A Self-consistent, Time-dependent Treatment of Dynamical Friction: New Insights Regarding Core Stalling and Dynamical Buoyancy

Uddipan Banik and Frank C. van den Bosch

Department of Astronomy, Yale University, P.O. Box 208101, New Haven, CT 06520, USA; uddipan.banik@yale.edu

Received 2021 February 3; revised 2021 February 25; accepted 2021 March 1; published 2021 May 3

Abstract

Dynamical friction is typically regarded as a secular process, in which the subject (“perturber”) evolves very slowly (secular approximation) and has been introduced to the host over a long time (adiabatic approximation). These assumptions imply that dynamical friction arises from the LBK torque with nonzero contribution only from purely resonant orbits. However, dynamical friction is only of astrophysical interest if its timescale is shorter than the age of the universe. In this paper we therefore relax the adiabatic and secular approximations. We first derive a generalized LBK torque, which reduces to the LBK torque in the adiabatic limit, and show that it gives rise to transient oscillations due to nonresonant orbits that slowly damp out, giving way to the LBK torque. This is analogous to how a forced, damped oscillator undergoes transients before settling to a steady state, except that here the damping is due to phase mixing rather than dissipation. Next, we present a self-consistent treatment that properly accounts for time dependence of the perturber potential and circular frequency (memory effect), which we use to examine orbital decay in a cored galaxy. We find that the memory effect results in a phase of accelerated, super-Chandrasekhar friction before the perturber stalls at a critical radius, $R_{\text{crit}}$, in the core (core stalling). Inside $R_{\text{crit}}$ the torque flips sign, giving rise to dynamical buoyancy, which counteracts friction and causes the perturber to stall. This phenomenology is consistent with N-body simulations, but has thus far eluded proper explanation.

Unified Astronomy Thesaurus concepts: Dynamical friction (422); Orbital resonances (1181); Gravitational interaction (669); Stellar dynamics (1596); Galaxy dark matter halos (1880); Galaxy nuclei (609)

1. Introduction

Dynamical friction is an important ingredient of hierarchical structure formation. It is the dynamical process by which galaxies merge, and by which globular clusters and black holes sink to the centers of their host systems, where they can form bulges and binary black holes, respectively. In his seminal 1943 paper, Chandrasekhar showed that dynamical friction arises from the transfer of energy and momentum from the subject to the individual particles that make up the host system traversed by the subject. In particular, Chandrasekhar (1943) considered a subject of mass $M$ moving on a straight orbit through a uniform and isotropic sea of background (or “field”) particles of mass $m \ll M$. When a field particle encounters the subject, it experiences velocity changes $\Delta v_\perp$ and $\Delta v_\parallel$ in the directions perpendicular and parallel to the direction of the relative velocity. Chandrasekhar summed the velocity changes from the encounters between the subject mass and all field particles, treating them as independent two-body encounters, and showed that the net result is a frictional force acting on $M$ given by

$$ F_{\text{DF}} = \frac{4\pi G^2 M^2}{v^2} \ln\Lambda \frac{\rho(<v) \, v}{v}, \quad (1) $$

(see e.g., Mo et al. 2010). Here $v$ is the velocity of the subject mass, $\rho(<v)$ is the density of field particles with a speed less than $v = |v|$, and $\ln\Lambda = \ln[b_{\max}/b_{\min}]$ is the Coulomb logarithm, with $b_{\max}$ and $b_{\min}$ the maximum and minimum impact parameters of the encounters contributing to the drag.

Equation (1) is used routinely in astrophysics, even though it formally only applies to a uniform, isotropic background of field particles. While numerous studies have shown that it gives a reasonably accurate description of the orbital decay rates in galaxies and dark matter halos (e.g., Lin & Tremaine 1983; Cora et al. 1997; van den Bosch et al. 1999; Hashimoto et al. 2003; Boylan-Kolchin et al. 2008; Jiang et al. 2008), there are also cases in which it clearly fails. For instance, according to Equation (1) the drag force is proportional to the local density. Hence, a subject mass orbiting outside a galaxy or halo of finite extent should experience no drag. This is inconsistent with numerical experiments, which show that even in such cases the subject loses orbital angular momentum (Lin & Tremaine 1983). Another example where the standard treatment of dynamical friction fails is “core stalling,” the cessation of dynamical friction in the central constant-density core of a halo or galaxy (e.g., Read et al. 2006; Inoue 2011; Cole et al. 2012; Petts et al. 2015, 2016; Dutta Chowdhury et al. 2019).

Since the seminal work by Chandrasekhar, dynamical friction has been studied using a variety of different techniques and approaches. These include the Fokker–Planck method, in which dynamical friction arises from the momentum exchange described by the first-order diffusion coefficient (Rosenbluth et al. 1957; Binney & Tremaine 2008), stochastic approaches based on the fluctuation–dissipation theorem, in which dynamical friction arises from a correlation between the perturber’s velocity vector and the stochastic force it experiences from the field particles (Bekenstein & Maoz 1992; Maoz 1993; Nelson & Tremaine 1999; Fouvry & Bar-Or 2018), and a wide variety of methods that treat dynamical friction as a drag force arising from a “wake” or “polarization cloud” developing behind the perturber (Marochnik 1968; Kalnajs 1971; Mulder 1983; Weinberg 1989; Colpi & Pallavicini 1998). An excellent in-depth account of how all these methods relate to a generalized Landau equation derived from a truncation of the BBGKY hierarchy issued from the Liouville equation can be found in Chavanis (2013).

A shortcoming of many, though not all, of these methods is that they only treat dynamical friction as a local phenomenon
and/or that they have only been worked out for perturbers moving through homogeneous density distributions. The first study to overcome this, and to treat dynamical friction in a more realistic, nonuniform density distribution, is that by Tremaine & Weinberg (1984, hereafter TW84). Using a perturbative approach, in which the subject mass, $M$, is treated as a small, time-dependent perturbation on a circular orbit in a spherical system, they show that dynamical friction is entirely due to resonant orbits that give rise to a net retarding torque on the perturbing subject mass. In particular, by integrating the torque exerted by a single resonant field particle, multiplied by the (unperturbed) velocity distribution function along the orbits perturbed to second order and summing over all resonances, they obtain a torque that is second order in the perturber’s mass (i.e., proportional to $M^2$). This torque is known as the LBK torque, after Lynden-Bell & Kalnajs (1972), who first derived it in their treatment of angular momentum transport due to spiral structure in disk galaxies. Note that this resonance picture of dynamical friction gives a natural explanation for the nonzero drag experienced by a subject mass orbiting outside the galaxy, because it simply arises from the net torque due to resonant interactions with stars inside the galaxy.

A slightly different perturbative approach was recently taken by Kaur & Sridhar (2018, hereafter KS18); rather than perturbing the resonant orbits, they use the collisionless Boltzmann equation to compute the perturbed distribution function of field particles, which they integrate along the unperturbed resonant orbits. This once again yields a net torque that is second order in $M$, and consistent with the LBK torque obtained by TW84. Interestingly, KS18 then proceed to show that when the perturber enters the core region of a galaxy the number of low-order resonances (which dominate the torque) is suppressed and the strength of the remaining resonances is weakened. Hence, core stalling has a natural explanation in terms of the LBK torque. However, it fails to explain two related phenomena that have been identified in numerical simulations. The $N$-body simulations by Cole et al. (2012) manifest “dynamical buoyancy” in that perturbers initially placed near the center of a cored galaxy are found to be “pushed out.” Others have reported that when a perturber approaches a core, it first experiences a phase of strongly enhanced “super-Chandrasekhar” dynamical friction, followed by a “kickback” effect in which the perturber is pushed out again (Read et al. 2006; Goerdt et al. 2010; Zelnikov & Kuskov 2016). These simulation results have thus far eluded a proper explanation, and appear inconsistent with the notion that dynamical friction arises from the LBK torque, which is always retarding (at least in a spherical, isotropic system; see Section 2.3).

The notion that dynamical friction arises solely from resonant interactions can be traced back to the assumption that it is a secular process. This secular approximation implies that the actions of the perturber change only very slowly, on a timescale much longer than the dynamical time. In addition, it is assumed that the perturber is introduced to the system on a long timescale (i.e., the past evolution was also secular). We shall refer to this as the adiabatic approximation. Both the secular and adiabatic approximations underlie the treatments of TW84 and KS18, which are based on Hamiltonian perturbation theory and/or the fluctuation–dissipation theorem (e.g., Chavanis 2012, 2013; Heyvaerts et al. 2017; Fouvy & Bar-Or 2018).

It is important, though, to realize that the secular and adiabatic approximations are really only justified if the dynamical friction time, defined as the timescale on which the perturber sinks to the center of its host, is much longer than the dynamical time. Such cases, though, are of limited astrophysical interest. If instead we focus on systems for which the dynamical friction time is (significantly) shorter than the Hubble time, we are unavoidably in a regime for which the secular and adiabatic approximations may no longer be justified. In this paper we examine the impact of relaxing both these approximations. We do this by properly accounting for the past orbital evolution of the perturber in a self-consistent way. The resulting “self-consistent” torque differs from the standard LBK torque in two important ways. First of all, the self-consistent torque makes it explicitly clear that the dynamical friction torque arises from both resonant and near-resonant orbits. Second, while the exact resonances always exert a retarding torque, the near-resonant orbits can exert both retarding and enhancing torques. As long as the orbital decay rate is slow, the self-consistent torque can be written as the sum of two terms: (i) an “instantaneous” torque, which is the torque experienced by a perturber introduced abruptly to the host galaxy, and (ii) a “memory” torque, which depends on the entire orbital history of the perturber. The instantaneous torque builds up slowly, and then starts to oscillate in amplitude. Over time these “transient” oscillations damp out due to phase mixing, after which the instantaneous torque reduces to the LBK torque due to the purely resonant orbits. The memory torque, which always has a nonzero contribution from both resonant and near-resonant orbits, starts out subdominant, but becomes the dominant contributor to the total torque when the perturber approaches the core region of a galaxy. When this happens the orbital decay enters a phase of accelerated, super-Chandrasekhar infall, which ceases after the perturber crosses a critical radius at which the torque flips sign. Inside this radius the torque is enhancing, giving rise to “dynamical buoyancy.” Hence, we argue that core stalling occurs at or near this critical radius, as a manifestation of a delicate balance between dynamical friction outside and buoyancy within.

