Galaxy stability within a self-interacting dark matter halo

Curtis J. Saxton
Mullard Space Science Laboratory, University College London, Holmbury St Mary, Surrey RH5 6NT, UK

ABSTRACT
This paper investigates spheroidal galaxies comprising a self-interacting dark matter (SIDM) halo plus de Vaucouleurs stellar distribution. These are coupled only via their shared gravitational field, which is computed consistently from the density profiles. Assuming conservation of mass, momentum and angular momentum, perturbation analyses reveal the galaxy’s response to radial disturbance. The modes depend on fundamental dark matter properties, the stellar mass, and the halo’s mass and radius. The coupling of stars and dark matter stabilizes some haloes that would be unstable as one-fluid models. However the centrally densest haloes are unstable, causing radial flows of SIDM and stars (sometimes in opposite directions). Depending on the dark microphysics, some highly diffuse haloes are also unstable. Unstable galaxies might shed their outskirts or collapse. Observed elliptical galaxies appear to exist in the safe domain. Halo pulsations are possible. The innermost node of SIDM waves may occur within 10 half-light radii. Induced stellar ripples may also occur at detectable radii if higher overtones are excited. If any SIDM exists, observational skotoseismology of galaxies could probe dark matter (DM) physics, measure the sizes of specific systems, and perhaps help explain peculiar objects (e.g. some shell galaxies, and the growth of red nuggets).

Key words: galaxies: elliptical and lenticular, cD – galaxies: kinematics and dynamics – galaxies: structure – dark matter.

1 INTRODUCTION
For decades it has been observed that visible matter in galaxies and clusters is insufficient to bind them gravitationally (e.g. Zwicky 1937; Freeman 1970; Roberts & Rots 1973). The extra gravity of invisible ‘dark matter’ (DM) is invoked, but its properties are uncertain, with many particle physics conjectures offering candidates. On the Earth, DM particle detection experiments differ in their findings: some report annually modulated signals (e.g. Bernabei et al. 2008; Aalseth et al. 2011a,b; Angloher et al. 2012); others see nothing above their expected background (e.g. Akerib et al. 2006; Angle et al. 2008; Ahmed et al. 2009; Aprile et al. 2010). Reconciling them may require extra physics: inelastic or composite forms of DM, or extra internal degrees of freedom (e.g. Smith et al. 2001; Chang et al. 2009; Alves et al. 2010; Kaplan et al. 2010, 2011).

Large-scale cosmic filaments and voids emerge in N-body computer simulations using gravity alone (e.g. Frenk, White & Davis 1983; Melott et al. 1983; Springel, Frenk & White 2006) but there is observational tension concerning smaller, self-bound systems. The predicted abundances of dwarf galaxies exceed those seen in many types of environments (e.g. Klypin et al. 1999; Moore et al. 1999; D’Onghia & Lake 2004; Tikhonov & Klypin 2009; Carlberg, Sullivan & Le Borgne 2009; Zavala et al. 2009; Peebles & Nusser 2010). The biggest simulated subhaloes appear too massive to match real satellite galaxies (Boylan-Kolchin, Bullock & Kaplinghat 2011, 2012). Collisionless DM haloes develop singular central density cusps (e.g. Gurevich & Zybin 1988; Dubinski & Carlberg 1991; Navarro, Frenk & White 1996b). On the contrary, observations of diverse galaxy classes indicate or allow a dark density profile that is smooth and nearly uniform within a kpc-scale inner core (e.g. Flores & Primack 1994; Moore 1994; Burkert 1995; Salucci & Burkert 2000; Kelson et al. 2002; Klyna et al. 2003; Gentile et al. 2004; de Blok 2005; Thomas et al. 2005; Goerdt et al. 2006; Kuzio de Naray et al. 2006; Gilmore et al. 2007; Oh et al. 2008; Weijmans et al. 2008; Inoue 2009; Nagino & Matsushita 2009; de Blok 2010; Pu et al. 2010; Murphy, Gebhardt & Adams 2011; Memola, Salucci & Babić 2011; Walker & Peñarrubia 2011; Aghello & Evans 2012; Jardel & Gebhardt 2012; Lora et al. 2012; Salucci et al. 2012). Galaxy cluster haloes should be centrally compressed by the steep density profile of cooling gas, but nonetheless many examples are definitely or possibly cored (e.g. Ettori et al. 2002; Sand, Treu & Ellis 2002; Halkola, Seitz & Pannella 2006; Voigt & Fabian 2006; Rzepecki et al. 2007; Halkola et al. 2008; Sand et al. 2008; Newman et al. 2009, 2011; Zitrin & Broadhurst 2009; Richtler et al. 2011).

To remedy (or defer) these DM crises, many astrophysicists invoke energetic ‘feedback’ by starbursts or active galaxies: explosions and outflows violent enough to flatten the dark cusps and denude small subhaloes of baryons (e.g. Navarro, Eke & Frenk 1996a; Gnedin & Zhao 2002; Mashchenko, Couchman &
Wadsley 2006; Peirani, Kay & Silk 2008; Governato et al. 2010). These events would have happened early enough to be difficult to observe today. The details would involve radiative shocks in magnetized multiphase interstellar media that are thermally and dynamically unstable. The relevant scales are numerically unresolved in present simulations. Results depend on phenomenological recipes for sub-grid processes, and the implementations are more varied and elaborate than any DM physics. For reasons of resolution alone, the proposition that feedback cures the DM crises cannot be decisively evaluated for at least another decade. Regardless of this fact, feedback may be energetically incapable of removing cusps whilst leaving realistic stellar populations in dwarf galaxies (Peharrubia et al. 2012).

An alternative approach extends standard particle physics phenomenologies: taking the DM crises as signatures of dynamically important physics in the dark sector. In self-interacting dark matter (‘SIDM’) theories, DM is self-interacting yet electromagnetically neutral. Pressure support or thermal conduction could prevent cusps, while drag and ablation reduce subhaloes (e.g. Flores & Primack 1994; Moore 1994; Firmani et al. 2000; Spergel & Steinhardt 2000; Rocha et al. 2012; Vogelsberger, Zavala & Loeb 2012). SIDM interactions entail direct inter-particle scattering (with specific cross-section $\sigma$) possibly enhanced by short-range Yukawa-type forces, or a long-range ‘dark electromagnetism’ (e.g. Ackerman et al. 2009; Loeb & Weiner 2011). Collisionless plasma supported by ‘dark magnetism’ is conceivable. Alternatively, the SIDM might consist of boson condensates of a scalar field (e.g. Peebles 2000). Plausible early-Universe evolution and cosmic structures emerge in cosmonogies with fluid-like SIDM (e.g. Moore et al. 2000; Stieie, Boeckel & Schaffner-Bielich 2010a,b) or scalar field DM (e.g. Matos & Arturo Ureña-López 2001; Suárez & Matos 2011). There were early fears that SIDM haloes would form sharper cusps via gravothermal collapse (e.g. Burkert 2000; Kochanek & White 2000). Later analyses put this beyond a Hubble time, and found two safe regimes that suppress thermal conductivity: weaker scattering (rarely per orbit) and stronger (fluid-like) cases with tiny mean-free paths (Balberg, Shapiro & Inagaki 2002; Ahn & Shapiro 2005; Koda & Shapiro 2011). We will refer to the $\zeta < 1 \text{cm}^{-1} \text{g}^{-1}$ and $\zeta \gtrsim 200 \text{cm}^{2} \text{g}^{-1}$ regimes as ‘weak-SIDM’ and ‘strong-SIDM’, respectively.

The ubiquity of dark cores supports SIDM, but there have been circumstantial counter-arguments. It was argued that observed halo ellipticities imply low collisionality (e.g. Buote et al. 2002; Miralda-Escudé 2002). Such limits cannot apply to systems that are elliptical due to tidal disturbance, rotation or non-radial pulsations. Newer SIDM simulations reject the spherical prediction and imply a looser limit $\zeta < 1 \text{cm}^{2} \text{g}^{-1}$ (Peter et al. 2012). Other papers inferred tight limits on SIDM collisionality, from the growth of supermassive black holes (SMBHs; $\zeta < 0.02 \text{cm}^{2} \text{g}^{-1}$; Ostriker 2000; Hennawi & Ostriker 2002), and the survival of cluster members ($\lesssim 0.35$ or $\gtrsim 10^{2} \text{cm}^{2} \text{g}^{-1}$; Gnedin & Ostriker 2001). However these investigations were in a way self-fulfilling, because they assumed cuspy initial conditions, with a strong thermal inversion at the centre. This choice was inconsistent with cored profiles expected from SIDM.

Gravitational lensing models of merging galaxy clusters give conflicting signs about DM physics: sometimes the inferred DM seems collisionless, following galaxies and separating from gas ($\zeta < 1.25 \text{cm}^{2} \text{g}^{-1}$; see e.g. Clowe et al. 2006; Randall et al. 2008). In other systems DM seems collisional, following the gas or separating from galaxies (e.g. Mahdavi et al. 2007; Williams & Saha 2011; Jee et al. 2012). The derived limits depend on assumed projected geometries and histories of unique systems; these are each open to specific counter-explanations and rebuttals. For instance, if the acclaimed ‘Bullet Cluster’ were truly a direct-hit merger then ram pressure and shocks would quench star formation. This is not observed (Chung et al. 2009).

Saxton & Wu (2008) modelled galaxy clusters in which the halo could possess extra heat capacity (e.g. composite SIDM particles). Stationarity in the presence of gas inflows requires a central accretor exceeding some minimum mass. For the range of DM properties giving plausible cluster and core profiles, a part of the inner halo terrors towards local gravitational instability, implying likely collapse to form SMBHs of realistic sizes. Saxton & Ferreras (2010) modelled elliptical galaxies comprising a polytropic halo plus coterminous collisionless stars. Observed stellar kinematics of elliptical galaxies were fitted by the same class of DM that gives the most realistic clusters.

This paper is a further theoretical study of spheroidal galaxies. Their stability properties will imply constraints on the end-points of galaxy evolution. If galaxies do possess SIDM haloes then their oscillatory modes will open the future possibility of skotoseismology – acoustic diagnosis of DM haloes that pulsate. The paper’s organization is as follows. Section 2 formulates the stationary model and perturbations; Section 3 describes the oscillatory solutions and instabilities; Section 4 discusses applications and interpretations; and Section 5 concludes.

## 2 FORMULATION

This paper investigates spherical galaxies comprising self-interacting DM and collisionless stars in near-equilibrium configurations. The SIDM halo is polytropic, adiabatic, non-singular and radially finite (Section 2.1). The system is isolated: unconfined and unexcited by any external medium. Stellar orbits are isotropically distributed, in an observationally motivated density profile (Section 2.2). Oscillations and instabilities are calculated in response to radial perturbations (Sections 2.4–2.6). Mass displacements are non-uniform, and at any location the two constituents may be displaced with differing amplitudes and phases. The perturbed variables are non-singular at the origin, and the outer boundary of the halo is free (Section 2.7).

### 2.1 Dark matter halo profile

For each mode in which ideal collisional gas particles are free to move, the energy per particle is $\frac{1}{2}kT = \frac{1}{2}\mu\sigma_{v}^{2}$ where $k$ is Boltzmann’s constant, $T$ is temperature, $\mu$ is the particle mass and $\sigma_{v}$ is a characteristic (one-dimensional) velocity dispersion. If SIDM gas has $F$ effective thermal degrees of freedom that are locally in equipartition (due to collisionality or collective effects), then the specific heat capacities at constant volume and pressure are $\frac{C_{V}}{k_{B}} = \frac{1}{2}Fk$ and $C_{P} = C_{V} + k$. The ratio of specific heats

$$\gamma \equiv \frac{C_{P}}{C_{V}} = 1 + \frac{2}{F}. \quad (1)$$

From the first law of thermodynamics, adiabatic processes imply a polytropic equation of state for the dark matter:

$$P_{\sigma} = \rho_{\sigma} \sigma_{\sigma}^{2} = s \rho_{\sigma}^{\gamma}, \quad (2)$$

or (equivalently) the mass density is $\rho_{\sigma} = Q \sigma_{\sigma}^{F}$ where $s$ is a pseudo-entropy, and $Q = s^{-F/2}$ is a generalized phase-space density. In adiabatic processes, $s$ and $Q$ remain constant.

