A Comprehensive Perturbative Formalism for Phase Mixing in Perturbed Disks. I. Phase Spirals in an Infinite, Isothermal Slab

Uddipan Banik1, Martin D. Weinberg2, and Frank C. van den Bosch1

1 Department of Astronomy, Yale University, P.O. Box 208101, New Haven, CT 06520, USA; uddipan.banik@yale.edu
2 Department of Astronomy, University of Massachusetts at Amherst, 710 N. Pleasant St., Amherst, MA 01003, USA

Received 2022 June 2; revised 2022 July 2; accepted 2022 July 7; published 2022 August 23

Abstract

Galactic disks are highly responsive systems that often undergo external perturbations and subsequent collisionless equilibration, predominantly via phase mixing. We use linear perturbation theory to study the response of infinite isothermal slab analogs of disks to perturbations with diverse spatiotemporal characteristics. Without self-gravity of the response, the dominant Fourier modes that get excited in a disk are the bending and breathing modes, which, due to vertical phase mixing, trigger local phase-space spirals that are one- and two-armed, respectively. We demonstrate how the laterally streaming motion of slab stars causes phase spirals to damp out over time. The ratio of the perturbation timescale ($\tau_p$) to the local, vertical oscillation time ($\tau_v$) ultimately decides which of the two modes is excited. Faster, more impulsive ($\tau_p < \tau_v$) and slower, more adiabatic ($\tau_p > \tau_v$) perturbations excite stronger breathing and bending modes, respectively, although the response to very slow perturbations is exponentially suppressed. For encounters with satellite galaxies, this translates to more distant and more perpendicular encounters triggering stronger bending modes. We compute the direct response of the Milky Way disk to several of its satellite galaxies and find that recent encounters with all of them excite bending modes in the solar neighborhood. The encounter with Sagittarius triggers a response that is at least 1–2 orders of magnitude larger than that due to any other satellite, including the Large Magellanic Cloud. We briefly discuss how ignoring the presence of a dark matter halo and the self-gravity of the response might impact our conclusions.

Unified Astronomy Thesaurus concepts: Disk galaxies (391); Milky Way disk (1050); Milky Way dynamics (1051); Perturbation methods (1215); Galaxy dynamics (591); Gravitational interaction (669); Galaxy encounters (592); Galaxy stellar disks (1594)

1. Introduction

The relaxation or equilibration of self-gravitating systems is a ubiquitous astrophysical phenomenon that drives the formation and evolution of star clusters, galaxies, and cold dark matter (DM) halos. In quasi-equilibrium, the phase-space density of such collisionless systems can be well characterized by a distribution function (DF), which, according to the strong Jeans theorem, is a function of the conserved quantities or actions of the system. When such a system is perturbed out of equilibrium by a time-dependent gravitational perturbation, either external (e.g., encounter with another galaxy) or internal (e.g., bars or spiral arms), the original actions of the stars are modified, and the system has to reestablish a new (quasi-)equilibrium. Since disk galaxies are highly ordered, low-entropy (i.e., cold) systems, they are extremely responsive. Even small gravitational perturbations can induce oscillations in the disk, which manifest as either standing or propagating waves (see Sellwood 2013, for a detailed review). Such oscillations consist of an initially coherent response of stars to a gravitational perturbation. This coherent response is called collective if its self-gravity is included. Over time, though, the coherence dissipates, which manifests as relaxation or equilibration and drives the system toward a new quasi-equilibrium, free of large-scale oscillations. Equilibration in galactic disks is dominated by collisionless effects, including purely kinematic processes like phase mixing (loss of coherence in the response due to different orbital frequencies of stars) and self-gravitating or collective processes like Landau damping (loss of coherence due to nondissipative damping of waves by wave–particle interactions; Lynden-Bell 1962) and violent relaxation (loss of coherence due to scrambling of orbital energies in a time-varying potential; Lynden-Bell 1967). It is noteworthy to point out that without phase mixing neither Landau damping (Maoz 1991) nor violent relaxation (see Sridhar 1989) would result in equilibration. A final equilibration mechanism is chaotic mixing, the loss of coherence resulting from the exponential divergence of neighboring stars on chaotic orbits (e.g., Merritt & Valluri 1996; Daniel & Wyse 2015; Banik & van den Bosch 2022). As long as most of the phase space is foliated with regular orbits (i.e., the Hamiltonian is near integrable), chaotic mixing should not make a significant contribution, and phase mixing may thus be considered the dominant equilibration mechanism.

Disk galaxies typically reveal out-of-equilibrium features due to incomplete equilibration. These may appear in the form of bars and spiral arms, which are large-scale perturbations in the radial and azimuthal directions, responsible for a slow, secular evolution of the disk. In the vertical direction, disks often reveal warps (Binney 1992). In the case of the Milky Way (MW) disk, which can be studied in much greater detail than any other system, recent data from astrometric and radial velocity surveys such as SEGUE (Yanny et al. 2009), RAVE (Steinmetz et al. 2006), GALAH (Bland-Hawthorn et al. 2019), LAMOST (Cui et al. 2012), and above all Gaia (Gaia Collaboration et al. 2016, 2018a, 2018b) have revealed a variety of additional vertical distortions. At large galactocentric
oscillations and vertical asymmetries have also been reported in mixing, sometimes referred to as “feathers” (e.g., Price-Whelan et al. 2015; Thomas et al. 2019; Laporte et al. 2022). Similar oscillations and vertical asymmetries have also been reported in the solar vicinity (e.g., Widrow et al. 2012; Williams et al. 2013; Yanny & Gardner 2013; Gaia Collaboration et al. 2018b; Quillen et al. 2018; Bennett & Bovy 2019; Carrillo et al. 2019). One of the most intriguing structures is the phase-space spiral discovered by Antoja et al. (2018) and studied in more detail in subsequent studies (e.g., Bland-Hawthorn et al. 2019; Li 2021; Li & Widrow 2021; Gandhi et al. 2022). Using data from Gaia DR2 (Gaia Collaboration et al. 2018a), Antoja et al. (2018) selected ~900,000 stars within a narrow range of galactocentric radius and azimuthal angle centered around the Sun. When plotting the density of stars in the $\{z, v_z\}$-plane of vertical position, $z$, and vertical velocity, $v_z$, they noticed a faint, unexpected spiral pattern, which became more enhanced when color-coding the $\{z, v_z\}$-“pixels” by the median radial or azimuthal velocities. The one-armed spiral makes two to three complete wraps, resembling a snail shell, and is interpreted as a signature of phase mixing in the vertical direction following a perturbation, which Antoja et al. (2018) estimate to have occurred between 300 and 900 Myr ago. More careful analyses in later studies (e.g., Bland-Hawthorn et al. 2019; Li 2021) have nailed down the interaction time to ~500 Myr ago.

