THE NATURE OF DARK MATTER

Joel R. Primack

Physics Department, University of California, Santa Cruz, CA 95064 USA

ABSTRACT

Cold Dark Matter (CDM) has become the standard modern theory of cosmological structure formation. Its predictions appear to be in good agreement with data on large scales, and it naturally accounts for many properties of galaxies. But despite its many successes, there has been concern about CDM on small scales because of the possible contradiction between the linearly rising rotation curves observed in some dark-matter-dominated galaxies vs. the $1/r$ density cusps at the centers of simulated CDM halos. Other CDM issues on small scales include the very large number of small satellite halos in simulations, far more than the number of small galaxies observed locally, and problems concerning the angular momentum of the baryons in dark matter halos. The latest data and simulations have lessened, although not entirely resolved, these concerns. Meanwhile, the main alternatives to CDM that have been considered to solve these problems, self-interacting dark matter (SIDM) and warm dark matter (WDM), have been found to have serious drawbacks.¹

¹This paper is a significantly updated revision of ker.
1 Introduction

Sometimes a theory is proposed in relatively early stages of the development of a scientific field, and this theory turns out to be not only a useful paradigm for the further development of the field — it also survives confrontation with a vast amount of data, and becomes accepted as the standard theory. This happened with General Relativity \textsuperscript{2}, and it seems to be happening now with general relativistic cosmology. It appears that the universe on the largest scales can indeed be described by three numbers:

- \( H_0 \equiv 100h\text{km}\text{s}^{-1}\text{Mpc}^{-1} \), the Hubble parameter (expansion rate of the universe) at the present epoch,
- \( \Omega_m \equiv \rho/\rho_c \), the density of matter \( \rho \) in units of critical density \( \rho_c \equiv 3H_0^2(8\pi G)^{-1} = 2.78 \times 10^{11}h^2\text{M}_\odot\text{Mpc}^{-3} \), and
- \( \Omega_\Lambda \equiv \Lambda(3H_0^2)^{-1} \), the corresponding quantity for the cosmological constant.

The currently measured values of these and other key parameters are summarized in the Table below. It remains to be seen whether the “dark energy” represented by the cosmological constant \( \Lambda \) is really constant, or is perhaps instead a consequence of the dynamics of some fundamental field as in “quintessence” theories \textsuperscript{3}.

In particle physics, the first unified theory of the weak and electromagnetic interactions \textsuperscript{4} had as its fundamental bosons just the carriers of the charged weak interactions \( W^+, W^- \), and the photon \( \gamma \). The next such theory \textsuperscript{5} had a slightly more complicated pattern of gauge bosons — a triplet plus a singlet, out of which came not only \( W^+, W^- \), and \( \gamma \), but also the neutral weak boson \( Z^0 \), and correspondingly an extra free parameter, the “Weinberg angle.” It was of course this latter SU(2)\( \times \)U(1) theory which has now become part of the Standard Model of particle physics. During the early 1970s, however, when the experimental data were just becoming available and some of the data appeared to contradict the SU(2)\( \times \)U(1) theory, many other more complicated theories were proposed, even by Weinberg \textsuperscript{6}, but all these more complicated theories ultimately fell by the wayside.
The development of theories of dark matter may follow a similar pattern. By the late 1970s it was becoming clear both that a great deal of dark matter exists and that the cosmic microwave background (CMB) fluctuation amplitude is smaller than that predicted in a baryonic universe. The first non-baryonic dark matter candidate to be investigated in detail was light neutrinos — what we now call “hot dark matter” (HDM). This dark matter is called “hot” because at one year after the big bang, when the horizon first encompassed the amount of matter in a large galaxy like our own (about $10^{12} M_\odot$) and the temperature was about 1 keV, neutrinos with masses in the eV range would have been highly relativistic.

It is hardly surprising that HDM was worked out first. Neutrinos were known to exist, after all, and an experiment in Moscow that had measured a mass for the electron neutrino $m(\nu_e) \approx 20$ eV (corresponding to $\Omega_m \approx 1$ if $h$ were as small as $\sim 0.5$, since $\Omega_\nu = m(\nu_e)/(92 h^2 \text{eV})^{-1}$) had motivated especially Zel’dovich and his colleagues to work out the implications of HDM with a Zel’dovich spectrum ($P_p(k) = A k^n$ with $n = 1$) of adiabatic primordial fluctuations. But improved experiments subsequently have only produced upper limits for $m(\nu_e)$, currently about 3 eV, and the predictions of the adiabatic HDM model are clearly inconsistent with the observed universe.

