Cold dark matter: Controversies on small scales

David H. Weinberg,a,b James S. Bullockc, Fabio Governatod, Rachel Kuzio de Narayd, and Annika H. G. Petera,b

Cold dark matter (CDM) cosmological models have been remarkably successful in explaining cosmic structure over an enormous span of redshift, but it has faced persistent challenges from observations that probe the innermost regions of dark matter halos and the properties of the Milky Way’s dwarf galaxy satellites. We review the current observational and theoretical status of these “small-scale controversies.” Cosmological simulations that incorporate only gravity and collisionless CDM predict halos with abundant substructure and central densities that are too high to match constraints from galaxy dynamics. The solution could lie in baryonic physics: Recent numerical simulations and analytical models suggest that gravitational potential fluctuations tied to efficient supernova feedback can flatten the central cusps of halos in massive galaxies, and a combination of feedback and low star formation efficiency could explain why most of the dark matter subhalos orbiting the Milky Way do not host visible galaxies. However, it is not clear that this solution can work in the lowest mass galaxies, where discrepancies are observed. Alternatively, the small-scale conflicts could be evidence of more complex physics in the dark sector itself. For example, elastic scattering from strong dark matter self-interactions can alter predicted halo mass profiles, leading to good agreement with observations across a wide range of galaxy mass. Gravitational lensing and dynamical perturbations of tidal streams in the stellar halo provide evidence for an abundant population of low-mass subhalos in accord with CDM predictions. These observational approaches will get more powerful over the next few years.

Dark matter | cosmology | galaxy formation

The cold dark matter (CDM) hypothesis—that dark matter consists of a weakly interacting particle whose velocity dispersion in the early universe was too small to erode structure on a galactic or subgalactic scale—emerged in the early 1980s and quickly became a central element of the theory of cosmic structure formation. Influential early papers include Peebles’ calculation of cosmic microwave background (CMB) anisotropies and the matter power spectrum (1), discussions of galaxy formation with particle dark matter by Bond et al. (2) and Blumenthal et al. (3, 4), and Davis et al.’s (5) numerical simulations of galaxy clustering. By the mid-1990s, the simplest CDM model with scale-invariant primordial fluctuations and a critical matter density of $\Omega_0 = 1$ had run afoul of multiple lines of observational evidence, including the shape of the galaxy power spectrum, estimates of the mean matter density from galaxy clusters and galaxy motions, the age of the universe inferred from estimates of the Hubble constant, and the amplitude of matter clustering extrapolated forward from the fluctuations measured in the CMB. Many variants on “canonical” CDM were proposed to address these challenges, and by the turn of the century, the combination of supernova evidence for cosmic acceleration and CMB evidence for a flat universe had selected a clear winner: $\Lambda$CDM, incorporating CDM, a cosmological constant ($\Lambda$), and inflationary initial conditions. Today, the $\Lambda$CDM scenario has a wide range of observational successes, from the CMB to the Lyman-$\alpha$ forest to galaxy clustering to weak gravitational lensing, and it is generally considered the “standard model” of cosmology.

However, as the resolution of cosmological $N$-body simulations improved in the mid- to late 1990s, they revealed two tensions with observations that have remained thorns in the side of the CDM hypothesis. First, simulations showed that CDM collapse leads to cuspy dark matter halos whose central density profiles rise as $r^{-\beta}$ with $\beta \sim 1 - 1.5$, whereas observed galaxy rotation curves favored constant density cores in the dark matter distribution (6–9). Second, simulated halos retained a large amount of substructure formed by earlier collapses on smaller scales, predicting hundreds or thousands of subhalos in contrast to the ~10 “classical” satellites of the Milky Way (10, 11). These two conflicts are often referred to as the “cusp-core problem” and the “missing satellites problem.” We will argue below that these two problems have largely merged into one, and that the most puzzling aspect of the Milky Way’s satellite galaxies is not their number but their low central matter densities, which again imply mass profiles shallower than the naive CDM prediction.