The discovery of all these oscillations in the MW disk has ushered in a new, emerging field of astrophysics, known as galactoseismology (Widrow et al. 2012; Johnston et al. 2017). Similar to how the timbre of musical notes reveals characteristics of the instrument that produced the sound, the “ringing” of a galactic disk can (in principle) reveal its structure (both stellar disk plus DM halo). And similar to how the timbre can tell us whether the string of a violin was plucked (pizzicato) or bowed (arco), the ringing of a galactic disk can reveal information about the perturbation that set the disk ringing. Phase spirals are especially promising in this regard: their structure holds information about the gravitational potential in the vertical direction (in particular, the vertical frequency as a function of the vertical action; Antoja et al. 2018) and about the type of perturbation that triggered the phase spiral (e.g., bending mode vs. breathing mode; see Widrow et al. 2014; Darling & Widrow 2019a; see also Section 3 below). In addition, by unwinding the phase spiral, one can in principle determine how long ago the vertical oscillations were triggered. By studying phase spirals at multiple locations in the disk, one may even hope to use some form of triangulation to infer the direction or location from which the perturbation emerged (assuming, of course, that the phase spirals at different locations were all triggered by the same perturbation).

However promising galactoseismology may seem, many questions remain: What kind of perturbation can trigger a phase spiral? How long do phase spirals remain detectable, and what equilibration mechanism(s) causes their demise? Can we really constrain the vertical potential of the disk, or does self-gravity of the perturbation make it difficult to achieve? What kind of constraints can we infer regarding the perturber that triggered the phase spiral? Is galactoseismology likely to be confusion limited, i.e., should we expect that each location in the disk experiences oscillations arising from multiple, independent perturbations? If so, how does this impact our ability to extract useful information? Answering these questions necessitates a deep understanding of how the MW disk and disk galaxies in general respond to perturbations.

To date, these questions have mainly been addressed using numerical $N$-body simulations or fairly simplified analytical approaches. In particular, numerous studies have investigated how the MW disk responds to interactions with the Sagittarius (Sgr) dwarf galaxy (e.g., Gómez et al. 2013; D’Onghia et al. 2016; Laporte et al. 2018; Khanna et al. 2019; Hunt et al. 2021). While simulations like these have demonstrated that the interaction with Sgr can indeed spawn phase spirals in the solar vicinity (Antoja et al. 2018; Binney & Schönrich 2018; Darling & Widrow 2019b; Bland-Hawthorn et al. 2019; Laporte et al. 2019; Bennett et al. 2021; Hunt et al. 2021), none of them have been able to produce phase spirals that match those observed in the Gaia data. As discussed in detail in Bennett et al. (2021) and Bennett & Bovy (2021), this seems to suggest that the amplitude and shape of the “Gaia snail” cannot be produced by Sgr alone. An alternative explanation, explored by Khoperskov et al. (2019), is that the Gaia snail was created by buckling of the MW’s bar. However, this explanation faces its own challenges (see, e.g., Laporte et al. 2019; Bennett & Bovy 2021). Triggering the Gaia snail with a spiral arm (Faure et al. 2014) is also problematic, in that it requires the spiral arms to have unusually large amplitude (Quillen et al. 2018). Clearly, then, despite a large number of studies, pinpointing the origin of the phase spiral in the solar vicinity still remains an unsolved problem.

Although simulations have the obvious advantage that they can probe the complicated response of a perturbed disk to a realistic perturbation, which often is analytically intractable, especially if the response is large (nonlinear), there are also clear disadvantages. Foremost, reaching sufficient resolution to resolve the kind of fine structure that we can observe with data sets like Gaia requires extremely large simulations with $N > 10^8 - 10^9$ particles (Weinberg & Katz 2007; Binney & Schönrich 2018; Hunt et al. 2021). Although such simulations are no longer beyond our reach (see, e.g., Bédorf et al. 2014; Fujii et al. 2019; Hunt et al. 2021; Petersen et al. 2022), it is clear that using such simulations to explore large areas of parameter space remains a formidable challenge. To overcome this problem, a seminal approach called the backward-integrating restricted $N$-body method was developed originally in the context of perturbation by bars (e.g., Leeuwin et al. 1993; Vauterin & Dejonghe 1997; Dehnen 2000) and later used by Hunt & Bovy (2018) and Hunt et al. (2019) to study nonequilibrium features in the MW caused by transient spiral arms. This method is effectively a Lagrangian formalism to solve the collisionless Boltzmann equation (CBE) by integrating test particles in the perturbed potential in a restricted $N$-body framework, i.e., without self-consistently developing the potential perturbation from the DF perturbation. Although appropriate for studying the local kinematic distribution of particles, this approach becomes too expensive to study the global equilibration of a system. Hence, it is important to consider alternative analytical methods that can be used to investigate the global response of a disk.

In this vein, this paper presents a rigorous, perturbative, Eulerian formalism to compute the response of a disk to perturbations. In order to gain valuable insight into the physical mechanism of phase mixing, without resorting to the computational complexity involved in modeling a realistic
disk, which we postpone to Paper II (U. Banik et al. 2022, in preparation), in this first paper in the series we consider perturbations of an infinite slab with a vertical profile, but homogeneous in the lateral directions. Although a poor representation of a realistic galactic disk, this treatment captures most of the essential features of how disks respond to gravitational perturbations. We study the response of the slab to perturbations of various spatial and temporal scales, with a focus on the formation and dissolution of phase spirals resulting from the vertical oscillations and phase mixing of stars.

This paper is organized as follows. Section 2 describes the application of perturbation theory to our infinite, isothermal slab. Section 3 then uses these results to work out the response to an impulsive, single-mode perturbation, which nicely illustrates how phase spirals originate from vertical oscillations and how they damp out owing to lateral mixing. Sections 4 and 5 generalize this to responses to localized (wave packet) and nonimpulsive perturbations, respectively. In Section 6, we investigate the response to satellite encounters and examine which satellite galaxies in the halo of the MW can trigger bending and/or breathing modes strong enough to trigger phase spirals at the solar radius (still approximating the MW disk as an infinite, isothermal slab). We summarize our findings in Section 7.

