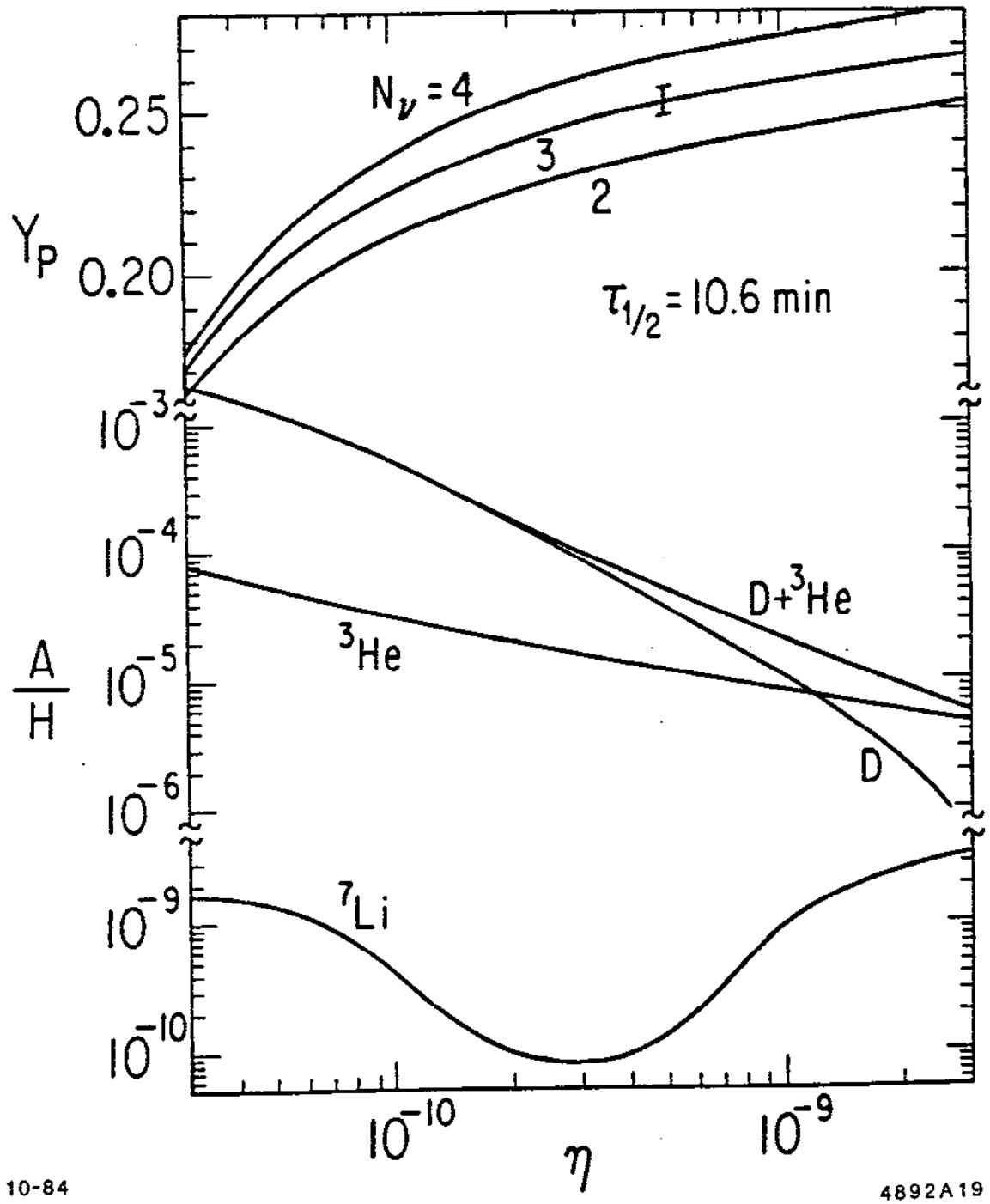

$N_\nu = 4$
3
2
1
$\tau_{1/2} = 10.6$ min
Y_P
0.25
0.20
10^{-3}
10^{-4}
10^{-5}
10^{-6}
10^{-9}
10^{-10}
$\frac{A}{H}$
$D+^3He$
^3He
D
^7Li
10^{-10}
10^{-9}
η
10-84
4892A19