In this brief article, based on our panel discussion at the 2012 Sackler Symposium on Dark Matter, we attempt to summarize the current state of the CDM controversies at a level that will be useful to readers not immersed in the field. The key question is whether the conflicts between $N$-body predictions and observed galaxy properties can be resolved by “baryonic physics”—gas cooling, star formation, and associated feedback—or whether they require different properties of the dark matter itself.

Cores, Cusps, and Satellites

Fig. 1 illustrates the “cusp-core” problem. To set the scene, Fig. 1 (Left) superposes an optical image of the low surface brightness galaxy F568-3 onto a numerical simulation of a CDM halo. As shown in Fig. 1 (Right), the inner mass profile of a dark matter halo can be probed by rotation curve measurements; for circular motions in a spherical matter distribution, the rotation speed is simply $v_c(r) = \sqrt{GM(r)/r}$, where $M(r)$ is the mass interior to radius $r$. Points with error bars show the measured rotation curve of F568-3 (from ref. 13). The dotted curve shows the $v_c(r)$ expected from the gravity of the stellar and gas components of the galaxy, which are subdominant even in the central regions. The solid curve shows the predicted rotation curve, including the contribution of an isothermal dark matter halo with a constant density core, which fits the data well. The dashed curve instead incorporates a halo with a Navarro–Frenk–White (NFW; 8) profile and a concentration typical for galaxy mass halos. When normalized to match the observed rotation at large radii, the
NFW halo overpredicts the rotation speed in the inner few kiloparsecs (kpc), by a factor of 2 or more.

Early theoretical discussions of the cusp-core problem devoted considerable attention to the predicted central slope of the density profiles and to the effects of finite numerical resolution and cosmological parameter choices on the simulation predictions [a recent state-of-the-art discussion is provided by Ludlow et al. (14)]. However, the details of the inner profile shape are not essential to the conflict; the basic problem is that CDM predicts too much dark matter in the central few kpc of typical galaxies, and the tension is evident at scales where $v_c(r)$ has risen to $\sim 1/2$ of its asymptotic value (e.g., refs. 15, 16). On the observational side, the most severe discrepancies between predicted and observed rotation curves arise for fairly small galaxies, and early discussions focused on whether beam smearing or noncircular motions could artificially suppress the measured $v_c(r)$ at small radii. However, despite uncertainties in individual cases, improvements in the observations, sample sizes, and modeling have led to a clear overall picture: A majority of galaxy rotation curves are better fit with cored dark matter profiles than with NFW-like dark matter profiles, and some well-observed galaxies cannot be fit with NFW-like profiles, even when one allows halo concentrations at the low end of the theoretically predicted distribution and accounts for uncertainties in modeling the baryon component (e.g., ref. 13). Resolving the cusp-core problem therefore requires modifying the halo profiles.

**Fig. 1.** Cusp-core problem. (Left) Optical image of the galaxy F568-3 (inset, from the Sloan Digital Sky Survey) is superposed on the dark matter distribution from the “Via Lactea” cosmological simulation of a Milky Way-mass CDM halo (12). In the simulation image, intensity encodes the square of the dark matter density, which is proportional to the annihilation rate and highlights the low-mass substructure. (Right) Measured rotation curve of F568-3 (points) compared with model fits assuming a cored dark matter halo (blue solid curve) or a cuspy dark matter halo with a Navarro–Frenk–White (NFW; 8) profile (red dashed curve, concentration $c = 9.2$, $V_{200} = 110$ km·s$^{-1}$). The dotted green curve shows the contribution of baryons (stars + gas) to the rotation curve, which is included in both model fits. An NFW halo profile overpredicts the rotation speed in the inner few kpc. Note that the rotation curve is measured over roughly the scale of the 40-kiloparsec (kpc) image (inset, Left).

**Fig. 2.** Missing satellite and too big to fail problems. (Left) Projected dark matter distribution (600 kpc on a side) of a simulated, $10^{12} M_\odot$ CDM halo (18). As in Fig. 1, the numerous small subhalos far exceed the number of known Milky Way satellites. Circles mark the nine most massive subhalos. (Right) Spatial distribution of the classical satellites of the Milky Way. The central densities of the subhalos (Left) are too high to host the dwarf satellites (Right), predicting stellar velocity dispersions higher than observed. (Right) Diameter of the outer sphere is 300 kpc; relative to the simulation prediction (and to the Andromeda galaxy), the Milky Way’s satellite system is unusually centrally concentrated (19).
of typical spiral galaxies away from the profiles that N-body simulations predict for collisionless CDM.