2. Linear Perturbation Theory for Collisionless Systems

2.1. Linear Perturbative Formalism

Let the unperturbed steady-state DF of a collisionless stellar system be given by \( f_0 \) and the corresponding Hamiltonian be \( H_0 \). The quantity \( f_0 \) satisfies the CBE,

\[ [f_0, H_0] = 0, \tag{1} \]

where the square brackets correspond to the Poisson bracket. Now let us introduce a small time-dependent perturbation in the potential, \( \Phi(t) \), such that the perturbed Hamiltonian becomes

\[ H = H_0 + \Phi(t) + f_1(t), \tag{2} \]

where \( \Phi \) is the gravitational potential sourced by the response density, \( \rho_1 = \int f_1 d^3v \), via the Poisson equation,

\[ \nabla^2 \Phi = 4\pi G \rho, \tag{3} \]

Here \( f_1 \) is the linear-order perturbation in the DF, i.e., the linear response of the system to the perturbation in the potential. The perturbed DF can thus be written as

\[ f = f_0 + f_1. \tag{4} \]

Assuming that the perturbations are small such that linear perturbation theory holds, the time evolution of \( f_1 \) is governed by the following linearized version of the CBE:

\[ \frac{\partial f_1}{\partial t} + [f_1, H_0] + [f_0, \Phi(t)] + [f_0, \Phi_1] = 0. \tag{5} \]

In this paper we shall neglect the self-gravity of the disk, i.e., neglect the polarization term, \( [f_0, \Phi_1] \), on the left-hand side of the linearized CBE. We briefly discuss the impact of self-gravity in Section 6.2, leaving a more detailed analysis including self-gravity to a forthcoming publication.

2.2. Hybrid Perturbative Formalism for an Infinite Slab

We consider the simplified case of perturbations in an infinitely extended slab, uniform in \((x, y)\), but characterized by a vertical density profile \( \rho(z) \). Although a rather poor approximation of a realistic galactic disk, this idealized case serves to highlight some of the main characteristics of disk response. We consider perturbations that can be described by a profile in the vertical \( z \)-direction and by a superposition of plane waves along the \( x \)-direction, such that \( \Phi_0 \) and \( f_1 \) are both independent of \( y \). After making a canonical transformation from the phase-space variables \((z, v_z)\) to the corresponding action-angle variables \((I_z, w_z)\), Equation (5) becomes

\[ \frac{\partial f_1}{\partial t} + \frac{\partial H_0}{\partial I_z} \frac{\partial f_1}{\partial w_z} + \frac{\partial H_0}{\partial w_z} \frac{\partial f_1}{\partial I_z} - \frac{\partial \Phi_0}{\partial w_z} \frac{\partial f_0}{\partial I_z} - \frac{\partial \Phi_0}{\partial I_z} \frac{\partial f_0}{\partial w_z} = 0. \tag{6} \]

The unperturbed Hamiltonian \( H_0 \) can be written as

\[ H_0 = \frac{v_x^2}{2} + \frac{v_y^2}{2} + v_z^2 + \Phi(z), \tag{7} \]

where \( v_x \), \( v_y \), and \( v_z \) are the unperturbed velocities of stars along \( x, y, \) and \( z \), respectively, and \( \Phi(z) \) is the unperturbed potential that dictates the oscillatory vertical motion of the stars. We expand \( \Phi_0 \) and \( f_1 \) as Fourier series that are discrete along \( z \) but continuous along \( x \):

\[ \Phi_0(z, x, t) = \sum_{n=-\infty}^{\infty} \int dk \exp[i(nw_z + kx)] \Phi_0(I_z, t), \]

\[ f_1(z, v_z, x, v_x, v_y, v_z, t) = \sum_{n=-\infty}^{\infty} \int dk \exp[i(nw_z + kx)] f_{1b}(I_z, v_z, v_x, v_y, v_z, t). \tag{8} \]

Here \( z \) can be expressed as the following implicit function of \( w_z \) and \( I_z \):

\[ w_z = \Omega_z \int_0^z \frac{dz'}{\sqrt{2[E_z(L_z) - \Phi_z(z')]}}, \tag{9} \]

where \( \Omega_z = \Omega_z(L_z) \) is the vertical frequency of stars with vertical action \( I_z \), given in Equation (11) below.

Here and throughout this paper we express any dependence on the continuous wavenumber \( k \) with an index rather than an argument, i.e., \( \Phi_0(I_z, t) \) rather than \( \Phi_0(k, I_z, t) \). This implies that any function that carries \( k \) as an index is in Fourier space.

We express the perturber potential and the DF perturbation or response as linear superpositions of Fourier modes. Since we do not take into account the self-gravity of the response itself, i.e., do not self-consistently solve the Poisson equation along with the CBE, these are not dynamical or normal modes of the system. In other words, the oscillation frequencies of the Fourier modes are just the unperturbed frequencies, \( \Omega_z \), and do not follow a dispersion relation as in the self-gravitating case. To aid the visualization of the various Fourier modes, Figure 1 illustrates what the \( n = 0, n = 1, \) and \( n = 2 \) modes for one particular value of the wavenumber \( k \) look like. The figure also indicates the direction of the velocity impulses resulting from an instantaneous perturbation of each mode.
Substitution of the above expressions in Equation (6) yields the following evolution equation for $f_{1nk}$:

$$\frac{\partial f_{1nk}}{\partial t} + i(n\Omega_c + kv_x)f_{1nk} = i \left( \frac{n}{\partial L_z} + k \frac{\partial}{\partial v_x} \right) \Phi_{nk},$$

(10)

where we have used that

$$\Omega_c = \frac{\partial H_0}{\partial L_z}, \quad v_x = \frac{\partial H_0}{\partial v_x}. \quad (11)$$

The above first-order differential equation in time is easily solved using the Green’s function technique. With the initial condition, $f_{1nk}(t_i) = 0$, we obtain the following integral form for $f_{1nk}$ for a given time dependence of the perturber potential:

$$f_{1nk}(l_z, v_x, v_y, t) = i \left( \frac{n}{\partial L_z} + k \frac{\partial}{\partial v_x} \right) \Phi_{nk}(l_z, \tau) \times \int_{t_i}^{t} d\tau \exp [-i(n\Omega_c + kv_x)(t - \tau)] \Phi_{nk}(l_z, \tau).$$

(12)

This solution is analogous to the particular solution for a forced oscillator with natural frequencies, $n\Omega_c$, and $kv_x$, which is being forced by an external perturber potential, $\Phi_{nk}$. The time dependence of this external perturbation ultimately dictates the temporal evolution of the perturbation in the DF, $f_{1nk}$. A net response requires gradients in the (unperturbed) DF with respect to the actions and/or velocities. Similar solutions for the response of perturbed, collisionless systems have been derived in a number of previous studies (e.g., Lynden-Bell & Kalnajs 1972; Tremaine & Weinberg 1984; Carlberg & Sellwood 1985; Weinberg 1989, 1991, 2004; Kaur & Sridhar 2018; Banik & van den Bosch 2021a; Chiba & Schönrich 2022; Kaur & Stone 2022), often in the context of phenomena like angular momentum transport, radial migration, or dynamical friction.