This paper is organized as follows. In Section 2 we first relax the adiabatic approximation. We use Hamiltonian perturbation theory to derive an expression for the “generalized LBK torque,” and discuss how it differs from the standard LBK torque using the analogy of a forced, damped oscillator. In Section 3 we subsequently also relax the secular approximation and derive an expression for the “self-consistent torque,” which self-consistently accounts for the orbital evolution (decay) of the perturber. We use this torque in Section 4 to discuss the orbital decay of a perturber in a cored background galaxy, providing new insight regarding core stalling, dynamical buoyancy, and super-Chandrasekhar dynamical friction. We summarize our results in Section 5.

2. Hamiltonian Perturbation Theory and the Generalized LBK Torque

Throughout this paper we follow TW84 and KS18, and consider a rigid perturber\footnote{Throughout we ignore potential mass loss of the perturber due to the tidal field of the host.} of mass $M_p$ moving on a circular orbit.

---

1. Throughout this paper we will use “subject” and “perturber” without distinction.
orbit in a spherical host potential (hereafter the “galaxy”) with mass profile $M_G(R)$. The host is made up of a large number of “stars” or “field particles” of mass $m$, and we have that $m \ll M_p \ll M_G$.

### 2.1. Hamiltonian Dynamics in the Corotating Frame

Since the total, perturbed gravitational potential, and hence the Hamiltonian for each field particle, is time-variable, energy is not a conserved quantity. And due to the lack of spherical symmetry, neither is angular momentum. However, as is well known (see, e.g., Binney & Tremaine 2008), the Jacobian Hamiltonian

\[ H_2 = E - \boldsymbol{\Omega}_p \cdot \mathbf{L} \tag{2} \]

is a conserved quantity (if we ignore time evolution of $\Omega_p$, $E$ and $\mathbf{L}$ are, respectively, the perturbed energy and angular momentum of the field particle in the nonrotating, inertial frame, given by

\[ E = E_0 + \Phi'_p = \frac{1}{2} \mathbf{r}^2 + \Phi_G(r) + \Phi'_p(r), \tag{3} \]

\[ \mathbf{L} = \mathbf{r} \times \mathbf{p}. \tag{4} \]

Here $\mathbf{r}$ is the position vector of the field particle with respect to the galactic center and $\dot{\mathbf{r}}$ is the velocity of the field particle in the inertial frame. The angular frequency of the galaxy–perturber system is given by $\Omega_p = (0, 0, \Omega_p)$, where

\[ \Omega_p = \sqrt{\frac{G [M_G(R) + M_p]}{R^3}}, \tag{5} \]

with $R$ the galactocentric radius of the perturber. $E_0$ is the unperturbed energy, i.e., the part of the Hamiltonian without the perturber potential, $\Phi_G$ is the gravitational potential due to the galaxy, and $\Phi'_p$ is the perturber potential, which consists of both direct and indirect terms and is given by

\[ \Phi'_p = \Phi_p + GM_p \frac{\mathbf{r} \cdot \mathbf{R}}{R^3}, \tag{6} \]

with $\Phi_p = -GM_p/|\mathbf{r} - \mathbf{R}|$ for a point perturber. The first term is the direct term, while the second term is the indirect term, which accounts for the fact that the galaxy center (the origin) is rotating about the common center of mass with the perturber.

In reality the perturbation in the potential also includes a gravitational “polarization” term that arises from the perturbation in the stellar distribution function induced by the perturber (also known as the “wake”). That term manifests the collective effects due to the self-gravity of the stars and significantly complicates the analysis. As first shown by Weinberg (1989), using Hamiltonian perturbation theory, and more recently by Chavanis (2012), Heyvaerts et al. (2017), and Fouvré & Bar-Or (2018) using the fluctuation–dissipation theorem, the collective effects primarily “dress” the perturber potential $\Phi'_p$ by introducing a prefactor that is the gravitational equivalent of the dielectric function in plasma physics. In particular, as nicely summarized in Fouvré & Bar-Or (2018), taking collective effects into account in the stochastic picture yields an inhomogeneous Landau–Balescu equation in which the diffusion coefficients involve the dressed potential.\(^3\)

\(^3\) Upon using the bare potential instead (i.e., ignoring collective effects due to self-gravity of the field particles), these diffusion coefficients reduce to those of the inhomogeneous Landau equation, which in turn implies a dynamical friction force consistent with the LBK torque.

formidable challenge in computing the dressed potentials, and given that the impact of collective effects is likely less important than for plasmas (see Chavanis 2013, for detailed discussion) we follow TW84 and KS18, and neglect the effects of self-gravity for the sake of simplicity.

### 2.2. Perturbation Analysis

The dynamics of the unperturbed system in the corotating frame is governed by the Jacobi Hamiltonian $H_2 = E_0 - \Omega_p \cdot \mathbf{L}$, which is a conserved quantity and therefore commutes with the unperturbed distribution function $f_0$, i.e.,

\[ [f_0, H_2] = 0. \tag{7} \]

where $[A, B]$ denotes the Poisson bracket of $A$ and $B$. This is nothing but the steady-state form of the collisionless Boltzmann equation (hereafter CBE) for the unperturbed galaxy in the corotating frame. When we introduce a perturber, the system is no longer in equilibrium and its dynamics is governed by the perturbed Hamiltonian. The perturber potential $\Phi'_p$ gives rise to a perturbation in the distribution function $f_1$, which in turn exerts a torque on the perturber. This is ultimately responsible for dynamical friction.\(^4\)

\(^4\) This is the key idea behind linear response theory.

In what follows, we shall, in the spirit of KS18, perturb the CBE up to linear order to obtain the expression for $f_1$ and use it to compute the torque on the perturber.

The perturbation in the distribution function can be computed by perturbing the CBE

\[ \frac{\partial f}{\partial t} + [f, H_2] = 0. \tag{8} \]

Up to linear order, we write $f$ and $H_2$ as the following perturbative series:

\[ f = f_0 + f_1, \]

\[ H_2 = H_{02} + \Phi'_p \tag{9} \]

We follow TW84 and KS18 and let the external perturbation grow as $\Phi'_p(R, t) = g(t) \Phi'_p(R)$, where the growth function

\[ g(t) = \begin{cases} e^{\gamma t}, & t < 0 \\ 1, & t \geq 0, \end{cases} \tag{10} \]

and $\gamma > 0$. This indicates that the perturber grows its mass exponentially from $t \to -\infty$ to $t = 0$ on a characteristic timescale $\tau_{\text{grow}} \equiv 1/\gamma$, while remaining at a fixed host-centric radius $R$. We do not consider this realistic, but before we consider an alternative we first aim to clarify the implications of this assumption. Note also that we have neglected the implicit time independence of $\Phi'_p$ and $\Omega_p$ through $R(t)$. This constitutes the secular approximation that $R$ changes on a timescale much longer than the dynamical time.

Substituting the series expansions given in Equation (9) in the CBE of Equation (8), we obtain the following evolution equation for $f_1$ up to linear order:

\[ \frac{\partial f_1}{\partial t} + [f_1, H_{02}] + [f_0, \Phi'_p] = 0. \tag{11} \]

In general, one can obtain the solution for $f_1$ once $f_0$ is known. The unperturbed distribution function $f_0$ is a solution of the...
can be different from $g$.

Make a canonical transformation from $(r, p)$ phase space to $(w, I)$ action–angle space spanned by the action vector $I = [I_1, I_2, I_3]$ and the corresponding angle vector $w = [\omega_1, \omega_2, \omega_3]$. Recall that $f_0$ and $H_0$ are both functions of only $I_1$, $I_2$, and $I_3$ and do not depend on the angles, while $f_1$ and $\Phi_\ell$ are functions of both actions and angles. Therefore, in action–angle space the Poisson brackets in the above equations become

$$[f_1, H_0] = \frac{\partial f_1}{\partial w_k} \frac{\partial H_0}{\partial I_k}, \quad (12)$$

$$[f_0, \Phi_\ell] = -\frac{\partial f_0}{\partial I_k} \frac{\partial \Phi_\ell}{\partial w_k}. \quad (13)$$

Here and throughout, the Einstein summation convention is implied, and indices $k$ and $l$ run from 1 to 3 and 1 to 2, respectively. In action–angle space, $[f_1, H_0]$ reduces to

$$[f_1, H_0] = \Omega_k \frac{\partial f_1}{\partial w_k}, \quad (14)$$

where the frequencies $\Omega_k$ are given by

$$\Omega_1 = \frac{\partial H_0}{\partial I_1} = \frac{\partial E_0}{\partial I_1},$$

$$\Omega_2 = \frac{\partial H_0}{\partial I_2} = \frac{\partial E_0}{\partial I_2},$$

$$\Omega_3 = \frac{\partial H_0}{\partial I_3} = -\Omega_p. \quad (15)$$

Here $I_3 = L_z$, and $I_1$ and $I_2$ are linear combinations of $I_1$ and $L$, respectively. Since $\Phi_\ell$ is a spherically symmetric potential, $E_0$ is a function of $I_1$ and $L$ only, or in other words of only $I_1$ and $L_2$. And since $f_0$ is a function of $E_0$ and $L$ only, it has a similar dependence on the actions, i.e., $\partial f_0 / \partial I_3 = 0$.