Fig. 2 illustrates the “missing satellite” problem. Fig. 2 (Left) shows the projected dark matter density distribution of a 10^12 M_☉ CDM halo formed in a cosmological N-body simulation. Because CDM preserves primordial fluctuations down to very small scales, halos today are filled with enormous numbers of subhalos that collapse at early times and preserve their identities after falling into larger systems. Before 2000, there were only nine dwarf satellite galaxies known within the ∼ 250-kpc virial radius of the Milky Way halo (Fig. 2, Right), with the smallest having stellar velocity dispersions ∼ 10 km s⁻¹ (10), and Moore et al. (11) predicted a factor of ∼ 5—20 more subhalos above a corresponding velocity threshold in their simulated Milky Way halos. Establishing the “correspondence” between satellite stellar dynamics and subhalo properties is a key technical point (17) that we will return to below, but a prima facie comparison suggests that the predicted satellite population far exceeds the observed one.

Fortunately (or perhaps unfortunately), the missing satellite problem seems like it could be solved fairly easily by baryonic physics. In particular, the velocity threshold at which subhalo and dwarf satellite counts diverge is close to the ∼ 30 km s⁻¹ value at which heating of intergalactic gas by the UV photoionizing background should suppress gas accretion onto halos, which could plausibly cause these halos to remain dark (20—22). Alternatively, supernovae and stellar winds from the first generation of stars could drive remaining gas out of the shallow potential wells of these low-mass halos. Complicating the situation, searches using the Sloan Digital Sky Survey (SDSS) have discovered another ∼ 15 “ultrafaint” satellites with luminosities of only 10^11—10^10 L_☉ (e.g., refs. 23, 24). The high-latitude SDSS imaging covered only ∼ 20% of the sky, and many of the newly discovered dwarfs are so faint that they could only be seen to 50—100 kpc (25, 26); thus, extrapolating to the full volume within the Milky Way virial radius suggests a population of several hundred faint dwarf satellites (27).

Estimates from stellar dynamics imply that the mass of dark matter in the central 0.3 kpc of the host subhalos is M_☉ ≈ 10^7 M_☉ across an enormous range of luminosities, L ∼ 10^11—10^10 L_☉ (encompassing the classical dwarf spheroidals as well as the SDSS dwarfs), which suggests that the mapping between halo mass and luminosity breaks down. Additionally, semianalytic models predict that L ∼ 10^11 L_☉ dwarfs have softening lengths near the threshold (28). The luminosity function of the faint and ultrafaint dwarfs can be explained by semianalytical models invoking photoionization and stellar feedback (e.g., refs. 29, 30), although the efficiency of converting baryons to stars remains surprisingly low (< 0.1%—1%), well above the photoionization threshold, and it is unclear which, if any, of the ultrafaint dwarfs are “fossils” from before the epoch of reionization (31, 32). Despite the gaps in understanding, it seems reasonable for now to regard the relation between low-mass subhalos and ultrafaint dwarfs as a puzzle of galaxy formation physics rather than a contradiction of CDM.

Instead, attention has focused recently on the most luminous satellites. Circles in Fig. 2 mark the nine most massive subhalos in the simulation, which one would expect to host galaxies like the Milky Way’s classical dwarf satellites. However, the mass in the central regions of these subhalos exceeds the mass inferred from stellar dynamics of observed dwarfs, by a factor of ∼ 5 (33—36). Although it is possible, in principle, that these massive subhalos are dark and that the observed dwarfs reside in less massive hosts, this outcome seems physically unlikely: in the spirit of the times, Boylan-Kolchin et al. (33) titled this conflict “too big to fail.” The degree of discrepancy varies with the particular realization of halo substructure and with the mass of the main halo, but even for a halo mass at the low end of estimates for the Milky Way, the discrepancy appears too large to be a statistical fluke, and a similar conflict is found in the satellite system of the Andromeda galaxy (37). Although missing satellites in low-mass subhalos may be explained by baryonic effects, the too big to fail problem arises in more massive systems whose gravitational potential is dominated by dark matter. In its present form, therefore, the satellite puzzle looks much like the cusp-core problem: Numerical simulations of CDM structure formation predict too much mass in the central regions of halos and subhalos. Indeed, Walker and Penarrubia (38), Amorisco et al. (39), and others have reported evidence that the Milky Way satellites Fornax and Sculptor have cored density profiles.