Fig. 3.2

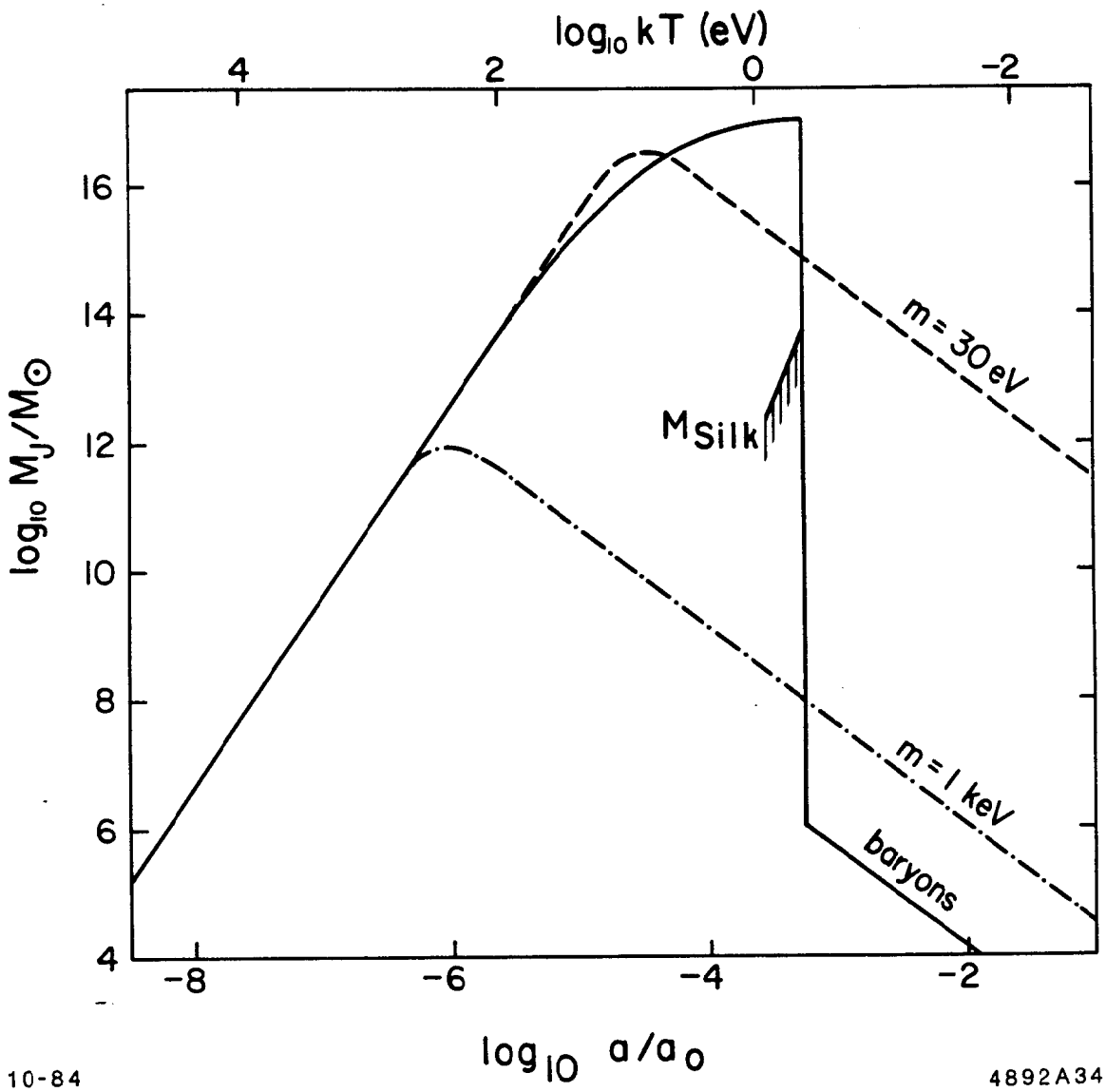

log10 kT (eV)
4
2
0
-2
16
14
12
10
8
6
4
log10 MJ/M⊙
m = 30 eV
MSilk
m = 1 keV
baryons
-8
-6
-4
-2
log10 a/a0
10-84
4892A34

