New dark matter physics: Clues from halo structure

Craig J. Hogan and Julianne J. Dalcanton

Astronomy Department, University of Washington, Seattle, Washington 98195-1580

(Received 17 February 2000; published 21 August 2000)

We examine the effect of primordial dark matter velocity dispersion and/or particle self-interactions on the structure and stability of galaxy halos, especially with respect to the formation of substructure and central density cusps. Primordial velocity dispersion is characterized by a ''phase density'' $Q \equiv \rho/\langle v^2 \rangle^{3/2}$, which for relativistically decoupled relics is determined by particle mass and spin and is insensitive to cosmological parameters. Finite Q leads to small-scale filtering of the primordial power spectrum, which reduces substructure, and limits the maximum central density of halos, which eliminates central cusps. The relationship between Q and halo observables is estimated. The primordial Q may be preserved in the cores of halos and if so leads to a predicted relation, closely analogous to that in degenerate dwarf stars, between the central density and velocity dispersion. Classical polytrope solutions are used to model the structure of halos of collisional dark matter, and to show that self-interactions in halos today are probably not significant because they destabilize halo cores via heat conduction. Constraints on masses and self-interactions of dark matter particles are estimated from halo stability and other considerations.

PACS number(s): 95.35.+d

I. HOW COLD AND HOW COLLISIONLESS IS THE DARK MATTER?

The successful concordance of predictions and observations of large scale structure and microwave anisotropy vindicates many assumptions of standard cosmology, in particular the hypothesis that the dark matter is composed of primordial particles which are cold and collisionless [1]. At the same time, there are hints of discrepancies observed in the small-scale structure within galaxy haloes, which we explore as two related but separate issues, namely the predictions of excessive substructure and sharp central cusps in dark matter halos.

The first ''substructure problem'' is that cold dark matter (CDM) predicts excessive relic substructure [2,3]: much of the mass of a CDM halo is not smoothly distributed but is concentrated in many massive sublumps, like galaxies in galaxy clusters. The model predicts that galaxy halos should contain many dwarf galaxies which are not seen, and which would disrupt disks even if they are invisible. The substructure problem appears to be caused by the ''bottom-up'' hierarchical clustering predicted by CDM power spectra; fluctuations on small scales collapse early and survive as dense condensations. Its absence hints that the small scale power spectrum is filtered to suppress early collapse on subgalactic scales.

The second ''cusp problem'' is that CDM also predicts [4–9] a universal, monotonic increase of density towards the center of halos which is not seen in close studies of dark-matter-dominated galaxies [10–12] (although the observational issue is far from settled [14,15]). The formation of central cusps has been observed for many years in simulations of collapse of cold matter in a wide variety of circumstances; it may be thought of as low-entropy material sinking to the center during halo formation. Simulations suggest that dynamical ''pre-heating'' of CDM by hierarchical clustering is not enough to prevent a cusp from forming—that some material is always left with a low entropy and sinks to the center. If this is right, the central structure of halos might provide clues to the primordial entropy which are insensitive to complicated details of nonlinear collapse.

It may be possible to explain these discrepancies in a CDM framework [16], for example by using various baryonic contrivances. It is also possible that the observations can be interpreted more sympathetically for CDM; we explore this possibility in more detail in a separate paper [15]. However it is also possible that the problems with halo structure are giving specific quantitative clues about new properties of the dark matter particles. By examining halo structure and stability, in this paper we make a quantitative assessment of the effects of modifications of the two main properties of CDM—the addition of primordial velocity dispersion, and/or the addition of particle self-interactions. In particular we focus on aspects of halo structure which provide the cleanest ''laboratories'' for studying dark matter properties. The ultimate goal of this exercise is to measure and constrain particle properties from halo structure.

Endowing the particles with non-zero primordial velocity dispersion produces two separate effects: a filter in the primordial power spectrum which limits small-scale substructure, and a phase packing or Liouville limit which produces halo cores. Both effects depend on the same quantity, the ''phase density'' which we choose to define using the most observationally accessible units, $Q \equiv \rho/\langle v^2 \rangle^{3/2}$, where ρ is the density and $\langle v^2 \rangle$ is the velocity dispersion. The definitions of these quantities depend on whether we are discussing fine-grained or coarse-grained Q.[1] For collisionless particles, the fine-grained Q does not change but the coarse-grained Q can decrease as the sheet occupied by particles folds up in phase space. The coarse-grained Q can be estimated directly from astronomical observations, while the fine-grained Q relates directly to microphysics of dark matter particles. For particles which decouple when still relativistic, the initial microscopic phase density Q_0, which for nondis-

---

[1] For a uniform monatomic ideal gas, Q is related in a straightforward way to the usual thermodynamic entropy; for N particles, $S = -kN[\ln(Q) + \text{const}]$.

sipationless collisionless particles is the maximum value for all time, can be related to the particle mass and type, with little reference to the cosmology. The most familiar examples are the standard neutrinos, but we include in our discussion the more general case which yields different numerical factors for bosons and for particles with a significant chemical potential.

The physics of both the filtering and the phase packing in the collisionless case closely parallels that of massive neutrinos [17,18], the standard form of ''hot'' dark matter. Dominant hot dark matter overdoes both of these effects—the filtering scale is too large to agree with observations of galaxy formation (both in emission and quasar absorption) and the phase density is too low to agree with observations of giant-galaxy halos [19]. However one can introduce new particles with a lower velocity dispersion (''warm'' dark matter [20–25,13]), which is the option we consider here.

Although warm dark matter has most often been invoked as a solution to fixing apparent (and no longer problematic [26,1]) difficulties with predictions of the CDM power spectrum for matching galaxy clustering data, a spectrum filtered on smaller scales may also solve several other classic problems of CDM on galactic and subgalactic scales [27,28] which are sometimes attributed to baryonic effects. The main effect in warm models is that the first nonlinear objects are larger and form later, suppressing substructure and increasing the angular momentum of galaxies [29]. This improves the predictions for dwarf galaxy populations [30], baryon-to-dark-matter ratio, disk size and angular momentum, and quiet flows on the scale of galaxy groups. If the filtering is confined to small scales the predictions are likely to remain acceptable for Lyman-α absorption during the epoch of galaxy formation at $z \approx 3$ [31–33].

Liouville's theorem tells us that dissipationless, collisionless particles can only decrease their coarse-grained phase density, and we conjecture that halo cores on small scales approximately preserve the primordial phase density. The universal character of the phase density allows us to make definite predictions for the scaling of core density and core radius with halo velocity dispersion. These relations are analogous to those governing nonrelativistic degenerate-dwarf stars: more tightly bound (i.e. massive) halos should have smaller, denser cores. A survey [15] of available evidence on the phase density of dark matter cores on scales from dwarf spheroidal galaxies to galaxy clusters shows that the phase density needed to create the cores of rotating dwarf galaxies is much lower than that apparently present in dwarf spheroidal galaxies [34–40]—so at least one of these populations is not probing primordial phase density. Translating into masses of neutrino-like relics, the spheroidals prefer masses of about 1 keV (unless the observed stars occupy only a small central portion of an implausibly large, massive and high-dispersion halo), and the disks prefer about 200 eV.

The larger phase density is also preferred from the point of view of filtering. If we take $\Omega \approx 0.3$ (instead of 1 as in most of the original warm scenarios—which reduces the scale for a given mass, because it lowers the temperature and therefore the number of the particles), the filtering scale for 1 keV particles is at about $k = 3$ Mpc^{-1}—small enough to pre-

serve the successful large-scale predictions of CDM but also large enough to impact the substructure problem. Galaxy halo substructure therefore favors a primordial phase density corresponding to collisionless thermal relics with a mass of around 1 keV. In this scenario the densest dwarf spheroidals might well preserve the primordial phase density and in principle could allow a measurement of the particle mass.[2] (Conversely, a mass as large as 1 keV can only solve the core problem in disks with additional nonlinear dynamical heating, so that the central matter no longer remains on the lowest adiabat, or with the aid of baryonic effects [16].)

To have the right mean density and phase density today, relativistically-decoupling particles of this phase density must have separated out at least as early as the QCD era, when the number of degrees of freedom was much larger than at classical weak decoupling. Their interactions with normal standard model particles must therefore be ''weaker than weak,'' ruling out not only standard neutrinos but many other particle candidates. The leading CDM particle candidates, such as weakly interacting massive particles (WIMPs) and axions, form in standard scenarios with much higher phase densities, although more elaborate mechanisms are possible to endow these particles with the velocities to dilute Q. We review briefly some of the available options for making low-Q candidates, such as particles decaying out of equilibrium.