**Solutions in Baryonic Physics?**

When the cusp-core problem was first identified, the conventional lore was that including baryonic physics would only exacerbate the problem by adiabatically contracting the dark matter density distribution (6, 40). Navarro et al. (41) proposed a scenario, which seemed extreme at the time, for producing a cored dark matter distribution: Dissipative baryons draw in the dark matter orbits adiabatically by slowly deepening the gravitational potential and then release them suddenly when the supernova feedback of a vigorous starburst blows out a substantial fraction of the baryonic material, leaving the dark matter halo less concentrated than the one that would have formed in the absence of baryons. Since then, hydrodynamic simulations have greatly improved in numerical resolution and in the sophistication with which they model star formation and supernova feedback. With the combination of a high gas density threshold for star formation and efficient feedback, simulations successfully reproduce the observed stellar and cold gas fractions of field galaxies. The ejection of low angular momentum gas by feedback plays a critical role in suppressing the formation of stellar bulges in dwarf galaxies (42), another long-standing problem in early simulations of galaxy formation. The episodic gas outflows also produce rapid fluctuations of the gravitational potential, in contrast to the steady growth assumed in adiabatic contraction models.

Fig. 3, based on work by Governato et al. (43), illustrates the impact of this episodic feedback on the dark matter density profile. In Fig. 3 (Left), the black dotted-dashed curve shows the final halo profile of an N-body simulation run with gravity and dissipationless matter only. Other curves show the evolution of the dark matter density profile in a hydrodynamic simulation with star formation and feedback, from the same initial conditions. Over time, the central dark matter density drops and the cuspy profile is transformed to one with a nearly constant density core (Fig. 3, Left, black solid curve). Pontzen and Governato (44) present an analytical model that accurately describes this transformation (and its dependence on simulation assumptions); essentially, the rapid fluctuations in the central potential pump energy into the dark matter particle orbits, so that they no longer penetrate to the center of the halo. The simulations of Governato et al. (43) use smoothed particle hydrodynamics, and the same flattening of dark matter cusps is found in adaptive mesh refinement simulations that have similarly episodic supernova feedback (45).

Fig. 3 (Right) compares the density profile slopes of simulated galaxies with observational estimates from 21-cm measurements of nearby galaxies (46) with predictions for an NFW dark matter halo. The reduced central density slopes agree well with observations for galaxies with a stellar mass of M_* > 10^9 M_☉. Strong gas outflows are observed in a wide variety of galaxies, including the likely progenitors of M_* ∼ 10^9—10^10 M_☉ dwarfs observed at z ∼ 2 (47). However, for galaxies with M_* below ∼ 10^9 M_☉, analytical models suggest that with so few stars, there is not enough energy in supernovae alone to create dark matter cores of ∼ 1 kpc (48). More generally, Garrison-Kimmel et al. (49) used idealized, high-resolution simulations to model potential fluctuations of the type expected in episodic feedback models and concluded that the energy required for solving the too big to fail problem...
Baryonic effects on CDM halo profiles in cosmological simulations. Other curves show the evolution of the dark matter profile in a simulation from the same initial conditions, which include gas dynamics, star formation, and efficient feedback. By $z = 0$ (black solid curve), the perturbations from the fluctuating baryonic potential have flattened the inner profile to a nearly constant density core. (Right) Logarithmic slope of the dark matter profile $\alpha$ measured at 0.5 kpc, as a function of galaxy stellar mass. Crosses show results from multiple hydrodynamic simulations. Squares show measurements from rotation curves of galaxies observed by The HI Nearby Galaxy Survey (THINGS). The black curve shows the expectation for pure dark matter simulations, computed from NFW profiles with the appropriate concentration. For $M_* > 10^7 M_\odot$, baryonic effects reduce the halo profile slopes to agree with observations. Both panels reprinted from ref. 43.