The micro-physical meanings of 'F' deserve further remarks. Point-like particles moving in three spatial dimensions have
$F = 3$, due to translational kinetic energy. Composite particle substructure raises $F$ (e.g. rotational and vibrational modes). A diatomic gas has $F = 5$ because two parts are moving in 3D space, at a fixed separation. Incompressible fluids are more constrained and $F = 0$. For a relativistic or radiation dominated plasma, $F = 6$. If particles were massive enough (short wavelength) to traverse hidden spatial dimensions, then extra translational modes would raise $F$. An isothermal medium has $F \to \infty$.

Polytropic fluids also emerge in some non-classical theories of DM. Scalar field and boson condensate models usually imply $F = 2$, but possibly other $F$ values depending on an index in the interaction potential (e.g. Goodman 2000; Peebles 2000; Arbe\-\-y, Lesgour\-\-gues & Salati 2003; Böhmer & Harko 2007; Chavanis & Delfini 2011; Harko 2011a,b; Robles & Matos 2012). This paper is agnostic about which particle theory causes the value of $F$, which should ultimately be constrained by observations.

The radial structure of a spherical halo is determined by its momentum equation,

$$\rho_{\text{d}} \frac{d^2 r}{dr^2} + \frac{d P_{\text{d}}}{dr} = - \frac{G \left( m_\odot + m_\bullet \right) \rho_{\odot}}{r^2}. \tag{3}$$

The stellar mass enclosed at radius $r$ is $m_\odot = m_\odot(<r)$ (see Section 2.2). The corresponding enclosed dark mass is obtained by numerical integration of

$$\frac{dm}{dr} = 4\pi r^2 \rho_{\text{d}}. \tag{4}$$

In hydrostatic equilibrium, condition (3) becomes

$$\frac{d \sigma^2}{dr} = - \frac{2}{F + 2} \frac{G \left( m_\odot + m_\bullet \right)}{r^2}. \tag{5}$$

Neglecting any compact central mass (e.g. black hole), a non-singular boundary condition applies at the origin ($d \sigma^2/dr|_{r=0} = 0$). The equilibrium dark halo is adiabatic, and its pseudo-entropy $s$ is spatially uniform. This may occur in nature if a galaxy halo is well mixed, or if $s$ is set by universal constants of the SIDM particle.

For $-2 < F < 10$ a halo of finite mass truncates ($\rho_{\text{d}} \to 0$) at a finite outer radius $R$. For other $F$, any finite mass equilibrium model has infinite radius: it boils off into the void, or fails to condense from the cosmic background in the first place. The high limit ($F = 10$) is a generalized Plummer (1911) model, with infinite radius and finite mass. Haloes may also seem radially infinite in some SIDM simulations with $F = 3$ (e.g. Rocha et al. 2012). This may be due to ongoing supersonic infall of neighbouring objects, or because the scattering is weak enough to leave the outskirts collisionless. If the SIDM ‘strength’ is increased (greater scattering or, long-range ‘dark magnetism’ effects) then polytropic behaviour affects more of the halo. Beyond some threshold strength, a definite surface must occur. Rocha et al. (2012) note a temperature drop at large radii in their largest $\zeta$ simulation. This is attributable to incipient truncation.

The value of $F$ determines the natural ratio of core to system radii: for greater $F$ the core is smaller (e.g. compare cases in Fig. 1). If a dark mass dominates, the circular orbit velocity profile ($V = \sqrt{Gm_\odot/r}$) peaks at a radius $R_b$, which occurs at $R_b/R \approx 0.750, 0.0782$ when $F = 3$ and $F = 9$, respectively. In these cases, the dark density index drops to $d \ln \rho_{\text{d}}/d \ln r = -1$ at radii $R_1/R \approx 0.379, 0.0285$; and index $-2$ at $R_2/R \approx 0.520, 0.0459$. In other words, a strong-SIDM core with $F = 9$ is radially smaller than a tenth of an $F = 3$ core. Appendix A lists standard radii for other $F$ cases. Adding stars distorts these ratios but the general trend with $F$ persists. Fig. 1 (right-hand panel) shows how a stellar component can shrink and steepen the dark core. For comparison, the “scale radius” of simulated CDM haloes (Navarro et al. 1996b) is $R_s$, and $R_s \approx 2.163R_b$. For the Burkert (1995) cored profile with scale radius $R_b$, the key radii are $R_1 \approx 0.657R_b$, $R_2 \approx 1.521R_b$, and $R_0 \approx 3.244R_b$. Simulated SIDM cores fitted by Rocha et al. (2012) ($R_0 \approx 0.7R_b$) appear consistent with Burkert’s parameterization. The ratios $R_1: R_2: R_0$ are similar for Burkert models and $F = 3$ polytropes, but less similar for other $F$.

These ideal core sizes are upper limits. In collisional SIDM, weaker $\zeta$ delays the formation of full-sized cores. Current SIDM simulations focus on the $F = 3$ case, which predicts cores larger than those observed, unless the scattering cross-section is kept small ($\zeta < 1 \text{ cm}^2 \text{ g}^{-1}$; e.g. Yoshida et al. 2000; Aradajis, Bautz & Garmire 2002; Katzerg, Biviano & Mazure 2004; Rocha et al. 2012; Vogelsberger et al. 2012). The alternative is to vary $F$: the range of $7 \lesssim F < 10$ results in galaxy clusters with realistic $\sim 10^1 \sim 10^2 \text{ kpc}$

![Image of radial density profiles](https://academic.oup.com/mnras/article-abstract/430/3/1578/977143)
cores (Saxton & Wu 2008). In principle, haloes with \( F \geq 7 \) could produce cores of realistic radius for arbitrarily large \( \varepsilon \). Such \( F \) values also provide \( R_s/R \) small enough to be consistent with observations of large flat regions in disc galaxies’ rotation curves. The choice of \( F \geq 7 \) is therefore an astrophysically preferable subdomain, but for the sake of generalization this paper considers a wider interval, \( 2 \leq F \leq 9 \).

2.2 Stellar mass profile

The assumed stationary distribution of the stars resembles the empirical Sérsic (1968) form, where the surface density at projected radius \( R \) is

\[
\Sigma_\star(X) \propto \exp \left[-b(X/R_\circ)^{1/n}\right],
\]

and \( n \) is the ‘Sérsic index’. The ‘effective radius’ \( R_\circ \) encircles half the starlight. Assuming that the stellar mass-to-light ratio is constant, there exist exact deprojections of (6) that express the 3D density profile (e.g. Mazure & Capelato 2002), but these involve exotic special functions. We will instead use the approximate deprojection of Prugniel & Simien (1997), with density and mass profiles

\[
\rho_\star = \rho_c \exp \left[b - b(r/R_\circ)^{1/n}\right] \left(r/R_\circ\right)^{-p} \quad \text{and}
\]

\[
m_\star(<r) = 4\pi n b^{(p-3)} e^b \rho_c R_\circ^3 \Gamma \left[n(3-p), b(r/R_\circ)^{1/n}\right],
\]

where \( \Gamma(a, u) \) is the lower incomplete gamma function. The index \( n \) determines \( b \) and \( p \) (Ciotti & Bertin 1999; Lima Neto, Gerbal & Márquez 1999; Márquez et al. 2000). Adopting \( n = 4 \) to match the classic de Vaucouleurs (1948) profile, the values are \( b \approx 7.669 \) and \( p \approx 0.8510 \). For this standard choice, the total stellar mass is \( M_\star = 45.79\rho_c R_\circ^3 \), and that within the half-light radius is \( m_\star(R_{1/2}) = 19.03\rho_c R_\circ^3 \).

The frequency scale of the stellar arms will be abbreviated as

\[
\omega_c = \sqrt{G \rho_c}.
\]

Given the properties of large elliptical galaxy like M87, \( \rho_c \sim 0.04 M_\odot \text{pc}^{-3} \) and \( \omega_c \approx 0.01 \text{Myr}^{-1} \) (Murphy et al. 2011). For a compact elliptical such as M32, \( \rho_c \sim 14 M_\odot \text{pc}^{-3} \) and \( \omega_c \approx 0.3 \text{Myr}^{-1} \) (Mateo 1998; Choi, Guhathakurta & Johnston 2002). For ‘red nuggets’ at high redshift, \( \rho_c \gtrsim 3 M_\odot \text{pc}^{-3} \) and \( \omega_c \lesssim 0.05 \text{Myr}^{-1} \) (e.g. van de Sande et al. 2011; Toft et al. 2012). If stars were the only component present, the Keplerian period at \( R = R_\circ \) would be \( t_K = 1.440/\omega_c \), and the angular frequency \( \kappa_c = 4.363 \omega_c \).

It is convenient to write the density gradient index as

\[
\alpha_\star \equiv \frac{d \ln \rho_\star}{d \ln r} = p + \frac{b(r/R_\circ)^{1/n}}{n},
\]

noting that \( \alpha_\star \to p \) near the origin. In central regions a modified radial coordinate, \( \zeta \equiv r^{1-p} \), is useful to soften terms involving the density.

\[
\frac{d}{d \zeta} \rightarrow \frac{r^p}{1 - p} \frac{d}{d r} \ldots \ldots ;
\]

as the \( r^p \) factor cancels the singular part of \( \rho_\star \) in any ODEs.

The stellar velocity dispersions may, in principle, differ in the three orthogonal directions: radial (\( \sigma_r \)), polar (\( \sigma_\theta \)) and azimuthal (\( \sigma_\phi \)). We will consider spherical models in which the transverse velocity dispersions are equal, \( \sigma_\theta = \sigma_\phi \equiv \sigma_\perp \). The radial and transverse velocity dispersions are related by defining an anisotropy profile,

\[
\beta = \frac{\sigma_r^2}{\sigma_\perp^2} \equiv 1 + \frac{\sigma_r^2}{\sigma_\perp^2}
\]

(e.g. Osipkov 1979; Merritt 1985; Mamon & Łokas 2005). The radial ‘pressure’ of stars is

\[
P_\star = \rho_\star \sigma_r^2,
\]

and its profile is calculated from a spherical Jeans equation for momentum, expressed in a frame comoving with the mean radial velocity (\( v_\phi^\star \)) of a mass ‘packet’ of stars (Jeans 1915; Binney & Tremaine 1987),

\[
\frac{dP_\star}{dr} = g \rho_\star \frac{\sigma_\star^2}{r} \left(2\sigma_\theta^2 - \sigma_\phi^2 - \sigma_r^2\right) - \rho_\star \frac{d^2 \rho_\star}{dr^2},
\]

where \( g \equiv -G(m_\star + m_\bullet)/r^2 \) is the gravitational field.

2.3 Acoustic and orbital time-scales

The acoustic crossing time of the DM from the origin to the outer surface is obtained by radial integration of

\[
\frac{d^2 \rho}{dr^2} = \frac{1}{\sqrt{1 - \rho_0^2}}.
\]

For the stars, the analogous ‘acoustic’ time is

\[
\frac{d \omega_\star}{dr} \equiv \frac{\sqrt{d\rho_\star/dP_\star}}{dP_\star} \equiv \left(\frac{d\rho_\star}{dr}\right) \left(\frac{dP_\star}{dr}\right).
\]

Assuming stationarity, then substituting (10) and (14) simplifies (16): e.g. if \( \beta = 0 \) then \( \omega_\star = \sqrt{\sigma_\phi^2/8\pi\rho_0} \). From dimensional considerations, the crossing time should scale with the halo radius as \( t_c \sim R^{1/2} \). Appendix B shows examples of the exact relations. When calculating the natural modes of the galaxy, we should expect oscillations to scale in proportion to the angular frequencies \( \Omega_\alpha = 2\pi t_\alpha \). Once the eigenfrequencies (\( \omega \)) of a galaxy are known, one can compare them to local Keplerian frequencies,

\[
\kappa = \sqrt{\frac{G(m_\star + m_\bullet)}{r^3}}.
\]

In regions where \( \kappa \gg \omega \), perturbations should only affect stellar orbits adiabatically. Wherever \( \kappa \ll \omega \), tidal shocking may occur when disturbances exceed the linear regime.