### 2.3. Perturbation in an Isothermal Slab

The infinite slab has a nonuniform (uniform) density profile along the vertical (horizontal) direction. Therefore, the unperturbed motion of the stars is only vertically bounded by a potential but is unbounded horizontally. This implies that the unperturbed DF, $f_0$, involves a potential $\Phi_z$ only along $z$. For simplicity, we assume it to be isothermal but with different velocity dispersions in the vertical direction, $\sigma_z$, and the in-plane directions, $\sigma_x = \sigma_y \equiv \sigma$, i.e.,

$$f_0(v_x, v_y, E_z) = \frac{\rho_0}{(2\pi)^{3/2}\sigma_x\sigma_y} \exp \left[ -\frac{E_z}{\sigma_z^2} \right] \times \exp \left[ -\frac{v_x^2 + v_y^2}{2\sigma^2} \right].$$

(13)
where
\[ E_z = \frac{1}{2} \dot{v}_z^2 + \Phi_z(z) \] (14)

is the energy involving the \( z \)-motion. The corresponding density and potential profiles in the vertical direction are given by
\[ \rho_z(z) = \rho_c \text{sech}^2(z/h_c), \quad \Phi_z(z) = 2\sigma_2^2 \ln[\cosh(z/h_c)], \] (15)

where \( h_c \) is the vertical scale height (Spitzer 1942; Camm 1950).

The vertical action, \( I_z \), can be obtained from the unperturbed Hamiltonian, \( E_z \), as follows:
\[ I_z = \frac{1}{2\pi} \oint v_z dz = \frac{1}{2} \int_0^{\tau_{\text{max}}} \sqrt{2[E_z - \Phi_z(z)]} dz, \] (16)

where \( \Phi_z(z_{\text{max}}) = E_z \), i.e., \( \tau_{\text{max}} = h_c \cosh^{-1}(\exp{[E_z/2\sigma_2^2]}) \). The time period of vertical oscillation is given by
\[ T_z = \oint v_z dz = 4 \int_0^{\tau_{\text{max}}} \frac{dz}{\sqrt{2[E_z - \Phi_z(z)]}}, \] (17)

and the vertical frequency is \( \Omega_z = 2\pi/T_z \). Throughout this paper, to compute the perturbative response of the slab, we shall use typical MW parameter values, i.e., \( h_c = 0.4 \text{ kpc}, \sigma_2 = 23 \text{ km s}^{-1}, \) and \( \sigma_1 = 1.5 \sigma_c = 35 \text{ km s}^{-1} \) (McMillan 2011).

Substituting the above form for \( f_0 \) (Equation (13)) in Equation (12) and using that \( \Omega_z = \Omega_{z,I} = \partial E_z/\partial \Omega_z \) yields the following closed integral form for \( f_{1ik} \):
\[ f_{1ik}(I_z, v_x, v_y, t) = -i \left( \frac{n\Omega_z}{\sigma_z^2} + \frac{k v_y}{\sigma_z} \right) f_0(v_x, v_y, E_z) \times \int_0^{\tau_{\text{max}}} d\tau \exp[-i(n\Omega_z + k v_y)(t - \tau)] \Phi_{nk}(I_z, \tau). \] (18)

### 2.4. Perturber Potential

The slab response depends on the spatiotemporal nature of the perturber. In this paper we consider two different functional forms of the perturber potential described below.

#### 2.4.1. Separable Potential

In order to capture the essential physics of perturbative collisionless dynamics without much computational complexity, we shall consider the following separable form for the perturber potential:
\[ \Phi_p(z, x, t) = \Phi_N Z(z) X(x) T(t), \] (19)

where \( \Phi_N \) has the units of potential and \( Z, X, \) and \( T \) are dimensionless functions of \( z, x, \) and \( t \), respectively, that specify the spatiotemporal profile of \( \Phi_p \). Thus, the Fourier transform of \( \Phi_p \) can also be written in the following separable form:
\[ \Phi_{nk}(I_z, t) = \Phi_N Z_n(I_z) X_k T(t). \] (20)

Here \( Z_n(I_z) \) is the \( n \)th Fourier coefficient in the discrete Fourier series expansion of \( Z(z) \) in the vertical angle, \( w_z \), given by
\[ Z_n(I_z) = \frac{1}{2\pi} \int_0^{2\pi} dv_z Z(z) \exp[-i n w_z], \] (21)

where we have used the implicit expression for \( z \) in terms of \( w_z \) and \( I_z \) given in Equation (9). \( X_k \) is the Fourier transform of \( \chi(x) \), given by
\[ \chi_k = \frac{1}{2\pi} \int_{-\infty}^{\infty} dx \chi(x) \exp[-ikx]. \] (22)

In the following sections, we investigate the slab response to perturbers with various functional forms for \( \chi(x) \) and \( T(t) \), while keeping the form for \( Z(z) \) arbitrary. We start in Section 3 with an impulsive \( (T(t) = \delta(t)) \) single-mode \( (\chi(x) = \exp[ikx]) \) perturbation, followed in Section 4 by a perturbation that is temporally impulsive but spatially localized \( \exp[-x^2/\Delta^2] \). In Section 5 we consider the same spatially localized perturbation, but this time temporally extended \( (T(t) = \exp[-\omega^2 t^2]) \).

#### 2.4.2. Satellite Galaxy along Straight Orbit

As a practical astrophysical application of our perturbative formalism, we also study the response of an isothermal slab to a satellite galaxy or DM subhalo undergoing an impact along a straight orbit with a uniform velocity \( v_p \) at an angle \( \theta_p \) (with respect to the disk normal). We model the impacting satellite as a point perturber, whose potential is given by
\[ \Phi_p(z, x, t) = \frac{-GM_p}{\sqrt{(z - v_p \cos \theta_p t)^2 + (x - v_p \sin \theta_p t)^2}}. \] (23)

In this case the spatial and temporal parts are coupled, and thus the slab response needs to be evaluated by performing the \( \tau \) integral before the \( w_z \) integral, as shown in Appendix B.