Fig. 3. Baryonic effects on CDM halo profiles in cosmological simulations. (Left) Black dotted-dashed curve shows the cuspy dark matter density profile resulting from a collisionless N-body simulation. Other curves show the evolution of the dark matter profile in a simulation from the same initial conditions, which include gas dynamics, star formation, and efficient feedback. By $z = 0$ (black solid curve), the perturbations from the fluctuating baryonic potential have flattened the inner profile to a nearly constant density core. (Right) Logarithmic slope of the dark matter profile $\alpha$ measured at 0.5 kpc, as a function of galaxy stellar mass. Crosses show results from multiple hydrodynamic simulations. Squares show measurements from rotation curves of galaxies observed by The HI Nearby Galaxy Survey (THINGS). The black curve shows the expectation for pure dark matter simulations, computed from NFW profiles with the appropriate concentration. For $M_* > 10^7 M_\odot$, baryonic effects reduce the halo profile slopes to agree with observations. Both panels reprinted from ref. 43.

Exceeds that available from supernovae in galaxies with stellar masses below $\sim 10^7 M_\odot$. The low-mass galaxies in Fig. 3 (from ref. 43) are consistent with this expectation, with density profile slopes that are negligibly affected by feedback at the 0.5-kpc scale. On the other hand, high-resolution simulations of luminous satellites in the halo of Milky Way-like hosts do show reduced central dark matter densities from a combination of early feedback effects with ram pressure stripping and tidal heating by the host halo and disk, processes that can extract energy from the host galaxy’s gravitational potential (50–52). Alternatively, Kuhlen et al. (53) argue that the regulation of star formation by molecular hydrogen cooling may make the stellar content of galaxies highly stochastic at a halo mass as high as $10^{10} M_\odot$ (also ref. 54), so that even the Milky Way’s most massive subhalos are not too big to fail. Ram pressure in the galactic halo could then remove the gas from the dark subhalos.

These arguments point to isolated, low-mass galaxies with $M_* \sim 10^6 - 10^7 M_\odot$ as ideal laboratories for testing the predictions of CDM-based models. Dwarfs that are far separated from a giant galaxy must rely on their own (modest) supernova reservoirs for energy injection. Ferrero et al. (55) have studied a population of $\sim 10^6 - 10^7 M_\odot$ field galaxies and argue that the central density problem persists even for relatively isolated dwarfs of this size. If this result holds up in further investigations, it will become a particularly serious challenge to CDM.

Solutions in Dark Matter Physics?

Instead of complex baryonic effects, the cusp-core and satellite problems could indicate a failure of the CDM hypothesis itself. One potential solution is to make dark matter “warm,” so that its free-streaming velocities in the early universe are large enough to erode primordial fluctuations on subgalactic scales. For a simple thermal relic, the ballpark particle mass is $m \approx 1$ keV, although details of the particle physics can alter the relation between mass and the free-streaming scale, which is the important quantity for determining the fluctuation spectrum. Alternatively, the small-scale fluctuations can be suppressed by an unusual feature in the inflationary potential (56). Although collisionless collapse of warm dark matter (WDM) still leads to a cuspy halo profile, the central concentration is lower than that of CDM halos when the mass scale is close to the spectral cutoff (e.g., ref. 57), thus allowing a better fit to observations of galaxy rotation curves and dwarf satellite dynamics. The mass function of halos and subhalos drops at low masses because there are no small-scale perturbations to produce collapsed objects, so the subhalo mass function can be brought into agreement with dwarf satellite counts. There have been numerous numerical simulations of structure formation with WDM (recent examples include refs. 58–63).