Cold Dark Matter (CDM) was worked out as the problems with HDM were beginning to become clear. CDM assumes that the dark matter is mostly cold — i.e., with negligible thermal velocities in the early universe, either because the dark matter particles are weakly interacting massive particles (WIMPs) with mass $\sim 10^2$ GeV, or alternatively because they are produced without a thermal distribution of velocities, as is the case with axions. The CDM theory also assumes, like HDM, that the fluctuations in the dark matter have a nearly Zel’dovich spectrum of adiabatic fluctuations. Considering that the CDM model of structure formation in the universe was proposed almost twenty years ago, its successes are nothing short of amazing. As I will discuss, the $\Lambda$CDM variant of CDM with $\Omega_m = 1 - \Omega_\Lambda \approx 0.3$ appears to be in good agreement with the available data on large scales. Issues that have arisen on smaller scales, such as the centers of dark matter halos and the numbers of small satellites, have prompted people to propose a wide variety of alternatives to CDM, such as self-interacting dark matter (SIDM). It remains to be seen whether such alternative theories with extra parameters ac-
ually turn out to be in better agreement with data. As I will discuss below, it now appears that SIDM is probably ruled out, while the small-scale predictions of CDM may be in better agreement with the latest data than appeared to be the case as recently as a year ago.

In the next section I will briefly review the current observations and the successes of ΛCDM on large scales, and then I will discuss the possible problems on small scales.

2 Cosmological Parameters and Observations on Large Scales

The table below summarizes the current observational information about the cosmological parameters, with estimated 1σ errors. The quantities in brackets have been deduced using at least some of the ΛCDM assumptions. It is apparent that there is impressive agreement between the values of the parameters determined by various methods, including those based on ΛCDM. In particular, (A) several different approaches (some of which are discussed further below) all suggest that $\Omega_m \approx 0.3$; (B) the location of the first acoustic peak in the CMB angular anisotropy power spectrum, now very well determined independently by the BOOMERANG and MAXIMA1 balloon data, and by the DASI interferometer at the South Pole, implies that $\Omega_m + \Omega_\Lambda \approx 1$; and (C) the data on supernovae of Type Ia (SNIa) at redshifts $z = 0.4 - 1.2$ from two independent groups imply that $\Omega_\Lambda - \frac{3}{4}\Omega \approx \frac{1}{3}$. Any two of these three results then imply that $\Omega_\Lambda \approx 0.7$. The 1σ errors in these determinations are about 0.1.

Questions have been raised about the reliability of the high-redshift SNIa results, especially the possibilities that the SNIa properties at high redshift might not be sufficiently similar to those nearby to use them as standard candles, and that there might be “grey” dust (which would make the SNIa dimmer but not change their colors). Although the available evidence disfavors these possibilities, additional observations are needed on SNIa at high redshift, both...

---

2Further discussion and references are given in 16).

3For example, SNIa at $z = 1.2$ and $\sim 1.7$ apparently have the brightness expected in a ΛCDM cosmology but are brighter than would be expected with grey dust, and the infrared brightness of a nearer SNIa is also inconsistent with grey dust.
to control systematic effects and to see whether the dark energy is just a cosmological constant or is perhaps instead changing with redshift as expected in “quintessence” models\(^3\). Such data could be obtained by the proposed SuperNova Acceleration Probe (SNAP) satellite\(^2\), whose Gigapixel optical camera and other instruments would also produce much other useful data. But it is important to appreciate that, independently of (C) SNIa, (A) cluster and other evidence for \(\Omega_m \approx 0.3\), together with (B) \(\sim 1^\circ\) CMB evidence for \(\Omega_m + \Omega_\Lambda \approx 1\), imply that \(\Omega_\Lambda \approx 0.7\).

All methods for determining the Hubble parameter now give compatible results, confirming our confidence that this crucial parameter has now been measured robustly to a 1\(\sigma\) accuracy of about 10%. The final result\(^2\) from the Hubble Key Project on the Extragalactic Distance Scale is 72\(\pm\)8 km s\(^{-1}\) Mpc\(^{-1}\), or \(h = 0.72 \pm 0.08\), where the stated error is dominated by one systematic uncertainty, the distance to the Large Magellanic Cloud (used to calibrate the Cepheid period-luminosity relationship). The most accurate of the direct methods for measuring distances \(d\) to distant objects, giving the Hubble parameter directly as \(H_0 = d/v\) where the velocity is determined by the redshift, are (1) time delays between luminosity variations in different gravitationally lensed images of distant quasars, giving \(h \approx 0.65\), and (2) the Sunyaev-Zel’dovich