### 3. Response to an Impulsive Perturbation

In order to gain some insight into the perturbative response of the slab, we start by solving Equation (18) for an instantaneous impulse at \( t = 0 \), i.e., \( \delta(t) \). Here the normalization factor \( \Phi_{nk} \) has the units of potential times time. With the initial time \( t_i < 0 \), the integral over \( \tau \) yields
\[ f_{1ik}(I_z, v_x, v_y, t) = \int_{-\infty}^{\infty} dv_y \int_{-\infty}^{\infty} dv_x f_{1ik}(I_z, v_x, v_y, t). \]

Further integrating \( f_{1ik} \) over \( v_x \) and \( v_y \) and summing over all \( n \) modes yields the following form for any given \( k \) mode of the perturbed DF for a given action \( I_z \) and angle \( w_z \):
\[ f_{ik}(I_z, w_z, t) = \frac{1}{\sqrt{2\pi} \sigma_z^2} \int_{-\infty}^{\infty} dv_y \int_{-\infty}^{\infty} dv_x f_{1ik}(I_z, v_x, v_y, t) \]
\[ = A_{\text{norm}} D_k(t) R_k(I_z, w_z, t), \] (24)

where
\[ A_{\text{norm}} = \frac{\theta_p}{\sqrt{2\pi} \sigma_z^2} \exp[-E_2/\sigma_z^2] \] (25)

is a normalization factor reflecting the vertical structure of the unperturbed disk.
\[ D_k(t) = \exp\left[-\frac{k^2 \sigma_z^2 t^2}{2}\right] \] (26)
is a damping term that describes the temporally Gaussian decay of the response by lateral mixing, and
\[
R_k(l_z, w_z, t) = -\Phi_N \mathcal{X}_k \sum_{n=-\infty}^{\infty} Z_n(l_z) 
\times \left( k^2 t + i \frac{n \Omega_z}{\sigma_z^2} \right) \exp \left[ i n (w_z - \Omega_z t) \right]
\] (27)
is a (linear) response function that includes vertical phase mixing.

Equation (24) is the basic “building block” for computing the response of our infinite isothermal slab to a perturbation mode \( k \) in the impulsive limit. Using the canonical transformation from \( (w_z, l_z) \) to \( (v_z, c) \), i.e., using Equations (9) and (14), we can transform \( f_{1d}(l_z, w_z, t) \) to \( f_{2d}(v_z, c, t) \). Upon multiplying this by \( \exp[inkx] \) and integrating over \( k \), and then integrating further over \( v_z \) at a fixed \( z \), one obtains the response density as a function of both time and position:
\[
\rho_1(\zeta, x, t) = -\frac{\rho_0 \Phi_N}{\sqrt{2\pi} \sigma_z} \sum_{n=-\infty}^{\infty} \int_0^L dt \frac{\Omega_z}{\sqrt{2[E_z - \Phi_\zeta(z)]}} \exp \left[ -E_z/\sigma_z^2 \right] \exp \left[ i n (w_z - \Omega_z t) \right] Z_n(l_z) 
\times \int dk \exp[inkx] \exp \left[ -\frac{k^2 \sigma_z^2}{2} \right] \left( k^2 t + i \frac{n \sigma_z}{\sigma_z^2} \right) \mathcal{X}_k, \tag{28}
\]
where \( \mathcal{X}_k \) is the solution of \( E_z(l_z) = \Phi_\zeta(z) \) and \( w_z \) is the solution for \( w_z(l_z, l_z) \) from Equation (9).

In order to gain insight into the slab response for a particular \( l_z \) and \( w_z \), let us start by analyzing Equation (24) for the \( n = 0 \) mode, an in-plane density wave, for which the perturbation causes an in-plane velocity impulse as depicted in Figure 1. The response is a standing, longitudinal oscillation in density. The response function for this mode is
\[
R_k(l_z, w_z, t) = \Phi_N Z_0(l_z) \mathcal{X}_k k^2 t, \tag{29}
\]
indicating that the amplitude of oscillation initially grows linearly with time. However, this growth is inhibited by the Gaussian damping function \( D_k(t) = \exp[-\nu k^2 \sigma_z^2 t^2] \), which describes lateral damping due to the nonzero velocity dispersion of stars in the \( k \)-direction. The Gaussian form of this temporal damping term owes its origin to the assumed Gaussian/Maxwellian form of the unperturbed velocity distribution along the plane. Hence, following the perturbation, the \( n = 0 \) mode starts to grow linearly with time but then rapidly damps away; the response loses its coherence owing to mixing in the direction of the wavevector. In the cold slab limit \((\sigma \rightarrow 0)\), without any lateral streaming motion to damp it out, the response will grow linearly in time until it eventually becomes nonlinear. This is because in an infinite, laterally homogeneous slab there is no restoring force in the lateral directions, causing the stars to stream uninhibited toward (away from) the minima (maxima) of \( \Phi_\zeta \) owing to the initial velocity impulse induced. This leads to over- and underdensity spikes that cannot be treated using linear theory. Hence, Equation (24) can only adequately describe the response to an instantaneous \( n = 0 \) mode at early times, or if the damping time \( \tau_D \sim (\sigma k)^{-1} \) is shorter than the timescale of formation of density spikes. The latter is roughly the time needed to cross one-quarter of the perturbation’s wavelength with the velocity impulse triggered at the zeros of \( \Phi_\zeta \). Therefore, in order for linear theory to be valid, we require that \( \sigma > (2/\pi) \Delta v \) [maximum].

The Astrophysical Journal, 935:135 (20pp), 2022 August 20

Banik, Weinberg, & van den Bosch

words, \( \lambda < \lambda_0 \approx \sqrt{\pi/2G\rho_0} \), where \( k_i \) and \( \lambda_0 = 2\pi/k_i \) refer to the Jeans wavenumber and Jeans wavelength, respectively. In the \( \sigma \rightarrow 0 \) limit, the Jeans wavelength \( \lambda_0 \rightarrow 0 \), and thus the \( n = 0 \) mode becomes globally unstable. Hence, the condition of Jeans stability requires an additional constraint on \( \sigma \): \( \sigma > \sqrt{4\pi G\rho_0}/k \).