Fig. 3.3

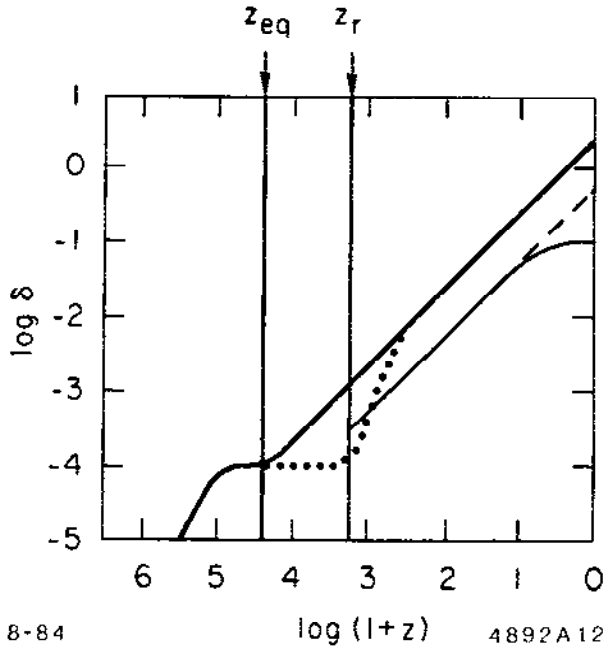


Fig. 3.4

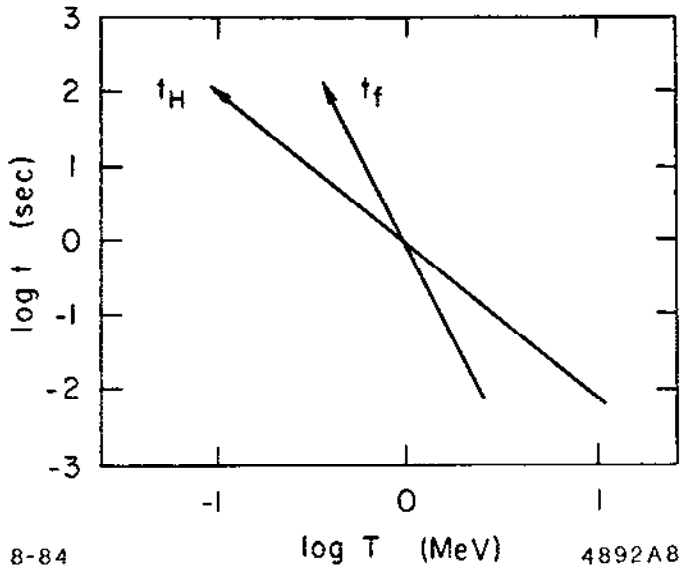
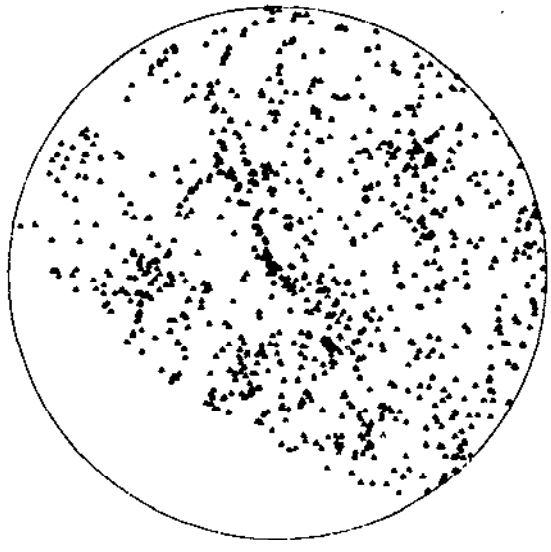
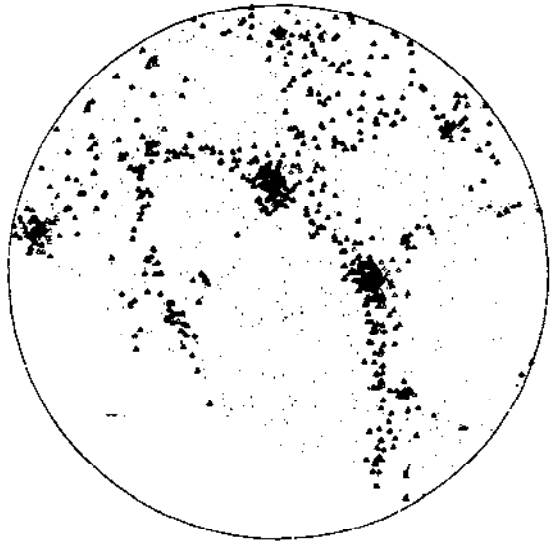
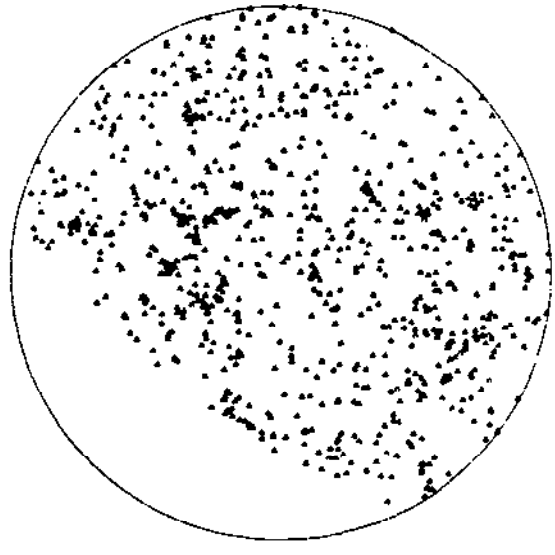


Fig. 3.5



(a)





10-84 (b) (c) 4892A36

Fig. 3.6

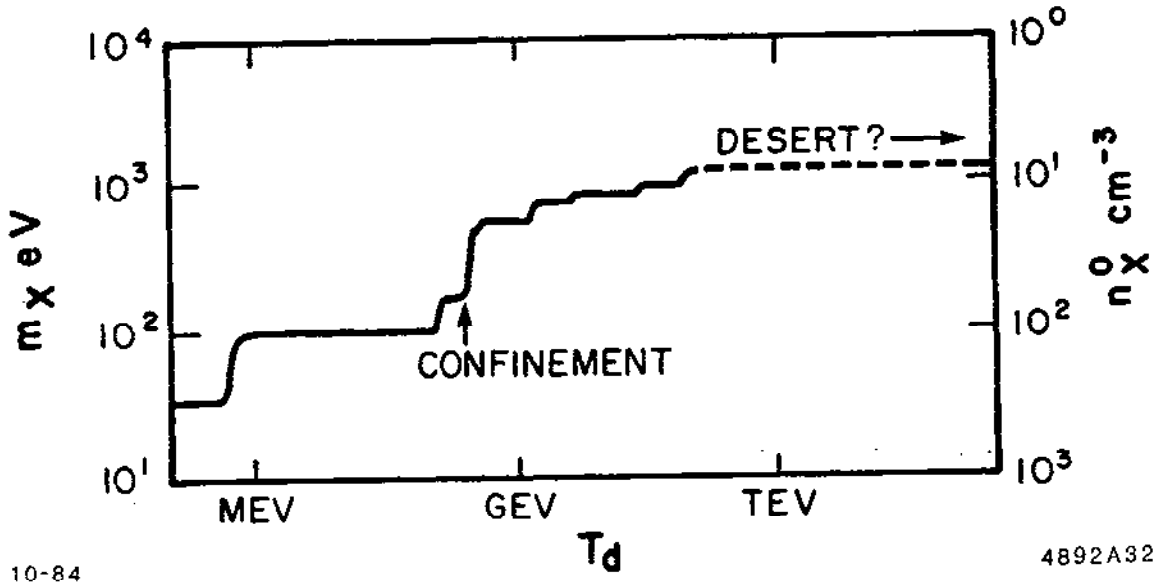

10^4
10^3
10^2
10^1
m x eV
10^0
10^1
10^2
10^3
n o cm^{-3}
DESERT?
CONFINEMENT
MEV
GEV
TEV
T_d
10-84
4892A32