2.4 Perturbative displacements and gravitation

For each of the variables describing local properties of the DM or stars, we define a first order perturbation. For instance for a density variable we have

\[
\rho = \rho_0 + \bar{\rho} e^{i \omega t} = \rho_0 \left(1 + \lambda_\rho e^{i \omega t}\right)
\]

where the subscript ‘0’ denotes the stationary solution; the ‘~’ part describes the response to perturbation; and the \( \lambda \) function describes the perturbation in relative terms (e.g. \( \lambda_\theta = \bar{\rho}/\rho_0 \)). Perturbed variables such as \( \lambda_\rho \) and \( \bar{\rho} \) are complex. The small constant \( \varepsilon \) is the scale of perturbation; higher order terms are discarded (\( \varepsilon^2 \to 0 \)). The frequency of perturbation is complex (\( \omega = \omega_h + i \omega_i \)) with the real part describing oscillations and the imaginary part describing the tendency to growth or decay.

As in Eddington (1918) we shall consider Lagrangian perturbations. For a mass shell residing at radius \( r \), in equilibrium, we write its displacement and velocity under perturbation:
The DM and stars at any $r_o$ may be displaced differently, so we treat them in two slightly different Lagrangian frames, and distinguish their coordinates carefully:

$$r_d = r_o + \tilde{r} e^{i \omega t},$$

$$v = i \tilde{r} \omega e^{i \omega t}.$$ \hfill (19)

The presence of the stars is felt through the gravitational perturbation $\tilde{g}_s$ obtained from (29).

Expressions for the perturbation of each radial gradient term depend on whether we consider the comoving frame of the stars or the DM. Let us abbreviate ‘+’ and ‘−’ for ‘$\star$’ and ‘d’ subscripts. Mass conservation requires that $\rho_\pm r_\pm^2 d r_\pm = \rho_\pm r_0^2 d r_0$, and then in the ± frame we have the operators

$$\frac{d}{d r} = \frac{d}{d r_0} + (\lambda_{\mp} + 2 \lambda_\pm) \epsilon e^{i \omega t} \frac{d}{d r_0}.$$ \hfill (25)

The gravitational field provides the only coupling between components. Consistent evaluation of its perturbations (in both frames) requires special care. The stellar and dark shells that rest at $r_o$ in the stationary solution will feel the same field strength there ($g_0$) but the perturbed versions of $g$ differ since the shells oscillate through displacements with different phase and amplitude. Define $g_\pm = g_0 + g_\pm e^{i \omega t}$. \hfill (27)

For Newtonian gravity, we have $g_0 = -G m_0 / r_0^2$ with $m_0 = m_{\ast,0} < r_{\ast,0} = m_{\ast,0} + m_0$ being the total mass contained within $r_0$ in equilibrium. Poisson’s equation applies at all orders of perturbation, which implies

$$\frac{d g_\pm}{d r_\pm} = -\frac{2 g_\pm}{r_\pm} - 4 \pi G (\rho_\pm + \rho_\ast).$$ \hfill (28)

Combining (26) and (27) gives the perturbed part of $g_\ast$ as felt by the ± component. Denoting $g_\pm \equiv -G \tilde{m}_\pm / r_\pm^2$ gives

$$G \frac{d \tilde{m}_\pm}{d r_0} = -2 \tilde{g}_0 r_0 (\lambda_{\mp} + 3 \lambda_\pm)$$

$$- 4 \pi G r_0^2 [2 \rho_{\mp} \lambda_{\mp} + \rho_\ast \lambda_\mp - \lambda_{\pm} + 3 \lambda_\pm] \right),$$ \hfill (29)

which has the advantage of being non-singular and $\tilde{m}_+ = \tilde{m}_- = 0$ at the origin.

2.5 Dark matter perturbations

The DM perturbations are considered adiabatic ($\epsilon$ is spatially and temporally constant). The equation of state links the pressure and density perturbations:

$$\lambda_{\pm} \equiv \frac{\tilde{P}_\pm}{\tilde{\rho}_\pm} = \gamma \lambda_{\mp} \equiv \gamma \frac{\tilde{P}_\mp}{\tilde{\rho}_\ast}.$$ \hfill (30)

Mass conservation then yields

$$\frac{d \tilde{r}_\pm}{d r_0} = -\lambda_{\pm} - 2 \lambda_\ast = -\frac{\tilde{P}_\pm - \tilde{\rho}_\pm}{\gamma \tilde{\rho}_\pm} - \frac{\tilde{r}_\pm}{r_0}. \hfill (31)$$

$$\frac{d \lambda_{\pm}}{d r_0} = -\frac{1}{r_0} \left( \lambda_{\pm} + 3 \lambda_\ast \right).$$ \hfill (32)

The Lagrangian equation for momentum (3) decomposes under (26) to give the hydrostatic condition,

$$\frac{d \tilde{P}_\pm}{d r_0} = \rho_{\pm} g_0,$$ \hfill (33)

and a perturbation equation, writable in two forms:

$$\frac{d \tilde{P}_\pm}{d r_0} = \rho_{\pm} \left( \alpha^2 r_o \lambda_{\pm} - 2 g_0 \lambda_{\ast} + \tilde{g}_s \right) = \rho_{\pm} g_0 \Lambda_{\pm},$$ \hfill (34)

$$\frac{d \lambda_{\pm}}{d r_0} = \frac{g_0}{\alpha^2 \Lambda_{\pm}} (\Lambda_{\ast} \pm - \lambda_{\pm}),$$ \hfill (35)

with the abbreviation

$$\Lambda_{\pm} \equiv \frac{\alpha^2 r_o \lambda_{\pm} - 2 g_0 \lambda_{\ast} + \tilde{g}_s}{g_0} = m_{\pm} \left( 2 + \frac{\alpha^2}{\kappa^2} \right) \lambda_{\pm}.$$ \hfill (36)

The presence of the stars is felt through the gravitational perturbation $\tilde{g}_s$ obtained from (29).

2.6 Stellar perturbations

For the collisionless stellar component, the perturbed mass conservation equation resembles the DM version, but without an equation of state to link density and pressure.

$$\frac{d \tilde{r}_\ast}{d r_0} = -\tilde{P}_\ast - 2 \tilde{r}_\ast = -\lambda_{\ast} - 2 \lambda_\ast;$$ \hfill (37)

$$\frac{d \lambda_{\ast}}{d r_0} = -\frac{1}{r_0} \left( \lambda_{\ast} + 3 \lambda_\ast \right).$$ \hfill (38)

Next, we define perturbed forms of the radial and transverse velocity dispersions of the stars,

$$\sigma_\parallel^2 = \sigma_\ast^2 + \alpha^2 \epsilon e^{i \omega t},$$ \hfill (39)

$$\sigma_\perp^2 = (1 - \beta) \sigma_\ast^2 + \alpha^2 \epsilon e^{i \omega t},$$ \hfill (40)

where $\sigma_\parallel^2$ and $\sigma_\perp^2$ are the perturbations of the radial and transverse velocity dispersions, respectively. The stationary models in this paper assume isotropy ($\beta = 0$ everywhere) although orbital anisotropies occur transiently during perturbations. Nonetheless let us retain $\beta$ terms in the general formulation. If the perturbations are radial then there is no torque. If the specific angular momentum of a mass element is conserved during a radial displacement then each local value of $r^2 \sigma_\parallel^2$ stays constant, eliminating a variable:

$$\tilde{\sigma}_\parallel = \frac{2 (1 - \beta) \tilde{\sigma}_\ast}{r_0} \sigma_\parallel^2 = -2 (1 - \beta) \sigma_\ast^2 \lambda_{\ast}.$$ \hfill (41)

Expressing the radial pressure perturbation as $\tilde{P} = \rho_\ast \sigma_\parallel^2 + \rho_\ast \sigma_\perp^2$ provides a convenient relation,

$$\tilde{\sigma}_\parallel^2 = \sigma_\ast^2 \left( \lambda_{\ast} - \lambda_{\ast} \right).$$ \hfill (42)

The Jeans equation (14) separates into stationary and perturbed parts,

$$\frac{d \tilde{P}_\ast}{d r_0} = g_\ast \rho_\ast - \frac{2 \tilde{r}_\ast}{r_0}.$$ \hfill (43)
\[ \frac{d\lambda_{P*}}{dr} = \left( \omega^2 \tilde{r}_* + \tilde{g}_* - 2g_\omega \lambda_{P*} - g_\sigma \lambda_{P*} \right) \nonumber \]
\[ \quad - \frac{2}{r_0} \left[ (2 - 5\beta)\lambda_{\tau*} + (1 - \beta)\lambda_{P*} - \lambda_{\tau*} \right]. \quad (44) \]

Given another Jeans equation for evolution of the radial stellar pressure (Appendix C),
\[ 0 = \frac{dP_\star}{dt} + 2\nu \frac{d^2P_\star}{dr^2} + \left( 3P_\star + 2\rho_\star v_\star^2 \right) \frac{dP_\star}{d\tau} \nonumber \]
\[ + \frac{2\rho_\star v_\star}{r} \left( v_\star^2 + 2\sigma_\tau^2 - \sigma_\nu^2 \right), \quad (45) \]
and performing the perturbation expansion yields an equation in the stationary variables (which reduces to a tautology, \( \theta = 0 \)) plus an equation in perturbed variables. The latter, first order equation is
\[ 0 = i\omega \tilde{P}_\star + 3i\omega_0 \frac{d\tilde{P}_\star}{d\tau} + 2i\omega(1 + \beta)P_\star \lambda_{\tau*}. \quad (46) \]

Substituting (37) into (46) leads to an algebraic relation,
\[ 0 = \lambda_{P*} - 3\lambda_{\nu*} + (2\beta - 4)\lambda_{\nu*} \quad (47) \]
which eliminates the \( \lambda_{\nu*} \) terms in (37), (38) and (44).

### 2.7 Boundary conditions

At the origin, by symmetry, there cannot be any oscillatory displacement of mass, and therefore
\[ \lim_{r_0 \to 0} \tilde{r}_* = \lim_{r_0 \to 0} \tilde{r}_0 = 0. \quad (48) \]
In relative terms the offset is free; any small non-zero value of \( \lambda_{\tau*} \) is acceptable at the inner boundary. Regularity of the mass equations (32) and (38) then gives
\[ \lim_{r_0 \to 0} \lambda_{P*} = -3\lambda_{\tau*} ; \quad (49) \]
\[ \lim_{r_0 \to 0} \lambda_{P*} = -3\nu \lambda_{\nu*}. \quad (50) \]
Combining (47) and (49) implies
\[ \lim_{r_0 \to 0} \lambda_{P*} = -(5 + 2\beta)\lambda_{\tau*}, \quad (51) \]
and \( P_\star = P_{0,0} \lambda_{P*} \) using the central pressure from the stationary model. Analogously, the central pressure oscillation of DM is \( \tilde{p}_d = P_{0,0} \lambda_{P*} \). Assuming that the stars and DM are coving in the dense inner regions requires \( \lambda_{\tau*} = \lambda_{\tau0} \) at the origin. This completes the set of inner boundary conditions.

As \( \sigma_{\nu0} \to 0 \) at the halo surface, (35) usually tends towards singularity. To avoid this, the natural modes must satisfy an outer boundary condition:
\[ \lim_{r \to R} (\lambda_{P0} - \Lambda_{P0}) = 0. \quad (52) \]
Condition (52) is only satisfied at \( \omega \) eigenfrequencies. The stellar mass profile lacks any finite outer boundary. Therefore there are no outer boundary conditions applicable to the \( \star \) perturbed variables.

### 2.8 Numerical integration and solution

To calculate the stationary profiles, trial values are set for the pseudo-entropy constant \( s \) and the central velocity dispersion of DM (\( \sigma_{0}^2 \mid_{r=0} \)). The system of coupled differential equations (4) and (5) is integrated numerically as an initial value problem in radius (using the RKF45MP and RKF89P integrators from the GNU SCIENTIFIC LIBRARY1). In practice, it is necessary to change the independent variable (the integration coordinate) partway through the integral. Near the origin a softened radial coordinate (11) improves accuracy in the stellar cusp. In the outskirts it is more practical to use \( \sigma_{0}^2 \) or \( \ln(\sigma_{0}^2) \) as the independent variable, and integrate to a standard limit at \( \sigma_{0}^2 = 10^{-8} \) or \( 10^{-12} \). (The chosen limit does not change any results, as long as it is consistent across all runs.) At the outer boundary (radius \( R \)) the total dark mass is recorded (\( M_r \equiv M_{\star} \mid_{r=R} \)).