For \( n = 1 \), the perturbation is a standing, transverse wave on the slab, formally known as the bending wave. The perturbation induces velocity impulses in the direction perpendicular to the perturbation, as indicated in Figure 1. At the locations marked A and B, separated by a lateral distance of \( \pi/k \), these velocity impulses point in the positive and negative \( z \)-directions, respectively. The top panels of Figure 2 illustrate the impact this has at location A. The left panels indicate the velocity impulses (cyan arrows) in the \((v_z, z)\)-plane. Prior to the perturbation, due to the vertical restoring force from the slab, each star executes a periodic oscillation in this plane. The black and yellow contours indicate the corresponding phase-space trajectories for two values of \( l_z \), while the heat map indicates phase-space density (bluer colors indicate higher density). The top middle panel shows that immediately following the impulse the phase-space density is boosted (reduced) where \( v_z > 0 \) (\( v_z < 0 \)), resulting in a dipole pattern in the phase-space distribution of stars. After the impulse, the stars continue to execute periodic motion in the \((v_z, z)\)-plane, but starting from their new position (corresponding to a modified action \( l_z \)). The angular frequency of this periodic motion is \( \Omega_z \), which is a function of the (modified) action, and hence stars with different actions oscillate in the \((v_z, z)\)-plane at different frequencies. As a consequence, the perturbed phase-space density shown in the middle panels is wound up into a phase spiral of over- and underdensities as depicted in the right panels of Figure 2. The bottom panels of Figure 2 show what happens following the impulsive perturbation at location B. Since the velocity impulses are now reversed in direction, the phase spiral that emerges is exactly the opposite of that at location A.

The creation of phase spirals is an outcome of phase mixing in the \( z \)-direction and is described by the oscillatory factor, \( \exp[i n (w_z - \Omega_z t)] \), that is part of the response function \( R_k(l_z, w_z, t) \). It consists of two terms: a term that scales as \( k^0 \), which describes the lateral streaming motion of stars due to the nonzero velocity impulses in the lateral directions (see Figure 1), and a term that scales as \( n \Omega_z / \sigma_z^2 \), which purely describes the vertical oscillations. As in the case of the \( n = 0 \) mode, the lack of a restoring force in the lateral directions \( z \) causes the perturbation to grow linearly with time in the absence of lateral streaming (for a cold disk with \( \sigma \approx 0 \)). Meanwhile, the phase spirals continue to wind up, which implies that the vertical bending loses its coherence. Over time, phase mixing in the vertical direction will ensure that the disk regains mirror symmetry with respect to the midplane, but with a scale height, \( h_z \), that would be a periodic function of \( x \), with a wavelength equal to \( \pi/k \) (i.e., half the wavelength of the original perturbation).

5 If accounting for self-gravity of the response density, there will be nonzero forces in the lateral direction, but these will promote growth rather than act as a restoring force. This ultimately leads to exponential growth (according to linear theory) of unstable modes and Landau damping of stable modes, which occurs exponentially, i.e., more slowly than the Gaussian lateral mixing in the absence of self-gravity.
positive and negative, respectively, of the phase spiral amplitude. Response at A reveals a one-armed phase spiral that is exactly opposite of that at location B, i.e., they exactly cancel each other. Hence, lateral mixing causes damping of the phase spiral amplitude at different locations in phase space. Note that, in the case of the density of stars in the vertical direction. Stars that received an impulse $\Delta v_z > 0$ create phase spirals that are exactly the inverse of those created by neighboring stars for which the impulse was negative. Thus, lateral mixing between neighboring points on the slab causes a damping of the phase spiral amplitude at any location, a process that is captured by the damping function $D_k(t)$. The lateral mixing timescale is $\tau_P \sim 1/k\sigma$, indicating, as expected, that small-scale perturbations (larger $k$) mix faster, and that mixing is more efficient for larger velocity dispersion in the lateral direction. After a few mixing timescales, the slab will once again be completely homogeneous (laterally), with a scale height $h_z$ that is independent of location. In addition, the density of stars in the $(z, v_z)$-plane will once again be perfectly symmetric without any trace of a phase spiral. The slab has completely equilibrated, and the only impact that remains of the impulsive perturbation is that the new scale height is somewhat larger than it was originally, i.e., the impulsive perturbation has injected energy into the disk, which causes it to puff up in the vertical direction. Hence, the final outcome is as envisioned in the impulsive-heating scenario discussed in the seminal study of Toth & Ostriker (1992). This persistent effect in the vertical density profile is, however, only captured in perturbation theory at second order (e.g., Carlberg & Sellwood 1985); to first order the perturbation simply phase-mixes away in the impulsive limit considered here.

For $n = 2$, the perturbation triggers a breathing mode, as depicted in Figure 1, i.e., at a given location A on the slab, the velocity impulses for this mode are positive (negative) for positive (negative) $z$. As evident from Figure 3, this leads to a quadrupole pattern for the initial perturbed phase-space distribution of stars, which becomes a two-armed phase spiral over time, as opposed to the one-armed phase spiral resulting from the $n = 1$ mode. This reveals an important lesson: the structure of a phase spiral depends, among others, on which perturbation mode(s) are triggered. The phase spirals in regions A and B are each other’s additive inverse. Hence, once again lateral mixing will cause damping of the phase spiral’s amplitude, as described by the damping function $D_k(t)$. Hunt et al. (2021) have shown, using $N$-body simulations, that two-armed phase spirals can indeed arise from breathing-mode oscillations and that both bending and breathing modes can be excited at different locations on the MW disk by satellite-induced perturbations such as the passage of Sagittarius (see Section 6.1 for detailed discussion).