A new wrinkle on this story comes if we endow the particles with self-interactions [41–46]. We consider a simple parametrized model of particle self-interactions based on massive intermediate particles of adjustable mass and coupling, and explore the constraints on these parameters from halo structure. Self-interactions change the filtering of the power spectrum early on, and if they are strong enough they qualitatively change the global structure and stability of halos.

In the interacting case, linear perturbations below the Jeans scale oscillate as sound waves instead of damping by free streaming—analogous to a baryon plasma rather than a neutrino gas. This effect introduces a filter which is sharper in k than that from streaming, and also on a scale about ten times smaller than the streaming for the same rms particle velocity—about right to reconcile the appropriate filtering scale with the Q needed for phase-density-limited disk cores. These self-interactions could be so weak that the particles are effectively collisionless today as in standard CDM.

On the other hand stronger self-interactions have major effects during the nonlinear stages of structure formation and on the structure of galaxy halos [44]. We consider this possibility in some detail, using Lane-Emden polytropes as fiducial models for collisional halos. Their structures are close analogs of degenerate dwarf stars and we call them ''giant dwarfs.'' We find that these structures are subject to an in-

---

[2]This raises another unresolved issue: whether the filtering actually prevents systems as small as dwarf spheroidals from forming at all. The predictions of warm dark models are not yet worked out enough to answer this question.

stability caused by heat conduction by particle diffusion.[3] Although a little of this might be interesting (e.g. leading to the formation of central black holes [49] or to high-density, dwarf spheroidal galaxies), typical halos can only be significantly collisional if they last for a Hubble time; for this to be the case, the particle interactions must be so strong that diffusion is suppressed, which in turn requires a fluid behavior for all bound dark matter structures. This option is not very attractive from a phenomenological point of view [42,44]; for example, dwarf galaxies or galaxies in clusters tend to sink like rocks instead of orbiting like satellites, and the collapse of cores occurs most easily in those low-dispersion halos where we seek to stabilize them.

II. PARTICLE PROPERTIES AND PHASE DENSITIES

We adopt the hypothesis that some dark matter cores are real and due to dark matter rather than baryonic physics or observational artifacts. At present this interpretation is suggested rather than proven by observations. We also conjecture that the heating which sets the finite central phase density is primordial, part of the physics of the particle creation rather than some byproduct of hierarchical clustering. At present this is a conjecture suggested rather than proven by simulations.

In the clustering hierarchy, more higher-entropy material is created as time goes on, but numerical experiments indicate that this heated material tends to end up in the outer halo. This is the basic reason why CDM halos always have central cusps: there is always a little bit of material which remembers the low primordial entropy and sinks to the center. The halo center contains the lowest-entropy material, which we conjecture is a relic of the original entropy of the particles—or equivalently, their original phase density, which is most directly related to measurable properties of halo dynamics. We begin by relating the phase density to particle properties in some simple models.

A. Phase density of relativistically-decoupled relics

Consider particles of mass m originating in equilibrium and decoupling at a temperature $T_D \gg m$ or chemical potential $\mu \gg m$. The original distribution function is [51]

$$f(\vec{p}) = (e^{(E-\mu)/T_D} \pm 1)^{-1} \approx (e^{(p-\mu)/T_D} \pm 1)^{-1} \qquad (1)$$

with $E^2 = p^2 + m^2$ and $\pm$ applies to fermions and bosons respectively. The number density and pressure of the particles are [52]

$$n = \frac{g}{(2\pi)^3} \int f d^3 p \qquad (2)$$

$$P = \frac{g}{(2\pi)^3} \int \frac{p^2}{3E} f d^3 p \qquad (3)$$

where g is the number of spin degrees of freedom. Unless stated otherwise, we adopt units with $\hbar = c = 1$.

With adiabatic expansion this distribution is preserved with momenta of particles varying as $p \propto R^{-1}$, so the density and pressure can be calculated at any subsequent time [53]. For thermal relics $\mu = 0$, we can derive the density and pressure in the limit when the particles have cooled to be nonrelativistic:

$$n = \frac{g T_0^3}{(2\pi)^3} \int \frac{d^3 p}{e^p \pm 1} \qquad (4)$$

$$P = \frac{g T_0^5}{(2\pi)^3 3m} \int \frac{p^2 d^3 p}{e^p \pm 1} \qquad (5)$$

where the pseudo-temperature $T_0 = T_D (R_D / R_0)$ records the expansion of any fluid element relative to its initial size and temperature at decoupling R_D, T_D.

It is useful to define a ''phase density'' $Q \equiv \rho / \langle v^2 \rangle^{3/2}$ proportional to the inverse specific entropy for nonrelativistic matter, which is preserved under adiabatic expansion and contraction. For nondissipative particles Q cannot increase, although it can decrease due to shocks (in the collisional case) or coarse-graining (in the collisionless case, e.g. in ''violent relaxation'' and other forms of dynamical heating). Combining the above expressions for density and pressure and using $\langle v^2 \rangle = 3P/nm$, we find

$$Q_X = q_X g_X m_X^4. \qquad (6)$$

The dimensionless coefficient for the thermal case is

$$q_T = \frac{4\pi}{(2\pi)^3} \frac{\left[\int dp (p^2/e^p \pm 1) \right]^{5/2}}{\left[\int dp (p^4/e^p \pm 1) \right]^{3/2}} = 0.0019625, \qquad (7)$$

where the last equality holds for thermal fermions. An analogous calculation for the degenerate fermion case ($T=0, \mu_D \gg m_X$) yields the same expression for Q but with a different coefficient,

$$q_d = \frac{4\pi}{(2\pi)^3} \frac{\left[\int_0^1 p^2 dp \right]^{5/2}}{\left[\int_0^1 p^4 dp \right]^{3/2}} = 0.036335. \qquad (8)$$

To translate from $\hbar = c = 1$ into more conventional astronomers' units,

---

[3]Degenerate dwarf stars are not subject to this instability because they are supported without a temperature gradient; the same stabilization could occur in halo cores only if the dark matter is fermionic and degenerate (e.g., [47,48]). The instability we discuss here is essentially what happens in a thermally-supported star with no nuclear reactions, except that the conduction is by particle diffusion rather than by radiation. This effect may have already been observed numerically [50].

$$(100 \ \mathrm{eV})^4/c^5 = 12808 \frac{(M_\odot/\mathrm{kpc}^3)}{(\mathrm{km\ s}^{-1})^3} = 12.808 \frac{(M_\odot/\mathrm{pc}^3)}{(100 \ \mathrm{km\ s}^{-1})^3}. \tag{9}$$

The phase density in this situation depends on the particle properties but not at all on the cosmology; the decoupling temperature, the current temperature and density do not enter. The numerical factors just depend on whether the particles are thermal or degenerate, bosons or fermions, which makes the quantity Q a potentially precise tool for measuring particle properties. Many scenarios envision thermal relics so we adopt this as a fiducial reference in quoting phase densities in m^4 units—bearing in mind that the actual mass may be different in cases such as degenerate sterile neutrinos [48,49], and that for the astrophysical effects discussed below, it is the phase density that matters. For a neutrino-like ($g=2$), thermal relic,

$$Q_T = 5 \times 10^{-4} \frac{(M_\odot/\mathrm{pc}^3)}{(\mathrm{km\ s}^{-1})^3} (m_X/1 \ \mathrm{keV})^4. \tag{10}$$

B. Space density of thermal relics

For a standard, relativistically-decoupled thermal relic, the mean density of the particles can be estimated [52] from the number of particle degrees of freedom at the epoch T_D of decoupling, g_{*D}; the ratio to the critical density is

$$\Omega_X = 78.3 h^{-2} [g_{eff}/g_{*D}](m_X/1 \ \mathrm{keV})$$
$$= 2.4 h_{70}^{-2} (m_X/1 \ \mathrm{keV})(g_{eff}/1.5)(g_{*D}/100)^{-1} \tag{11}$$

where g_{eff} is the number of effective photon degrees of freedom of the particle ($=1.5$ for a two-component fermion). For standard neutrinos which decouple at around 1 MeV, $g_{*D} = 10.75$.

Current observations suggest that the dark matter density $\Omega_X \approx 0.3$ to 0.5, hence the mass density for a warm relic with $m_X \gtrsim 200$ eV clearly requires a much larger g_{*D} than the standard value for neutrino decoupling. Above about 200 MeV, the activation of the extra gluon and quark degrees of freedom (24 and 15.75 respectively including uds quarks) give $g_{*D} \approx 50$; activation of heavier modes of the standard model above ≈ 200 GeV produces $g_{*d} \approx 100$; this gives a reasonable match for $m_X \approx 200$ eV and $\Omega_X \lesssim 0.5$, as suggested by current evidence. Masses of the order of 1 keV can be accommodated by somewhat earlier decoupling ($\approx$ TeV) and including many extra (e.g., supersymmetric or extra-dimensional) degrees of freedom. Alternatively a degenerate particle can be introduced via mixing of a sterile neutrino, combined with a primordial chemical potential adjusted to give the right density [48]. In any of these cases, the particle must interact with standard model particles much more weakly than normal weak interactions, which decouple at ≈ 1 MeV.