Fig. 3.7

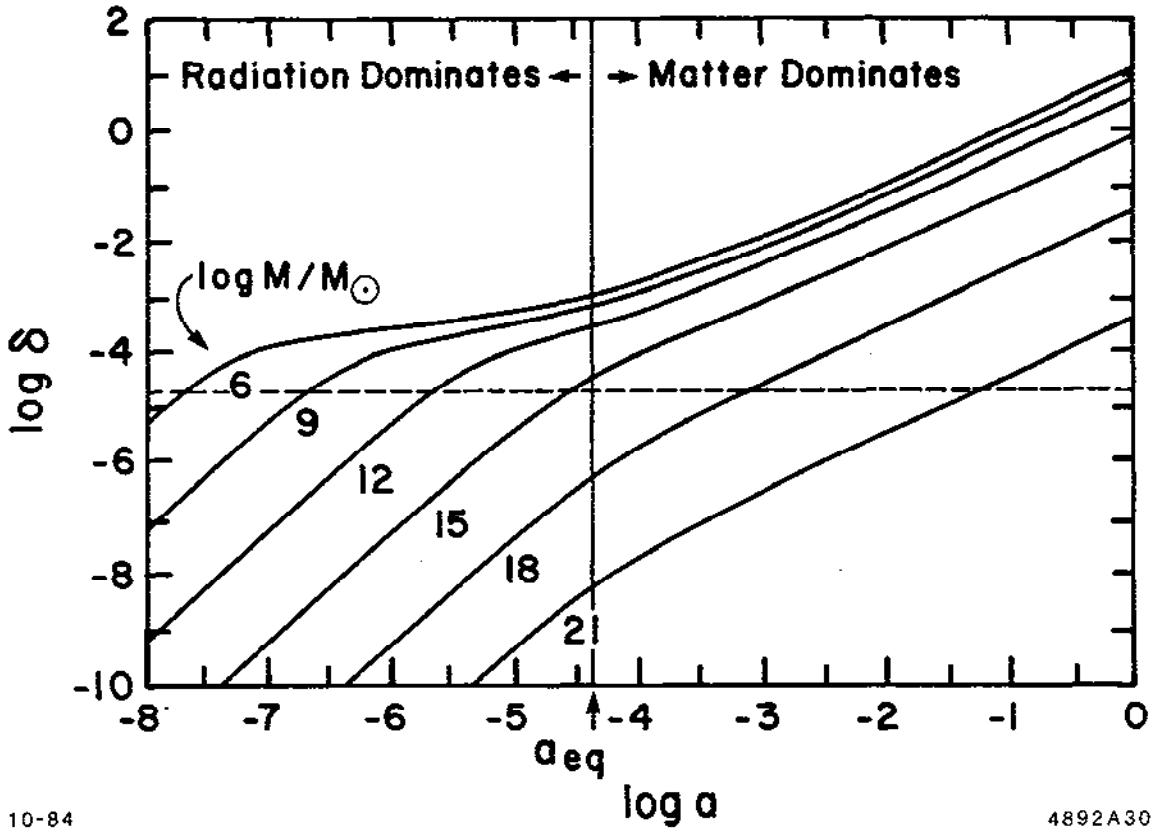

Radiation Dominates
Matter Dominates
log M/M⊙
6
9
12
15
18
21
log δ
log a
a eq
2
0
-2
-4
-6
-8
-10
-8
-7
-6
-5
-4
-3
-2
-1
0
10-84
4892A30