\[\begin{align*}
H_0 &= 100\ km\ s^{-1}\ Mpc^{-1}\ ,\ h = 0.7 \pm 0.08 \\
t_0 &= 13 \pm 2\ \text{Gyr (from globular clusters)} \\
&= [12 \pm 2\ \text{Gyr from expansion age, ACDM model}] \\
\Omega_b &= (0.039 \pm 0.006)h_{70}^2\ \text{from D/H} \\
&> [0.035h_{70}^{-2}\ \text{from Ly}\alpha\ \text{forest opacity}] \\
\Omega_m &= 0.4 \pm 0.2\ \text{(from cluster baryons)} \\
&= 0.34 \pm 0.1\ \text{from Ly}\alpha\ \text{forest} P(k) \\
&= 0.4 \pm 0.2\ \text{(from cluster evolution)} \\
&\approx \frac{2\Omega_A - \frac{1}{3}}{\Omega_m + \Omega_\Lambda}
\end{align*}\n
\[
\Omega_m + \Omega_\Lambda = 1.02 \pm 0.06\ \text{from CMB peak location} \\
\Omega_A &= 0.73 \pm 0.08\ \text{(from previous two lines)} \\
&< 0.73\ \text{(2\(\sigma\) from radio QSO lensing)} \\
\Omega_\nu &\approx 0.001\ \text{(from SuperKamiokande data)} \\
&\approx [0.1\ \text{in ACDM-type models}]
effect (Compton scattering of the CMB by the hot electrons in clusters of galaxies), giving $h \approx 0.63$ [23, 24]. For the rest of this article, I will take $h = 0.7$ whenever I need to use an explicit value, and express results in terms of $h \tau_0 = H_0/70 \text{ km s}^{-1} \text{ Mpc}^{-1}$.

For a $\Lambda$CDM universe with $\Omega_m = (0.2)(0.3(0.4, 0.5)$, the expansion age is $t_0 = (15.0)13.47(12.41, 11.61)h^{-1}_70 \text{ Gyr}$. Thus for $\Omega_m \approx 0.3 - 0.4$ and $h \approx 0.7$, there is excellent agreement with the latest estimates of the ages of the oldest globular cluster stars in the Milky Way, both from their Main Sequence turnoff luminosities [20], giving $12 - 13 \pm 2 \text{ Gyr}$, and using the thorium and uranium radioactive decay chronometers [27], giving $14 \pm 3 \text{ Gyr}$ and $12.5 \pm 3 \text{ Gyr}$, respectively.

The simplest and clearest argument that $\Omega_m \approx 1/3$ comes from comparing the baryon abundance in clusters $f_b \equiv M_b/M_{\text{tot}}$ to that in the universe as a whole $\Omega_b/\Omega_m$, as emphasized by White et al. [28]. Since clusters are evidently formed from the gravitational collapse of a region of radius $\sim 10 \text{ Mpc}$, they should represent a fair sample of both baryons and dark matter. This is confirmed in CDM simulations [24]. The fair sample hypothesis implies that

$$\Omega_m = \frac{\Omega_b}{f_b} = 0.3 \left( \frac{\Omega_b}{0.04} \right) \left( \frac{0.13}{f_b} \right).$$

We can use this to determine $\Omega_m$ using the baryon abundance $\Omega_b h^2 = 0.019 \pm 0.003$ (95% C.L.) from the measurement of the deuterium abundance in high-redshift Lyman limit systems [30, 31]. Using X-ray data from an X-ray flux limited sample of clusters to estimate the baryon fraction $f_b = 0.075h^{-3/2}$ gives $\Omega_m = 0.25h^{-1/2} = 0.3 \pm 0.1$ (using $h = 0.70 \pm 0.08$). Estimating the baryon fraction using Sunyaev-Zel’dovich measurements of a sample of 18 clusters gives $f_b = 0.077h^{-1}$ [25], and implies $\Omega_m = 0.25h^{-1} = 0.36 \pm 0.1$.

There is another way to use clusters to measure $\Omega_m$, which takes advantage of the fact that the redshift at which structures form depends strongly on $\Omega_m$. This happens because in a low-density universe the growth rate of fluctuations slows when, on the right hand side of the Friedmann equation,

$$H^2 = \frac{8\pi G\rho}{3} - \frac{k}{R^2} + \frac{\Lambda}{3},$$

the first (matter) term becomes smaller than either the second (curvature) term (for the case of an open universe) or the third (cosmological constant) term.
As I have already discussed, the $\Lambda$ term appears to be dominant now; note that if we evaluate the Friedmann equation at the present epoch and divide both sides by $H_0^2$, the resulting equation is just

$$1 = \Omega_m + \Omega_k + \Omega_\Lambda .$$

(3)

Therefore, if we normalize the fluctuation power spectrum $P(k)$ for an $\Omega_m = 1$ (Einstein-de Sitter) cosmology and for a $\Lambda$CDM one by choosing $\sigma_8$ so that each is consistent with COBE and has the same abundance of clusters today, then at higher redshifts the low-$\Omega_m$ universe will have a higher comoving number density of clusters. Probably the most reliable way of comparing clusters nearby with those at higher redshift uses the cluster X-ray temperatures; the latest results, comparing 14 clusters at an average redshift of 0.38 with 25 nearby clusters, give $\Omega_m = 0.44 \pm 0.12$ (90\% CL) \cite{35}. There is greater leverage in this test if one can use higher redshift clusters, but the challenge is to find large samples with well understood cluster selection and properties. The largest such sample now available is from the Las Companas Distant Cluster Survey, which goes to redshifts $\sim 1$, from which the preliminary result is $\Omega_m = 0.30 \pm 0.12$ (90\% CL) \cite{35}.