Note that warm dark matter particles have low densities compared with photons and other species at 1 MeV so they do not strongly affect nucleosynthesis. However, their effect is not entirely negligible since they are relativistic at early times and add considerably more to the mean total density in the radiation era than standard CDM particles. They add the equivalent of $(T_X/T_\nu)^3 = 10.75/g_{*d}$ of an effective extra neutrino species, which leads to a small increase in the predicted primordial helium abundance for a given η. Because the phase density fixes the mean density at which the particles become relativistic, it also fixes this effect on nucleosynthesis (independent of the other particle properties, thermal or degenerate etc.) This effect might eventually become detectable with increasingly precise measurements of cosmic abundances.

C. Decaying WIMPs and other particle candidates

Thermally decoupled relics are the simplest way to obtain the required finite phase density, but they are not the only way. Heavier particles can be produced with a kinetic temperature higher than the radiation, accelerated by some non-thermal process. Weakly interacting massive particles, including the favored lightest supersymmetric particles, can reduce their phase density if they form via out-of-equilibrium particle decay. A small density of heavy unstable particles (X1) can separate out in the standard way, then later decay into the present-day (truly stable) dark matter particles (X2). In a supersymmetric scheme one can imagine for example a gravitino separating out and decaying into neutralino dark matter.

In the normal Lee-Weinberg scenario for WIMP generation, the particle density is in approximate thermal equilibrium until $T \approx m_X/20$. The particles thin out by annihilation until their relic density freezes out when the annihilation rate matches the Hubble rate, $n_X \langle \sigma_{ann} v \rangle \approx H$. The density today is then

$$\Omega_X \approx T_{\gamma 0}^3 H_0^{-2} m_{Planck}^{-3} \langle \sigma_{ann} v \rangle^{-1}$$
$$\approx (m_W/100 \ \mathrm{GeV})^2 (m_W/m_X)^2 \tag{12}$$

where we have used the typical weak annihilation cross section $\sigma_{ann} \approx \alpha^2 m_X^2/m_W^4$ determined by the mass of the W. The kinetic temperature of the WIMPs freezes out at about the same time as the abundance, so they are very cold today, with typical velocities $v \approx \sqrt{20 T_0}/m_X \approx 10^{-14} (m_X/100 \ \mathrm{GeV})^{-1}$. This of course endows them with small velocities and an enormous phase density.

A smaller phase density can be produced if these particles decay at some point into the particles present today. If the secondary particles are much lighter than the first, they can be generated with relativistic velocities at relatively late times as we require. Suppose the primary $X1$ particles decay into secondary $X2$ particles at a temperature T_{decay}. To produce particles with the velocity ≈ 0.4 km/sec today (characteristic of a fiducial 200 eV thermal relic phase density), or $v \approx c$ at $T \approx /300$ eV,

$$m_{X1} \approx m_{X2} T_{decay}/300 \ \mathrm{eV}. \tag{13}$$

We also want to get the right density of $X2$ particles. Suppose the density of $X1$ is determined by a Lee-Weinberg freezeout, such that $n_{X1}(T_\gamma \approx m_{X1}/20) \sigma_{ann} v \approx H$. In order to have $\rho_{X2} \approx \rho_{rad}/600$ at $z_{nr} \approx 10^6$, $\rho_{X1} \approx \rho_{rad}/600$ at T_{decay}, and then

$$m_{X2}^2 T_{decay} \approx \frac{600 \times 20 m_W^4}{\alpha^2 m_{Planck}} \approx (100 \text{ MeV})^3. \qquad (14)$$

Thus we obtain

$$m_{X1} m_{X2} \approx (30 \text{ GeV})^2. \qquad (15)$$

A simple example would be a more or less standard 50 GeV WIMP primary which decays at $T_{decay} \approx 1$ keV into marginally relativistic 20 GeV secondaries. Alternatively the primary could be heavier than this and the secondary lighter. Such scenarios have to be crafted to be consistent with various constraints, such as the long required lifetime for $X1$ (in the example just given, a week or so) and the decay width of the Z (which must not notice the existence of $X2$); although not compelling, they are not all ruled out.[4]

The other perennial favorite dark matter candidate is the axion. The usual scenario is to produce these by condensation, which if homogeneous produces dark matter even colder than the WIMPs—indeed, as bosons in a macroscopic coherent state. However, it is natural to contemplate modifications to this picture where the condensing fields are not uniform but have topological defects or Goldstone excitations, produced by the usual Kibble mechanism during symmetry breaking (e.g. [54,55]). In this case the axions are produced with relativistic velocities and could in principle lead to the desired velocity dispersion.

III. CORES FROM FINITE PRIMORDIAL PHASE DENSITY

We have shown several examples of how particle properties determine primordial phase density. Here we explore how the phase density affects the central structure of dark matter halos.

A. Core radius of an isothermal halo

Consider the evolution of classical dissipationless, collisionless particles in phase space. Truly cold dark matter is formed with zero velocity dispersion occupying a three dimensional subspace (determined by the Hubble flow $\vec{v} = H\vec{r}$) of six dimensional phase space. Subsequent nonlinear evolution wraps up the phase sheet so that a coarse-grained average gives a higher entropy and a lower phase density. In general a small amount of cold material remains which naturally sinks to the center of a system. There is in principle no limit to the central density; the phase sheet can pack an arbitrary number of phase wraps into a small volume.

By contrast, with warm dark matter the initial phase sheet has a finite thickness. The particles do not radiate so the phase density can never exceed this initial value. In the non-

---

[4]It is also possible to reduce the scale of filtering of linear perturbations for a given phase density by arranging for the decay relatively late, and for the decay products to be nonrelativistic. This option seems even more contrived and we will not pursue it in detail here.

linear formation of a halo, the phase sheet evolves as an incompressible fluid in phase space. The outer parts of a halo form in the same way as CDM by wraps of the sheet whose thickness is negligible, but in the central parts the finite thickness of the sheet prevents arbitrarily close packing—it reaches a ''phase packing'' limit. For a given velocity dispersion at any point in space, the primordial phase density of particles imposes an upper limit on their density ρ, corresponding to adiabatic compression. Thus warm dark matter halos cannot form the singular central cusps predicted by cold dark matter but instead form cores with a maximum limiting density at small radius, determined by the velocity dispersion.

We estimate the structure of the halo core by conjecturing that the matter in the central parts of the halo lies close to the primordial adiabat defined by Q. This will be a good assumption for cores which form quietly without too much dynamical heating. Simulations indicate this to be the case in essentially all CDM halos, although in principle it could be that warm matter typically experiences more additional dynamical heating than cold matter, in which case the core could be larger. This question can be resolved with warm simulations, including a reasonable sampling of the particle distribution function during nonlinear clustering [56]; for the present we derive a rigorous upper limit to the core density for a given velocity dispersion, and conjecture that this will be close the actual central structure.

A useful model for illustration and fitting is a standard isothermal sphere model for the halo. The spherical case with an isotropic distribution of velocities maximizes the central density compatible with the phase density limit. The conventional definition of core size in an isothermal sphere [57] is the ''King radius''

$$r_0 = \sqrt{9\sigma^2/4\pi G\rho_0} \qquad (16)$$

where σ denotes the one-dimensional velocity dispersion, and ρ denotes the central density. Making the adiabatic assumption, $\rho_0 = Q(3\sigma^2)^{3/2}$, we find

$$r_0 = \sqrt{9\sqrt{2}/4\pi 3^{3/2}} (QGv_{c\infty})^{-1/2} = 0.44(QGv_{c\infty})^{-1/2} \qquad (17)$$

where $v_{c\infty} = \sqrt{2}\sigma$ denotes the asymptotic circular velocity of the halo's flat rotation curve. [Note that aside from numerical factors this is the same mass-radius relation as a degenerate dwarf star; the galaxy core is bigger than a Chandrasekhar dwarf of the same specific binding energy by a factor $(m_{proton}/m_X)^2$. The collisional case treated below is even closer to a scaled version of a degenerate dwarf star.]