Fig. 4.1

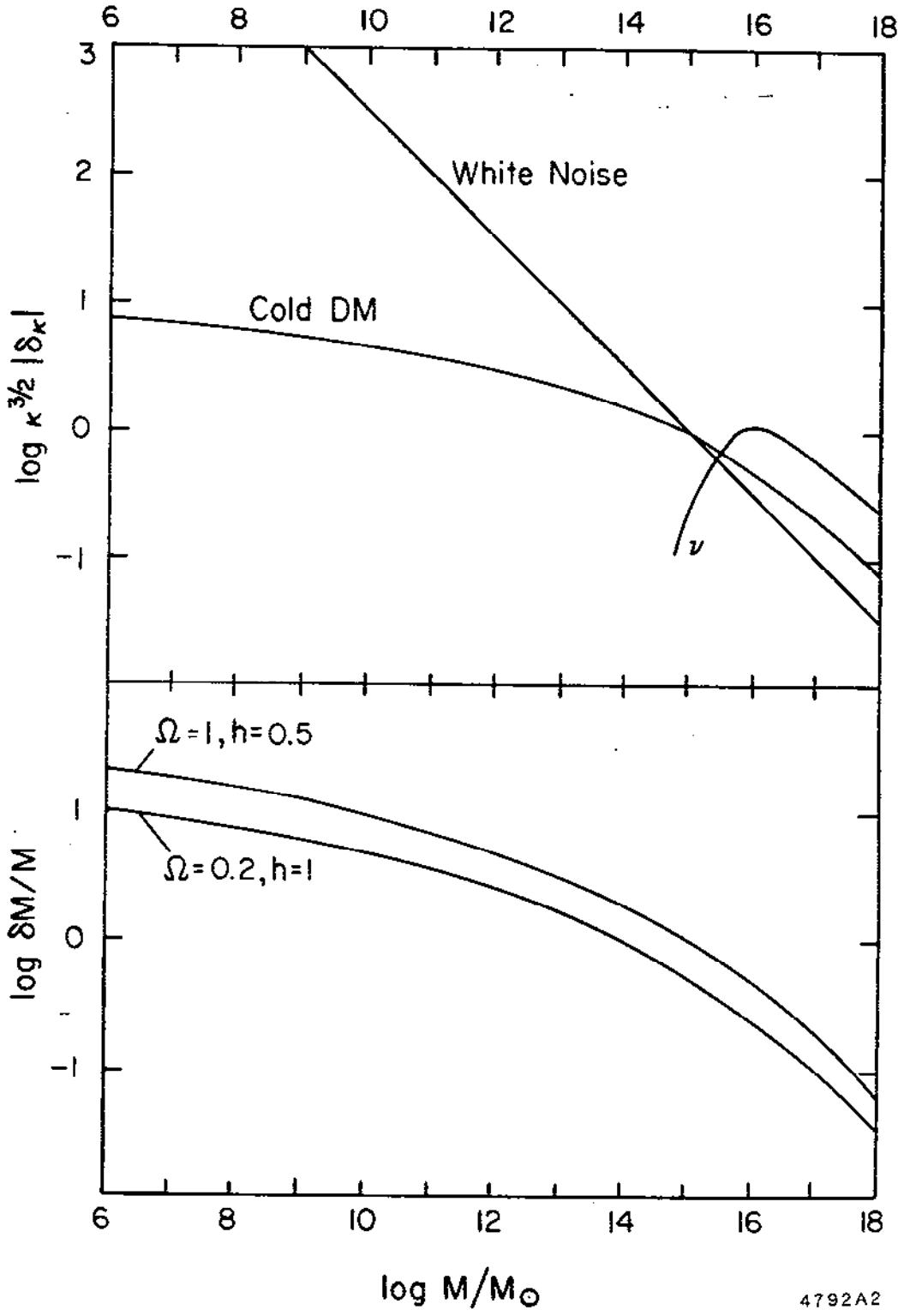

6
8
10
12
14
16
18
3
2
1
0
−1
White Noise
Cold DM
ν
log κ^{3/2} |δ_κ|
Ω = 1, h = 0.5
Ω = 0.2, h = 1
log δM/M
log M/M⊙
4792A2

Fig. 4.2

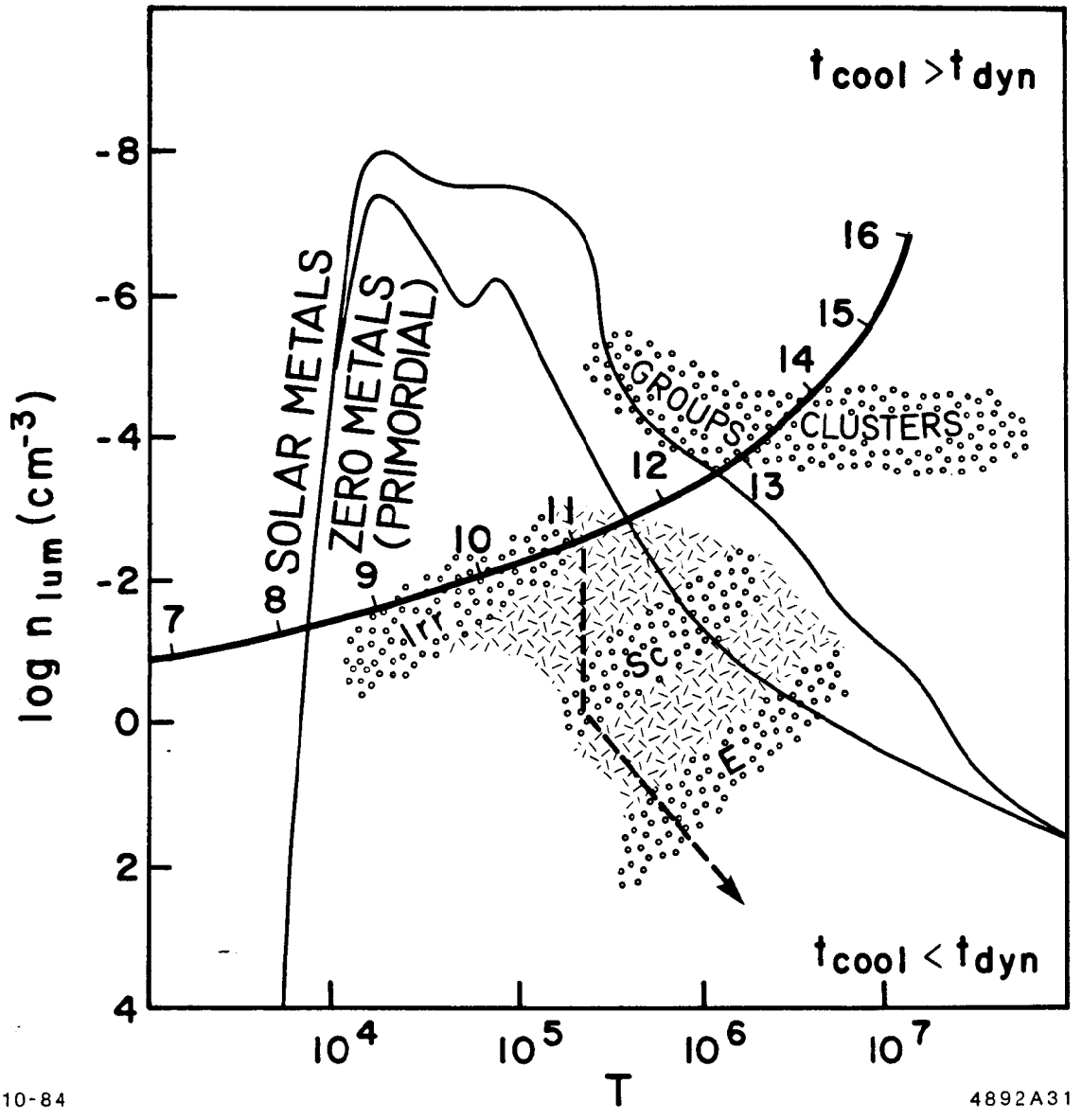

$t_{cool} > t_{dyn}$
$t_{cool} < t_{dyn}$
SOLAR METALS
ZERO METALS (PRIMORDIAL)
GROUPS
CLUSTERS
Irr
Sc
E
7
8
9
10
11
12
13
14
15
16
-8
-6
-4
-2
0
2
4
10^4
10^5
10^6
10^7
T
log n_{lum} (cm^{-3})
10-84
4892A31