At a higher level of the program, an amoeba search routine iterates this integration for different trial values of \( (s, T_0) \), seeking to minimize the difference between the obtained \( (R, M_r) \) and target values chosen by the user. Usually these targets are set to collect a sequence of solutions of different \( R \) but equal \( M_r \), i.e. describing galaxies of identical mass composition but different compactness.

Once the \( (s, T_0) \) values are known for a desired stationary model, the stellar pressure profile (43) is integrated inwards from infinity and tabulated finely. Trial values are chosen for the complex frequency of perturbation, \( \omega \). The stationary model’s ODEs are integrated again simultaneously with the coupled ODEs for the perturbed variables: (29), (31), (32), (34), (35), (37), (38), (43) and (44). When the halo surface is reached \( (r \to R) \) the outer boundary condition (52) is tested there to assess the closeness of \( \omega \) to an eigenvalue. Using a complex test score,
\[ Z = Z_R + iZ_k = (\lambda_{P0} - \Lambda_{P0}) \]
(53)
a root-finding routine refines the \( \omega \) trial values until \( Z \to 0 \). To seek eigenfrequencies on the real and imaginary axes of \( \omega \), a bisection root-finder is sufficient. To find eigenfrequencies elsewhere in the complex plane, a two-dimensional amoeba search is unleashed on the \( Z = Z(\omega) \) landscape. Due to symmetries of the ODEs in their \( \sigma_{0}^2 \) factors, the eigenfrequencies can be: purely imaginary (Section 3.1); purely real (Section 3.2); or occurring in complex conjugate pairs (Section 3.3).

If numerical integration traverses many orders of magnitude in \( r \) then unfortunately there are instances in which one or more of the \( \lambda \) variables overflow or underflow the computer variables. In practice, numerical integration towards the outer boundary produces diverging \( |\lambda_{P0}| \), if \( \omega \) is poorly chosen. This must be tackled through a practical numerical trick. One can pick the worst offending variable \( \lambda_j = \lambda_j + \lambda_k \) and assign its magnitude to a new dependent variable, \( \hat{\xi} \equiv (\lambda_k^2 + \lambda_j^2)^{1/2} \). The code tracks the logarithmic profile of this quantity, by numerically integrating an extra ODE simultaneously with the other ODEs:
\[ \frac{d\ln \hat{\xi}}{dr} = \frac{1}{\lambda_k^2 + \lambda_j^2} \left( \lambda_k \frac{d\lambda_k}{dr} + \lambda_j \frac{d\lambda_j}{dr} \right). \quad (54) \]
All the \( \lambda \) variables can then be expressed in a locally normalized form, \( \lambda_k \equiv \lambda_k \hat{\xi} \). If each corresponding ODE \( (k) \) is abbreviated as a function \( \lambda_k (\xi) = \xi \hat{\lambda}_k \frac{d\ln \hat{\xi}}{dr} \) then
\[ \frac{d}{dr} \hat{\lambda}_k = \frac{\hat{h}_k(\lambda_k, \lambda_j, \ldots)}{\hat{\xi}} - \hat{\lambda}_k \frac{d\ln \hat{\xi}}{dr} \]
\[ = \hat{h}_k(\hat{\lambda}_k, \hat{\lambda}_j, \ldots) - \hat{\lambda}_k \frac{d\ln \hat{\xi}}{dr}. \quad (55) \]
The simplification \( h_k(\lambda) \hat{\xi} = h_k(\hat{\lambda}) \) depends on the linearity of the perturbations. Evolving (54) for \( \xi = \xi (r) \) in conjunction with the

1 http://www.gnu.org/software/gsl/ (last accessed 2012 December 11)
ODEs of the modified form (55) keeps all the $\lambda$ variables normalized to moderate values that do not overflow the floating-point representations. Keeping the running values small also avoids driving the numerical integrators into stiff or unreliable domains. Near the outer boundary, where $\lambda_{F*}$ tends to become large, the running values of $\lambda_{F*}$ and $\hat{\lambda}_{F*}$ can be kept on the order of unity or smaller. This prevents numerical noise in the sign of a score $Z \equiv Z / \xi$ that is based on (53). For presentational purposes, the $\lambda$ profiles can be rescaled by $\xi$ afterwards in post-processing. For brevity and computational ease, we choose physical units such that $G = 1$, along with the stellar half-light radius $R_e = 1$ and the stellar density there $\rho_0 = 1$. Results can be scaled into metric or astronomical units by inserting fiducial values of these constants for a known galaxy.

Families of models will be labelled by their total stellar mass fractions, $F_* \equiv M_*/(M_* + M_\odot)$. This paper presents galaxy models for a variety of choices of $F_*$. A fiducial ‘cosmic’ baryon fraction $\approx 0.163$ is defined as in Saxton & Wu (2008) and Saxton & Ferreras (2010). For poorly understood reasons, observed densities of baryons (e.g. Persic & Salucci 1992; Bell et al. 2003; Xue et al. 2008; Anderson & Bregman 2010; McGaugh et al. 2010) and DM (Karachentsev 2012) differ from cosmic mean estimates (except in rich galaxy clusters). There may be an unseen ambient sea of both materials, free from self-bound objects. Matching the abundances of observed galaxies to simulated CDM halo distributions implies baryon fractions $\lesssim 0.04$ depending on galaxy mass (e.g. Guo et al. 2010; Papastergis et al. 2012). This may mean that galaxies lost most of their primordial gas endowments, or it may indicate that real halo populations differ from CDM. Cosmological initial conditions are beyond the scope of the present analyses, so $F_*$ is for now a free parameter. Nonetheless, Section 3.1 shows that spherical SIDM galaxies have stability limits that may influence the emergent distribution of $F_*$. In collisionless CDM, the outer halo density drops gradually without a well-defined boundary. These simulated objects are conventionally circumscribed by the ‘virial radius’, containing some multiple of the cosmic critical density, e.g. $\rho_0 = 100\rho_{\text{crit}} \approx 1.4 \times 10^3 M_\odot$ kpc$^{-3}$. If the natural density and radius scaling of SIDM haloes are not too different from CDM then we can estimate upper limits to the halo radii of representative galaxies, $R^3 \lesssim 3 M_*/4\pi\rho_0 F_*$. For the example of M87, $R/R_e \lesssim 30/\sqrt{F_*/3}$; for M32, $R/R_e \lesssim 200/\sqrt{F_*/3}$; for a red nugget, $R/R_e \lesssim 120/\sqrt{F_*/3}$; and for the Fornax dwarf spheroidal galaxy (dSph) $R/R_e \lesssim 10/\sqrt{F_*/3}$ (data: Walker et al. 2009). It is, however, unclear whether this rule-of-thumb applies to SIDM, which warrants new cosmic structure formation calculations.

3 RESULTS

3.1 Stability properties

Classical arguments based on secular conditions suggested that $F > 6$ polytropes should be unstable to runaway expansion or collapse (e.g. Ritter 1878; Emden 1907; Chandrasekhar 1939). This tendency manifests itself (in the present formulation) as a conjugate pair of modes with imaginary $\omega$ values. These ‘collapse modes’ or ‘growing modes’ involve no oscillatory motion, only exponential growth or decay. In practice, such modes are found by searching along the imaginary axis of the $\omega$ Argand plane.

Next, a polytropic halo can be tested in a static stellar background (acting as a fixed potential, with $\lambda_{\hat{F}*} = \hat{\lambda}_{F*} = \hat{m}_{\hat{*}} = 0$). The stability properties can be compared to the generic system where both the stars and DM are enabled to move radially. Fig. 2 illustrates modes of a particular model with $F = 8$, comparing the results if the stars are fixed or free. With stars held fixed, a collapse mode appears. With the stars enabled to move, the collapse mode vanishes from the model shown. This stabilization is typical of galaxies where the halo extends farther than $R > 10R_e$.

The stabilization can be explained qualitatively as follows. SIDM fluid carries pressure waves, and $F > 6$ models are collapsible because pressure may not rise strongly enough to counteract compressive perturbations. Stars, however, tend to persist on their orbits unless tidally disturbed, and do not conduct true sound waves. They deafen the propagation of overdensities, and exert a local damping and stabilizing effect. For many galaxy models this suffices to confer global stability. However, as we shall see below, in extreme cases the stabilization is insufficient (Section 3.1.1) or a new type of instability emerges (see Section 3.1.2 and 3.1.4).

The rest of this paper assumes that the DM and stars are gravitationally coupled and both respond to radial perturbations. Exhaustive numerical surveys have located the collapse modes as functions of: the equation of state parameter ($F$), the radial extent ($R$) of the halo and the global richness of stars ($F_*$). Fig. 3 illustrates the $|\omega|$ collapse frequencies as functions of $F$ and $R$ (colours), for models where DM dominates (small $F_*$, in respective panels). Fig. 4 compares the modes when $DM$ is present in the ‘cosmic’ abundance, and a case that is DM-poor. Realistic haloes could have radii of tens or hundreds of kpc, perhaps a range $10 \lesssim R/R_e \lesssim 10^3$ depending on the galaxy. Ultra-dense and ultra-diffuse cases are also plotted to clarify some theoretically notable modes. Two distinct types of modes occur in particular domains.

3.1.1 Slow-growing modes

If $R$ is large (a very diffuse halo) and $F > 6$ then an instability occurs. This is a re-emergence of the instability of classic one-component polytropes with $F > 6$. In these cases the stars are unable to suppress instability because the stellar density is too low in the outer halo. This happens over a larger domain of $R$ if $F_*$ is lower (DM dominates stars). As $R$ is increased, the mode turns on suddenly: for $R$ below this threshold, a galaxy is safe; for slightly
Galaxy halo stability

Figure 3. Frequencies of the halo collapse modes as a function of the outer radius of the halo, for models dominated by dark mass. Curves are coloured in spectral order to show the SIDM thermal degrees of freedom: from violet ($F = 2$) to red ($F = 9$). Panels portray different values of the global stellar mass fraction. As $\mathcal{F}_* \approx 0.6$ is increased, ‘fast’ modes rise in frequency, and ‘slow’ modes (bottom-right curves) recede to affect only the most radially extended haloes.

larger $R$ the mode operates on its shortest time-scale; for very large $R$ the eigenfrequency diminishes ($|\omega| \sim R^{-3/2}$). In a survey of configurations with $7 \leq F \leq 9$ in the range $0.0001 \leq \mathcal{F}_* \leq 0.64$, the frequency scale of these ‘slow’ growth modes is lower than the effective stellar frequency scale, $|\omega| < 0.02 \Omega_*$. In other words, the $e$-folding duration of the growth/collapse would last several tens of orbits for most stars. The slow mode straddles the acoustic frequency scale: always $|\omega| \lesssim 8 \Omega_*$ for $F = 9$ haloes of the largest radii considered, but $|\omega| \lesssim 1 \Omega_*$ more usually. The $|\omega|/\Omega_*$ ratios are lower for $F = 8$ than $F = 9$, and lowest for $F = 7$.

The perturbed mass flux density through a sphere of given radius is $\rho \vec{v} = i \omega \rho_0 \vec{r}$, so the quantity $\rho_0 \vec{r}$ is a useful expression of the radial flows of stars and DM. Fig. 5 depicts the radial flow structure during ‘slow mode’ evolution. Amplitudes for stars tend to be highest in the centre and decline radially. The dark flow has its minimum in the centre. The stars tend to move in the same direction everywhere. DM has alternating ingoing and outgoing layers, separated by nodes at rest. In the interior, DM moves the same way as the stars. Then there is a node at some radius around $\gtrsim 0.1 R_e$ to $\sim 10 R_e$ (farther out for larger $F$). Outside this node, the dark flow is in antiphase with the stars. Some models have further nodes and flow reversals at larger radii in the outer halo (e.g. two far outer nodes in the $F = 9$ example in Fig. 5). The stability analyses are symmetric under time reversal, so it is impossible to say whether a particular sign implies collapse or expulsion of matter. The slow mode could imply contraction of all stars and the inner DM while an external dark layer escapes. It could also mean radial expansion of stars and inner halo, while dark outskirts converge inwards.
3.1.2 Fast-growing modes

If \( R \) is sufficiently small then one or two growing modes appear. They occur for all \( F \) choices. The affected models have hot and dense haloes, with much of the dark mass residing inside the half-light radius. These ‘fast’ modes occur at high frequencies: usually \(|\omega| \gg \omega_s\) (and often \(|\omega| \gg 10^6 \omega_s\) for larger \( F \) cases). Where there are two ‘fast’ eigenvalues, the uppermost is quick compared to the acoustic crossing of the halo \(|\omega| > \Omega_s\), in many models by a few dex). The lower eigenvalue is faster than \( \Omega_s \) if \( F \geq 7 \), but can be slower than \( \Omega_s \) for \( F < 7 \).