For the thermal and degenerate phase densities derived above,

$$r_{0,thermal} = 5.5 \text{ kpc}(m_X/100 \text{ eV})^{-2}(v_{c\infty}/30 \text{ km s}^{-1})^{-1/2} \qquad (18)$$

$$r_{0,degenerate} = 1.3 \text{ kpc}(m_X/100 \text{ eV})^{-2}(v_{c\infty}/30 \text{ km s}^{-1})^{-1/2}, \qquad (19)$$

where we have set $g=2$. The circular velocity in the central core displays the harmonic behavior $v_c \propto r$; it reaches half of its asymptotic value at a radius $r_{1/2} \approx 0.4 r_0$.

Instead of fitting an isothermal sphere to an entire rotation curve, in some situations we might opt to measure the central density directly by fitting the linear inner portion of a rotation curve if it is well-resolved in the core:

$$\begin{aligned}
v_c/r &= \sqrt{4\pi G\rho/3} \\
&= 2.77 G^{1/2} Q^{1/2} v_{c\infty}^{3/2} \\
&= 6.71 \ \text{km s}^{-1} \ \text{kpc}^{-1} (m_X/100 \ \text{eV})^2 \\
&\quad \times (v_{c\infty}/30 \ \text{km s}^{-1})^{3/2}.
\end{aligned} \tag{20}$$

B. Comparison with galaxy and cluster data

In a separate paper [15] we review the current relevant data in more detail, including a consideration of interpretive ambiguities. Here we offer a summary of the situation.

The relationship of core radius or central density with halo velocity dispersion is a simple prediction of the primordial phase density hypothesis, which can in principle be tested on a cosmic population of halos. In particular if phase packing is the explanation of dwarf galaxy cores, the dark matter cores of giant galaxies and galaxy clusters are predicted to be much smaller than for dwarfs, unobservably hidden in a central region dominated by baryons. There is currently at least one well-documented case of a galaxy cluster with a large core (≈ 30 kpc) as measured by a lensing fit [58], which cannot be explained at all by phase packing with primordial phase density. On the other hand more representative samples of relaxed clusters do not show evidence of cores [15,59].

The favorite laboratories for finding evidence of dark matter cores are dwarf disk galaxies which display a central core even at radii where the baryonic contribution is negligible [12,10,11]. Rotation curves allow a direct estimate of the enclosed density as a function of radius, right out to a fairly flat portion which allows an estimate of the dark matter velocity dispersion—all the information we need to estimate a phase density for a core. Three of the best-resolved cases [15] yield estimates of $Q \approx 10^{-7}-10^{-6} (M_\odot/\text{pc}^3)/(\text{km s}^{-1})^3$. The sensitive dependence of Q on particle mass means that m_X is reasonably well bounded even from just a handful of such cases; a thermal value of $m_X \gtrsim 300$ eV does not produce large enough cores to help at all (that is, one must seek unrelated explanations of the data), while values $m_X \lesssim 100$ eV produce such large cores that they conflict with observed rotation curves of normal giant galaxies [19] and LSB galaxies [60]. This is why we adopt a fiducial reference value of 200 eV for dwarf disk cores.

Dwarf spheroidal galaxies do not have gas on circular orbits so their dynamics is studied with stellar velocity dispersions [35–36,40]. Here we have an estimate of the mean density in the volume encompassed by the stellar test particles, but we do not know the velocity dispersion of the dark matter halo particles (which may be larger than that of the stars if the latter occupy only the harmonic central portion of

a large dark matter core) so estimates of the phase density are subject to other assumptions and modeling constraints [37,38]. If we assume that the stars are not much more concentrated than the dark matter, we get the largest estimate[5] of the phase density, which in the largest case [15] is about $Q \approx 2 \times 10^{-4} (M_\odot/\text{pc}^3)/(\text{km s}^{-1})^3$ corresponding to a thermal relic of mass $m_X \approx 800$ eV. The apparent phase densities estimated for dwarf spheroidals are thus much larger than for dwarf disks, even at the same radius. The mass-to-light ratio in the most extreme of these systems is about 100 in solar units, an order of magnitude more than that found for purely baryonic, old stellar populations in elliptical galaxies [61], so there is little doubt that they are dominated by dark matter. The CDM prediction is that there are other, more weakly bound halos in which gas was unable to cool and form stars, and which therefore have an even higher mass-to-light ratio.

IV. FILTERING OF SMALL-SCALE FLUCTUATIONS

The non-zero primordial velocity dispersion naturally leads to a filtering of the primordial power spectrum. The transfer function of warm dark matter is almost the same as cold dark matter on large scales, but is filtered by free-streaming on small scales. The characteristic wave number for filtering at any time is given by $k_X = H/\langle v^2 \rangle^{1/2}$, the inverse distance travelled by a particle at the rms velocity in a Hubble time. The detailed shape of the transfer function depends on the detailed evolution of the Boltzmann equation, and in particular whether the particles are free-streaming or collisional.

In the current application, we are concerned with H during the radiation-dominated era ($z \gtrsim 10^4$), so that $H^2 = 8\pi G\rho_{rel}/3 \propto (1+z)^4$, where ρ_{rel} includes all relativistic degrees of freedom. For constant Q, $\langle v^2 \rangle^{1/2} = (\rho_X/Q)^{1/3} \propto (1+z)$ as long as the X particles are nonrelativistic. For particles with a small velocity dispersion today, the comoving filtering scale [52] is thus approximately independent of redshift over a considerable interval of redshift (see Fig. 1). The ''plateau'' scale is independent of H_0:

$$k_{X,comov} = H_0 \Omega_{rel}^{1/2} v_{X0}^{-1} = 0.65 \ \text{Mpc}^{-1} (v_{X0}/1 \ \text{km s}^{-1})^{-1} \tag{21}$$

where $\Omega_{rel} = 4.3 \times 10^{-5} h^{-2}$ is the density in relativistic species and $v_{X0} = (Q/\bar{\rho}_{X0})^{-1/3}$ is the rms velocity of the particles at their present mean cosmic density $\bar{\rho}_{X0}$. For the thermal case, in terms of particle mass, we have

$$\begin{aligned}
v_{X0,thermal} &= 0.93 \ \text{km s}^{-1} h_{70}^{2/3} (m_X/100 \ \text{eV})^{-4/3} \\
&\quad \times (\Omega_X/0.3)^{1/3} (g/2)^{-1/3},
\end{aligned} \tag{22}$$

---

[5]This is the largest value of the mean phase density of material in the region enclosed by the stellar velocity tracers; there is no real observational upper limit for the maximum phase density. Without the rotation curve information, these systems are consistent with singular isothermal spheres or other cuspy profiles for the dark matter.
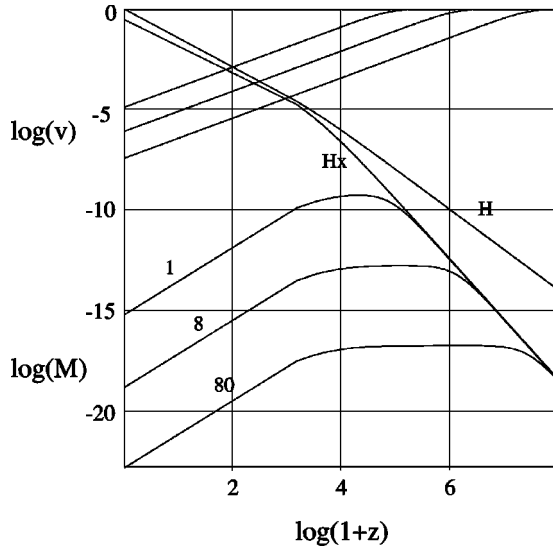
FIG. 1. Characteristic masses and velocities as a function of inverse scale factor $(1+z)$, for a cosmological model with $\Omega_X = 0.3$, $\Lambda = 0.7$. Mass and velocity are plotted in units with $H_0 = \bar{\rho} = c = 1$, or $M = 0.3\rho_{crit}c^3H_0^{-3} = 1.56 \times 10^{21}h_{70}M_\odot$. The total rest mass of dark matter in a volume H^{-3} is denoted by Hx; total mass-energy of all forms in the same volume is denoted by H. Characteristic rms velocities and streaming masses (rest mass of X in a volume k_X^{-3}) are also shown, for dark matter with three different phase densities. The cases plotted correspond to relativistically-decoupled thermal relics decoupling at three different effective degrees of freedom, corresponding to 1, 8, and 80 times that for a single standard massive neutrino—''hot,'' ''warm,'' and ''cool.'' (For $h = 0.7$, the corresponding masses are 13, 108, and 1076 eV respectively, and the rms velocities at the present epoch are 1.3 $\times 10^{-5}$, 7.9×10^{-7}, and 3.6×10^{-8}, respectively.) Note the long flat period with nearly constant comoving k_X for the cool particles, during the period when the universe is radiation-dominated but X is nonrelativistic. The difference between streaming and collisional behavior during this period has a significant effect on the scale of filtering in the transfer function, with a sharper cutoff and a smaller scale (for fixed k_X) in the collisional Jeans limit.