Fig. 4.3

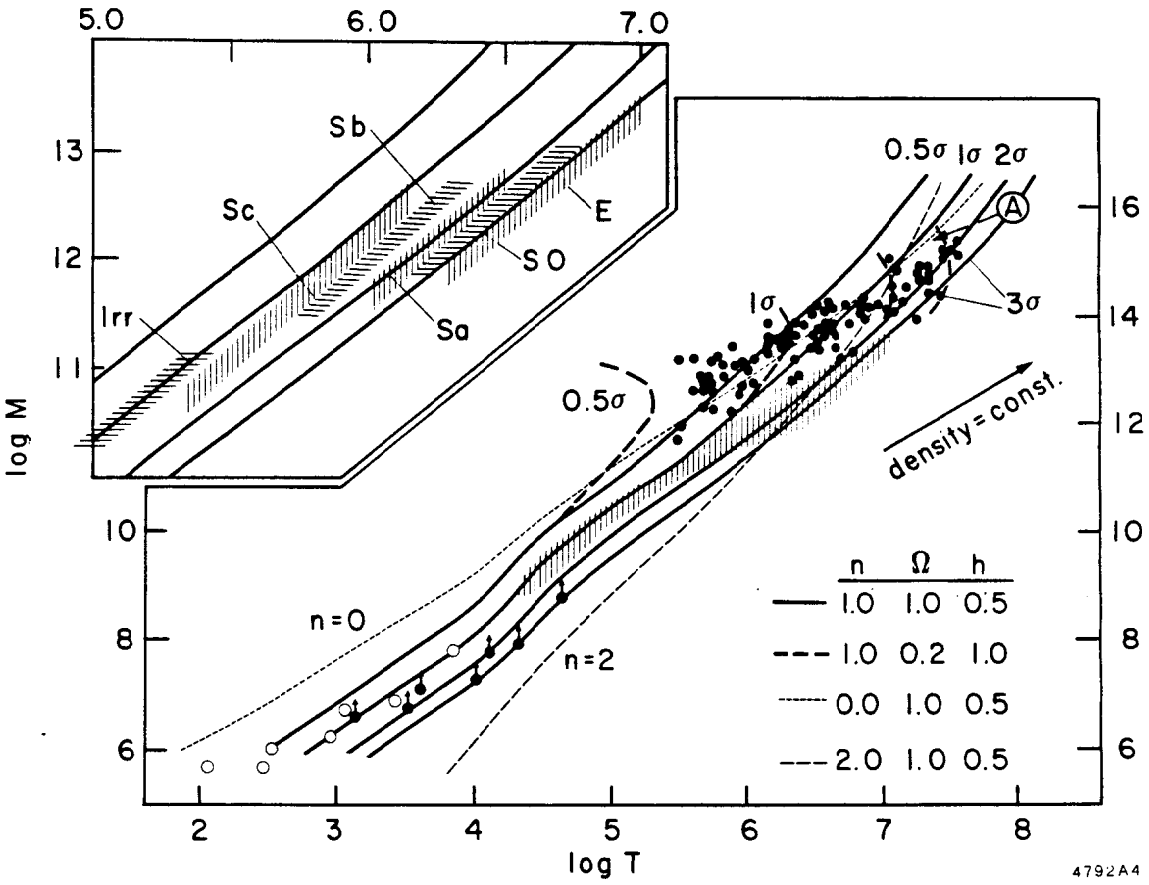

5.0
6.0
7.0
Sb
Sc
E
S0
Sa
Irr
0.5σ
1σ
2σ
3σ
A
density = const.
n=0
n=2
n Ω h
1.0 1.0 0.5
1.0 0.2 1.0
0.0 1.0 0.5
2.0 1.0 0.5
log M
log T
4792A4