**Figure 2.** The formation of a one-armed phase spiral due to an impulsive $n = 1$ bending-mode perturbation. The color-coding in the left panels shows the unperturbed DF $f_d(z, v_z)$ (Equation (13)) in the isothermal slab at neighboring locations A (top) and B (bottom), separated by a lateral distance of $\pi/k$, with blue (red) indicating a higher (lower) phase-space density. Locations A and B coincide with extrema in the perturbation mode as depicted in Figure 1. The black and yellow contours indicate the phase-space trajectories for two random values of $E_1$ (or, equivalently, $I_1$). The cyan arrows indicate the velocity impulses resulting from the instantaneous perturbation at different locations in phase space. Note that, in the case of the $n = 1$ mode considered here, at the extrema A and B all velocity impulses $\Delta v_z$ are positive and negative, respectively (see Figure 1). The middle panels indicate the response $f_1$ immediately following the instantaneous response (at $t = 0$), with blue (red) indicating a positive (negative) response density. Finally, the right panels show the response after some time $t$, computed using Equation (24). Note how the response at A reveals a one-armed phase spiral that is exactly opposite of that at location B, i.e., they exactly cancel each other. Hence, lateral mixing causes damping of the phase spiral amplitude.
To summarize, we see that, in the case of our infinite slab, equilibration after an impulsive perturbation is driven by a combination of phase mixing in the vertical direction and free-streaming damping in the horizontal direction. While the former gives rise to phase spirals in the $(z, w_z)$ plane, the latter causes them to damp away by lateral mixing. Due to vertical phase mixing, the phase spiral will continue to wrap itself up into a more and more tightly wound pattern, until its structure can no longer be discerned observationally owing to finite-$N$ noise (Beraldo e Silva et al. 2019a, 2019b) and measurement errors in the actions and angles of individual stars (this is an example of coarse-grain mixing). Hence, even without lateral mixing, phase spirals are only detectable for a finite duration.

4. Response to a Localized Perturbation

In the previous section we investigated the slab response to an external disturbance with a single wavenumber $k$. Realistic perturbations are, however, localized in space and thus consist of many wavenumbers. In this section we shall look into what happens when the slab is hit by an impulsive perturbation that is spatially localized.

For simplicity, we assume that the external perturber behaves as a Gaussian packet with half-width $\Delta_x$, along the $x$-direction, i.e., $\Phi_p$ is given by Equation (19) with

$$\lambda(x) = \exp[-x^2/2\Delta_x^2].$$  \hspace{1cm} (30)

The $Z(z)$ term in Equation (19) denotes the vertical structure of the perturber potential, which is part of what dictates the relative strength of bending- and breathing-mode oscillations. We shall see in the next section, though, that the relative strength of the modes is mostly dictated by the form of $T(t)$. For simplicity, we only consider localization along the $x$- and $z$-directions; along the $y$-direction the perturbation is assumed to extend out to infinity.

We emphasize, though, that this assumption does not impact the essential physics of phase mixing and lateral mixing discussed below.

The Fourier transform of the perturber potential, $\Phi_{nk}$, is given by Equation (20), with

$$\lambda_k = \frac{\Lambda_k}{\sqrt{2\pi}} \exp[-k^2\Delta_x^2/2].$$ \hspace{1cm} (31)

Upon substituting the above expression for $\lambda_k$ in Equation (24), we obtain the response for a single $k$ mode, $f_{1k}$. After multiplying this by $\exp[ikx]$, integrating over all $k$ and summing over all $n$ modes, we obtain the following final form for the slab response density in the case of a (laterally) Gaussian perturber:

$$f_{1}(l_z, w_z, x, t) = \sum_{n=-\infty}^{\infty} \exp[inw_z] \times \int_{-\infty}^{\infty} dk \exp[ikx] f_{1k}(l_z, w_z, t) = A_{\text{norm}} D(x, t) R(l_z, w_z, x, t),$$ \hspace{1cm} (32)
where
\[ A_{\text{norm}} = \frac{\rho_{0}}{\sqrt{2\pi}\sigma_{z}} \exp\left[-E_{z}/\sigma_{z}^{2}\right] \] (33)

is the same normalization factor as in Equation (24),
\[ D(x, t) = \frac{\Delta_{x}}{\sqrt{\Delta_{x}^{2} + \sigma^{2}t^{2}}} \exp\left[-\frac{x^{2}}{2(\Delta_{x}^{2} + \sigma^{2}t^{2})}\right] \] (34)
is a factor that captures the decay of the response by lateral mixing, and
\[
R(I_z, w_z, x, t) = -\Phi_N \sum_{n=-\infty}^{\infty} Z_n(I_z) \\
\times \left[ \frac{t}{\Delta_{x}^{2} + \sigma^{2}t^{2}} \left( 1 - \frac{x^{2}}{\Delta_{x}^{2} + \sigma^{2}t^{2}} \right) + \frac{i n\Omega_{z}}{\sigma_{z}^{2}} \right] \\
\times \exp\left[in(w_z - \Omega_z t)\right],
\] (35)

with \( Z_n(I_z) \) given by Equation (21), corresponds to the remaining part of the response that includes vertical phase mixing.

The above expression (Equation (32)) for the slab response to a localized disturbance has several important features. First, the profile of the slab response is nearly Gaussian in \( x \) since we assumed a Gaussian form (along \( x \)) for the perturber potential. Second, the \( D(x, t) \) factor describes the decay of the response amplitude and widening of the response profile due to mixing by lateral streaming. The mixing in this case occurs as a power law in time rather than like a Gaussian as for a single \( k \) mode (see Equation (24)), since the power spectrum of the Gaussian perturber is dominated by small \( k \) that mix very slowly, at a timescale \( \tau_{D} \sim 1/\sigma \). Third, the \( R \) factor captures two important effects: (i) a transient response reflecting an initial linear growth due to the perturber-induced velocity impulse, followed by a subsequent decay by lateral mixing; and (ii) vertical oscillations of stars (for \( n \neq 0 \)) at different frequencies resulting in phase mixing over time and the formation of phase spirals as described in detail in Section 3. The \( n = 0 \) modes, i.e., perturbations confined to the slab, damp out faster than the nonzero \( n \) modes that manifest the vertical oscillations of stars. Since the perturber was introduced impulsively by means of a Dirac delta function in time, the higher-order oscillations are stronger for the same value of \( Z_n(I_z) \), as the corresponding changes in the vertical actions have larger amplitude. Typically, for \( n \geq 2 \), \( Z_n(I_z) \) gets smaller with larger \( n \); hence, the \( n = 2 \) breathing mode turns out to be the dominant mode of oscillation for impulsive disturbances. The response characteristics, however, change as we move to nonimpulsive or more temporally extended perturbers in the next section.