3 Further Successes of $\Lambda$CDM

We have already seen that $\Lambda$CDM correctly predicts the abundances of clusters nearby and at $z < 1$ within the current uncertainties in the values of the parameters. It is even consistent with $P(k)$ from the Ly$\alpha$ forest \cite{36} and from CMB anisotropies. Low-$\Omega_m$, CDM predicts that the amplitude of the power spectrum $P(k)$ is rather large for $k < 0.02h$/Mpc$^{-1}$, i.e. on size scales larger ($k$ smaller) than the peak in $P(k)$. The largest-scale surveys, 2dF and SDSS, should be able to measure $P(k)$ on these scales and test this crucial prediction soon; preliminary results are encouraging \cite{37}.

The hierarchical structure formation which is inherent in CDM already explains why most stars are in big galaxies like the Milky Way \cite{34}: smaller galaxies merge to form these larger ones, but the gas in still larger structures takes too long to cool to form still larger galaxies, so these larger structures — the largest bound systems in the universe — become groups and clusters instead of galaxies.
What about the more detailed predictions of ΛCDM, for example on the spatial distribution of galaxies. On large scales, there appears to be a pretty good match. In order to investigate such questions quantitatively on the smaller scales where the best data is available it is essential to do N-body simulations, since the mass fluctuations $\delta \rho / \rho$ are nonlinear on the few-Mpc scales that are relevant. My colleagues and I were initially concerned that ΛCDM would fail this test, since the dark matter power spectrum $P_{dm}(k)$ in ΛCDM, and its Fourier transform the correlation function $\xi_{dm}(r)$, are seriously in disagreement with the galaxy data $P_g(k)$ and $\xi_g(r)$. One way of describing this is to say that scale-dependent antibiasing is required for ΛCDM to agree with observations. That is, the bias parameter $b(r) \equiv [\xi_g(r)/\xi_{dm}(r)]^{1/2}$, which is about unity on large scales, must decrease to less than 1/2 on scales of a few Mpc.

This was the opposite of what was expected: galaxies were generally thought to be more correlated than the dark matter on small scales. However, when it became possible to do simulations of sufficiently high resolution to identify the dark matter halos that would host visible galaxies, it turned out that their correlation function is essentially identical with that of observed galaxies. This is illustrated in Fig. 1.

Jim Peebles, who largely initiated the study of galaxy correlations and first showed that $\xi_g(r) \approx (r/r_0)^{-1.8}$ with $r_0 \approx 5h^{-1}$Mpc, thought that this simple power law must be telling us something fundamental about cosmology. However, it now appears that the power law $\xi_g$ arises because of a coincidence – an interplay between the non-power-law $\xi_{dm}(r)$ (see Fig. 1) and the decreasing survival probability of dark matter halos in dense regions because of their destruction and merging. But the essential lesson is that ΛCDM correctly predicts the observed $\xi_g(r)$.

The same theory also predicts the number density of galaxies. Using the observed correlations between galaxy luminosity and internal velocity, known as the Tully-Fisher and Faber-Jackson relations for spiral and elliptical galaxies respectively, it is possible to convert observed galaxy luminosity functions into approximate galaxy velocity functions, which describe the number of galaxies per unit volume as a function of their internal velocity. The velocity function of dark matter halos is robustly predicted by N-body simulations for CDM-type theories, but to connect it with the observed internal velocities of bright galaxies it is necessary to correct for the infall of the baryons in these galaxies.
When we did this it appeared that $\Lambda$CDM with $\Omega_m = 0.3$ predicts perhaps too many dark halos compared with the number of observed galaxies with internal rotation velocities $V \approx 200$ km s$^{-1}$ [14, 17]. While the latest results from the

![Figure 1: Bottom panel: Comparison of the halo correlation function in an $\Lambda$CDM simulation with the correlation function of the APM galaxies [22]. The dotted curve shows the dark matter correlation function. Results for halos with maximum circular velocity larger than 120 km s$^{-1}$, 150 km s$^{-1}$, and 200 km s$^{-1}$ are presented by the solid, dot-dashed, and dashed curves, respectively. Note that at scales $>0.3h^{-1}$ Mpc the halo correlation function does not depend on the limit in the maximum circular velocity. Top panel: Dependence of bias on scale and maximum circular velocity. The curve labeling is the same as in the bottom panel, except that the dotted curve now represents the bias of halos with $V_{max} > 100$ km s$^{-1}$. From Colin et al. [11].]
big surveys now underway appear to be in better agreement with these ΛCDM predictions\(^\text{49-50}\), this is an important issue that is being investigated in detail\(^\text{51}\).