and hence

$$k_{X,comov} = 15 \ \mathrm{Mpc}^{-1}h_{70}^{-2/3}(m_X/1 \ \mathrm{keV})^{-4/3}$$
$$\times (\Omega_X/0.3)^{-1/3}(g/2)^{1/3}. \tag{23}$$

In the case of free-streaming, relativistically-decoupled thermal particles, the transfer function has been computed precisely [21,29]; the characteristic wave number where the square of the transfer function falls to half the CDM value is about $k_{1/2,stream} \approx k_{X,comov}/5.5$. The simple streaming case only works for high phase densities $m_X \gtrsim 1$ keV, that is, comparable to that observed in dwarf spheroidals. For example, to produce an acceptable number of galaxies at a dwarf galaxy scale without invoking disruption, Press-Schechter theory [62] implies a spectral cutoff at about $k = 3h_{70}$ Mpc^{-1}, requiring a thermal relic mass of about 1100 eV. Hydrodynamic simulations show that the same cutoff scale preserves the large scale success of CDM and probably improves the CDM situation on galaxy scales in ways men-

tioned previously [29]. Although the typical uncertainty on the phenomenologically best filtering scale is at least a factor of two, it is clear that the smallest phase density compatible with standard streaming filtering is too large to have a direct impact on the core problem in dwarf disk galaxies.

On the other hand the discrepancy is only a factor of a few in mass, less than an order of magnitude in linear damping scale. We have already mentioned two modifications which could reconcile these scales. It could be that warm models turn out to be sometimes more effective at producing smooth cores than we have guessed from the minimal phase-packing constraint, due to more efficient dynamical heating than CDM; this would produce a nonlinear amplifier of the primordial velocities, probably with a large variation depending on dynamical history (an especially good option if cores turn out to be common in galaxy clusters.) Another possibility is that the primordial velocities are introduced relatively late (nonrelativistically) by particle decay.

Still another possibility is a different relationship of $k_{1/2}$ and k_X from the standard collisionless streaming behavior. For example, if the particles are self-interacting, then the free streaming is suppressed and the relevant scale is the standard Jeans scale dividing growing behavior from acoustic oscillations, $4\pi G\rho_{total} - k_J^2 c_S^2 = 0$. This comes out to $k_J = \sqrt{3}H/c_S = \sqrt{27/5}k_X$, 13 times shorter than $k_{1/2,stream}$ at a fixed phase density. (An intuitive view of this numerical factor is that during the long period when k_X is flat, streaming particles continue to move and damp on larger scales, whereas the comoving Jeans scale just remains fixed, sharply dividing oscillating from growing behavior.) The acoustic case is similar to the behavior of fluctuations in high-density, baryon-dominated models, which have a sharp cutoff at the Jeans scale [26]. We conclude that some particle self-interactions may be desirable to reconcile the scale of the transfer function of primordial perturbations with the phase packing effect on disk cores.

V. COLLISIONAL DARK MATTER

We now turn to the case where the dark matter particles are not collisionless, but scatter off of each other via a new intermediate force. Self-interactions of dark matter have been motivated from both an astrophysical and a particle physics point of view [41–45]. Our goal here is again to relate the properties of the new particles to the potentially observable properties of dark matter halos. In addition to the single parameter Q considered for the collisionless case, we can use halo properties to constrain fundamental parameters of the particles—the masses of the dark matter particles and intermediate bosons carrying the interactions, as well as a coupling constant.

Such self-interactions lead to modifications in several of the previous arguments. As we have seen, self-interactions can have observable effects via the transfer function even if they are negligible today. Stronger self-interactions also affect the structure and stability of halos; collisional matter has a fluid character leading to equilibrium states of self-gravitating halos much like those of stars. These systems are quite different from collisionless systems. Although entropy
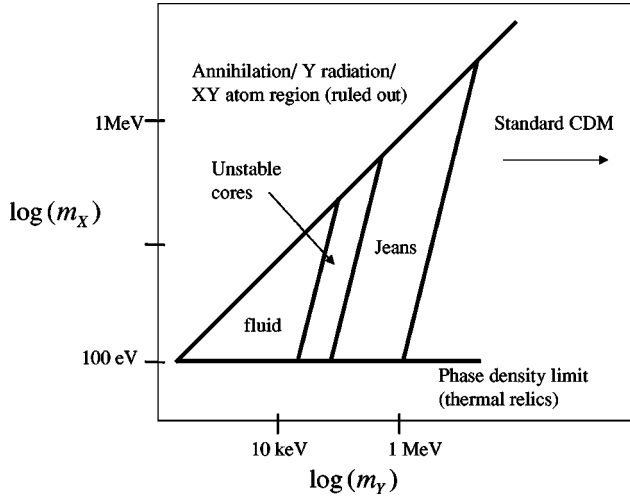
FIG. 2. A sketch of the principal constraints from halo structure arguments on the masses of collisional dark matter particle X and particle mediating its self-interactions, Y. This plot assumes a coupling constant $e=0.1$. The rightmost region is indistinguishable from standard collisionless CDM. The region labled ''Jeans'' is essentially collisionless today, but collisional before t_{eq} and consistent with other constraints; in this regime the particles are no longer free-streaming, and the filtering scale and the shape of the transfer function are significantly modified by self-interactions. Somewhat stronger interactions lead to a conductive instability in halos; the ''unstable cores'' constraint is ruled out if we require stability down to halo velocities of 30 km s^{-1}. The leftmost region (''fluid'') produces halos which are so collisional they are stable against conduction for a Hubble time, but is probably ruled out by the unusual fluid-dynamical behavior this would cause in the trajectories of satellite galaxies and galaxies in clusters. The upper constraint comes from suppression of the annihilation channel (by the inability to radiate Y); if this does not apply (that is, if there no $\bar{X}$ around) then parallel, somewhat higher constraints come from suppressing dissipation by Y radiation, or from the prohibition against bound X atoms. The bottom constraint corresponds to a phase-packing limit for giant galaxies; this last constraint on mass applies for relativistically-decoupled light relics only, and is ten times higher if we use the limit from dwarf spheroidals.

must increase outwards for stability against convection (which naturally happens due to shocks in the hierarchy), it cannot increase too rapidly and remain hydrostatically stable; in particular, stable solutions have a minimum non-negligible temperature gradient, and the isothermal case is no longer a stable static solution as it is for collisionless matter. Since collisional matter conducts heat between fluid elements, these solutions are all unstable on some time scale.

A. Particles and interactions

We now apply several simple physical arguments to constrain properties of the dark matter candidate and its interactions. Some of these have been considered previously [44]. The most important constraints are summarized in Fig. 2.

Suppose that the dark matter X particles with mass m_X, which may be either fermions or bosons, interact via massive bosons Y whose mass m_Y determines the range of the inter-

actions, and a coupling constant e. These may be considered analogous to strong interaction scatterings where we regard pions as Yukawa scalar intermediates, or electroweak interactions with W,Z as vector intermediates. The interactions must be elastic scatterings to avoid a net energy loss, although ''dissipative'' three-body encounters are permitted as long as the energy does not leave the XY subsystem nor travel far in space. For most purposes even the sign of the interaction does not matter—it may be attractive or repulsive, as happens with vectors and like charges. The Y particles at tree level interact only with X, although the X may (as is usual with dark matter candidates) be allowed some much weaker interactions with ordinary matter. In this model the collision cross section for strong scattering is about

$$\sigma \approx m_Y^{-2} \min\left[e^4\left(\frac{m_Y}{m_X v^2}\right)^2, \ e^4\left(\frac{m_X}{m_Y}\right)^2, \ 1\right] \quad (24)$$

where the first case is coupling-limited (and depends on the particle velocity and coupling strength, like electromagnetic scattering of electrons), the second case holds for $m_Y>m_X$ (like neutrino neutral-current interactions) and the third is the range-limited, strong interaction limit (like neutron scattering).

There are several simple constraints on the particle masses. If the dark matter is collisional, the rate of net annihilations of X must be highly suppressed compared to the scattering rate, or the mass of the halo would quickly radiate away as Y particles. Either there is a primordial asymmetry (so the number of $\bar{X}$ is negligible), or

$$m_Y > 2m_X, \quad (25)$$

suppressing what would otherwise be a rapid channel for X to annihilate and radiate Y. (Recall that in this model, there is no direct route to annihilate into anything else.) In any case the Y must not be too light or the typical inelastic collisions will radiate them; for particles with relative velocities $v \approx 10^{-3}$ typical of dark matter in galaxies, we must have

$$m_Y > m_X v^2 \approx 10^{-6} m_X, \quad (26)$$

so that the energy of collisions is typically insufficient to create a real Y. In addition, if attractive, the range of the interactions must be less than the ''Bohr radius'' for these interactions, requiring

$$m_Y > e^2 m_X, \quad (27)$$

in order not to form bound ''atoms.'' The close analogy with Y is the pion, which is just light enough to allow a bound state of deuterium. Bound states would be a disaster since they would behave like nuclear reactions in stars. Such states would add an internal source of energy in the halos, creating winds or other energy flows which would unbind large amounts of matter. All of these constraints eliminate the upper left region of Fig. 2, with details depending on the coupling strength and halo velocity.