Fig. 4.4

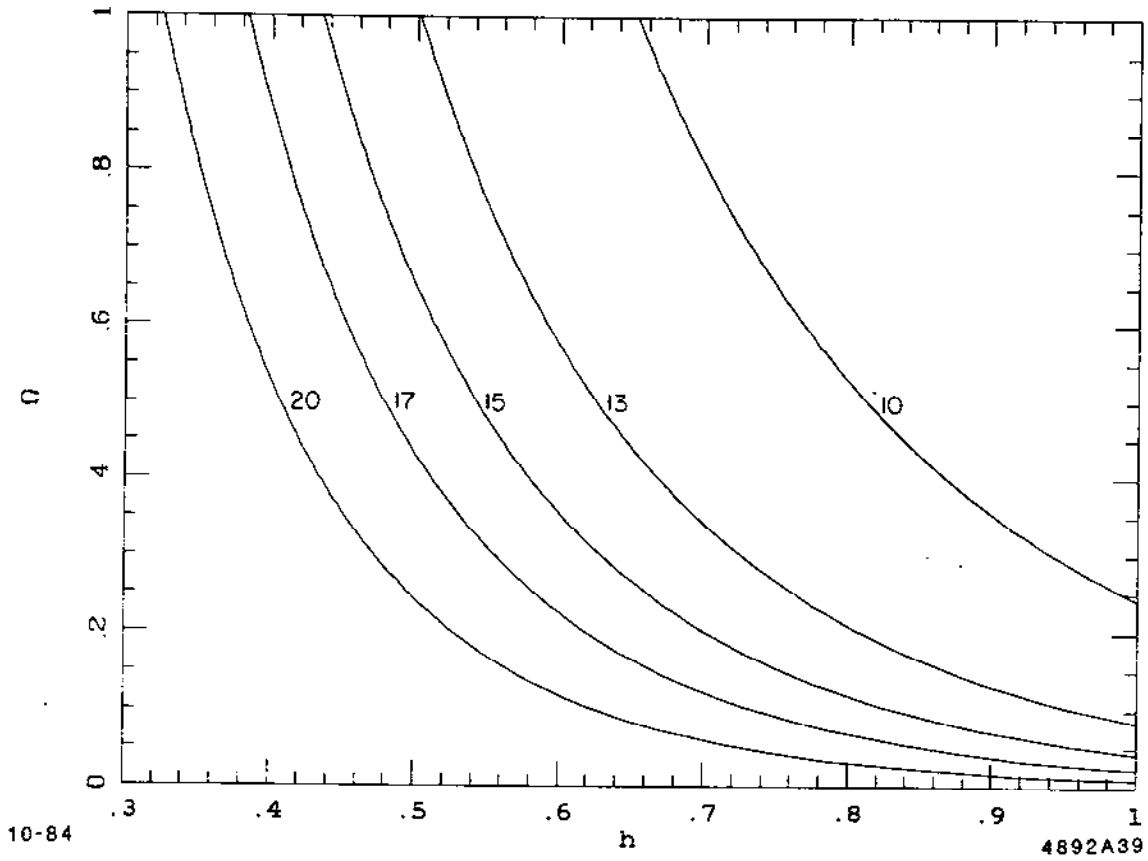


Fig. 4.5

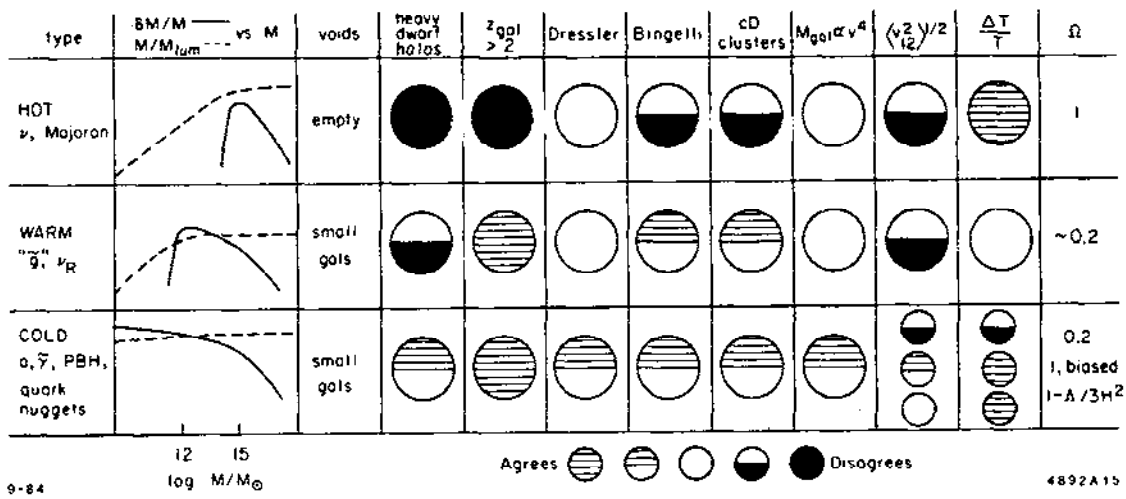

type
δM/M ——
M/M_lum - - - vs M
voids
heavy dwarf halos
z_gal > 2
Dressler
Bingelli
cD clusters
M_gal ∝ v^4
⟨v²_12⟩^1/2
ΔT/T
Ω
HOT
ν, Majoron
empty
1
WARM
"g", ν_R
small gals
~0.2
COLD
a, γ̃, PBH, quark nuggets
small gals
0.2
1, biased
1−Λ/3H²
12
15
log M/M⊙
Agrees
Disagrees
9-84
4892A15