The problem just mentioned of accounting for baryonic infall is just one example of the hydrodynamical phenomena that must be taken into account in order to make realistic predictions of galaxy properties in cosmological theories. Unfortunately, the crucial processes of especially star formation and supernova feedback are not yet well enough understood to allow reliable calculations. Therefore, rather than trying to understand galaxy formation from full-scale hydrodynamic simulations (for example\(^\text{52}\), more progress has been made via the simpler approach of semi-analytic modelling of galaxy formation (initiated by White and Frenk\(^\text{53-55}\), recently reviewed and extended by Rachel Somerville and me\(^\text{56}\)). The computational efficiency of SAMs permits detailed exploration of the effects of the cosmological parameters, as well as the parameters that control star formation and supernova feedback. We have shown\(^\text{56}\) that both flat and open CDM-type models with \(\Omega_m = 0.3-0.5\) predict galaxy luminosity functions and Tully-Fisher relations that are in good agreement with observations. Including the effects of (proto-)galaxy interactions at high redshift in SAMs allows us to account for the observed properties of high-redshift galaxies, but only for \(\Omega_m \approx 0.3-0.5\)\(^\text{57}\). Models with \(\Omega_m = 1\) and realistic power spectra produce far too few galaxies at high redshift, essentially because of the fluctuation growth rate argument mentioned above.

In order to tell whether ΛCDM accounts in detail for galaxy properties, it is essential to model the dark halos accurately. The Navarro-Frenk-White (NFW)\(^\text{58}\) density profile \(\rho_{\text{NFW}}(r) \propto r^{-1}(r+r_s)^{-2}\) is a good representation of typical dark matter halos of galactic mass, except possibly in their very centers (§4). Comparing simulations of the same halo with numbers of particles ranging from \(\sim 10^3\) to \(\sim 10^6\), my colleagues and I have also shown\(^\text{59}\) that \(r_s\), the radius where the log-slope is -2, can be determined accurately for halos with as few as \(\sim 10^3\) particles. Based on a study of thousands of halos at many redshifts in an Adaptive Refinement Tree (ART)\(^\text{60}\) simulation of the ΛCDM cosmology, we\(^\text{61}\) found that the concentration \(c_{\text{vir}} \equiv R_{\text{vir}}/r_s\) has a log-normal distribution, with \(1\sigma \Delta(\log c_{\text{vir}}) = 0.14\) at a given mass\(^\text{62-63}\). This scatter in concentration results in a scatter in maximum rotation velocities of \(\Delta V_{\text{max}}/V_{\text{max}} = 0.12\); thus the distribution of halo concentrations has as
large an effect on galaxy rotation curves shapes as the well-known log-normal distribution of halo spin parameters $\lambda$. Frank van den Bosch showed, based on a semi-analytic model for galaxy formation including the NFW profile and supernova feedback, that the spread in $\lambda$ mainly results in movement along the Tully-Fisher line, while the spread in concentration results in dispersion perpendicular to the Tully-Fisher relation. Remarkably, he found that the dispersion in $\Lambda$CDM halo concentrations produces a Tully-Fisher dispersion that is consistent with the observed one.

4 Halo Centers

Already in the early 1990s, high resolution simulations of individual galaxy halos in CDM were finding $\rho(r) \sim r^{-\alpha}$ with $\alpha \approx 1$. This behavior implies that the rotation velocity at the centers of galaxies should increase as $r^{1/2}$, but the data, especially that on dark-matter-dominated dwarf galaxies, instead showed a linear increase with radius, corresponding to roughly constant density in the centers of galaxies. This disagreement of theory with data led to concern that CDM might be in serious trouble.

Subsequently, NFW found that halos in all variants of CDM are well fit by the $\rho_{\text{NFW}}(r)$ given above, while Moore’s group proposed an alternative $\rho_M(r) \propto r^{-3/2}(r + r_M)^{-3/2}$ based on a small number of very-high-resolution simulations of individual halos. Klypin and collaborators (including me) initially claimed that typical CDM halos have shallow inner profiles with $\alpha \approx 0.2$, but we subsequently realized that the convergence tests that we had performed on these simulations were inadequate. We now have simulated a small number of galaxy-size halos with very high resolution, and find that they range between $\rho_{\text{NFW}}$ and $\rho_M$. Actually, these two analytic density profiles are almost indistinguishable unless galaxies are probed at scales

Actually, this was the case with the dispersion in concentration $\Delta(\log c_{\text{vir}}) = 0.1$ found for relaxed halos by Jing, while we found the larger dispersion mentioned above. However Risa Wechsler, in her dissertation research with me, found that the dispersion in the concentration at fixed mass of the halos that have not had a major merger since redshift $z = 2$ (and could thus host a spiral galaxy) is consistent with that found by Jing. We also found that the median and dispersion of halo concentration as a function of mass and redshift are explained by the spread in halo mass accretion histories.
smaller than about 1 kpc, which is difficult but sometimes possible.