Following KS18 we expand $f_1$ and $\Phi_\ell$ as Fourier series in $w$ using

$$f_1(w, I, t) = \sum_{\ell} \hat{f}_\ell(I, t) e^{i \omega_\ell t},$$

$$\Phi_\ell(w, I) = \sum_{\ell} \hat{\Phi}_\ell(I) e^{i \omega_\ell t}. \quad (16)$$

where the summation is over all integer triplets $\ell = (\ell_1, \ell_2, \ell_3)$. Note that, since both $\hat{f}_\ell$ and $\hat{\Phi}_\ell$ are real, we have that $\hat{f}_{-\ell} = \hat{f}_\ell^*$ and $\hat{\Phi}_{-\ell} = \hat{\Phi}_\ell^*$, where $A^*$ indicates the complex conjugate of $A$. Substituting the Fourier series in Equation (11) yields the following evolution equation for $\hat{f}_\ell$:

$$\frac{\partial \hat{f}_\ell}{\partial t} + i \Omega_k \hat{f}_\ell = g(t) i \ell \frac{\partial f_0}{\partial I_k} \hat{\Phi}_\ell. \quad (17)$$

At this point in their analysis KS18 assume that the perturbation in the distribution function evolves in a way similar to the external perturber, i.e., $f_1 \propto e^{\gamma' t}$ with $\gamma' = \gamma$.

Under this assumption, the above differential equation becomes a simple algebraic equation that can be solved for $\hat{f}_\ell$. KS18 thus assume that the response density builds up on the same timescale as that on which the perturber is introduced. This is a fair assumption as long as $\gamma$ is sufficiently small, such that there is sufficient time for the host to respond. However, if dynamical friction is very efficient, then $\gamma'$ can be different from $\gamma$. In fact, in general the perturbation does not have a single growth rate. Different parts of the phase space respond to the perturber at different rates. Therefore, in what follows we will not make any a priori assumption about the growth rate of the response density due to the perturber. Rather, we solve the differential equation for $\hat{f}_\ell$ using the Green’s function technique with the initial condition that $\hat{f}_\ell(t \to -\infty) = 0$. We find the Green’s function to be $e^{-i \Omega_k \gamma t}$, which can be used to obtain the following particular solution for $\hat{f}_\ell$:

$$\hat{f}_\ell(I, t) = \frac{i \ell \Omega_k \hat{f}_\ell(I) e^{-i \Omega_k \gamma t}}{\gamma + i \ell \Omega_k}. \quad (18)$$

This can be integrated to yield

$$\hat{f}_\ell(I, t) = \frac{i \ell \Omega_k \hat{f}_\ell(I) e^{-i \Omega_k \gamma t}}{\gamma + i \ell \Omega_k} + \frac{1 - e^{-i \Omega_k \gamma t}}{i \ell \Omega_k}. \quad (19)$$

Note that the solution for $t < 0$ is identical to that obtained by KS18.

2.3. The Generalized LBK Torque

As shown in KS18, the torque on the perturber by a field particle is given by $\partial \Phi_1 / \partial \phi$. Hence, the total torque on the perturber can be computed by weighting $\partial \Phi_{1, \text{ext}} / \partial \phi$ by the perturbed distribution function and then integrating over all of phase space as follows:

$$\mathcal{T} = \int \mathcal{d}r \int \mathcal{d}p \frac{\partial \Phi_{1, \text{ext}}}{\partial \phi}(f_0 + f_1)$$

$$= \int \mathcal{d}r \int \mathcal{d}p \frac{\partial \Phi_{1, \text{ext}}}{\partial \phi} f_1. \quad (20)$$

Note that, since $f_0$ is independent of $\phi$ and $\int \mathcal{d}p \phi (\partial \Phi_{1, \text{ext}} / \partial \phi) = 0$, the leading-order contribution to the torque comes from $f_1$. And since both $\Phi_{1, \text{ext}}$ and $f_1$ are first order in $M_p$, the torque itself is second order in the mass of the perturber.

To evaluate the torque we follow KS18 and note that $\partial \Phi_{1, \text{ext}} / \partial \phi = -[\Omega_1, \Phi_{1, \text{ext}}] = -[L_\phi, \Phi_{1, \text{ext}}]$. And since the Poisson bracket is invariant under canonical transformation, we thus have that $\partial \Phi_{1, \text{ext}} / \partial \phi = -[L_\phi, \Phi_{1, \text{ext}}] = \partial \Phi_{1, \text{ext}} / \partial \phi_{1, \text{ext}}$. Moreover, the volume element $\mathcal{d}r \mathcal{d}p$ is also invariant under canonical transformation and becomes $\mathcal{d}w \mathcal{d}\ell$. Therefore, using Equation (10), the torque can be written in action–angle space.
as
\[ T = g(t) \int dw \int dl \frac{d\Phi_{\ell^*}}{dw^3} f_{\ell^*}, \]  
(21)

After substituting the Fourier expansions of \( \Phi_{\ell^*} \) and \( f_{\ell^*} \), and using the reality condition, i.e., \( \hat{f}_{-\ell^*}(I, t) = f_{\ell^*}(I, t) \), we obtain that
\[ T = g(t) \sum_\ell \sum_{\ell'} \ell_3 \int dl \hat{\Phi}_{\ell'}(I) \hat{\Phi}_{\ell^*}(I, t) \int dl w \ e^{i(d-\ell')w}, \]  
(22)

which can be integrated over \( w \) using the following identity for the Dirac delta function,
\[ \delta^3(x) = \frac{1}{(2\pi)^3} \int dl \ e^{iwx}, \]  
(23)

and summed over the \( \ell' \) indices to yield
\[ T = (2\pi)^3 g(t) \sum_\ell \ell_3 \int dl \hat{\Phi}_{\ell}(I) \hat{\Phi}_{\ell^*}(I, t). \]  
(24)

Substituting \( \hat{f}_{\ell^*}(I) \) given by Equation (19) in the above expression, the second-order torque can be written as
\[ T_2 = T_{\text{gen}} = 16\pi \sum_{\ell} \ell_3 \int dl |\mathcal{J}(I, \ell_4, t)| \frac{\partial \Phi_{\ell^*}}{\partial \ell^*} |\Phi_{\ell^*}(I)|^2, \]  
(25)

where
\[ \sum_{\ell} = \sum_{\ell_4=-\infty}^{\infty} \sum_{\ell_5=-\infty}^{\infty} \sum_{\ell_6=-\infty}^{\infty}, \]  
(26)

is the “reduced” summation over the positive-\( \ell_3 \) hemisphere of \((\ell_1, \ell_2, \ell_3)\)-space, which arises from applying the mirror symmetry operation \( \ell \to -\ell \) and retaining the symmetric part of the integrand. The function \( \mathcal{J}(I, \ell_4, t) \) is given by
\[ \mathcal{J}(I, \ell_4, t) = \frac{\gamma}{\gamma^2 + (\ell_4 \Omega_4)^2} \times \begin{cases} 2e^{it}, & t < 0 \\ \cos \ell_4 \Omega_4 t + \frac{\sin \ell_4 \Omega_4 t}{\ell_4 \Omega_4}, & t \geq 0. \end{cases} \]  
(27)

Note that this torque contains the contribution from all orbits: resonant, near-resonant, and nonresonant. We therefore refer to this as the generalized LBK torque, which is based on the secular approximation but not the adiabatic one (i.e., we did not take the limit \( \gamma \to 0 \)).

The amplitude of the torque scales as the Lorentzian-like function \( \gamma/|\gamma|^2 + (\ell_4 \Omega_4)^2 | \), which peaks at the resonances, where the commensurability condition of the frequencies,
\[ \ell_4 \Omega_4 = \ell_5 \Omega_5 (I_1, I_2) + \ell_2 \Omega_2 (I_1, I_2) - \ell_3 \Omega_3 = 0, \]  
(28)

is satisfied. The width of the Lorentzian is proportional to \( \gamma \) and determines the relative contributions to the torque from resonant and near-resonant orbits, with larger values of \( \gamma \) (i.e., smaller \( \gamma_{\text{grow}} \)) resulting in a more dominant contribution from the near-resonant orbits.

An important feature of the generalized LBK torque is that it can be either retarding (\( T < 0 \)) or enhancing (\( T > 0 \)), depending on the sign of \( \mathcal{J} \). Using that \( \frac{\xi \partial f_0}{\partial \ell} = \xi \Omega_f \partial f_0 / \partial E_0 < 0 \) for stable distribution functions (e.g., Doremus et al. 1971; Binney & Tremaine 2008), we see that a coherent retarding (enhancing) torque corresponds to \( \mathcal{J} > 0 (\mathcal{J} < 0) \). To get some insight, we start by examining the generalized LBK torque in the limits of both small and large \( \gamma \), which correspond to adiabatic growth and instantaneous introduction of the perturber, respectively.

### 2.3.1. Adiabatic Growth of Perturber Potential

Adiabatic growth of the perturber potential implies that the perturber has to be introduced on a timescale that is long compared to all other relevant dynamical times in the problem. Only then can the distribution function \( f_0(I_4, I_5) \), expressed according to the Jeans theorem as a function of its actions, remain perfectly invariant. The longest timescales of relevance are the libration times \( T_{\text{lib}} \sim 1/\ell_4 \Omega_4 \), which become infinitely long for orbits that satisfy the commensurability condition. Hence, assuring strict adiabatic invariance requires that we take the limit \( \gamma \to 0 \). In this limit, \( \mathcal{J}(I_4, \ell_4, t) \) converges according to
\[ \lim_{\gamma \to 0} \mathcal{J}(I_4, \ell_4, t) = \pi \delta(\ell_4 \Omega_4), \]  
(29)

where we have used that the Lorentzian function becomes a Dirac delta function in the limit of vanishing width. Substituting this expression for \( \mathcal{J} \) in Equation (25) yields the standard LBK torque
\[ T_2 = T_{\text{LBK}} = 16\pi \sum_{\ell} \ell_3 \int dl \delta(\ell_4 \Omega_4) (\ell_5 \Omega_5 + \ell_2 \Omega_2 \frac{\partial f_0}{\partial E_0} |\Phi_{\ell^*}(I)|^2, \]  
(30)

which has a nonzero contribution from only the exact resonances. And since \( \partial f_0 / \partial E_0 < 0 \) for a stable distribution function, we see that \( T_{\text{LBK}} < 0 \) for all resonances. In other words, the standard second-order LBK torque is always retarding in nature.