Fig. 6 depicts the mass flows due to the lower-\( \omega \) fast mode for illustrative models with halo radius \( R = 10R_e \) and DM enriched over ‘cosmic’ composition. In some ways these profiles resemble the waveforms of slow modes: the stellar amplitude is high in the interior; the DM amplitude is peaked farther out. The DM has one node near the outer boundary, well outside the half-light radius of stars. Except in the outskirts, the dark mass is flowing in the same direction as the stars. For the corresponding higher-[\( \omega \)] fast modes (Fig. 7) there is again just one node in the DM, but both constituents show a dramatic rise in amplitude towards the outer boundary. The rise is steeper for larger \(|\omega|\) and larger \( F \). These fast modes act mainly in the diffuse outskirts; rapidly driving the stars and DM in opposite directions. The core is essentially unaffected.

Since the fast instability occurs for all \( F \), its main causes must be traits of the stellar profile. We might suspect the stellar cusp and its coldness \((\rmd \sigma /\rmd r > 0 \text{ at the origin})\); however tests show fast modes afflicting non-cuspy models with a Plummer (1911) profile of \( \rho_\ast \). If the exponential stellar fringe is a factor, the fast modes might differ in tidally truncated models (generalizing King 1966) which will be analysed in a future paper. In the present model, the index of the stellar gravitational field \((\rmd \ln g_\ast /\rmd \ln r)\) curves significantly near \( r \sim 10R_e \) but settles to the asymptote \(-2\) farther out. If the halo surface occurs at \( r \lesssim 30R_e \), then the oscillatory restoring force varies in form during the cycle. Fast instabilities may result from the inadequacy of restoring forces for haloes that truncate in this region.

3.1.3 Safe configurations

Fast modes and slow modes are absent throughout a special intermediate range of \( R \). In these cases the interaction between stars and the halo suppresses the collapse mode (e.g. in the ‘coupled’ case of Fig. 2). These haloes with \( F > 6 \) are stabilized by the mingling of collisionless stars, contrary to naïve textbook expectations. The three radial domains – defined by the ‘fast’ instability, stable zone and ‘slow’ instability – imply limits on the compactness of the stable end-points of galaxy evolution. The fast-growing modes mean that models are unstable unless \( R \) exceeds some lower limit (which is a slowly varying function of \( F \) and \( F_\ast \)). The slow mode only afflicts \( F > 6 \) cases, and it applies an upper limit on \( R \). This upper limit is a sensitive function of \( F_\ast \). For models with \( F > 6 \) that are poor in stars \((F_\ast \lesssim 0.001)\), the slow- and fast-mode domains leave only a narrow band of safe \( R \). This safe band becomes wider as \( F_\ast \) increases. A numerical survey indicates that the slow mode (upper limit) vanishes for sufficiently high \( F_\ast \). (It persists...
till $F_{\star} \lesssim 0.0004$ for $F = 7$, $F_{\star} \lesssim 0.004$ for $F = 8$ and $F_{\star} \lesssim 0.01$ for $F = 9$.) From this point till $F_{\star} \gtrsim 0.4$, galaxies with arbitrarily wide outer radii $R$ are stable. Fig. 8 illustrates the stability limits on the central dark density ($\rho_{d} |_{r=0}$) as a function of $F_{\star}$. Fig. 9 shows the corresponding limits on the dark mass fraction within the half-light radius [$f_{d} (<1R_{e})$]. The latter quantity can be measured through gravitational lensing and measurements of kinematic tracers. Significantly, it appears that a SIDM halo around an isolated de Vaucouleurs type galaxy is only stable if the dark mass within $1R_{e}$ is less than $\approx 90$ per cent of the total, or less depending on $F_{\star}$ and $F$.

Mass modelling of observed elliptical galaxies often indicates $f_{d} (<1R_{e})$ values in the range of a few tens of per cent (Loewenstein & White 1999; Ferreras, Saha & Williams 2005; Thomas et al. 2005, 2007; Cappellari et al. 2006; Bolton et al. 2008; Tortora et al. 2009, 2012; Weijmans et al. 2010; Saxton & Ferreras 2010; Grillo 2010; Memola et al. 2011; Murphy et al. 2011; Norris et al. 2012; Wegner et al. 2012). If these galaxies are stable within polytropic haloes, we could broadly infer $0.005 \lesssim F_{\star} \lesssim 0.3$ independently of $F$. Observational estimates of the stellar mass are consistent with this range. No large nearby galaxy is yet known to challenge the density limit of the fast instability. A few exceptional galaxies appear DM-poor at large radii (Méndez et al. 2001; Romanowsky et al. 2003; Salinas et al. 2012). If their impoverishment is not an illusion due to orbital anisotropies, then we might explain them in terms of runaway halo loss via either of the instabilities (Section 4.3).

3.1.4 Fast modes in DM-poor systems

Entering the domain where stellar mass is globally dominant (Fig. 4) a new fast mode appears. Beyond some threshold ($F_{\star} \gtrsim 0.4$) an island of the fast modes begins to affect large-$R$ haloes with $F > 6$. This implies that DM-poor systems have a maximum stable radius, just like the DM-rich systems that are subject to slow modes. The $\omega$ values of these fast modes are slowly varying functions of $R$ (horizontal curves in bottom panel, Fig. 4) which implies they arise from the stars rather than the DM crossing time.

As $F_{\star}$ is increased further, the fast mode affects $F > 6$ haloes with all possible $R$ values, and there are no stable configurations with $F > 6$. For example when $F_{\star} = 0.64$ there are no stable models with $F = 8$ or $F = 9$. The fast modes that afflict these DM-poor models are found to have steeply rising waveforms (mainly acting at the halo surface) resembling those in Fig. 7. Results were not sought beyond $F_{\star} = 0.64$, because the system of equations becomes stiff in ways that make the $\omega$ root-finding slow and arduous.

3.2 Neutrally stable oscillations

As in single-fluid asteroseismology (e.g. Eddington 1918; Cox 1980; Pijpers 2006; Aerts, Christensen-Dalsgaard & Kurtz 2010) each polytropic halo exhibits an infinite sequence of radial pulsation modes at discrete real values of $\omega$. These radial modes are neutrally stable: neither growing nor decaying. This is to be expected, as the present analysis is purely adiabatic: omitting any dissipation or mixing processes (e.g. shocks in the DM; violent relaxation of the stars; chaotic migration of stars).
3.2.1 Flow profiles of standing waves

Fig. 10 illustrates examples of eigenfunction waveforms expressed in terms of oscillating mass flux density, $(\rho \dot{r})$ for the DM (coloured curves) and stars (black curves). For the stars, flow amplitudes are higher in central regions and decrease with radius. For the DM, the peak amplitudes occur at intermediate radii. Such profiles are generic to the radial pulsations of all models with the $(F, F_\star, R)$ values inspected so far. Near the centre, the DM and stars oscillate in phase, but in outer regions they can oscillate in anti-phase.

For each constituent there are nodes at specific radii, where the matter remains locally at rest, and where there is a local reversal in the flow direction. The number of nodes of the DM oscillations is a harmonic number, $N \geq 0$. Values of $N$ uniquely label the eigen-modes. As $N$ increments, additional nodes appear in the outer fringe of the halo (e.g. between $50R_e$ and $100R_e$ for the $N = 1, \ldots, 4$ modes in Fig. 10). Most of the nodes appear far enough out that the Keplerian orbital frequency is locally less than the pulsation frequency: $\kappa < \omega$ (marked with arrows in Fig. 10). If wave amplitudes grew to the non-linear regime, some tidal shocking and orbital migration of stars could occur. Such dissipative effects scale as $\propto \rho v \cdot g$ which is of non-linear order $\sim \varepsilon^2$, while the radial displacements are of the order of $\sim \varepsilon$. Inwards/outwards pulsational displacements are an immediate effect, but orbital heating (and net outward expansion) could become significant gradually as more cycles elapse.

Depending on $N$ and the galaxy model, a number of nodes can also occur in the stellar profile ($N_\star$). For a galaxy with $F = 9$, $F_\star = 0.163$ and $R = 100R_e$, the node numbers are $(N, N_\star) = (1, 0), (2, 0), (3, 1), (4, 1), (5, 1), (6, 2), \ldots, (9, 3), \ldots, (18, 5)$ and so on. Fig. 11 shows the dark and stellar node positions for the first 30 oscillatory modes of this model, and for equivalent models where $F = 3$ and $F = 6$. It appears that $N_\star$ never decreases as $N$ increases, but there is no explicit link between these numbers. The sequence of $(N, N_\star)$ and their radii provide a fingerprint of a particular galaxy, and must be obtained numerically.

Fig. 11 also shows that the dark nodes are more crowded near the outer boundary ($R$) in halo models with larger $F$. This is a consequence of slower wave propagation in the fringes, since small-cored (large $F$) models have proportionally lower $\sigma$ profiles in the
Galaxy halo stability

Figure 10. Radial wave profiles of the first four oscillatory modes in a galaxy model with \( F = 9 \) halo of radius 100\( R_e \), and \( F \approx 0 \). The vertical axis is \( \log_{10}(\rho) \). DM is coloured; the stellar profile is black. Curves are solid or faded in the parts where the sign is positive or negative, respectively. Arrows indicate the radius beyond which the Keplerian frequency is less than the radial oscillation frequency (\( \kappa < \omega \)).

Outskirts. Stellar nodes are more subtly affected by \( F \). For models with small \( F \), the stellar nodes are more numerous and reach smaller radii for given \( N \). Haloes with large \( F \) tend to have fewer stellar nodes, and they are distributed farther out. Varying the stellar mass fraction has slight effects on the positions of the nodes in the outskirts. Across a wide domain \( 0.002 \leq F_* \leq 0.48 \) (but with fixed \( N, F \) and \( R \)), the stellar nodes do not shift significantly. The outer nodes of the DM appear more affected: they occur at radii a few per cent farther out if \( F_* \) is very small.

For waves that have at least one DM node (\( N \geq 1 \)), the radius of the innermost node varies more widely than any other node: from \( \sim 0.1R_e \) to \( \approx 10R_e \) in the domain of \( 0.002 \leq F_* \leq 0.48, 2 \leq F \leq 9 \) and \( 10R_e \leq R \leq 100R_e \). Its location varies only subtly with harmonic number \( N \). Fig. 12 depicts the detailed distribution of nodes in the inner, astronomically observable regions (\( r < 20R_e \)) of galaxies with \( F = 3 \) and \( F = 9 \) where the halo radius is \( R/R_e = 10, 100 \) and the DM content is rich or ‘cosmic’ (\( F_* = 0.002, 0.163 \)). All else being equal, the innermost dark node tends to occur at smaller radii if the galaxy is DM-poor (large \( F_* \)), and farther out if the galaxy is DM-rich (small \( F_* \)). For DM-rich haloes, a larger radius \( R \) shifts the inner dark node to larger radii. For some DM-poorer haloes, larger \( R \) shifts the inner dark node inwards. Comparison of the \( F = 3, 9 \) cases in Fig. 12 also shows how the lower-\( F \) models have more stellar nodes at radii small enough that ripples might be observable in the profiles of starlight and stellar velocities.

3.2.2 Frequency spectrum

Fig. 13 illustrates the oscillatory spectra for various baryon fractions (\( F_* \)) and thermal degrees of freedom \( F = 3 \) or \( F = 9 \). Among these and other results, it is usually found that most eigenfrequencies \( \omega \) are approximately linearly spaced in relation to \( N/t_d \) (where \( t_d \) is the acoustic crossing time). Typically \( \Delta \omega \approx 0.5/t_d \) (or \( \Delta \omega \approx 0.5 \Omega_0 \)) independently of \( (F, F_*, R) \). If the mode spacing of a real galaxy halo were measurable, this information would provide a diagnostic of the halo radius via its crossing time (see Appendix B). The confirmed detection or inference of the lowest mode would constrain \( (F, F_*, R) \) more tightly.