Meanwhile, the observational situation is improving. The rotation curves of dark matter dominated low surface brightness (LSB) galaxies were measured with radio telescopes during the 1990s, and the rotation velocity was typically found to rise linearly at their centers \cite{70,71,72}. But a group led by van den Bosch \cite{73} showed that in many cases the large beam size of the radio telescopes did not adequately resolve the inner parts of the rotation curves, and they concluded that after correcting for beam smearing the data are on the whole consistent with expectations from CDM. Similar conclusions were

\begin{figure}[h]
\centering
\includegraphics[width=0.5\textwidth]{rotation_curves.png}
\caption{High resolution H\alpha rotation curves (filled circles, solid lines) and HI rotation curves for the same galaxies (open circles, dotted lines) from Ref. 64. The horizontal bar shows the FWHM beam size of the HI observations. From Swaters, Madore, and Trewhella \cite{74}.}
\end{figure}
reached for dwarf galaxies [4]. Swaters and collaborators showed that optical (Hα) rotation curves of some of the LSB galaxies rose significantly faster than the radio (HI) data on these same galaxies [5] (see Fig. 2), and these rotation curves (except for F568-3) appear to be more consistent with NFW [6]. At a conference in March 2000 at the Institute for Theoretical Physics in Santa Barbara, Swaters also showed a Hα rotation curve for the nearby dwarf galaxy DDO154, which had long been considered to be a problem for CDM [65, 66]; but the new, higher-resolution data appeared consistent with an inner density profile $\alpha \approx 1$.

Very recently, a large set of high-resolution optical rotation curves has been analyzed for LSB galaxies, including many new observations [7]. The first conclusion that I reach in looking at the density profiles presented is that the NFW profile often appears to be a good fit down to about 1 kpc. However, some of these galaxies appear to have shallower density profiles at smaller radii. Of the 48 cases presented (representing 47 galaxies, since two different data sets are shown for F568-3), in a quarter of the cases the data do not probe inside 1 kpc, and in many of the remaining cases the resolution is not really adequate for definite conclusions, or the interpretation is complicated by the fact that the galaxies are nearly edge-on. Of the dozen cases where the inner profile is adequately probed, about half appear to be roughly consistent with the cuspy NFW profile (with fit $\alpha \gtrsim 0.5$), while half are shallower. This is not necessarily inconsistent with CDM, since observational biases such as seeing and slight misalignment of the slit lead to shallower profiles [8]. Perhaps it is significant that the cases where the innermost data points have the smallest errors are cusplier.

I think that this data set may be consistent with an inner density profile $\alpha \sim 1$ but probably not steeper, so it is definitely inconsistent with the claims of the Moore group that $\alpha \gtrsim 1.5$. But very recent work by Navarro and collaborators [9] has shown that Moore’s simulations did not have adequate resolution to support their claimed steep central cusp; the highest-resolution simulations appear to be consistent with NFW, or even shallower with $\alpha \approx 0.75$. Further simulations and observations, including measurement of CO rotation curves [10], may help to clarify the nature of the dark matter.

---

5Swaters (private communication) and Hoffman have subsequently confirmed this with better data, which they are preparing for publication.
It is something of a scandal that, after all these years of simulating dark matter halos, we still do not have a quantative — or even a qualitative — theory explaining their radial density profiles. In her dissertation research \cite{63}, Risa Wechsler found that the central density profile and the value of $r_s$ are typically established during the early, rapidly merging phase of halo evolution, and that, during the usually slower mass accretion afterward, $r_s$ changes little. The mass added on the halo periphery increases $R_{\text{vir}}$, and thus the concentration $c_{\text{vir}} \equiv R_{\text{vir}}/r_s$. Now we want to understand this analytically. Earlier attempts to model the result of sequences of mergers (e.g., \cite{81, 82}) led to density profiles that depend strongly on the power spectrum of initial fluctuations, in conflict with simulations (e.g. \cite{83}). Perhaps it will be possible to improve on the simple analytic model of mass loss due to tidal stripping during satellite inspiral that we presented in \cite{101}. Avishai Dekel and his students have recently shown that including the tidal puffing up of the inspiralling satellite before tidal stripping can perhaps account for the origin of the cusp seen in dissipationless simulations, independent of the power spectrum. They argue that the profile must be steeper than $\alpha = 1$ as long as enough satellites make it into the halo inner regions, simply because for flatter profiles the tidal force causes dilation rather than stripping. The proper modeling of the puffing and stripping in the merger process of CDM halos may also provide a theoretical framework for understanding the observed flat cores as a result of gas processes; work on this by Ari Maller and Dekel is in progress. Reionization and feedback into the baryonic component of small satellites would make their cores puff up before merging. This could cause them to be torn apart before they penetrate into the halo centers, and thus allow $\alpha < 1$ cores.

Another possible explanation for flatter central density profiles involving the baryonic component in galaxies has recently been proposed \cite{84}, in which the baryons form a bar that transfers angular momentum into the inner parts of the halo. It is not clear, however, that this effect could be very important in dark matter dominated dwarf and LSB galaxies that have small or nonexistent bulge components.