This makes it clear that the LBK torque is ultimately an outcome of taking the adiabatic limit (\( \gamma \to 0 \)). This should not come as a surprise: in the limit where the perturber takes infinitely long to present itself, the only contribution to a net torque that is not phase-mixed away (see Section 2.5 and Figure 2) is that from orbits in perfect resonance with the perturber. However, to what extent the LBK torque is relevant for dynamical friction ultimately depends on the timescale on which phase mixing removes the transient contributions. The rate at which an orbit and the perturber get out of phase depends on the incommensurability between their frequencies. If large, phase mixing is fast, and the contribution to the total torque vanishes rapidly. However, phase mixing of the contribution from the near-resonant orbits can easily take many dynamical times. And since a typical galaxy or dark matter halo is only a few dynamical times old (at least in its outskirts), we are not a priori justified in assuming that dynamical friction is dominated by the LBK torque.

### 2.3.2. Instantaneous Introduction of the Perturber

Idealized numerical simulations that examine dynamical friction (e.g., Lin & Tremaine 1983; White 1983; Cora et al. 1997; van den Bosch et al. 1999; Jiang & Binney 2000; Hashimoto et al. 2003; Boylan-Kolchin et al. 2008; Inoue 2009, 2011; Tamfal et al. 2020) typically do not adiabatically grow the perturber potential over time, but rather introduce it instantaneously to the host system. We can use our expression for the generalized torque to examine such
a scenario. Instantaneous introduction of the perturber corresponds to taking \( \gamma \to \infty \), for which

\[
\lim_{\gamma \to \infty} \mathcal{J}(\ell_k \Omega_k, t) = \begin{cases} 
0, & t < 0, \\
\sin \ell_k \Omega_k t / \ell_k \Omega_k, & t \geq 0.
\end{cases}
\]

This yields the following expression for the torque:

\[
\mathcal{T}_2 = \mathcal{T}_{\text{inst}} \equiv \frac{16\pi^2}{7} \sum_{\ell_1} \int \frac{d\Omega_1}{\ell_1 \Omega_1} \sin \frac{\ell_1 \Omega_1 t}{\ell_1 \Omega_1} \frac{\partial F}{\partial E_0} \left| \psi'(t) \right|^2.
\]

We shall hereafter refer to this as the instantaneous torque. At small \( t \), i.e., shortly after the instantaneous introduction of the perturber, we have that \( \mathcal{T}_{\text{inst}} \propto t \), indicating that the second-order torque builds up linearly with time. Note that the contribution to the torque from the nonresonant orbits can be either retarding or enhancing. In particular, for modes that are sufficiently far away from resonance, \( \sin \ell_k \Omega_k t \) can become negative a short time after the instantaneous introduction of the perturber, which can result in an enhancing torque. In addition, since \( \ell_1 \Omega_1 + \ell_2 \Omega_2 \) can be either positive or negative, it is even possible (at least in principle) for \( \mathcal{T}_{\text{inst}} \) to be enhancing when \( \sin \ell_k \Omega_k t \) is positive.

Finally, note that in the \( t \to \infty \) limit, \( \sin(\ell_k \Omega_k t) / \ell_k \Omega_k \to \pi / \delta(\ell_k \Omega_k) \), and we recover the familiar LBK formula for the torque, with a nonzero contribution only from the exact resonances. Hence, in accordance with the analogy of the forced, damped oscillator discussed in Section 2.5, following the instantaneous introduction of the perturber, the torque initially builds up linearly with time, then undergoes oscillations (corresponding to transients) that slowly phase-mix away, after which only the LBK torque due to the perfect resonances remains.

### 2.4. Dynamical Friction in an Isochrone Sphere

In order to illustrate how the instantaneous torque differs from the LBK torque, we compare the two for the case of a point-mass perturber on a circular orbit in a spherical, isotropic isochrone galaxy (Henon 1959). This configuration was also considered by KS18 and has the advantage that (i) it is a fairly realistic representation of a galaxy, (ii) many of its physical quantities can be computed analytically, and (iii) it has a central constant-density core, which allows us to examine core stalling.

The gravitational potential of Henon’s isochrone sphere with mass \( M_G \) and scale radius \( b \) is given by

\[
\Phi_G = -\frac{GM_G}{b + \sqrt{b^2 + r^2}},
\]

and its corresponding density profile

\[
\rho_G(r) = \frac{M_G}{4\pi} \left\{ \frac{3(b^2 + r^2)(b + \sqrt{b^2 + r^2})}{(b + \sqrt{b^2 + r^2})^3} \right\}^{3/2}.
\]

falls off as \( r^{-4} \) at large \( r \), and asymptotes to a constant core value of \( 3M_G/16\pi b^3 \) as \( r \to 0 \). Following KS18, we adopt \( M_G = 1.6 \times 10^9 M_\odot \) and \( b = 1 \) kpc. These parameters were chosen by KS18 such that the isochrone sphere has the same core radius and central density as the Burkert (1995) sphere used in the high-resolution N-body simulation of Inoue (2011). Following both KS18 and Inoue (2011), we adopt a point-mass perturber of mass \( M_p = 2 \times 10^5 M_\odot \) (corresponding to a mass ratio \( M_p/M_G = 1.25 \times 10^{-4} \)), which we consider to be on a circular orbit. In what follows we shall refer to this setup as our fiducial example.

As detailed in the Appendix and KS18, the commensurability condition for this system can be written as

\[
\ell_k \Omega_k = n \Omega_{cs} + l \Omega_\delta - m \Omega_p
\]

where, following KS18, we have used \( \ell_1 = n, \ell_2 = l, \) and \( \ell_3 = m \). The frequencies \( \Omega_{cs} \) and \( \Omega_\delta \) are related to the radial and angular frequencies in the orbital plane, as described in the Appendix. Although the total (generalized) LBK torque is the sum over all \((n, l, m)\), KS18 have shown that the torque is dominated by the corotation resonances, which have \( m = n \). In addition, the torque is typically stronger for lower-order modes (\(|l| \lesssim 3m|\)). In what follows we therefore restrict ourselves to the \((m, l) = (m, l, m)\) modes with dominant LBK torque.

Figure 1 plots the LBK torque (computed using Equation (A11) as detailed in the Appendix) as a function of the galactocentric distance of the perturber, \( R \), for the 10 dominant \((m, l)\) modes. Note how the LBK torque is dominated by that due to the \((m, l) = (2, 2)\) resonance, and that the LBK torque dies out as the perturber approaches the “filtering radius”\( R_a = 0.22b = 220 \) pc, marked by the black vertical dashed line. As detailed in KS18, at this radius, the circular frequency \( \Omega_p \) equals \( \Omega_p = 0.5 \sqrt{GM_G/b^3} \), roughly the circular frequency of stars in a central core of the isochrone sphere. As a result, the phase space contributing to the resonances shrinks, strongly suppressing the contribution of the dominant, lower-order modes to the total torque.

The solid lines in Figure 2 plot the instantaneous torque of Equation (A9) as a function of time in units of \( T_{\text{orb}} = 2\pi/\Omega_\delta \) for six modes that dominate the total torque at either early and/or late times. Results are shown for three different radii of introduction of the perturber, \( R = 0.7b \) (left panel), \( 0.5b \) (middle panel), and \( 0.4b \) (right panel). For comparison, we also plot the corresponding LBK torques as horizontal dashed lines. All modes initially show a coherent, retarding torque, causing the torque to build up linearly with time, before undergoing oscillations about the corresponding LBK value that slowly damp away with time. This is a classic example of phase mixing in collisionless systems where all but the purely resonant responses in the distribution function are damped out. Note, though, that the transients from some modes take many orbital times to phase-mix away, especially at smaller radii (closer to the core). The reason is that the core region with a nearly constant density has a much narrower dynamic range in orbital frequencies, resulting in a more coherent response. Interestingly, the transient oscillations can even contribute a positive, enhancing torque at times. Therefore, while the late-time dynamics is governed by the LBK torque from the perfectly resonant orbits, the transient behavior following the introduction of the perturber is driven by the near-resonant orbits.

### 2.5. Analogy: The Forced, Damped Oscillator

It is insightful to compare the galaxy-plus-perturber system to a forced, damped oscillator. As is well known, the general
solution of a sinusoidally forced, damped oscillator is a sum of a transient solution that depends on the initial conditions and a steady-state solution that is independent of the initial conditions. With time, the transients damp away, causing the response to settle toward the steady-state solution, which is a sinusoidal oscillation with a frequency equal to that of the driver, and an amplitude that depends on the driving amplitude, the driving frequency, the eigenfrequency of the (undamped) oscillator, and the damping ratio.

This is similar to how a galaxy responds to a perturber, $M_p$, on a circular orbit with frequency $Ω_p$, which can be regarded as the forcing frequency. The galaxy, in turn, acts as a damped oscillator. In fact, the galaxy is an ensemble of many individual oscillators (the individual orbits), each with three frequencies (corresponding to the three actions). Although each of these orbital frequencies can be considered to correspond to an individual undamped oscillator, the collective response of all orbits acts as if damped. The source of damping is phase mixing; initially, when the perturber is introduced (i.e., the forcing commences), all near-resonant orbits respond in phase, and the galaxy response is dominated by transient behavior. Due to phase mixing, though, the responses of different orbits get out of phase, causing the transient behavior to die out. The only orbits that never get out of phase with the forcing are the resonant orbits, which due to the commensurability of their orbital frequencies with $Ω_p$, remain in phase, thereby resisting phase mixing. As a consequence, the “steady-state response” of the galaxy is the LBK torque due to the resonant orbits.