WDM is a “just-so” solution to CDM’s problems, requiring a particle mass (or free-streaming velocity) that is tuned to the particular scale of dwarf galaxy halos. However, the more serious challenge to WDM is observational, for two reasons. First, WDM does too good a job in eliminating power on small scales; for a thermal relic of mass $m = 2$ keV, there are too few subhalos in the Milky Way to host the known satellite galaxies (58). It also appears to be in conflict with observations of strong-lens systems, which show evidence for a significant subhalo fraction as well as the existence of small ($10^6 M_\odot$) subhalos (64–70). Second, suppressing primordial fluctuations on small scales alters the predicted structure of Lyman-$\alpha$ forest absorption toward quasars at high redshift, where these scales are still in the quasilinear regime (71). Recent studies of the Lyman-$\alpha$ forest set a lower limit on the dark matter particle mass of several kiloelectronvolts, high enough that the dark matter is effectively “cold” from the point of view of the cusp-core problem (refs. 72, 73, but a counterclaim of a lower minimum particle mass is made in ref. 74). Even setting these problems aside, it appears that WDM on its own does not fix the shape of rotation curves across the full range of galaxy masses, where conflict with CDM is observed (75). Although some uncertainties in the numerical simulations and observational data remain, it appears that WDM cannot solve the cuspy-core and missing satellite problems while remaining consistent with Lyman-$\alpha$ forest and substructure observations.

An alternative idea, made popular by Spergel and Steinhardt (76), is that CDM has weak interactions with baryons but strong self-interactions. The required scattering cross-section is roughly $(m/g)^{-1}$ cm$^2$, where $m$ is the particle mass; note that $1$ cm$^2$ g$^{-1}$ $\approx 1$ barn GeV$^{-1}$ is approximately a nuclear-scale cross-section. In this case, elastic scattering in the dense central regions of halos is frequent enough to redistribute energy and angular momentum among particles, creating an isothermal, round core of approximately constant density (77). Some early studies suggested that this idea was ruled out by gravitational lensing (78) or the by catastrophic gravitational core collapse found in a simulation of an isolated halo (79), but recent numerical studies show that these concerns are not borne out in fully cosmological simulations.
Instead, simulations show that there is a viable window of mass and cross-section where self-interacting dark matter (SIDM) can produce cored dark matter profiles and remain consistent with observational constraints (80, 81).

Fig. 4, based on the work of Rocha et al. (80), compares the structure and density profiles of halos formed from the same initial conditions with collisionless CDM and SIDM. Elastic scattering in the central regions, where an average particle experiences a few collisions per Hubble time, flattens the density cusp and reduces triaxiality. The scattering mechanism would operate across a wide range of halo masses, allowing SIDM to address both the rotation curves of Milky Way-like galaxies and the central densities of dwarf satellites. Because they are more weakly bound, SIDM subhalos are more easily subject to tidal disruption than CDM subhalos. However, the suppression of the low-mass subhalo count is not significant for allowed cross-sections, except in the innermost region of the host halo (80, 82). Thus, SIDM can solve the cusp-core problem while leaving enough subhalos to host Milky Way satellites, unlike WDM.

The prospects for SIDM appear much more hopeful than for WDM [although a summary of pro-WDM views is provided by Biermann et al. (83)]. Velocity-independent cross-sections in the range of \( \sim 0.1 - 0.5 \text{ cm}^2 \cdot \text{g}^{-1} \) create cores that are approximately the right size for Milky Way dwarf galaxies, spiral galaxies, and galaxy clusters (80, 84, 85), although leaving halos triaxial enough to match observations (81). Cross-sections in this range are also consistent with observations of merging galaxy clusters (86–88).

Moreover, particle model builders have recently focused attention on new classes of “hidden sector” models that generically produce SIDM particle candidates, although, in general, the elastic scattering cross-section has strong velocity dependence (89–93). For these models, strong self-interactions may only be present in a narrow range of halo mass, leaving halos on other scales effectively collisionless. Observationally, the goal is to either rule out or find evidence for SIDM cross-sections \( \sigma > 0.1 \text{ cm}^2 \cdot \text{g}^{-1} \), because for smaller cross-sections, the halo phenomenology is likely to be indistinguishable from CDM.