B. Parameters for collisional behavior

The properties of interacting particles define a characteristic column density, m_X/σ; a slab of X at this column is one mean free path thick. This is the quantity that specifies the degree of collisional or collisionless behavior of a system. In order to connect the halo astrophysics with dark matter properties we convert from units with $\hbar = c = 1$:

$$(1 \ \text{GeV})^3 = 4.6 \times 10^3 \ \text{g cm}^{-2} = 2.2 \times 10^7 \ \text{M}_\odot \ \text{pc}^{-2}. \tag{28}$$

For comparison, the average mass column density within radius r_{kpc}kpc for a halo with a circular velocity $v_{30} \times 30$ km sec^{-1} is

$$\Sigma_h = \frac{v^2}{\pi G r} = 0.014 \, v_{30}^2 r_{kpc}^{-1} \ \text{g cm}^{-2} = (15 \ \text{MeV})^3 v_{30}^2 r_{kpc}^{-1}. \tag{29}$$

A halo therefore enters the strongly-collisional regime—qualitatively different from classical CDM—if

$$m_Y^4 e^{-4} (15 \ \text{MeV})^{-3} v_{30}^{-2} r_{kpc}$$
$$< m_X < \min[(15 \ \text{MeV})^3 v_{30}^2 r_{kpc}^{-1} m_Y^{-2},$$
$$15 \ \text{MeV}(e/v)^{4/3}(v_{30}^2 r_{kpc}^{-1})^{1/3}]. \tag{30}$$

This criterion is shown in Fig. 2 as the right boundary of the ''unstable cores'' region; indeed this marginally-collisional case maximizes the rate of thermal conduction instability, as discussed below.

We also compute the criterion for non-streaming behavior in the early universe—the amount of self-interaction needed to affect the transfer function as discussed above. It is significantly less than that required for collisional behavior today:

$$\frac{\sigma}{m_X} \approx H(t_{eq})/n_X(t_{eq}) v_X(t_e q) m_X = \Sigma_0^{-1} \Omega_{rel}^{5/2} \Omega_X^1 v_X(t_0), \tag{31}$$

where eq refers to the epoch of equal densities in dark matter and relativistic species, and

$$\Sigma_0 \equiv c \rho_{crit}/H_0 = 0.1213 h_{70}^{-1} \ \text{g cm}^{-2} \tag{32}$$

is the characteristic cosmic column density today. Using the units conversion above we have

$$\frac{\sigma}{m_X} \approx (600 \ \text{MeV})^{-3}(v_{X0}/1 \ \text{km s}^{-1})^{-1}(\Omega_X/0.3)^2 h_{70}^4, \tag{33}$$

corresponding to a mass column for one expected scattering of 2×10^4 g cm^{-2}. Particles scattering off of each other more strongly than this no longer have streaming behavior at high redshift but support acoustic oscillations, much like baryons but with only their own pressure (that is, without the interaction with radiation pressure and without decoupling from it). We should bear in mind that a somewhat larger cross

section is needed to avoid diffusive (''Silk'') damping, but even at this level of interaction the scale of damping is significantly reduced from the streaming case. This criterion is shown in Fig. 2 as the right boundary of the ''Jeans'' region (although some acoustic behavior before t_{eq} occurs even to the right of this).

C. Polytropes

The equilibrium configurations of collisional dark matter correspond to those of classical self-gravitating fluids. The simplest cases to consider and general enough for our level of precision are classical polytrope solutions—stable configurations of a classical, self-gravitating, ideal gas with a polytropic equation of state [63]. In the absence of shocks or conduction, the pressure and density of a fluid element obey an equation of state $p = K_1 \rho^{\gamma_1}$. For an adiabatic, classical, nonrelativistic, monatomic gas, or for nonrelativistic degenerate particles, the adiabatic index $\gamma_1 = 5/3$ and different values of K_1 correspond to different entropy. If the entropy varies radially as a power-law, equilibrium self-gravitating configurations are given by classical Lane-Emden polytrope solutions. The radial variation of pressure and density obey $p(r) = K_2 \rho^{\gamma_2}(r)$; the second index γ_2 tracks the radial variation between different fluid elements in some particular configuration (that is, including variations in entropy). For gas on the same adiabat everywhere, $\gamma_1 = \gamma_2$; for the case of nonrelativistic degenerate or adiabatic matter, $\gamma_1 = \gamma_2 = 5/3$ applies and is a good model of degenerate dwarfs. If the entropy is increasing with radius, as would be expected if assembled in a cosmological hierarchy, then $\gamma_2 < \gamma_1$, conferring stability against convection.

The character of the solutions is well known [63]. As long as $\gamma_2 > 6/5$ the halo structure is like a star, with a flat-density core in the center, falling off in the outer parts to vanishing density at a boundary. If it is rotating, the structure is similar but rotationally flattened. These solutions describe approximately the structure of stars, especially degenerate dwarfs, and halos of highly collisional dark matter.[6] If $\gamma_2 < 6/5$ (and in particular for the isothermal case $\gamma_2 = 1$) there is a dy-

---

[6]It is worth commenting on some differences and similarities with collisionless halos with finite phase density material. The polytrope solutions are for collisional matter with an isotropic pressure and local balance of pressure gradient and gravity. Collisionless particles can fill phase space more sparsely, but this just means that at a given mass density they must have a larger maximum velocity; the collisional solution saturates the phase density limit and has the largest mass density for a given coarse-grained phase density. In this sense, once one is solving the cusp problem with finite phase density, nothing further is gained by making the particles collisional. Collisionless particles allow anisotropy in the momentum distribution function, and therefore a wider range of ellipsoidal figures, but cannot pack into tighter cores. For the same reason, the inner phase-density-limited core is expected to be close to spherical except for rotational support, whether the particles are collisional or not. The phase space is fully occupied and therefore the velocity distribution is close to isotropic wherever the local entropy approaches the primordial value.

namical instability and no stable solution; the system runs away on a gravitational time scale, with the center collapsing and the outer layers blowing off.

D. Giant dwarfs

At zero entropy the equilibrium configuration is the exactly soluble $\gamma=5/3$ polytrope, which we adopt as an illustrative example. That is, we model a dwarf galaxy core as a degenerate dwarf star, the only difference being a particle mass much smaller than a proton allowing a halo mass much bigger than a star. For total mass M and radius R, the Lane-Emden solution gives a central pressure $p_c=0.770GM^2/R^4$ and a central density $\rho_c=5.99\bar{\rho}=1.43M/R^3$. Using the above relation for the equation of state we obtain the standard degenerate dwarf solution, which has

$$R=4.5m_X^{-8/3}M^{-1/3}m_{Planck}^2$$
$$=0.98 \ \text{kpc}\left(\frac{m_X}{100 \ \text{eV}}\right)^{-8/3}\left(\frac{M}{10^{10}M_\odot}\right)^{-1/3}, \quad (34)$$

where $m_{Planck}=\sqrt{\hbar c/G}$ and $M_\odot=9.48\times10^{37}m_{Planck}$. This ''giant dwarf'' configuration is stable even at zero temperature up to the Chandrasekhar limit for X particles.[7]

Since the mass is not directly observable, it is more useful to consider the velocity of a circular orbit at the surface, $v_c=(GM/R)^{1/2}$. We then obtain the relation for a degenerate system,

$$m_X=4.5^{3/8}v_c^{-1/4}(r_cm_{Planck})^{-1/2}m_{Planck}, \quad (36)$$

or in more conventional astrophysical units,

$$m_X=87 \ \text{eV} \ (v_c/30 \ \text{km/s})^{-1/4}(r_c/1 \ \text{kpc})^{-1/2}. \quad (37)$$

Note that as in the collisionless case, no cosmological assumptions or parameters have entered into this expression.

For any adiabatic nonrelativistic matter the solution is similar. The absolute scale of the giant dwarf, determined by K_2, is fixed by the phase density Q. In general there is a range of entropy but once again the lowest-entropy material (which is densest at a given pressure) sinks to the center of a halo and forms an approximately adiabatic core. The rest of the halo forms a thermally-supported atmosphere above it. Once again cores are the places to look for signs of a primordial ceiling to phase density. However, as we see below the behavior changes if conduction or radiation are not negligible. As we know, a thermally supported star which conducts heat and has no nuclear or other source of energy is unstable.