This analogy shows that taking the limit $γ → 0$ corresponds to “skipping” the transient behavior, assuming that phase mixing has caused the response of all nonresonant orbits to die out. However, it is clear from the analogy that one can ignore the transients only after a sufficiently long time. If the perturber inspirals rapidly (i.e., $Ω_p$ is large), then the forcing with frequency $Ω_p$ may not last sufficiently long for phase mixing to nullify the net response of all nonresonant orbits. In particular, the near-resonant orbits, whose response takes the longest to phase-mix away, are expected to make a significant, if not dominant, contribution. The generalized LBK torque includes the transient response due to nonresonant orbits and presents a proper description of how dynamical friction builds up in idealized numerical simulations that introduce the perturber instantaneously. However, it is obtained by only relaxing the adiabatic approximation while still relying on the secular approximation, i.e., the perturber undergoes a slow inspiral under dynamical friction. In the following section we relax this assumption and develop a fully self-consistent treatment for dynamical friction that includes the “memory effect” due to the entire past orbital history of the perturber.

3. Self-consistent Computation of the Torque

The previous section has given some useful insight as to how transients that result from an instantaneous introduction of the perturber ($γ → ∞$) phase-mix away, ultimately giving rise to the LBK torque that one obtains in the adiabatic limit ($γ → 0$). However, that entire analysis is based on the generalized LBK torque (Equations (25)–(27)), which assumes that the perturber grows its mass exponentially, on a timescale $τ_{\text{grow}} = 1/γ$, while continuing to orbit at a fixed radius $R$. This is not realistic. The proper way to introduce the perturber is to self-consistently account for its past trajectory $R(t)$, from $t = 0$ to the present, which is what we tackle in this section.

The main shortcoming with the derivation of the generalized LBK torque, or with that of the standard LBK torque in TW84 and KS18, is that it ignores the time dependence of $Ω_p$ and $Ω_p$ in solving the evolution equation for $\tilde{f}_{\ell,\ell}$ (Equation (17)). In this section we are going to relax this assumption of secular evolution and compute $\tilde{f}_{\ell,\ell}$ and ultimately the torque in a fully self-consistent way. Here we assume that the perturber starts out at $t = 0$ at some large initial radius, $R_p$, and slowly makes its way inward following a circular orbit with time-dependent

![Figure 1](image-url)
radius $R(t)$. Therefore, both the external perturbation and its circular frequency depend implicitly on time according to

$$\Phi^\text{ext}_I = \Phi'_I(R(t)), \quad \Omega_p = \Omega_p(R(t)).$$

(35)

In what follows, for brevity, we simply write these dependences as $\Phi'_I(t)$ and $\Omega_p(t)$ and consider it understood that the time dependence enters implicitly via the perturber’s orbit, $R(t)$.

As in Section 2, we can perturb the CBE up to linear order and expand $f_I$ and $\Phi'_I$ as the Fourier series of Equation (16) to obtain the following evolution equation for $f^I$:

$$\frac{\partial f^I}{\partial t} + i[\ell^i \Omega_i - \ell^i \Omega_p(t)]f^I = i\ell^i \frac{\partial}{\partial t} f^I(I, t),$$

(36)

where, as before, the index $i$ runs from 1 to 2. The above equation can be solved using the Green’s function with the initial condition that $f^I(I, 0) = 0$ to yield the following form for $f^I(I, t)$:

$$f^I(I, t) = i\ell^i \frac{\partial}{\partial I} \int_0^t d\tau e^{-i\zeta(\tau)} \Phi'_I(I, t - \tau),$$

(37)

where

$$\zeta(\tau) = \ell^i \Omega_i \tau - \ell^i \int_0^\tau \Omega_p(t') dt'.$$

(38)

We substitute this expression into Equation (24) without the $g(t)$ factor to obtain the following form for the self-consistent torque:

$$T_2 = T_{\text{SC}} = 16\pi^3 \sum_{\ell} \ell \int dI \frac{\partial f^I}{\partial I} [J^I t(I, t) + J^H t(I, t)].$$

(39)

Here $J^I t(I, t)$ and $J^H t(I, t)$ are given by

$$J^I t(I, t) = \text{Re} \left[ \Phi'^*_I(I, t) \int_0^t d\tau \cos \zeta(\tau) \Phi'_I(I, t - \tau) \right],$$

$$J^H t(I, t) = \text{Im} \left[ \Phi'^*_I(I, t) \int_0^t d\tau \sin \zeta(\tau) \Phi'_I(I, t - \tau) \right],$$

(40)

where $\text{Re}(z)$ and $\text{Im}(z)$ are the real and imaginary parts of $z$, respectively. As we shall see shortly, $J^I t$ is generally the dominant term and $J^H t$ is subdominant.

Equation (39) is the most general form for the dynamical friction torque in the framework of linear perturbation theory in the absence of collective effects. This self-consistent torque differs from the instantaneous, generalized LBK torque of Equation (31) in two important ways. First of all, it modifies the resonances by introducing a time dependence to the circular frequency $\Omega_p$. Mathematically, this implies that the argument $\ell^i \Omega_i \tau - \ell^i \int_0^\tau \Omega_p(t') dt'$ of the sinusoidal function in the instantaneous torque is replaced by $\ell^i \Omega_i \tau - \ell^i \int_0^\tau \Omega_p(t') dt'$. Second, the self-consistent torque properly accounts for the fact that the perturber potential $\Phi'_I$ evolves as the perturber falls in. This implies that the $\Phi'_I(R(t))^2$ term in the instantaneous torque is replaced by a convolution term, i.e.,

$$\int_0^t d\tau \cos \zeta(\tau) \Phi'_I(I, t - \tau) \to \text{Re} \left[ \Phi'^*_I(I, t) \int_0^t d\tau \cos \zeta(\tau) \Phi'_I(I, t - \tau) \right].$$

(41)

In the secular limit where the temporal evolution is very slow, such that $\Phi'_I(t - \tau) \approx \Phi'_I(t)$ and $\Omega_p(t') \approx \Omega_p(t)$, we have that

$$J^I t(I, t) \approx |\Phi'^*_I(I, t)|^2 \int_0^t d\tau \cos \ell^i \Omega_k \tau = |\Phi'^*_I(I, t)|^2 \frac{\sin \ell^i \Omega_k \tau}{\ell^i \Omega_k},$$

(42)

while $J^H t \approx 0$. Substituting this in Equation (39), one recovers the expression for the instantaneous torque of Equation (31), as required.

While the instantaneous torque only depends on the current time $t$, the self-consistent torque takes into account the entire infall history of the perturber, thereby introducing temporal correlation into the system. This is reminiscent of how in the linear response theory of Colpi & Pallavicini (1998) the overdensity along the trail marked by the perturber exerts a retarding torque on it, i.e., dynamical friction originates from a memory effect involving the stars along the path of the
perturber. The self-consistent torque properly accounts for this memory effect, which is ignored in both the instantaneous torque and the LBK torque. While Colpi & Pallavicini (1998) compute the torque for an impulsive, straight orbit of the perturber through a homogeneous medium, our self-consistent torque (Equation (39)) describes dynamical friction for the more realistic case of a circular orbit in an inhomogeneous background. We also emphasize that even the inhomogeneous Lenard–Balescu equation derived by Heyvaerts (2010), Chavanis (2012), and Fouvrý & Bar-Or (2018), which is considered the most complete kinetic theory for gravitational systems to date, accounting for both inhomogeneity and collective effects, ignores the memory effect by assuming, in the computation of the diffusion coefficients, that the motion of the perturber is given by the mean-field limit with time-invariant actions (the secular approximation).

3.1. Orbital Decay

Under the assumption that the evolution of $R(t)$ is governed by the second-order dynamical friction torque, we have that

$$\frac{dR}{dt} = \frac{dR}{dL_p} \frac{dL_p}{dt} = \left( M_p \frac{d}{dR} \left[ R^2 \Omega_p(R) \right] \right)^{-1} T_2,$$  \hspace{1cm} (43)

where $L_p = M_p R^2 \Omega_p$ is the angular momentum of the perturber. And since the torque $T_2$ itself depends on time both explicitly and implicitly (through $R(t)$), we have that the evolution of $R$ is governed by the following integro-differential equation:

$$\frac{dR}{dt} = 16 \pi^2 \left( M_p \frac{d}{dR} \left[ R^2 \Omega_p(R) \right] \right)^{-1} \times \sum \frac{1}{t} \int dt_0 \left[ \mathcal{J}_{t0}(I, t) + \mathcal{J}_{t0}(I, t) \right].$$ \hspace{1cm} (44)

Solving this equation is rather challenging. However, we can obtain some powerful insight by expanding $R(t)$ using a Taylor series expansion. As long as the rate of infall, $dR/dt$, varies sufficiently slowly (i.e., $\frac{d^2 R}{d t^2}$ is small), we have that, to good approximation, $R(t - \tau) \approx R(t) - \tau \frac{dR}{dt}$. Since $\frac{dR}{dt} \approx M_p/M_G$ and is therefore typically small, $\hat{\mathcal{J}}_{t0}(I, t - \tau)$ can be expanded as a Taylor series and truncated at the leading order to obtain

$$\hat{\mathcal{J}}_{t0}'(I, t - \tau) \approx \hat{\mathcal{J}}_{t0}(I, t - \tau \frac{dR}{dt}) \approx \hat{\mathcal{J}}_{t0}(t) - \frac{d\hat{\mathcal{J}}_{t0}}{dt} \tau,$$ \hspace{1cm} (45)

where we remind the reader that the time dependence of $\hat{\mathcal{J}}_{t0}$ only enters through $R(t)$. Next we note that $\zeta(\tau) = \tau [\ell_1 \Omega_1 + \ell_2 \Omega_2 - \ell_\Omega \Omega_p(\tau)]$, where

$$\tilde{\Omega}_p(\tau) = 1/\tau \int \Omega_p(t') dt'$$ \hspace{1cm} (46)

is the time-averaged value of the circular frequency of the perturber. If we now make the assumption that $\tilde{\Omega}_p \approx \Omega_p$, i.e., we neglect the temporal evolution of $\Omega_p$, then $\zeta(\tau) \rightarrow \ell_\Omega \Omega_\tau$.