It would be interesting to see whether CDM can give a consistent account of the distribution of matter near the centers of big galaxies, but this is not easy to test. One might think that big bright galaxies like the Milky Way could help to test the predicted CDM profile, but the centers of such galaxies
are dominated by ordinary matter (stars) rather than dark matter.\footnote{Navarro and Steinmetz had claimed that the Milky Way is inconsistent with the NFW profile \cite{Navarro1995}, but they have now shown that $\Lambda$CDM simulations with a proper fluctuation spectrum are actually consistent with the data \cite{Navarro2011}.}

\section{Too Much Substructure?}

Another concern is that there are more dark halos in CDM simulations with circular velocity $V_c \lesssim 30$ km s$^{-1}$ than there are low-$V_c$ galaxies in the Local Group \cite{Bullock1998,Rubin1980}. A natural solution to this problem was proposed by Bullock et al. \cite{Bullock1999}, who pointed out that gas will not be able to cool in $V_c \lesssim 30$ km s$^{-1}$ dark matter halos that collapse after the epoch of reionization, which occurred perhaps at redshift $z_{\text{reion}} \approx 6$ \cite{Madau1997}. When this is taken into account, the predicted number of small satellite galaxies in the Local Group is in good agreement with observations \cite{Bullock1999,Bullock1999b}. It is important to develop and test this idea further, and this is being done by James Bullock and by Rachel Somerville and their collaborators; the results to date (e.g. \cite{Somerville2003}) look rather promising.

Other groups (e.g. \cite{Vogelsberger2010,Diemand2010}) now agree that astrophysical effects will keep most of the subhalos dark. As a result, theories such as warm dark matter (WDM), which solve the supposed problem of too many satellites by decreasing the amount of small scale power, may end up predicting too few satellites when reionization and other astrophysical effects are taken into account \cite{Vogelsberger2010}.

The fact that high-resolution CDM simulations of galaxy-mass halos are full of subhalos has also led to concerns that all this substructure could prevent the resulting astrophysical objects from looking like actual galaxies \cite{Bullock2000}. In particular, it is known that interaction with massive satellites can thicken or damage the thin stellar disks that are characteristic of spiral galaxies, after the disks have formed by dissipative gas processes. However, detailed simulations \cite{Diemand2007,Diemand2008} have shown that simpler calculations \cite{Somerville2004} had overestimated the extent to which small satellites could damage galactic disks. Only interaction with large satellites like the Large Magellanic Cloud could do serious damage. But the number of LMC-size and larger satellites is in good agreement with the number of predicted halos \cite{Bullock2000}, which suggests that preventing disk damage will not lead to a separate constraint on halo substructure.
6 Angular Momentum Problems

As part of James Bullock’s dissertation research, we found that the distribution of specific angular momentum in dark matter halos has a universal profile \(^1\). But if the baryons have the same angular momentum distribution as the dark matter, this implies that there is too much baryonic material with low angular momentum to form the observed rotationally supported exponential disks \(^2\). It has long been assumed (e.g. \(^3\), \(^4\)) that the baryons and dark matter in a halo start with a similar distribution, based on the idea that angular momentum arising from large-scale tidal torques will be similar across the entire halo. But as my colleagues and I argued recently, a key implication of our new picture of angular momentum growth by merging \(^5\) is that the DM and baryons will get different angular momentum distributions. For example, the lower density gas will be stripped by pressure and tidal forces from infalling satellites, and in big mergers the gaseous disks will partly become tidal tails. Feedback is also likely to play an important role, and Maller and Dekel (in preparation) have shown using a simple model that this can account for data on the angular momentum distribution in low surface brightness galaxies \(^6\).

A related concern is that high-resolution hydrodynamical simulations of galaxy formation lead to disks that are much too small, evidently because formation of baryonic substructure leads to too much transfer of angular momentum and energy from the baryons to the dark matter \(^7\). But if gas cooling is inhibited in the early universe, more realistic disks form \(^8\), more so in \(\Lambda\)CDM than in \(\Omega_m = 1\) CDM \(^9\). Hydrodynamical simulations also indicate that this disk angular momentum problem may be resolved if small scale power is suppressed because the dark matter is warm rather than cold \(^10\), which I discuss next.

7 Alternatives to \(\Lambda\)CDM?

Because of the concerns just mentioned that CDM may predict higher densities and more substructure on small scales than is observed, many people have proposed alternatives to CDM. Two of these ideas that have been studied in the greatest detail are self-interacting dark matter (SIDM) \(^11\) and warm dark matter (WDM).

Cold dark matter assumes that the dark matter particles have only weak
interactions with each other and with other particles. SIDM assumes that the dark matter particles have strong elastic scattering cross sections, but negligible annihilation or dissipation. The hope was that SIDM might suppress the formation of the dense central regions of dark matter halos, although the large cross sections might also lead to high thermal conductivity which drains energy from halo centers and could lead to core collapse \cite{109}, and which also causes evaporation of galaxy halos in clusters, resulting in violation of the observed “fundamental plane” correlations \cite{110}. But in any case, self-interaction cross sections large enough to have a significant effect on the centers of galaxy-mass halos will make the centers of galaxy clusters more spherical \cite{111,112} and perhaps also less dense \cite{113,114} than gravitational lensing observations \cite{115} indicate.