There are domains where the lowest mode deviates from the linear spectrum. For galaxies with sufficiently compact halo (small \( R \)) there exists a fundamental oscillation that has no DM nodes in the waveform (\( N = 0 \)). Keeping the same \( (F, F_*) \) but increasing the halo radius \( R \), there is a narrow range of \( R \) where the fundamental frequency detaches from the rest of the spectrum, and drops to zero frequency \( (\omega_0 \rightarrow 0) \). For radially larger haloes, the lowest mode is the first overtone (\( N = 1 \)). For \( F < 6 \) the fundamental mode drops out somewhere from a few \( 0.1R_e \) to a few \( 10^2R_e \). The dropout threshold \( R \) is greater in DM-rich systems (small \( F_* \)). For haloes

Figure 11. Node locations of DM (• coloured) and stars (♀ symbols) for the \( N \leq 30 \) radial pulsation modes of galaxy models with \( F_* = 0.163, R = 100R_e \) and \( F = 3, 6, 9 \) (left-hand, middle and right-hand panels, respectively). The black curve marks the radius where \( \kappa = \omega \).
Figure 12. Node locations for $N < 30$ pulsations in the inner regions of model galaxies with $F = 3$ (top row) and $F = 9$ (bottom row) and various halo radii and stellar mass fractions ($R = 10, 100R_e$, and $F_* = 0.002, 0.163$ as annotated). The innermost dark node occurs at a smaller radius if the stellar mass is relatively large (big $F_*$). Symbols and the $\kappa = \omega$ line are marked as in Fig. 11.

Figure 13. Oscillatory spectra for galaxies for various choices of the halo radius $R$ (horizontal axes). Frequencies (vertical axis) are scaled by the acoustic crossing time of the halo. Panels show cases where the halo has $F = 3$ (top row) and $F = 9$ (bottom row). The stellar mass fraction is $F_* = 0.002, 0.01, 0.04, 0.163$ in the columns from left to right, respectively. The frequencies are nearly linearly spaced, except in intervals of $R$ where one of the lower modes vanishes.
with $F \geq 6$ the $N = 0$ dropout occurs below $R < 0.1R_e$, so that all such galaxies of realistic size lack a fundamental mode (if $F \geq 6$). In the $(F, \mathcal{F}_e) = (9, 0.002)$ and $(9, 0.010)$ cases of Fig. 13, we see a further dropout of the $N = 1$ mode when $R$ is very large. The dropout range in $R$ is not in any obvious way related to the onset of fast and slow instabilities. Apart from influencing the dropout radius, the stellar mass fraction ($\mathcal{F}_e$) has little effect on the shape of the oscillatory spectrum.

For halo sizes that are realistic for elliptical galaxies ($10^4 \lesssim R/R_e \lesssim 10^3$) the fundamental mode is non-existent. If so then all oscillatory modes have at least one dark node. The innermost dark node turns out to occur at radii sufficiently small that there may be consequences for the visible galaxy. Since this node location is nearly independent of $N$, it seems possible that dark eigenmodes could excite each other via disturbances in the DM close to this region.

The stellar nodes are stationary points in the radial displacement of the stars. The mean radial velocities would be zero at these points of rest, and opposite signs on either side of each node. Rippling overdensities and underdensities of stars would appear between consecutive nodes. Such density variations might be observable if any stellar nodes occur within $r \lesssim 10R_e$, where the surface brightness is practically measurable. In the present model calculations, stellar nodes tend to occur farther out than the dark nodes, but some stellar nodes can appear at smaller radii if $F$ is small or if the oscillations are of high order. For observable ripples to appear, the order of the overtone needs to be large ($N \gg 30$).

### 3.3 Mixed complex modes

As emphasized in Section 2.8, the $\lambda$ ODEs only depend on the perturbation frequency via $\omega^2$ terms. For a one-component polytrope the eigenfrequencies should be purely real or purely imaginary. The two-component galaxy model, however, is an innately more complicated (double) oscillator, due to the $g_{\pm}$ gravitational couplings. Even if $\lambda_{\pm}$ and $\lambda_{\mp}$ are in phase at the inner boundary, they may be out of phase at other radii due to the different respective wave speeds. It is possible for some eigenfrequencies to occur off the real and imaginary axes in conjugate pairs ($\omega = \omega_h + i\omega_i$, $\omega^* = \omega_h - i\omega_i$).

In models where such modes do occur, they coexist with the ordinary oscillations and growth modes. Finding them requires extra effort to explore the $\omega$ plane in two dimensions. The landscape of outer boundary scores, $Z \equiv \lambda_{F_e} - \Lambda_{F_e}$, exhibits steep gradients ($\partial Z/\partial \omega$), which easily confounds ‘amoeba’ or ‘steepest descent’ solvers. It is more practical to map $Z$ on a fine rectangular grid. The desired solutions appear as contour rings or spikes on surface plots of log $|Z|$ (e.g. as in the analyses of shock instabilities; Toth & Draine 1993; Saxton et al. 1998; Saxton & Wu 2001) or by tracing the intersections of $Z = 0$ and $Z = 0$ contours. When an eigenfrequency is approximately located, an amoeba solver can be unleashed in that vicinity to refine the $\omega$ solution.

Fig. 14 locates the first few complex modes of a representative system. Complex modes occur in a minority of the galaxy models investigated so far. These tend to be radially large enough ($R \gtrsim 10^3 R_e$) that the crossing times are similar for the DM and stars, $t_* \approx t_\delta$ (see curves in Fig. B1 in the Appendices). Such huge halo radii (Mpc scale) are astronomically unrealistic, and so the mathematical existence of complex modes may be inconsequential to galaxies. Complex modes may be of more interest if they manifest in other analogous astrophysical systems: e.g. a star cluster possessing a gaseous atmosphere; or collisionless particles bound inside a fluid star.

Like the fast growth modes, the complex modes have wave amplitudes that are small in the interior and rise faster than a power law towards the halo’s surface. Fig. 15 shows the profile of the mass displacement ($\rho \delta r$) corresponding to the lowest-$|\omega|$ complex mode of a model with $F = 4$. As $\omega$ is complex the eigenfunction has real and imaginary parts too. Radial nodes are numerous for all complex modes, because $|\omega| \gg \Omega_\delta$. The innermost node for the DM motions is around $r \approx 0.13R_e$, similar to the innermost node of this.
C. J. Saxton

causes the complex but with \( F \) haloes, where the relatively high DM densities in the outskirts enable significant gravitational coupling between the components there. For fixed \((F, R)\) larger \( F \) causes the complex frequency rippling ought to induce violent relaxation and orbital mixing of stars. This would introduce non-linear damping, and may prevent the mode from growing to destructive amplitudes. Proving such damping effects would require higher-order perturbative analyses, including work terms \((\propto \rho v \cdot g \sim e^2)\).

4 DISCUSSION AND IMPLICATIONS

4.1 Reversibility

The collapse modes of the cores of giant stars are considered to cause some classes of supernovae. Are galaxy haloes susceptible to dark explosions if they possess the analogous collapse mode? The analogy should not be overstretched. Stellar core collapse is irreversible due to the accompanying nuclear reactions (and loss of neutrinos), and subsequently due to an emerging event horizon (e.g. Goldreich & Weber 1980). If these microphysical processes have counterparts in the SIDM domain then the convulsions of a galaxy-scale halo could sometimes have irreversible results. Otherwise, the DM collapse modes could produce repeated, reversible, non-linear bounces. At least four phenomena could incur some irreversibility.

(i) If the dark particles can self-scatter, then their self-annihilation and decay channels may also be non-negligible in some astrophysically attainable conditions (e.g. Buckley & Fox 2010; Feng, Kaplinghat & Yu 2010). High densities of SIDM might react to yield neutrinos, photons or other escaping byproducts. This is an ‘unknown unknown’, beyond the scope of the present paper.

(ii) Shock heating of the SIDM would raise the halo entropy (unless the dark shocks are reversible, as in a scalar field model of Peebles 2006). A shocked halo will eventually settle with a different radius \( R \) but with \((F, F_p)\) unchanged.

(iii) Tidal shocking and mixing of the stellar orbits could alter the orbit distribution irreversibly. This could, in principle, affect the evolution of luminosity verses radius relations of elliptical galaxies. Early-type galaxies in violently oscillating haloes might evolve lower stellar densities via tidal exchanges driving stellar diffusion radially.

(iv) Strong compressions of the inner halo could perhaps form a central black hole, or feed it significantly. The possibility of forming SMBHs via parsec-scale Jeans-unstable hydrodynamic ‘dark gulping’ was discussed by Saxton & Wu (2008). The masses involved are only an inner portion of a halo, and any collapse requires \( F > 6 \). Non-linear effects and long free-fall times would prevent the halo from vanishing into the horizon entirely.

Ostriker (2000), Hennawi & Ostriker (2002) and Balberg & Shapiro (2002) also considered growing black holes from SIDM, by a mechanism that depends on a gravothermal catastrophe in thermally conductive, weakly scattering haloes (not magnetic SIDM, nor scalar field SIDM). Balberg & Shapiro (2002) predict plausible black hole seeds in galaxy haloes, starting from a cored initial condition. Ostriker (2000) and Hennawi & Ostriker (2002) set initial conditions with a thermal inversion in a central density cusp (similar to the hypothetical cusps of collisionless CDM). From this they claimed tight upper limits on SIDM interactivity. Either way, gravothermal generation of a SMBH requires the SIDM scattering cross-section to be in a special range that ensures kpc-scale mean-free paths. If this requirement is unmet, collapse need not be imminent within a Hubble time (Balberg et al. 2002; Ahn & Shapiro 2005; Koda & Shapiro 2011).

4.2 Contrast to collisionless models

To calculate the discrete and continuous modes of entirely collisionless haloes is a complicated task, due to the higher dimensionality and freedom available to phase-space distribution functions; or due to a lack of closure of the Jeans equations for large-amplitude perturbations. Elaborate matrix formulations are normally used (Kalnajs 1977; Polyachenko & Shukhman 1981; Palmer & Papaloizou 1987; Saha 1991; Weinberg 1991). These involve choosing orthogonal basis functions, and practical applications require infinite matrices to be truncated approximately. More elegant methods exist for special systems (e.g. Sobouti 1984, 1985, 1986).

Despite the complication of a live halo as a second mass component, the present formulation is tractable because of several fortunate conditions. The existence of the dark halo’s boundary conditions constrains the freedom of the collisionless stellar component indirectly via their gravitational coupling and the \( \Lambda_p \) target. The finite halo radius ensures that the collective modes occur at discrete values of \( \omega \), rather than on continua. The assumption of small purely radial perturbations enables the zero-torque condition (41). (Analysis of non-radial modes would require a different approach.)
Much of the literature concerning the modes of collisionless galaxies has concentrated on the presence of ‘radial orbit instabilities’ which are non-spherical (l = 2) unstable modes of systems where stellar orbits are radially biased (Polyachenko & Shukhman 1981; Palmer & Papaloizou 1987; Weinberg 1991; Bertin et al. 1994; Trenti & Bertin 2006). In theory, non-spherical deformations could arise spontaneously in some objects that are stable against radial perturbations. Though this paper focuses on radial modes, we can at least say that radially unstable systems are likely to be unstable to non-radial motions too. Persistent non-radial pulsations are a possibility, and they could alter the instantaneous ellipticities and asymmetries of galaxies and their haloes. It is at this stage unknowable whether non-radial instabilities could destroy radially stable galaxies. A general non-radial analysis would be worthwhile, if another closure condition could be derived to replace the zero-torque assumption.

4.3 Pulsating galaxies

It was long ago recognized that collisionless galaxy models, if built from superpositions of integrable orbits, are capable of persistent pulsations (e.g. Louis & Gerhard 1988; Vandervoot 2003). These require sharp gaps in the phase-space distributions of stars. Unfortunately, the gaps are difficult to resolve numerically in N-body simulations, due to the innate limits of mass discretization. Perhaps because of this systematic effect, the possibility of pulsating galaxies is rarely mentioned in simulations and observational research. Orbital chaos and diffusion processes might eventually dampen these oscillations in real galaxies and star clusters.