We thus have that $\mathcal{J}_{t0} \tau \approx 0$, and

$$\mathcal{J}_{t0} = \left[ \hat{\mathcal{J}}_{t0}(I, t) \right] \frac{d}{dR} \left[ R^2 \Omega_p(R) \right]^{-1} \left( \frac{1}{2} \frac{d}{dt} \int_0^t d\tau \cos \ell_\Omega \Omega_\tau \right.$$

$$\left. - \frac{1}{2} \right] \frac{d}{dR} \left[ R^2 \Omega_p(R) \right]^{-1} \left( \frac{1}{2} \right] \frac{d}{dR} \left[ R^2 \Omega_p(R) \right]^{-1} \right) \times \left( \sin \ell_\Omega \Omega_p \right) \left( - \cos \ell_\Omega \Omega_p \right) \left( \ell_\Omega \Omega_p^2 \right).$$ \hspace{1cm} (47)

Substituting the above expression for $\mathcal{J}_{t0}$ in Equation (39), we can write the second-order self-consistent torque as

$$T_2 = T_{\text{inst}} + T_{\text{mem}},$$ \hspace{1cm} (48)

where

$$T_{\text{mem}} = -8 \pi^2 \frac{dR}{dt} \frac{1}{\frac{d}{dR}} \frac{d}{dR} - p_{\text{mem}}$$ \hspace{1cm} (49)

Therefore, for $\frac{d}{dR} \approx \frac{L_p}{R} + \frac{d}{dR} \ln \Omega_p$ is a momentum term associated with the orbital decay (i.e., “sinking”) of the perturber, and

$$\frac{dL_p}{dR} = \frac{L_p}{R} \left[ 2 + \frac{d}{dR} \ln \Omega_p \right]$$ \hspace{1cm} (51)

is related to the momentum of the perturber in the absence of orbital decay.

At small $t$, $p_{\text{mem}}$ is small compared to $dL_p/dR$ and the infall is driven by the instantaneous torque, which is subject to transients that slowly die out due to phase mixing. As time goes on, and orbital decay becomes significant, $p_{\text{mem}}$ which we find to be typically positive, becomes more and more important, causing the denominator to become smaller. This in turn enhances the orbital decay rate. Hence, the memory term of the self-consistent torque has a destabilizing effect on the orbital decay. This is similar to the destabilizing “dynamical feedback” discussed in TW84. As we will see below, this becomes particularly important when the perturber approaches a central constant-density core.

Note that Equation (50) indicates the potential presence of a singularity at a critical radius, $R_{\text{crit}}$, where $dL_p/dR = p_{\text{mem}}$. Whether such a radius exists or not depends on both the galaxy potential and the mass of the perturber. In the limit $R \rightarrow R_{\text{crit}}$, the orbital decay rate $dR/dt \rightarrow \infty$. Inside $R_{\text{crit}}$, the denominator flips sign and $dR/dt \rightarrow +\infty$ as $R \rightarrow R_{\text{crit}}$. Figure 3

---

5 This is a reasonable approximation for a low-mass perturber on a circular orbit inside or close to a constant-density core, which is the case of interest here.

6 In the absence of the perturber, for a cored profile, $dL_p/dR \sim R$ for small $R$. $p_{\text{mem}}$, on the other hand, typically increases with decreasing $R$. Therefore, for small enough $R$, $p_{\text{mem}}$ will overtake $dL_p/dR$ and $dR/dt$ will flip sign.
demonstrates this by plotting $dR/dt$ as a function of radius $R$ for our fiducial isochrone plus point-mass perturber (see Section 2.4 for details). Here we have assumed the asymptotic (long time) forms for both the instantaneous torque (which equates to the LBK torque) and for the “memory torque,” $F_{\text{mem}}$. The fact that $dR/dt$ flips sign when crossing $R_{\text{crit}}$ suggests that this radius must act as an attractor for the dynamical evolution of the perturber, and we therefore associate $R_{\text{crit}}$ with the “core stalling” radius.

4. Super-Chandrasekhar Dynamical Friction, Dynamical Buoyancy, and Core Stalling

For our fiducial example of a point-mass perturber of mass $M_p = 2 \times 10^7 M_\odot$ on a circular orbit in a spherical isochrone galaxy of mass $M_{\text{gal}} = 1.6 \times 10^{10} M_\odot$ and scale radius $b = 1$ kpc (see Section 2.4), we use Equation (50) and a fourth-order Runge–Kutta integrator to evolve the radius $R(t)$ of the perturber. As before, we only consider the contribution to the torque from the 10 dominant $(m, l)$ modes shown in Figure 1. We have verified that this sampling of only the dominant modes does not significantly impact the results; in fact, we obtain virtually identical results if we use only the eight most dominant modes. In order to avoid problems with the integrator close to the singularity at $R_{\text{crit}} \approx 0.29b = 290$ pc, we implement a maximum cutoff in $|dR/dt|$ (shown by red dotted lines in Figure 3).

The solid lines in Figure 4 plot the resulting orbital decay tracks, $R(t)$, obtained for six different initial radii, $R_0 = 700$ pc, 600 pc, ..., 200 pc, which bracket $R_{\text{crit}}$. For comparison, we also show for each case the orbital decay track obtained using the LBK torque (dashed lines) and the standard Chandrasekhar formalism (dotted lines), which are obtained by solving Equation (43) with $T_2 = T_{\text{LBK}}$ and $T_2 = R \times F_{\text{DF}}$, respectively. Here $F_{\text{DF}}$ is given by Equation (1), where we follow KS18 by setting $\ln \Lambda = \ln(R/a)$, with $a = 10$ pc the assumed scale radius for the perturber, and properly compute $\rho(<v)$ from the isotropic distribution function of the isochrone sphere (Equation (A1)).

When the initial radius of introduction, $R_0$, is large (compared to $R_{\text{crit}}$), the orbital decay is characterized by four distinct phases of infall.

1. Phase I: Following the introduction into the system, the perturber falls in at a slightly slower rate than what is predicted by the LBK torque alone. This is because it takes time for the torque to build up and saturate to the asymptotic LBK value.

2. Phase II: Once the transients have died out, and the torque has reached the steady LBK value, the infall rate of the perturber matches that predicted by the LBK torque.

3. Phase III: As it approaches $R_{\text{crit}}$, the perturber starts to fall in at an accelerated pace, much faster than predicted by either the LBK torque or the standard Chandrasekhar formalism. This enhancement of the torque occurs only in the core region of the galaxy and is known as super-Chandrasekhar dynamical friction (see, e.g., Read et al. 2006; Goerdt et al. 2010; Zelnikov & Kuskov 2016).

4. Phase IV: Finally, the perturber reaches the stalling radius $R_{\text{stall}} \approx R_{\text{crit}}$, about which it oscillates under the action of dynamical friction (retarding torque) outside and buoyancy (enhancing torque) within.

When the initial radius $R_0$ is close to $R_{\text{crit}}$, the transients due to the self-consistent torque become more pronounced, in agreement with Figure 2. Introducing the perturber inside the critical radius (i.e., $R_0 < R_{\text{crit}}$) results in it being pushed out to $R_{\text{crit}}$ (following initial transient oscillations). Following Cole et al. (2012), we refer to this as “dynamical buoyancy.”

This complicated behavior is in excellent qualitative agreement with numerical $N$-body simulations, which have revealed how perturbers, upon approaching a central core, undergo accelerated super-Chandrasekhar friction, followed by a “kickback” effect in which the perturber is pushed out again prior to it ultimately settling (stalls) at some radius, typically close to the core radius (see Section 1 for references).

None of this is predicted by the standard Chandrasekhar formalism, according to which the perturber continues to sink all the way toward the center, albeit at a rate that becomes smaller toward the core. The latter owes to the fact that both $\rho(<v)$ and the Coulomb logarithm $\ln \Lambda = \ln(R/a)$ decrease with decreasing $R$. However, the resulting decline in the Chandrasekhar torque is insufficient to result in stalling.

The LBK torque is more successful, in that it clearly predicts core stalling. As discussed in detail in KS18, the stalling is expected to occur at or near the “filtering radius,” $R_\ast$, defined as the radius where the circular frequency of the perturber is equal to that of stars in the central core region (see also Read et al. 2006). KS18 showed that the lower-order modes, which otherwise exert a strong torque on the perturber, drop out of resonance, causing a significant reduction in the amplitude of the LBK torque (see also Figure 1). This suppression of resonances arises from the fact that just outside $R_\ast$, the circular frequency of the perturber is just a little bit lower than that of the core stars. Indeed, as shown by the dashed curves in Figure 4, based on the LBK torque one predicts that the infall stalls just outside $R_\ast$ (indicated by the horizontal brown line). Note, though, that the LBK torque predicts neither a super-Chandrasekhar phase nor dynamical buoyancy (the LBK...
torque is always retarding). In fact, introducing a perturber at \( R_0 = 200 \text{ pc} < R_{\text{crit}} \), the formalism based on the LBK torque predicts that it remains at that radius (see the purple dashed line in Figure 4).

Our formalism based on the self-consistent torque predicts a much richer dynamics, including dynamical buoyancy and super-Chandrasekhar infall. It also implies an explanation for core stalling that is intriguingly different from that based on the LBK torque. Rather than resulting from a diminishing of the dynamical friction torque, core stalling is an outcome of a balance between friction and buoyancy. All of this owes to the memory torque, which becomes dominant over the instantaneous torque close to \( R_{\text{crit}} \), and which causes the total torque to flip sign upon crossing \( R_{\text{crit}} \). This is illustrated in Figure 5, which plots \( T_{\text{mem}}/T_2 \) as a function of time (left panel) and radius (right panel) for three different values of the initial radius \( R_0 \), as indicated. Initially, the torque is dominated by the instantaneous term, \( T_{\text{inst}} \), given by Equation (31). The memory torque slowly gains strength, while undergoing oscillations, and starts dominating when the perturber approaches \( R_\text{ext} \) (indicated by the vertical black line in the right panel). At this radius, \( T_{\text{mem}} \) (and thus also the total torque) flips sign and becomes enhancing, giving rise to dynamical buoyancy when \( R < R_{\text{crit}} \), even though the instantaneous torque remains retarding.