Warm dark matter arises in particle physics theories in which the dark matter particles have relatively high thermal velocities, for example because their mass is \( \leq 1 \text{ keV} \) \cite{116}, comparable to the temperature about a year after the Big Bang when the horizon first encompassed the amount of dark matter in a large galaxy. Such a velocity distribution can suppress the formation of structure on small scales. Indeed, this leads to constraints on how low the WDM particle mass can be. From the requirement that there is enough small-scale power in the linear power spectrum to reproduce the observed properties of the Ly\( \alpha \) forest in quasar spectra, it follows that this mass must exceed about 0.75 keV \cite{117}. The requirement that there be enough small halos to host early galaxies to produce the floor in metallicity observed in the Ly\( \alpha \) forest systems, and early galaxies and quasars to reionize the universe, probably implies a stronger lower limit on the WDM mass of at least 1 keV \cite{118}. Simulations \cite{119,120} do show that there will be far fewer small satellite halos with AWDM than \( \Lambda \text{CDM} \). However, as I have already mentioned, inclusion of the effects of reionization may make the observed numbers of satellite galaxies consistent with the predictions of \( \Lambda \text{CDM} \), in which case AWDM may predict too few small satellite galaxies. Lensing can be used to look for these subhalos \cite{121,122} and may already indicate that there are more of them than expected in AWDM \cite{123}. Thus it appears likely that WDM does not solve all the problems it was invoked to solve, and may create new problems. Moreover, even with an initial power spectrum truncated on small scales, simulations appear to indicate that dark matter halos nevertheless have density profiles
much like those in CDM \cite{124, 28, 89} (although doubts have been expressed about the reliability of such simulations because of numerical relaxation \cite{125}). But WDM does lead to lower concentration halos in better agreement with observed rotation velocity curves \cite{126, 127}.

One theoretical direction that does appear very much worth investigating is $\Lambda$CDM with a tilt $n \sim 0.9$ in the primordial power spectrum $P_p(k) \propto k^n$ \cite{128}. Such $\Lambda$CDM cosmology is favored by recent measurements of the power spectrum of the Ly $\alpha$ forest \cite{36} and appears to be consistent with the latest CMB measurements and all other available data \cite{129}. Our simple analytic model \cite{61} predicts that the concentration of halos in $\Lambda$CDM will be approximately half that in LCDM, which appears to be true in a trial simulation by A. Kravtsov. While this does not resolve the cusp problem, it is a step in the right direction which may lessen the conflict with galaxy rotation curves.

8 \hspace{1em} Outlook

The successes of the CDM paradigm are remarkable. Except possibly for the density profiles at the centers of dwarf and low surface brightness galaxies, the predictions of $\Lambda$CDM appear to be in good agreement with the available observations. The disagreements between predictions and data at galaxy centers appear to occur on smaller scales than was once thought, but as the data improve it is possible that the discrepancies on $\leq$1 kpc scales may ultimately show that CDM cannot be the correct theory of structure formation. However, it appears to be better than any alternative theory that has so far been studied, even though these alternative theories have more adjustable parameters.

This article started by discussing the analogy between the effort to understand dark matter and structure formation in modern cosmology and the effort to understand particle physics in the 1960s and 1970s. In both cases, the result was a “standard model” which has guided further work and led to great progress in both theory and observation/experiment. But in both cases, the standard model is not an ultimate theory, and the search is on for a better theory. In the case of particle physics, there is a leading candidate: supersymmetry, and perhaps ultimately string or M theory. Here the analogy fails, because I am not aware of any theory that has all the virtues of CDM but which avoids its possible failure at the centers of galaxies. The quest for such a theory is a worthwhile goal. But for many purposes, including studies of the
formation and evolution of galaxies and their large scale distribution, the CDM standard model may still remain very useful. And maybe it is even true.

9 Acknowledgements

I learned a great deal about the topics discussed here from my collaborators James Bullock, Avishai Dekel, Ricardo Flores, Anatoly Klypin, Andrey Kravtsov, Ari Maller, Rachel Somerville, and Risa Wechsler, and I also thank Frank van den Bosch, Volker Springel, Rob Swaters, and Simon White for very helpful discussions about the centers of galaxies. This work was supported by grants from NASA and NSF at UCSC, and by a Humboldt Award which supported my visits in fall 2001 to the Max Planck Institutes for Physics in Munich and Astrophysics in Garching, where I thank Leo Stodolsky and Simon White for hospitality. I also thank Aldo Morselli for inviting me to give these lectures at the ISSS, and for his patience waiting for me to send him these notes.

References


51. M.A. Pahre et al. (MALIGN survey), AAS Meeting 197, #96.01 (2000).