---

[7]Defined analogously to the Chandrasekhar limit for standard dwarfs (with $Z=A$ because there is just one kind of particle providing both mass and pressure, similar to a neutron star),
$$M_{CX}=3.15(m_{Planck}^3/m_X^2)=4.95\times10^{14}M_\odot(m_X/100 \ \text{eV})^{-2}. \quad (35)$$

E. Heat conduction time and halo stability

If the collision rates are not very high we must consider heat and momentum transport between fluid elements by particle diffusion. The most serious consideration for radial stability is the transport of heat. In all stable thermally supported solutions the dense inner parts are hotter; if conduction is allowed, heat is transported outwards. The entropy of the central material decreases, the interior is compressed to higher density and the outer layers spread to infinity, a manifestation of the gravothermal catastrophe. With conduction the inner gas falls in and the outer gas drifts out on a diffusion time scale, attempting to approach a singular isothermal sphere.

Consider the scenario [44] where the dark matter cross section is small enough to remain essentially noninteracting on large scales, preserving the successes of CDM structure formation simulations, but large enough to become collisional in the dense central regions of galaxies. Although this scenario was introduced to help solve the cusp problem, we will see that the conductive instability makes matters worse. If stable cores are to last for a Hubble time, the dark matter halos must either be effectively collisionless (standard dark matter), or very strongly interacting, so that the inevitable conduction is slow (or made of degenerate fermions so there is no temperature gradient).

Elementary kinetic theory [51] yields an estimate for the conduction of heat by particle diffusion; the ratio of energy flux to temperature gradient is the classical conduction coefficient $\kappa\approx\sigma^{-1}\sqrt{T/m}$. Assuming a halo in approximate virial equilibrium and profile $v(r)$, this yields a time scale for heat conduction,

$$t_{cond}\approx\frac{v\sigma}{2Gm_X}\frac{-d\log r}{d\log v} \quad (38)$$

where v is the typical particle velocity (which is about the virial velocity of the halo independent of the mass of the particles m_X). The first factor is essentially the time it takes a particle to random walk a distance r, $t_{diffuse}\approx r^2n\sigma/v$. The last factor characterizes the temperature and entropy gradient; dynamical stability prevents it from being very large, and in most of the matter it typically takes a value not much larger than unity.[8]

A halo with conduction therefore forms a kind of cooling flow, with the core collapsing and the envelope expanding. If

---

[8]The conductive destabilization probably happens faster than Spergel and Steinhardt estimated. They used the Spitzer formula describing core collapse in globular clusters, which takes about 300 times longer than the two-particle relaxation time. However, the large factor arises because in the globular cluster case the relaxation is entirely gravitational and is dominated by very long-range interactions with distant stars. In the present situation the interactions are strong and short-range, leading to significant exchange of both energy and momentum in each scattering. The transport of heat takes place on the same time scale as the diffusion of particles, with numerical factors of the order of unity as in standard solutions of the Boltzmann equation for gases.

it is hydrostatically quasi-stable (that is, if the core collapse is slow and regulated by the particle diffusion), we can use the Lane-Emden solutions to set bounds on the numerical factor $d \log r/d \log v$ governing the instability. The equation of state tells us that $v \propto \rho^{1/2n_2}$ where $n_2 = (\gamma_2 - 1)^{-1}$. The largest value of n_2 which corresponds to a quasi-stable solution is $n_2 = 5$. The density profile is steeper than isothermal ($n_2 = \infty$), $\rho \propto r^{-2}$; therefore $|d \log r/d \log v| \lesssim n_2 \lesssim 5$. In the rough estimates here we set these factors to unity.[9]

Conduction can be suppressed if the scattering is very frequent. For nondegenerate X, stable cores require that the conduction time exceeds the Hubble time H_0^{-1}. For stability over a Hubble time, the column density of a halo with velocity v must exceed $m_X/\sigma = Hv/G$; therefore the particles must satisfy

$$\frac{m_X}{\sigma} \gtrsim 1.0 \times 10^{-4} \ \ \text{g cm}^{-2} h_{70} v_{stable,30}, \tag{39}$$

where $v_{stable,30} \times 30$ km s^{-1} denotes the velocity in the lowest-velocity stable halo. Perhaps surprisingly, the mass and radius of the halo do not enter explicitly.

This condition constrains the particles to be highly interactive. Galaxy halos have slow conduction compared to H only above a critical velocity dispersion $v_{crit} \approx (G/H)(m_X/\sigma)$. Halos below this threshold should have collapsed cores, and above the threshold the core radius–mass relation is determined as before by the giant dwarf sequence for the particle's phase density. The existence of stable bound 30 km/s halos of highly-collisional dark matter requires

$$\frac{m_X}{\sigma} \lesssim (2.8 \ \text{MeV})^3 h_{70} v_{stable,30}, \tag{40}$$

shown in Fig. 2 as the right boundary of the "fluid" region.

The "thickness" of a halo with velocity $v_{30} \times 30$ km s^{-1}, in units of particle pathlengths, is

$$\frac{\Sigma_h \sigma}{m_X} \approx 10^2 v_{30}^2 r_{kpc}^{-1} h_{70}^{-1} v_{stable,30}^{-1} \tag{41}$$

so it is clear that all dark-matter-dominated structures, from small galaxies to galaxy clusters ($v_{30} \approx 1-30$, $r_{kpc} \approx 0.1 - 1000$), are highly collisional and their dark matter behaves as a fluid. Even for very diffuse matter at the mean cosmic density ($\Omega_X = 0.3$), the particle mean free path is at most $12 v_{core,30} h_{70}^{-1}$ Mpc, about the same as the scale of nonlinear clustering, so all bound dark matter structures act like fluids.

---

[9]Another interesting limit is that of small but nonzero self-interactions. The halo is essentially collisionless, but occasional scatterings still take place. The collisionless isothermal sphere, singular or not, is then an approximate solution, but still subject to a slow secular instability from heat conduction. It is also possible to set up situations where halos are evaporated by a hot external environment, heated from outside by collapse of the cosmic web.

Are other data consistent with the idea that essentially all dark matter acts like a fluid? This option has been considered previously [42] and while it is perhaps not definitively ruled out, it is not phenomenologically compelling. Serious problems arise for example from satellite galaxies which are thought [64] to have had several orbits without stopping and sinking as they would in a fluid, or from galaxies in clusters, at least some of which appear (from lens reconstruction mass maps) to have retained some of their dark matter halos. An intriguing possibility is that a small collision rate might contribute to enough instability to feed the formation of black holes [49]. However the rate of the instability is greatest in the lowest mass, lowest density galaxy cores, a trend not conspicuous in the demography of central black holes of galaxies [65].[10]

We conclude that dark matter self-interactions are likely to be negligible in galaxy halos, and that this places significant constraints on the particles. Figure 2 summarizes the constraints on the parameters m_X, m_Y of this interacting-particle model from the various constraints considered here.

VI. CONCLUSIONS

We have found that some halos might preserve in their inner structure observable clues to new dark matter physics, and that indeed some current observations already hint that the dark matter might be warm rather than cold. We conclude with a summary:

(1) Halo cores can be created by a "phase-packing limit" depending on finite initial phase density. They may provide a direct probe of primordial velocity dispersion in dissipationless dark matter.

(2) For relativistically-decoupled thermal relics, the phase density depends on the particle mass and spin but not on cosmological parameters.

(3) Rotation curves in a few dwarf disk galaxies indicate cores with a phase density corresponding to that of a 200 eV thermal relic or an rms velocity of about 0.4 km/sec at the current cosmic mean density. Velocity dispersions in dwarf spheroidal galaxies indicate a higher phase density, corresponding to a thermal relic mass of about 1 keV. At most one of these populations can be tracing the primordial phase density.

(4) Thermal relics in this mass range can match the mean cosmic density with a plausible superweak decoupling from standard model particles before the QCD epoch.

(5) Other very different particles are consistent with the

---

[10]We have to take note of another possibility: perhaps the dwarf spheroidals, which have the lowest velocity dispersions of all galaxies and are also the densest, have already collapsed by heat conduction. In this way we could use phase packing to give the cores of the dwarf disk galaxies and still explain why the dwarf spheroidals have such a large phase density. Note that this scheme also gives the right filtering scale since the particles are collisional at early times. The dwarf spheroidals need not of course collapse all the way to black holes, but they may well have singular dark matter profiles.

halo data, provided they have the about the same mean density and phase density. Examples include WIMPs from particle decay and axions from defect decay.