In the present paper, it is shown that galaxies that combine stars with a fluid-like dark halo can sustain coupled oscillations, without fine requirements on the stellar orbit distributions. A disturbed mass of SIDM would echo and juggle with its own internal sound waves. These darkquakes displace the visible swarm of stars in their orbits. Reciprocally, the stars affect the dark waves, e.g. removing the instability of many F ≥ 6 models. Concentric radial overdensities of apparently displaced stars, if observed in nature, could be a signature of galaxy pulsations. If so then these would be an interesting diagnostic of the dark physics, enabling studies of skotosseismology – a seismology of DM. The density ripples of high-overton darkquakes (with kiloparsec wavelengths) would in principle have subtle effects on the gravitational lensing profiles of galaxies.

‘Shell galaxies’ are observed to have sharp stellar overdensities in bands encircling the main host galaxy (e.g. Malin & Carter 1980; Quinn 1984; Hernquist & Quinn 1987; Prieur 1988; Turnbull, Bridges & Carter 1999; Wilkinson et al. 2000). These are conventionally thought to be phase-wrapped remnants of dwarf galaxies destroyed in minor mergers. Stars from the disrupted secondary galaxy are strewn preferentially in certain orbits about the primary, and migrate till they occupy shells (rather than narrow streams). This is an attractive model when the observed shells differ from the stellar populations of the host galaxy, and when shell radii interleave in steps between the left and right sides. If, however, any case were found in which shell stars match the host populations exactly, or if any shell encircled the galaxy completely, then we might alternatively propose that host stars were displaced radially by dark halo pulsations (of high enough order to reach radii ≤ 10R_c). If the shell intervals appear smaller at larger radii, this would signify proximity to the halo’s truncation edge, R_c.

Passive elliptical galaxies observed at high redshifts appear to have modern stellar masses (∼ 10^11 M_☉) but contained within half-light radii (R_e ∼ 1 kpc) much smaller than modern counterparts (Cimatti et al. 2004; Daddi et al. 2005; Toft et al. 2005; Trujillo et al. 2006; Longhetti et al. 2007; van Dokkum et al. 2008; Damjanov et al. 2009; Trujillo, Ferreras & de La Rosa 2011; van de Sande et al. 2011; Ferreras et al. 2012; Toft et al. 2012). A mysterious process drives these ‘red nuggets’ to inflate on Gyr timescales, but without appreciable addition of mass, nor any significant starbursts. Expulsion of gas by active nuclei or stellar activity might conceivably alter the galaxy’s potential (e.g. Fan et al. 2008; Damjanov et al. 2009). Gasless minor mergers might inflate the fringes of the stellar profile (e.g. Khochfar & Silk 2006; Naab, Johansson & Ostriker 2009). Now we might invoke an additional explanation in terms of halo seismology. For stable galaxies, large oscillations could be excited by flybys or disturbances beyond the visible stellar outskirts. Stochastic excitations might be effective in cosmic environments of any density. If higher overtone pulsations reach large enough amplitudes, the outer stars may suffer tidal shocking or orbital diffusion, leading to the observed radial expansion. However if the initial conditions of the red nuggets already transgress the stability limits (Figs 8 and 9) then mass will redistribute spontaneously on time-scales ∼ o^2_α^2, even without external stimulation. Comparing these theorized evolutionary channels to empirical redshift-growth trends is worth a dedicated study elsewhere.

4.4 Extremes of dark matter content

The stability limits on spheroidal galaxies with SIDM haloes are loosest for stellar mass fractions around $\mathcal{F}_* \sim 0.04$. The constraints are tighter for systems with very small or large $\mathcal{F}_*$ (DM-rich and DM-poor extremes). Thus the theory predicts limits on the stable end-points of galaxy evolution. These depend on intrinsic structural properties rather than traits passively inherited from cosmological initial conditions and halo merger histories.

The strongest finding is the exclusion of DM impoverished galaxies ($\mathcal{F}_* > 0.5$) when the halo has a high heat capacity ($F > 6$). If such a system were born, the ubiquitous fast modes (Section 3.1.4) would destroy it by one process or another. The flux profiles imply dramatic evolution in the outskirts: perhaps runaway loss of the halo (till only a bare star cluster remains). Runaway accretion of DM is another possibility (if an external reservoir is available) until $\mathcal{F}_*$ is lowered to a stable configuration. The observed non-existence of stationary isolated galaxy-scale systems with $\mathcal{F}_* \sim 0.5$ is consistent with SIDM predictions.

At the opposite extreme in DM richness, observed dSphs are aged, quiescent stellar systems with low luminosities $10^3 \lesssim L/L_\odot \lesssim 10^7$ and dark masses > $10^4 M_\odot$ within $R_e$ (e.g. Mateo 1998; Gilmore et al. 2007; Strigari et al. 2008; Walker et al. 2009; Walker & Peñarrubia 2011; Amorisco & Evans 2012; Jardel & Gebhardt 2012). For the Milky Way satellites, $0.80 \lesssim f_d (< 1 R_e) \lesssim 0.999$ which implies $\mathcal{F}_* < 0.2$ for the brightest and $\mathcal{F}_* < 10^{-2}$ for ultrafaint cases. It appears that dSph occupy the upper-left, low-$\mathcal{F}_*$ part of Figs 8 and 9, where their stability is problematic. Their stellar density profiles are not cuspy like (7) but rather cored like a Plummer (1911) model. Preliminary calculations with this profile reveal fast and slow instabilities resembling those reported above. Cuspiness is unimportant; the effect of tidally truncated outskirts is less clear. It may also be significant that satellite galaxies are not isolated (Section 4.5). Immersion of dSph in the host galaxy’s halo would reduce the unstable domains. The analysis of galaxies with
truncated King (1966) profiles and a stabilizing external pressure awaits a future paper.

4.5 Alternative boundaries and asymmetries

A real halo need not be isolated in vacuum. External medium confinement can stabilize polytropes (e.g. McCrea 1957; Bonnor 1958; Horedt 1970; Umemura & Ikeuchi 1986). This effect may stabilize clustered galaxies and dwarf satellites that would be unstable in the field. An ambient background of unbound DM could imply a warm halo surface ($\sigma^2 > 0$). If external DM accretes supersonically, signals cannot propagate upstream and the isolated halo model still holds. If the contact is somehow subsonic then open boundary conditions apply instead of (52): waves might escape rather than reflecting at the surface. If the interface has a high contrast (like the solar atmosphere) then the vacuum boundary calculations might approximate $\omega$ well enough. If the halo blends smoothly into its surroundings then the oscillatory frequency intervals might decrease to the external Jeans frequency ($\sqrt{G\rho_0}$).

If the outskirts are infinite and fluid-like but the asymptotic density is near-zero then the eigenmodes would merge into a continuum: a conditional relation between $\omega_0$ and $\omega_1$. This relation would be found by setting $\lambda$ values at infinity, integrating radially inwards and matching the inner boundary conditions (48)–(51). This top-down approach is used in the analyses of radiative shocks (e.g. Saxton, Wu & Pongracic 1997). The stellar profile (7) is radially infinite, though real galaxies may truncate at a tidal radius. If so, an outer boundary condition on $\lambda$ analogous to (52) would affect the modes. If the stellar boundary is cold ($\sigma = 0$) then stellar nodes would be bunched there, and ripples could appear more numerous than in Section 3.2.1. As long as either the SIDM or the stars are spatially finite, a discrete spectrum is expected.

In weak-SIDM models with effectively collisionless outskirts (e.g. Yoshida et al. 2000; Peter et al. 2012; Rocha et al. 2012; Vogelsberger et al. 2012), the outer halo would act as part of the ‘$\star$’ component in this paper. The fluid-like ‘d’ component would reside in isolation within a reduced boundary radius $R$, shortening the crossing time and raising the pulsation frequencies.

Unless a galaxy is well isolated, its halo may receive cosmological infall of smaller objects. If the outskirts are quasi-collisionless then these mergers may cause tidal disturbances that excite oscillations in the core. If the outskirts are in a strong-SIDM condition, then weak shocks occur, the impactor can ablate and dissolve, and ripples may propagate (akin to raindrops in a pond). Mergers and mixing raise the total halo entropy ($\propto M_d \ln\chi$), which may alter the final halo mass–radius trends.

Non-radial effects could complicate the oscillations and stability limits. Disc galaxies (which provide the best evidence of cores) require axisymmetric models to assess the role of rotation (e.g. Hachisu 1986). Tidal fields, cosmic filaments and mergers could excite non-radial pulsations directly. Non-radial instabilities might destroy some configurations that are radially stable. The directionally dependent crossing times of a triaxial halo would split the frequencies, giving a non-linear spectrum of $\omega$ modes with angular dependencies. Since $F > 6$ polytropes are less stable than the standard $F = 3$ case, we might expect larger shape distortions during non-radial pulsation. The greatest attainable transient triaxiality must depend on $F$ and the environment; the extent is unknowable without 3D simulations. It would be interesting to see the results of strong-SIDM $F > 6$ shape investigations extending Peter et al. (2012).

4.6 New thermostatistics, dark fluid or exotica?

Tsallis thermostatistics generalize the Boltzmann entropy in an attempt to describe systems where long-range interactions (such as gravity) are influential (Tsallis 1988; Plastino & Plastino 1993; Nunez et al. 2006; Zavala et al. 2006; Vignat, Plastino & Plastino 2011). If this conjecture applies to self-bound bodies of DM, then the optimal spherical configurations are polytropes, even without short-range interactions or scatterings among dark particles. The equilibrium solutions for spherical galaxies have the same density and potential profiles as for SIDM fluid polytropes. Saxton & Ferreras (2010) fitted polytropic halo models to observed elliptical galaxies without distinguishing between SIDM and Tsallis scenarios.

However in their time-dependent and dynamical properties, the SIDM polytrope is not entirely equivalent to the Tsallis ‘stellar polytropes’ (e.g. Chavanis 2002). Fluids and stellar polytropes differ in their phase-space distributions, and their microscopic physics. If collisionless, DM need not support or transmit pressure waves. For a collisionless dark halo (even with a polytropic profile) the necessary analytic methods would involve phase-space distribution functions or moments in Jeans equations, with all the problems that accompany the treatment of the stars. The complexities of two-species collisionless modelling are beyond the scope of the present paper. Such models might not exhibit the oscillations and stability properties calculated above. Observation of darkquakes would favour SIDM over the Tsallis interpretation.

Besides SIDM, various alternative DM particle theories may be seismically distinct. If haloes consist of neutral fermions, then degeneracy pressure may explain the cores while the fringes are collisionless (e.g. Munyaneza & Biermann 2005; Nakajima & Morikawa 2007). Other candidates include ultra-light particles with de Broglie lengths reaching kpc scales (e.g. Sin 1994; Ji & Sin 1994; Lee & Koh 1996; Hu, Barkana & Gruzinov 2000; Rindler-Daller & Shapiro 2012). Perturbations and collisions of these ‘fuzzy DM’ galaxies produce interference patterns (González & Guzmán 2011). Such exotic halo models deserve their own stability analyses in order to predict observable signatures distinguishing them from SIDM and CDM.

5 CONCLUSIONS

This paper formulates perturbative stability analyses for spheroidal self-gravitating systems comprising a fluid mingled with collisionless matter. The present application describes spheroidal galaxies comprising a halo of self-interacting DM and an empirical de Vaucouleurs distribution of stars (with isotropic orbits in equilibrium). The halo is polytropic, adiabatic, non-singular and radially finite. It is isolated, with a vacuum outer boundary. The perturbations from equilibrium are radial but non-uniform.

All radially finite halo models exhibit a spectrum of natural modes of radial pulsation. These modes are neutrally stable, at least in the absence of high-order dissipative effects and stellar orbital mixing. The frequencies are almost linearly distributed. The frequency interval scales inversely with the halo’s acoustic crossing time, which depends on the halo radius. The thermal degrees of freedom $F$ influence the frequency gap below the lowest mode. For haloes of plausible radius ($R \gg 10R_*)$ and central density, all oscillations have a dark node within 10 stellar half-light radii. If detected via its rippling displacement of gas and stars, these modes would help diagnose the actual radial extent and total mass of the dark halo. Nodes in the stellar oscillation are more likely to appear as detectable ripples and velocity patterns if the harmonic number $N$
is large, if the halo is compact (small $R$), or if there are few thermal degrees of freedom (small $F$).