There are alternative dark matter physics mechanisms that could reduce the central densities of halos, including particle decay and particle-antiparticle annihilation (94, 95) or the recently suggested possibility of escape from flavor-mixed quantum states (96).

Conclusions

Are the tensions between CDM predictions and observations on the scales of galactic cores and satellite halos telling us something about the fundamental properties of dark matter, or are they telling us something interesting about the complexities of galaxy formation? After two decades of debate, the current state of the field is an unsatisfying stalemate (or perhaps a draw by repetition). However, there are several directions for future progress that could resolve the question.

Developments of the past several years have focused the “small-scale controversies” down to one fairly specific issue: the influence of baryons on the dark matter halo profile in systems where the baryons are today greatly subdominant. A variety of studies have shown that baryonic effects can plausibly account for cores in halos occupied by high surface brightness galaxies and can plausibly suppress star formation in very low-mass halos. Improved simulations may show that baryonic effects can soften cusps even in galaxies that are now dominated by dark matter, or they may show that the energetics arguments summarized above do indeed point to a genuine problem for CDM that cannot be resolved by supernova feedback or galactic tides. Improved simulations of models with interacting dark matter may show that they can readily solve the small-scale problems, or they may show that cross-section parameters chosen to match one set of observations ultimately fail when confronted with another set. SIDM models might also be ruled out if they predict halo shapes that can be excluded by observations of stellar or gas dynamics. Improved measurements of stellar velocities in satellite galaxies, and discovery of new satellites from imaging surveys such as Pan-STARRS (the Panoramic Survey Telescope and Rapid Response System) and the Dark Energy Survey, may better delineate the satellite problem itself.

These developments will affect the credibility of baryonic and dark matter solutions to the CDM controversies, but they may not yield a definitive conclusion. More satisfactory would be a direct test of the CDM prediction that vast numbers of low-mass subhalos (\( \sim 20,000 \) with masses \( > 10^5 M_\odot \)) and masses above \( 10^4 M_\odot \) are orbiting within the virial radius of the Milky Way and similar galaxies. Flux anomalies in gravitational lenses have already provided important evidence for subhalos that collectively contain a few percent of the mass within their parent halos, a level roughly consistent with CDM predictions and an order of magnitude above that expected from luminous satellites alone (64, 97). These anomalies do not directly probe the mass spectrum of the subhalos, although at masses of \( M \sim 10^6 M_\odot \), they...
produce detectable astrometric deviations in addition to flux anomalies (e.g., ref. 68). Current constraints are derived from a small sample of lensed radio-loud quasars, because optical flux anomalies could be produced by stellar microlensing. Studies should improve dramatically with the advent of the Atacama Large Millimeter Array and the James Webb Space Telescope, which can resolve lenses at submillimetric and mid-IR wavelengths that are not affected by stellar microlensing because the quasar dust emission regions are too large. An alternative route is to study cold tidal streams in the Milky Way, which would be perturbed by the multitude of passing subhalos. Carlb erg and Gillmain (ref. 98 and references therein) argue that observed tidal streams already show evidence of these perturbations, and a combination of better numerical simulations, more streams, and more detailed density and dynamical measurements could yield definitive evidence for or against CD M’s predicted subhalo population. Further theoretical work is needed to determine whether lensing or stream perturbations can distinguish CD M from SIDM.

As emphasized throughout the Sackler Symposium, there are great hopes that underground detection experiments, γ-ray observations, or collider experiments will identify the dark matter particle within the next decade. Such detections might definitively demonstrate whether dark matter is cold and weakly interacting or if they will confirm the particle while yielding ambiguous answers to this question. In the meantime, astronomers will continue their decades-long practice of studying the dark sector by observing the visible.

ACKNOWLEDGMENTS. We thank our many collaborators whose work we have summarized here. We thank Chris Kochanek for helpful comments on lensing constraints. Our work on these topics is supported by NASA and the National Science Foundation, including Grants NNX10AJ95G, NNX08ABG4, AST-1099505, AST-1099793, and AST-0607819.