(6) Cores due to phase packing limited by primordial Q_0 predict a universal relation between core radius and halo velocity dispersion. The relation is not found in a straightforward interpretation of the data.

(7) Primordial velocity dispersion also suppresses halo substructure (and solves some other difficulties with CDM) by filtering primordial adiabatic perturbations. Estimates based on luminosity functions prefer filtering on a scale of about $k \approx 3$ Mpc^{-1}; for collisionless particles, this scale corresponds to a filter caused by streaming of about a 1 keV thermal relic.

(8) Weak self-interactions change from streaming to acoustic behavior, reducing the damping scale and sharpening the filter.

(9) Stronger self-interactions destabilize halos by thermal conduction, making the cusp problem worse (unless they are very strong—too strong for satellite-galaxy kinematics—or particles are degenerate, eliminating the central temperature gradient).

(10) A simulation which samples a warm distribution function reasonably well is strongly motivated, to determine whether primordial Q is preserved in the centers of halos, or whether nonlinear effects can amplify dynamical heating in such models to explain cores on all scales.

ACKNOWLEDGMENTS

[1] J. R. Primack, in *Formation of Structure in the Universe*, edited by A. Dekel and J. P. Ostriker (Cambridge University Press, Cambridge, England, 1997).

[2] S. Ghigna, B. Moore, F. Governato, G. Lake, T. Quinn, and J. Stadel, astro-ph/9910166, 1999.

[3] A. Klypin, A. V. Kravtsov, O. Valenzuela, and F. Prada, Astrophys. J. **523**, 32 (1999).

[4] J. Dubinski and R. Carlberg, Astrophys. J. **378**, 496 (1991).

[5] J. F. Navarro, C. S. Frenk, and S. D. M. White, Mon. Not. R. Astron. Soc. **462**, 563 (1996).

[6] B. Moore, F. Governato, T. Quinn, J. Stadel, and G. Lake, Astrophys. J. Lett. **499**, L5 (1998).

[7] B. Moore, S. Ghigna, F. Governato, G. Lake, T. Quinn, J. Stadel, and P. Tozzi, Astrophys. J. Lett. **524**, L19 (1999).

[8] B. Moore, T. Quinn, F. Governato, J. Stadel, and G. Lake, astro-ph/9903164, 1999.

[9] J. S. Bullock *et al.*, astro-ph/9908159, 2000.

[10] R. Swaters, Ph.D. thesis, Groningen, 1999.

[11] R. A. Swaters, B. A. Madore, and M. Trewhella, Astrophys. J. Lett. **531**, L107 (2000).

[12] C. Carignan, Astrophys. J., Suppl. Ser. **58**, 107 (1985).

[13] S. Colombi, S. Dodelson, and L. M. Widrow, Astrophys. J. **458**, 1 (1996).

[14] F. C. van den Bosch, B. E. Robertson, J. J. Dalcanton, and W. J. G. de Blok, Astron. J. **119**, 1579 (2000).

[15] J. Dalcanton and C. J. Hogan, astro-ph/0004381, 2000.

[16] F. van den Bosch *et al.*, astro-ph/9911372, 1999.

[17] S. S. Gerstein and Ya. B. Zeldovich, Zh. Éksp. Teor. Fiz., Pis'ma Red. **4**, 174 (1966) [JETP Lett. **4**, 120 (1966)].

[18] R. Cowsik and J. McClelland, Phys. Rev. Lett. **29**, 669 (1972).

[19] S. Tremaine and J. E. Gunn, Phys. Rev. Lett. **42**, 407 (1979).

[20] J. R. Bond, G. Efstathiou, and J. Silk, Phys. Rev. Lett. **45**, 1980 (1980).

[21] J. Bardeen *et al.*, Astrophys. J. **304**, 15 (1986).

[22] G. Blumenthal, H. Pagels, and J. R. Primack, Nature (London) **299**, 37 (1982).

[23] A. Melott and D. N. Schramm, Astrophys. J. **298**, 1 (1985).

[24] J. R. Primack and G. 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), pp. 163–183; reprinted in *Beyond the Standard Model*, edited by J. Tran Thanh Van (Editions Frontieres, Gif-sur-Yvette, France, 1983), pp. 445–464.

[25] S. Dodelson and L. M. Widrow, Phys. Rev. Lett. **72**, 17 (1994).

[26] J. A. Peacock, *Cosmological Physics* (Cambridge University Press, Cambridge, England, 1999).

[27] S. D. M. White and M. J. Rees, Mon. Not. R. Astron. Soc. **183**, 341 (1978).

[28] J. F. Navarro and M. Steinmetz, Astrophys. J. **528**, 607 (2000).

[29] J. Sommer-Larsen and A. Dolgov, astro-ph/9912166, 1999.

[30] P. Colin, V. Avila-Reese, and O. Valenzuela, astro-ph/0004115, 2000.

[31] R. A. C. Croft *et al.*, Astrophys. J. **520**, 1 (1999).

[32] R. Dave, L. Hernquist, N. Katz, and D. H. Weinberg, Astrophys. J. **511**, 521 (1999).

[33] M. White and R. A. C. Croft, astro-ph/0001247, 2000.

[34] M. Aaronson, Astrophys. J., Lett. Ed. **266**, L11 (1983).

[35] E. Olszewski and M. Aaronson, Astrophys. J. **94**, 657 (1987).

[36] S. Faber and D. Lin, Astrophys. J., Lett. Ed. **266**, L21 (1983).

[37] G. Lake, Astrophys. J. Lett. **356**, L43 (1990).

[38] O. E. Gerhard and D. N. Spergel, Astrophys. J. Lett. **389**, L9 (1992).

[39] J. P. Ralston and L. L. Smith, Astrophys. J. **367**, 54 (1991).

[40] M. Mateo, Annu. Rev. Astron. Astrophys. **36**, 435 (1998).

[41] E. D. Carlson, M. E. Machacek, and L. J. Hall, Astrophys. J. **398**, 43 (1992).

[42] A. de Laix, R. J. Scherrer, and R. K. Schaefer, Astrophys. J. **452**, 495 (1995).

[43] F. Atrio-Barandela and S. Davidson, Phys. Rev. D **55**, 5886 (1997).

[44] D. N. Spergel and P. J. Steinhardt, Phys. Rev. Lett. **84**, 3760 (2000).

[45] R. N. Mohapatra and V. L. Teplitz, astro-ph/0001362, 2000.

[46] H. Pagels and J. R. Primack, Phys. Rev. Lett. **48**, 223 (1982).

[47] G. Fuller (unpublished).

[48] X. Shi and G. M. Fuller, Phys. Rev. Lett. **82**, 2832 (1999).

[49] J. P. Ostriker, Phys. Rev. Lett. **84**, 5258 (2000).

[50] S. Hannestad, astro-ph/9912558, 1999.

[51] L. D. Landau and E. M. Lifschitz, *Statistical Physics* (Addison-Wesley, Reading, MA, 1958).

[52] E. W. Kolb and M. S. Turner, *The Early Universe* (Addison-Wesley, Redwood City, CA, 1990).

[53] P. J. E. Peebles, *Principles of Physical Cosmology* (Princeton University Press, Princeton, NJ, 1993).

[54] C. J. Hogan and M. J. Rees, Phys. Lett. B **205**, 28 (1988).

[55] R. A. Battye and E. P. S. Shellard, Nucl. Phys. **B423**, 260 (1994).

[56] J. Wadsley *et al.* (in preparation).

[57] J. Binney and S. Tremaine, *Galactic Dynamics* (Princeton University Press, Princeton, NJ, 1987).

[58] J. A. Tyson, G. P. Kochanski, and I. P. Dell'Antonio, Astrophys. J. Lett. **498**, L107 (1998).

[59] L. L. R. Williams, J. F. Navarro, and M. Bartelmann, Astrophys. J. **527**, 535 (1999).

[60] T. E. Pickering, J. H. van Gorkom, C. D. Impey, and A. C. Quillen, Astron. J. **118**, 765 (1999).

[61] M. Fukugita, C. J. Hogan, and P. J. E. Peebles, Astrophys. J. **503**, 518 (1998).

[62] M. Kamionkowski and A. R. Liddle, Phys. Rev. Lett. **84**, 4525 (2000).

[63] Ya. B. Zeldovich and I. D. Novikov, *Relativistic Astrophysics* (University of Chicago Press, Chicago, 1971), Vol. 1.

[64] K. V. Johnston, S. R. Majewski, M. H. Siegel, I. N. Reid, and W. E. Kunkel, Astron. J. **118**, 1719 (1999).

[65] J. Magorrian *et al.*, Astron. J. **115**, 2285 (1998).