Gravitational coupling of a collisionless stellar distribution with a polytropic dark halo significantly affects the collective stability of the galaxy. Stability depends on the dark heat capacity (via $F$), the stellar mass fraction ($\mathcal{F}_* \propto R$) and the halo radius ($R$). Elliptical galaxies are stable for plausible $\mathcal{F}_*$ and radii ($R > 6$) smaller than Mpc scale, even in cases where the classical pure polytrope would be unstable ($F > 6$). This includes the range of DM equations of state, $7 < F \lesssim 10$, favoured by observed galaxy kinematics, cluster core sizes and black hole masses (Saxton & Wu 2008; Saxton & Ferreras 2010). The DM fraction within the stellar half-light radius never exceeds $f_h(1 < R_e) \lesssim 0.9$ for any stable de Vaucouleurs type configuration. A spheroid born in an unstable state would presumably evolve towards one of the stable end-points.

The stability limits are defined by the onset of three types of growing modes. Slow-growing modes occur if the halo is radially large and $F > 6$. Slow modes resemble non-cyclic versions of the pulsation waveforms: maximum stellar flows near the centre, and maximum dark flows at intermediate radii, with DM and stars moving in opposite directions in some regions. Fast-growing modes (quicker than orbital periods at $1R_e$) afflict galaxies with very centrally concentrated DM (regardless of $F$). Isolated systems that are rich in DM are only stable if the central density is quite low and the halo radius is huge. If stars dominate the galaxy mass, haloes with $F > 6$ may suffer the fast instability regardless of the radius. Complex modes are a mixture of oscillation and exponential growth. They exist in principle but they only seem to afflict haloes of unrealistically vast radii. Fast-growing modes and complex modes entail eruption or collapse of mass layers mainly in the fringes of the halo, leaving the interior less disturbed.

The identification of these oscillations and instabilities (perhaps via visible stellar displacements or ripples in gravitational lens profiles) would place informative constraints on fundamental DM physics, and distinguish SIDM from other candidates. If proven, more detailed skotoseismic analyses would provide complementary probes of the mass and energy distribution surrounding individual galaxies. This model also suggests an alternative interpretive framework that may help to explain: the expansion of red nuggets; the possible loss of haloes from extreme galaxies; morphologies of some shell galaxies and perhaps some of the drivers of early-type galaxy scaling relations.

ACKNOWLEDGEMENTS

I thank I. Trujillo and I. Ferreras for the conversation that challenged me to undertake this study. I also thank K. Wu, R. Soria, I. Ferreras, M. Mehdipour, Z. Younsi and D. Kawata for discussions during the work, and the anonymous referee for rigorous suggestions leading to extensive improvements.

REFERENCES

Aalseth C. E. et al., 2011b, Phys. Rev. Lett., 107, 141301
Angile et al., 2008, Phys. Rev. Lett., 100, 021303
Aprile E. et al., 2010, Phys. Rev. Lett., 105, 131302
Cimatti A. et al., 2004, Nat, 430, 184
de Vaucouleurs G., 1948, Ann. d’Astrophys., 11, 247
Emden F., 1907, Gaskugeln. Verlag B. G. Teubner, Leipzig, Berlin
Ferrerias I., 2012, AJ, 144, 47
Galaxy halo stability

APPENDIX A: CORE SCALES

A spherical polytropic halo without other gravitating components is a classic Lane–Emden sphere (Lane 1870; Emden 1907). The core can be defined by locations where the logarithmic index of density reaches certain values,

$$\alpha_d = -\frac{d \ln \rho_d}{d \ln r},$$

(A1)

with $\alpha_d = 1$ at radius $R_1$, and $\alpha_d = 2$ at $R_2$. These occur in fixed ratios to the halo surface ($R$). The index $R_1$ is important because beyond this radius an isolated polytropic SIDM halo differs from the NFW collisionless halo, in which $\rho_d \sim r^{-3}$ at large radii (Navarro et al. 1996b). Table A1 shows standard radius ratios for various $F$. The radius where the circular orbit velocity is maximal is also given in terms of $R_d / R$. Other models’ key radii were stated in appendix D of Saxton & Wu (2008).

APPENDIX B: EQUILIBRIUM PROPERTIES

The acoustic crossing time of the DM halo (Section 2.3) scales with the halo radius as $t_{ac} = C R^{3/2}$, where the factor $C = C(R, F)$ varies slowly in $R$. For a fixed stellar mass fraction ($f_*\Sigma$) the factor $C$ is slightly larger for greater $F$. The global mass ratio has a greater effect on $C$ than $R$ or $F$ do. The left-hand panels of Fig. B1 show the range of values obtained for DM-dominated models, and cosmic mean composition. The crossing time associated with the stars is shorter than for DM, for all models except those with very large radii. This is a result of the coldness of the DM in layers near the halo outer surface, where the stellar density profile is assumed to continue outwards gradually. Within the half-light radius of the stars, the DM has a higher velocity dispersion. Consequently, disturbances traverse the central regions faster via DM than via stars; and faster via stars in the outskirts.

The right-hand panels of Fig. B1 show the dark mass fraction within $1R_e$ as a function of the outer radius of the halo. For models

<table>
<thead>
<tr>
<th>$F$</th>
<th>$R_1 / R_e$</th>
<th>$R_2 / R_e$</th>
<th>$R_3 / R_e$</th>
<th>$R_4 / R_e$</th>
<th>$R_5 / R_e$</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.0</td>
<td>0.500</td>
<td>0.646</td>
<td>0.729</td>
<td>0.781</td>
<td>0.873</td>
</tr>
<tr>
<td>3.0</td>
<td>0.379</td>
<td>0.520</td>
<td>0.614</td>
<td>0.681</td>
<td>0.750</td>
</tr>
<tr>
<td>4.0</td>
<td>0.288</td>
<td>0.413</td>
<td>0.507</td>
<td>0.583</td>
<td>0.623</td>
</tr>
<tr>
<td>5.0</td>
<td>0.215</td>
<td>0.319</td>
<td>0.407</td>
<td>0.486</td>
<td>0.499</td>
</tr>
<tr>
<td>6.0</td>
<td>0.155</td>
<td>0.237</td>
<td>0.313</td>
<td>0.389</td>
<td>0.381</td>
</tr>
<tr>
<td>7.0</td>
<td>0.106</td>
<td>0.165</td>
<td>0.225</td>
<td>0.293</td>
<td>0.271</td>
</tr>
<tr>
<td>8.0</td>
<td>0.0636</td>
<td>0.101</td>
<td>0.143</td>
<td>0.196</td>
<td>0.169</td>
</tr>
<tr>
<td>9.0</td>
<td>0.0285</td>
<td>0.0459</td>
<td>0.0668</td>
<td>0.0988</td>
<td>0.0782</td>
</tr>
<tr>
<td>9.5</td>
<td>0.0133</td>
<td>0.0216</td>
<td>0.0320</td>
<td>0.0495</td>
<td>0.0372</td>
</tr>
<tr>
<td>9.9</td>
<td>0.00250</td>
<td>0.00408</td>
<td>0.00610</td>
<td>0.00986</td>
<td>0.00706</td>
</tr>
</tbody>
</table>
with \( F < 6 \), these fractions decline rapidly with \( R \). For models with \( F > 6 \) the dark mass fraction within these central regions is not very sensitive to the outer extent of the halo; it is effectively determined by the global ratio of dark to stellar matter. In observed systems where the DM fraction has been estimated from the kinematics of stars, planetary nebulae or globular clusters, the range of model \( f_p(<1R_e) \) and \( f_p(<5R_e) \) might be used graphically to constrain the likeliest \((F, f_p, R)\) parameters of the galaxy as a preliminary step to seismic modelling.

APPENDIX C: JEANS EQUATION FOR STELLAR PRESSURE

The stars are described by a velocity distribution function in spherical coordinates, \( f = f(r, \theta, \phi, v_r, v_\theta, v_\phi) \), where mean observables at any radius \( r \) are obtained by integration over velocity-space, \( \langle q \rangle = \iiint f q \, d^3 v \). (This Appendix will omit the \( \ast \) subscripts.) In spherical coordinates, the collisionless Boltzmann equation is

\[
\frac{\partial f}{\partial t} + v_r \frac{\partial f}{\partial r} + \frac{v_\theta}{r} \frac{\partial f}{\partial \theta} + \frac{v_\phi}{r \sin \theta} \frac{\partial f}{\partial \phi} + \dot{v}_r \frac{\partial f}{\partial v_r} + \dot{v}_\theta \frac{\partial f}{\partial v_\theta} + \dot{v}_\phi \frac{\partial f}{\partial v_\phi} = 0
\]

(C1)

with the derivative terms

\[
\dot{v}_r = \frac{v_r^2 + v_\theta^2}{r} - \frac{\partial \Phi}{\partial r}
\]

(C2)

\[
\dot{v}_\theta = \frac{v_\theta^2 \cot \theta - v_\phi v_r}{r} - \frac{1}{r} \frac{\partial \Phi}{\partial \theta}
\]

(C3)

\[
\dot{v}_\phi = \frac{- v_\phi (v_r + v_\theta \cot \theta)}{r} - \frac{1}{r \sin \theta} \frac{\partial \Phi}{\partial \phi}.
\]

(C4)

Multiplying (C1) by an observable quantity \( q \) and taking means over the velocity-space yields moment equations of the type found by Jeans (1915).

Consider the special case where a spherically symmetric galaxy is non-rotating; \( \theta \) and \( \phi \) are irrelevant and the \( v_r \) and \( v_\theta \) velocity components are indistinguishable. Then we have \( f(r, v_r, v_\theta, v_\phi) = f(r, v_r, \pm v_\theta, \pm v_\phi) \). If we consider purely radial perturbations in which the distribution function remains even with respect to the transverse velocity components, then some velocity moments vanish, e.g. \( \langle v_\theta \rangle = 0, \langle v_\phi^2 \rangle \approx 0 \) and \( \langle v_r v_\theta \rangle = \langle v_r^2 \rangle \approx v_r^2 \). The gravitational potential \( \Phi \) is spherical, so \( \partial \Phi / \partial \theta = \partial \Phi / \partial \phi = 0 \), and the gravitational field is radial \( (g = -\partial \Phi / \partial r) \). Under these simplifications, multiplying (C1) by \( v_r^2 \) and integrating over velocity-space obtains a time-dependent spherical Jeans equation for the radial pressure, in an Eulerian frame:

\[
0 = \frac{\partial}{\partial t} \rho \left\langle v_r^2 \right\rangle + \frac{\partial}{\partial r} \left[ \rho \left\langle v_r^2 \right\rangle - 2 \rho \left\langle v_r \right\rangle g - \frac{2 \rho}{r} \left( \left\langle v_r v_\theta^2 \right\rangle - \left\langle v_r \right\rangle v_\theta^2 + \left\langle v_r^2 \right\rangle \cot \theta \right) \right] .
\]

(C5)

If we assume that our radial perturbations are small enough to preserve a zero skewness of the velocity distribution in respect to \( v_r \) (e.g. a Gaussian form for the stationary model), then \( \langle v_r^2 \rangle = v_r^2 + \sigma_r^2 \) and \( \langle v_r \rangle = \langle v^2 \rangle = \sigma_v^2 + 3\sigma_r^2 / 2 \). This step is similar to the truncation of moment equations in the one-dimensional slab models of Louis (1992), but more generally allows pressure anisotropy and radial flows. Then (C5) converts to the ‘Lagrangian frame’ (following the local mean velocity) by way of

\[
\frac{d\rho}{dt} = \frac{\partial}{\partial t} + v \cdot \nabla,
\]

(C6)

to obtain

\[
0 = \frac{\partial P}{\partial t} + 2 v^3 \partial P + (3P + 2v^2) \frac{\partial v}{\partial r} + \frac{2P}{r} \left( v_r^2 + 2\sigma_v^2 - \sigma_r^2 \right) .
\]

(C7)