4.1. Caveats and Outstanding Issues

Despite its success in reproducing previously unexplained aspects of dynamical friction observed in numerical simulations, in particular super-Chandrasekhar infall and dynamical buoyancy, the treatment of orbital decay based on the self-consistent torque presented above is subject to a few caveats.

First of all, we have ignored the time evolution of \( \Omega_p \) (i.e., we assumed that \( \Omega_p = \bar{\Omega}_p \)). Although this is likely to be a reasonable approximation close to a constant-density core, where \( \Omega_p(R) \) is nearly independent of radius, it remains to be seen how a proper treatment with a nonzero \( d\Omega_p/dt \) impacts the orbital decay. Unfortunately, since the temporal evolution of \( \Omega_p \), as quantified by \( \bar{\Omega}_p \) (Equation (46)), enters as an argument of the cosine and sine in the expressions for \( J_{\text{fl}}(I, t) \) and \( J_{\text{mem}}(I, t) \), respectively, numerically evaluating the corresponding integrals is nontrivial.

Second, in deriving the expression for \( dR/dt \) (Equation (50)) we expanded \( R(t) \) as a Taylor series that we truncated at first order. This is only valid as long as \( d^2R/dt^2 \) is sufficiently small. Unfortunately, this is likely to be violated during the super-Chandrasekhar phase, when the rate of infall rapidly accelerates. This caveat, which is also responsible for the singular behavior at \( R_{\text{crit}} \), can be overcome by using a higher-order truncation of the Taylor series or by trying to directly solve the integro-differential Equation (44). We leave this as an exercise for future investigations.

Finally, the entire formalism is based on perturbation theory, and therefore hinges on the assumption that the perturbation parameter \( |f_1^{\text{ext}}/H_0| \) is small. This assumption becomes questionable whenever the galaxy mass enclosed by the perturber, \( M_G(R) \), becomes comparable to the perturber mass.
Unfortunately, in numerical simulations core stalling often happens at a radius at which $M_G(R) \sim M_p$ (Pett et al. 2015, 2016; Dutta Chowdhury et al. 2019). In addition, when the perturber stalls at a fixed radius, the resonances no longer sweep by the stars fast enough to prevent nonlinear perturbations from developing (i.e., one is no longer in what TW84 refer to as the “fast regime”). These nonlinearities can even reverse the gradient of the distribution function near the resonances and contribute to an enhancing torque (which may counteract the retarding torque from the “fast” resonances and stall the infall) if the stars remain near-resonant for long enough ($dQ_p/dt$ is slow enough), as is the case near the stalling radius. Sellwood (2006) finds such an effect in N-body simulations of a rotating bar-like perturbation in a spherical galaxy.

All of this suggests that a proper treatment of core stalling may not be possible with perturbation theory. In a follow-up paper (U. Banik & F. C. van den Bosch, in preparation), we therefore examine dynamical friction, and core stalling in particular, using a non-perturbative, orbit-based approach. This reveals a family of (perturbed) orbits that exert a coherent, enhancing torque, thus contributing to dynamical buoyancy. When the perturber approaches the central core region, the nature of the near-resonant orbits changes, due to a bifurcation of the inner Lagrange points, causing buoyancy to become dominant over friction. Hence, the non-perturbative, orbit-based approach lends support to our conclusion that central core regions manifest dynamical buoyancy, something that was first noticed in the numerical simulations by Cole et al. (2012).

5. Conclusions

Various approaches to describe dynamical friction in inhomogeneous systems have shown that it ultimately arises from a torque that has a nonzero contribution only from stars in resonance with the perturber. Ultimately, this notion that only the resonances contribute to the torque has its origin in the assumption that the orbital decay rate of the perturber is (secular approximation) and always has been (adiabatic approximation) very slow compared to the dynamical time of the host. In the Hamiltonian perturbation theory of TW84 and KS18 the adiabatic approximation is enforced by multiplying the perturber potential by $e^{-t/\gamma}$ and taking the limit $\gamma \to 0$, while the secular approximation enters when the assumption is made that the orbital radius and circular frequency of the perturber are time-invariant over a dynamical time. This, in turn, implies that the equations of motion of both the perturber and the field particles are predominantly determined by the mean field, and thus characterized by slowly varying actions. Note that the same assumptions also underlie other approaches to dynamical friction in an inhomogeneous background, such as that based on the generalized Landau equation (e.g., Chavanis 2013) or the stochastic approach in action–angle space based on the fluctuation–dissipation theorem (e.g., Fouvry & Bar-Or 2018).

The secular and adiabatic approximations are justified when the mass of the perturber is sufficiently small. In that case, the dynamical friction time is much longer than the dynamical time. However, dynamical friction is mainly of astrophysical interest if the friction time is shorter than the Hubble time, which typically implies a perturber mass $M_p$ in excess of 1%–10% of the host mass. For such massive perturbers the dynamical friction time is no longer well separated from the dynamical time, and the secular and adiabatic assumptions are no longer justified. This breakdown is especially acute in the case of super-Chandrasekhar dynamical friction observed in numerical simulations when a massive perturber approaches a constant-density core.

In this paper we have examined implications of relaxing the adiabatic and secular assumptions. Using Hamiltonian perturbation theory similar to KS18, but without taking the limit $\gamma \to 0$ and without the assumption that the response density builds up on the same timescale as that on which the perturber is introduced, we first relaxed the adiabatic approximation and derived an expression for the generalized LBK torque

---

8 This is similar to Bogoliubov’s ansatz in plasma physics that the two-point correlation function relaxes much faster than the one-point distribution function.
(Equation (25)). This differs from the standard LBK torque in that it depends on the growth rate $\gamma$ and has contribution from all orbits, resonant and nonresonant. Taking the adiabatic limit $\gamma \rightarrow 0$, i.e., assuming an extremely slow growth of the perturber potential, we recovered the LBK torque with a nonzero contribution only from the pure resonances. The opposite limit, $\gamma \rightarrow \infty$, corresponding to an instantaneous introduction of the perturber as typically done in idealized numerical simulations, leads to a time-dependent torque with a nonzero contribution from the near-resonant orbits along with the purely resonant ones. This “instantaneous” torque builds up linearly with time before undergoing oscillations (“transients”) about the LBK value. Over time these oscillations damp out, and in the long term the generalized torque reduces to the LBK torque. This behavior is analogous to how a forced, damped oscillator undergoes transients before settling to a steady-state solution in which the frequency of the response matches the driving frequency. The main difference is that here the damping is due to phase mixing of the responses from the individual orbits (each with its own frequencies), rather than due to some dissipative processes. The timescale of relevance here is the timescale on which the transients damp away, which is proportional to the dynamic range in orbital frequencies of the field particles that make up the host. Typically this range is sufficiently large and phase mixing is very efficient, causing the generalized LBK torque to quickly transition to the LBK torque. This justifies the standard treatments of dynamical friction, in which only the resonances contribute, even when the adiabatic approximation is not necessarily justified. However, there is one important exception, which is the case when the perturber is introduced close to a central constant-density core. Here the dynamic range of frequencies is drastically suppressed, causing large transient oscillations that can take many orbital periods of the perturber to phase-mix away (see Figure 2).

Although the generalized LBK torque gives useful insight as to how transients that result from a non-adiabatic introduction of the perturber phase-mix away, it is still based on the unphysical ansatz that the perturber grows its mass exponentially over time, on a characteristic time $\tau_{\text{grow}} = 1/\gamma$, while remaining at a fixed host-centric radius, $R$. This time invariance of $R$ is a manifestation of the secular approximation. In order to improve on this, we next computed the torque in a self-consistent manner, in which we retained the information about the time dependence of the potential and circular frequency of the perturber throughout the entire evolution of the perturber’s orbital radius, $R(t)$ (we relaxed the secular approximation). This self-consistent torque differs from the (generalized) LBK torque in that the instantaneous circular frequency of the perturber is replaced by its time-averaged value, and that it includes a convolution term that embodies the temporal correlation of the perturber potential. As a consequence, the self-consistent torque always has a nonzero contribution from the near-resonant orbits, and depends on the entire infall history, $R(t)$, which in turn is dictated by the torque itself. A proper description of dynamical friction thus requires solving an integro-differential equation for $R(t)$ (Equation (44)).

While solving this equation in full generality is highly nontrivial, we obtained some valuable insight by Taylor-expanding $R(t)$ and truncating it at first order. This is valid as long as $\dot{d}^2R/dt^2$ is sufficiently small, i.e., the rate of infall, $dR/ dt$, varies slowly. If, in addition, we assume that the time dependence of the perturber’s frequency is small, which is a valid assumption in or near a central core region, we can write the self-consistent torque as a sum of two terms—the instantaneous torque, which depends on $R(t)$, and a memory torque, which is proportional to $dR/dt$ besides having an $R$ dependence. We used this simplified form of the self-consistent torque to evolve the radius $R(t)$ of a point-mass perturber in an isochrone galaxy (which has a central core). We found that the infall of the perturber occurs in four subsequent phases: (i) sub-LBK infall during the initial (linear) buildup of the torque, (ii) infall at the LBK rate as the instantaneous torque asymptotes to the LBK torque, (iii) accelerated super-Chandrasekhar infall due to a destabilizing effect of the memory torque, and (iv) kickback of the perturber from within a critical radius $R_{\text{crit}}$ due to buoyancy effects followed by stalling at that radius. The instantaneous torque dominates the early phase of the infall while the memory torque becomes dominant near the critical radius. It is responsible for the super-Chandrasekhar infall and flips sign at $R_{\text{crit}}$, causing the total torque to become enhancing for $R < R_{\text{crit}}$. When the perturber is introduced inside $R_{\text{crit}}$, it is consequently pushed out (dynamical buoyancy) to $R_{\text{crit}}$ by this enhancing memory torque.