Fig. 4.6

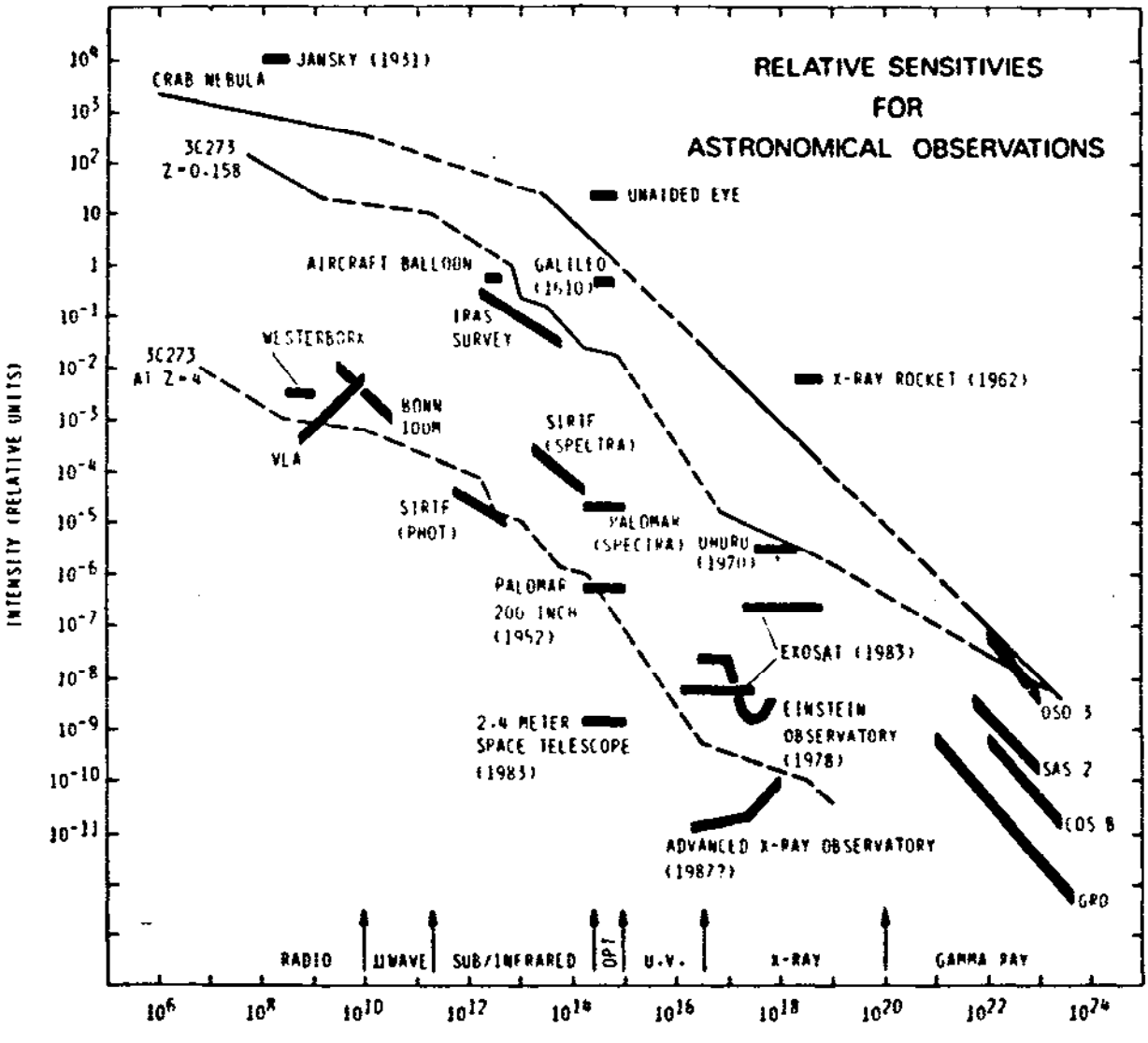

RELATIVE SENSITIVIES
FOR
ASTRONOMICAL OBSERVATIONS
INTENSITY (RELATIVE UNITS)
JANSKY (1931)
CRAB NEBULA
3C273
Z=0.158
UNAIDED EYE
AIRCRAFT BALLOON
GALILEO
(1610)
IRAS
SURVEY
X-RAY ROCKET (1962)
WESTERBORK
3C273
AT Z=4
BONN
100M
VLA
SIRTF
(SPECTRA)
SIRTF
(PHOT)
PALOMAR
(SPECTRA)
UHURU
(1970)
PALOMAR
200 INCH
(1952)
EXOSAT (1983)
EINSTEIN
OBSERVATORY
(1978)
2.4 METER
SPACE TELESCOPE
(1985)
ADVANCED X-RAY OBSERVATORY
(1987?)
OSO 3
SAS 2
COS B
GRO
RADIO
MMWAVE
SUB/INFRARED
OPT
U.V.
X-RAY
GAMMA RAY

Fig. 4.7

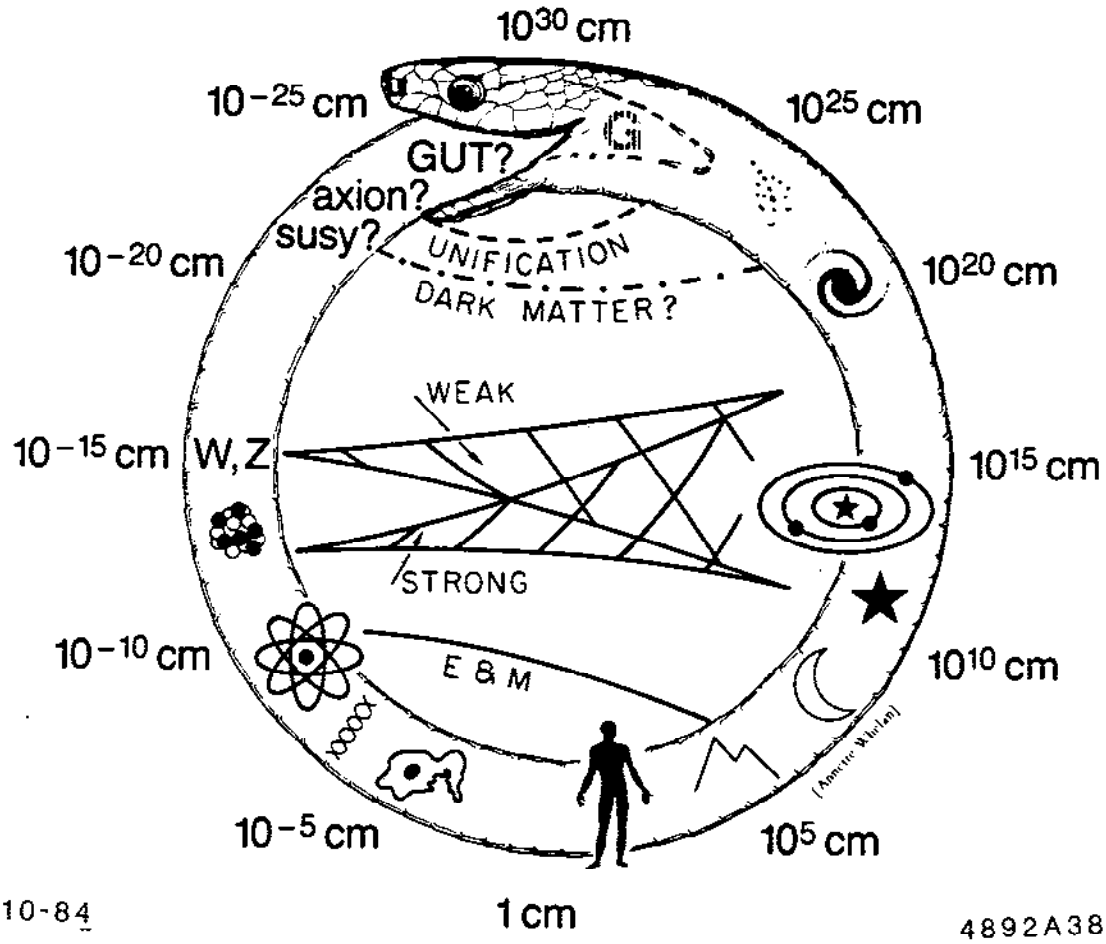

10^{30} cm
10^{-25} cm
10^{25} cm
GUT?
axion?
susy?
G
UNIFICATION
DARK MATTER?
10^{-20} cm
10^{20} cm
WEAK
10^{-15} cm
W, Z
10^{15} cm
STRONG
10^{-10} cm
E & M
10^{10} cm
10^{-5} cm
10^{5} cm
1 cm
10-84
4892A38

Fig. 4.8