These phenomena of super-Chandrasekhar infall followed by kickback and core stalling, as well as dynamical buoyancy inside central core regions, have been observed in numerous $N$-body simulations (Read et al. 2006; Goerdt et al. 2010; Inoue 2011; Cole et al. 2012; Dutta Chowdhury et al. 2019), but have thus far eluded a proper explanation (but see Read et al. 2006; Petts et al. 2015, 2016; Zelnikov & Kuskov 2016, for some phenomenological explanations). Although KS18 had shown that the LBK torque strongly diminishes as one approaches a core, which they advocated as an explanation for core stalling, they were unable to explain either super-Chandrasekhar infall or dynamical buoyancy. Based on our results, we argue that core stalling is ultimately a consequence of a subtle balance between dynamical friction (retarding torque) and buoyancy (enhancing torque), which is preceded by a phase of super-Chandrasekhar friction caused by the destabilizing effect of the memory torque that depends on the past infall history.

Finally, while completing this paper, we became aware of an unpublished study by Weinberg (2004), in which they also point out the problematic nature of the “time-asymptotic limit” (i.e., taking $\gamma \rightarrow 0$) used to derive the LBK torque. Using Hamiltonian perturbation theory similar to what is presented here, but using Laplace transforms rather than Green’s functions to solve for the response, they obtain a time-dependent torque (Equation (14) in their paper) that is identical to our self-consistent torque of Equation (39), except that it does not explicitly account for a time dependence of the perturber frequency. They then proceed to examine how the time-dependent torque differs from the LBK torque for the examples of a slowing bar and a decaying satellite. In the latter case, rather than calculating the orbital decay of the satellite self-consistently, as done here, they first compute the orbital decay $R(t)$ using the local Chandrasekhar formula, which is then substituted in the expression for the time-dependent torque. In agreement with our results, they show that massive perturbers, which decay rapidly, are significantly impacted by transients that are not accounted for in the LBK torque.

We are thankful to the anonymous referee for insightful comments and suggestions. We are grateful to Martin...
Weinberg for an enlightening discussion about the role of nonresonant orbits in dynamical friction, which ultimately set us on the path to explore a self-consistent treatment not knowing that he had worked on this topic before. F.v.d.B. is supported by the National Aeronautics and Space Administration through grant Nos. 17-ATP17-0028 and 19-ATP19-0059 issued as part of the Astrophysics Theory Program.

Appendix
The Isochrone Sphere

All specific examples presented in this paper correspond to a point-mass perturber, with mass $M_p$, moving on a circular orbit in an isotropic isochrone sphere, whose potential and density are given by Equations (32) and (33), respectively. In addition, the distribution function of the (unperturbed) isotropic isochrone sphere of mass $M_G$ is given by

$$f_0(\varepsilon) = \frac{M_G}{\sqrt{2(2\pi)^3(GM_Gb)^{3/2}}} \frac{\sqrt{\varepsilon}}{[2(1-\varepsilon)]^3/2} \times \left[ 27 - 66\varepsilon + 320\varepsilon^2 - 240\varepsilon^3 + 64\varepsilon^4 + 3(16\varepsilon^2 + 28\varepsilon - 9) \arcsin \sqrt{\varepsilon} \right],$$

(A1)

where $\varepsilon = -Ex_p/GM_G$ and covers the range $0 < \varepsilon < 1/2$.

In the absence of the perturber, the orbits of the field particles are characterized by four isolating integrals of motion: the energy $H_0$ and the three actions $(I_r, L, L_z)$. Following KS18 we make a canonical transformation from $(I_r, L, L_z)$ to $(I, L, L_z)$ where $I$ is given by $2I_r + L$, with $0 \leq L \leq I$ and $L \leq L_z \leq I$. In terms of $I$ and $L$, the orbital energy per unit mass is given by

$$E_0(I, L) = -\frac{2(GM_G)^2}{[I + \sqrt{I_r^2 + L^2}]}.$$  

(A2)

The fact that all these (unperturbed) frequencies can be expressed as simple algebraic functions is what makes the isochrone potential ideal for an analytical exploration of core stalling.

Following KS18, we ignore the torque from the stars outside the core, which allows us to truncate the integration over $I$ at a maximum value $I_{\text{max}} \ll I_0$. We follow KS18 and adopt $I_{\text{max}} = 0.1 I_0$. Under this approximation the expressions for the frequencies can be simplified as follows:

$$\Omega_w \approx \Omega_0 \left(1 - 3 \frac{L}{I_0}\right),$$

$$\Omega_g \approx \Omega_0 \frac{L}{I_0},$$

(A6)

where $\Omega_0 = 0.5 \sqrt{GM_G/b^3}$ is the frequency of stars in the central constant density core of the galaxy. Following the above expressions for the frequencies, we have the following expression for the resonance angle:

$$\ell_c \Omega_k = n\Omega_w + m\Omega_g - m\Omega_p = n\Omega_0 - m\Omega_p,$$

(A7)

where

$$s = s(I, L) \equiv \left[n \left(1 - 3 \frac{L}{I_0}\right) + L \frac{L}{I_0}\right].$$

(A8)

KS18 find that the corotation resonances with $n = m$ exert much stronger torque than the ones with $n \neq m$; therefore we shall only study corotation modes in this paper.

Substituting the above expressions in Equation (31), we arrive at the following form for the instantaneous torque for the $(m, l, m)$ mode:

$$T_{2ml} = 16\pi^3 m \Omega_0 \int_0^{I_{\text{max}}} dl \int_0^l dL \sin [s(\Omega_0 - m\Omega_p)l] s \Omega_0 - m\Omega_p,$$

(A9)

where $P_{\text{mln}}(I, L)$ is given by

$$P_{\text{mln}}(I, L) = \int_{-L}^{L} dL_z |\dot{\Phi}^\prime_{\text{mln}}(I, L, L_z)|^2.$$  

(A10)
We compute the Fourier coefficients $\hat{\Phi}_{\text{mlm}}(I, L, \ell_\phi)$ using the analytical expressions given in Appendix A of KS18.

The corresponding LBK torque is given by

$$
\mathcal{T}_{\text{LBK}} = 16\pi m^2 \Omega_p \int_0^{l_{\text{max}}} dl \\
\times \int_0^l dL \, \delta[l \Omega_b - m \Omega_p] \frac{\partial \phi}{\partial E_0} P_{\text{mlin}}(I, L).
$$

### ORCID iDs

Uddipan Banik @ https://orcid.org/0000-0002-9059-381X
Frank C. van den Bosch @ https://orcid.org/0000-0003-3236-2068

### References

Bekenstein, J. D., & Maoz, E. 1992, ApJ, 390, 79
Binney, J., & Tremaine, S. 2008, Galactic Dynamics (2nd ed.; Princeton, NJ: Princeton Univ. Press)
Boylan-Kolchin, M., Ma, C.-P., & Quataert, E. 2008, MNRAS, 383, 93
Burkert, A. 1995, ApJ, 447, L25
Chandrasekhar, S. 1943, ApJ, 97, 255
Chavanis, P.-H. 2012, PhysA, 391, 3680
Chavanis, P. H. 2013, A&A, 556, A93
Cole, D. R., Dehnen, W., Read, J. I., & Wilkinson, M. I. 2012, MNRAS, 426, 601
Colpi, M., & Pallavicini, A. 1998, ApJ, 502, 150
Cora, S. A., Muzzio, J. C., & Vergne, M. M. 1997, MNRAS, 289, 253
Doremus, J. P., Feix, M. R., & Baumann, G. 1971, PhRvL, 26, 725
Dutta Chowdhury, D., van den Bosch, F. C., & van Dokkum, P. 2019, ApJ, 877, 133
Fouvry, J.-B., & Bar-Or, B. 2018, MNRAS, 481, 4566
Goerdt, T., Moore, B., Read, J. I., & Stadel, J. 2010, ApJ, 725, 1707
Hashimoto, Y., Funato, Y., & Makino, J. 2003, ApJ, 582, 196
Henon, M. 1959, AnAp, 22, 126
Heyvaerts, J. 2010, MNRAS, 407, 355
Heyvaerts, J., Fouvry, J.-B., Chavanis, P.-H., & Pichon, C. 2017, MNRAS, 469, 4193
Inoue, S. 2009, MNRAS, 397, 709
Inoue, S. 2011, MNRAS, 416, 1181
Jiang, C. Y., Jing, Y. P., Faltenbacher, A., Lin, W. P., & Li, C. 2008, ApJ, 675, 1095
Jiang, I.-G., & Binney, J. 2000, MNRAS, 314, 468
Kalnajs, A. J. 1971, Ap&SS, 13, 279
Kaur, K., & Sridhar, S. 2018, ApJ, 868, 134
Lin, D. N. C., & Tremaine, S. 1983, ApJ, 264, 364
Lynden-Bell, D., & Kalnajs, A. J. 1972, MNRAS, 157, 1
Maoz, E. 1993, MNRAS, 263, 75
Marochnik, L. S. 1968, SvA, 11, 873
Mo, H., van den Bosch, F. C., & White, S. 2010, Galaxy Formation and Evolution (Cambridge: Cambridge Univ. Press)
Mulder, W. A. 1983, A&A, 117, 9
Nelson, R. W., & Tremaine, S. 1999, MNRAS, 306, 1
Petts, J. A., Gualandris, A., & Read, J. I. 2015, MNRAS, 454, 3778
Petts, J. A., Read, J. I., & Gualandris, A. 2016, MNRAS, 463, 858
Read, J. I., Goerdt, T., Moore, B., et al. 2006, MNRAS, 373, 1451
Rosenbluth, M. N., MacDonald, W. M., & Judd, D. L. 1957, PhRv, 107, 1
Sellwood, J. A. 2006, ApJ, 637, 567
Tamfal, T., Mayer, L., Quinn, T. R., et al. 2020, arXiv:2007.13763
Tremaine, S., & Weinberg, M. D. 1984, MNRAS, 209, 729
van den Bosch, F. C., Lewis, G. F., Lake, G., & Stadel, J. 1999, ApJ, 515, 50
Weinberg, M. D. 1989, MNRAS, 239, 549
Weinberg, M. D. 2004, arXiv:astro-ph/0404169
White, S. D. M. 1983, ApJ, 274, 53
Zelnikov, M. I., & Kuskov, D. S. 2016, MNRAS, 455, 3597