It takes time for the local response to propagate along the slab by lateral streaming. Initially the perturber’s gravity draws in stars toward the center of impact, \( x = 0 \). Thus, immediately following the impulse, the region near the center of impact has a larger concentration of stars, which laterally stream outward owing to nonzero velocity dispersion. This leads to a damping of the response amplitude at small \( x \) and growth at large \( x \), or equivalently damping and widening of the response profile, which occurs at the rate
\[ D_{s}(t) = \frac{d}{dt} \sqrt{\Delta_{x}^{2} + \sigma^{2}t^{2}} = \frac{\sigma^{2}t}{\sqrt{\Delta_{x}^{2} + \sigma^{2}t^{2}}}. \] (36)

This rate of outward streaming of slab material is initially equal to
\[ \lim_{t \to 0} D_{s}(t) = \frac{\sigma^{2}t}{\Delta_{x}}, \] (37)
but at later times it asymptotes to a constant value,
\[ \lim_{t \to \infty} D_{s}(t) = \sigma. \] (38)

To summarize, the response to a spatially localized perturbation can be understood in the context of that to a single mode plane-wave perturbation discussed in the previous section, but also be extended in time. In both cases, the response involves vertical oscillations that phase-mix away, thus giving rise to phase spirals. However, whereas the plane-wave response maintains its sinusoidal profile in the lateral direction with an overall Gaussian decay of the amplitude due to lateral mixing, the response profile in the case of localized perturbation changes its shape and undergoes both decay and widening. This is because in the latter case the response is an integral superposition of responses to many plane-wave perturbations with different \( k \), each decaying in amplitude over a timescale \( \tau_{D} \sim 1/\sigma \), due to lateral mixing by free streaming.

The above expression (Equation (32)) for the slab response to a localized disturbance has several important features. First, the profile of the slab response is nearly Gaussian in \( x \) since we assumed a Gaussian form (along \( x \)) for the perturber potential. Second, the \( D(x, t) \) factor describes the decay of the response amplitude and widening of the response profile due to mixing by lateral streaming. The mixing in this case occurs as a power law in time rather than like a Gaussian as for a single \( k \) mode (see Equation (24)), since the power spectrum of the Gaussian perturber is dominated by small \( k \) that mix very slowly, at a timescale \( \tau_{D} \sim 1/\sigma \). Third, the \( R \) factor captures two important effects: (i) a transient response reflecting an initial linear growth due to the perturber-induced velocity impulse, followed by a subsequent decay by lateral mixing; and (ii) vertical oscillations of stars (for \( n \neq 0 \)) at different frequencies resulting in phase mixing over time and the formation of phase spirals as described in detail in Section 3. The \( n = 0 \) modes, i.e., perturbations confined to the slab, damp out faster than the nonzero \( n \) modes that manifest the vertical oscillations of stars. Since the perturber was introduced impulsively by means of a Dirac delta function in time, the higher-order oscillations are stronger for the same value of \( Z_n(I_z) \), as the corresponding changes in the vertical actions have larger amplitude. Typically, for \( n \geq 2 \), \( Z_n(I_z) \) gets smaller with larger \( n \); hence, the \( n = 2 \) breathing mode turns out to be the dominant mode of oscillation for impulsive disturbances. The response characteristics, however, change as we move to nonimpulsive or more temporally extended perturbers in the next section.

It takes time for the local response to propagate along the slab by lateral streaming. Initially the perturber’s gravity draws in stars toward the center of impact, \( x = 0 \). Thus, immediately following the impulse, the region near the center of impact has a larger concentration of stars, which laterally stream outward owing to nonzero velocity dispersion. This leads to a damping of the response amplitude at small \( x \) and growth at large \( x \), or equivalently damping and widening of the response profile, which occurs at the rate
\[ D_{s}(t) = \frac{d}{dt} \sqrt{\Delta_{x}^{2} + \sigma^{2}t^{2}} \] (36)

This rate of outward streaming of slab material is initially equal to
\[ \lim_{t \to 0} D_{s}(t) = \frac{\sigma^{2}t}{\Delta_{x}}, \] (37)
but at later times it asymptotes to a constant value,
\[ \lim_{t \to \infty} D_{s}(t) = \sigma. \] (38)

5. Response to a Nonimpulsive Perturbation

Thus far we have only considered impulsive perturbations of our slab, with \( \mathcal{T}(t) = \delta(t) \). However, a realistic disturbance would not only have a spatial structure, the effects of which we studied in the previous section, but also be extended in time. In this section we investigate the effect of nonimpulsive or temporally extended disturbances on the slab oscillations. In particular, we broaden the Dirac delta pulse from the previous section into a Gaussian in time, i.e., \( \Phi_{p} \) is given by Equation (19) with \( \mathcal{T}(t) = \frac{1}{\sqrt{\pi}} \exp\left[-\omega_0^2 t^2\right] \), where \( \omega_0 \) is the pulse frequency. We define the pulse width or pulse time as \( \tau_{\text{pulse}} = \sqrt{2/\omega_0} \). We also assume that the pulse is localized and follows a Gaussian profile in \( x \) as in the previous section, i.e., \( \lambda(x) = \exp\left[-x^2/2\Delta_{\lambda}^2\right] \). As before, \( Z(z) \) in Equation (19) denotes some generic vertical profile. The (spatial) Fourier transform of this potential, \( \Phi_{nk} \), is provided in Equation (20), with \( \lambda_k \) given by Equation (31) and \( Z_k \) given by Equation (21).

We can substitute this in Equation (12) and perform the integration over \( \tau \) and \( v_z \) to obtain the following expression for the response for a single \( k \) mode:
\[ f_{ik}(I_z, w_z, t) = A_{\text{norm}} D_k(t) R_k(I_z, w_z, t), \] (39)
where
\[ A_{\text{norm}} = \frac{\rho_{0}}{\sqrt{2\pi}\sigma_{z}} \exp\left[-E_{z}/\sigma_{z}^{2}\right] \] (40)
is the same normalization factor as in Equation (24),
\[ D_k(t) = \frac{Q^3}{2\omega_0} \exp\left[-Q^2k^2\sigma^2/2\right] \]  
(41)
is a factor that describes the damping of the response due to lateral mixing, and
\[ R_k(I, z, t) = -\Phi_N X_k \exp\left\{i(k\omega_0 t - k\sigma z)\right\} \]
\[ \times \left[ S_{nk}(k^2 + i\omega_0 t - i\sigma z) \exp\left[2\nu_i(w_z - Q\Omega_z t)\right] - G_{nk}(w, t) \right], \]  
(42)
with \( Z_n(I) \) given by Equation (21), includes the vertical phase mixing of the response. Here \( Q \) is a factor that depends on the pulse frequency, \( \omega_0 \), and the wavenumber, \( k \), and is given by
\[ Q = Q(\omega_0, k\sigma) = \frac{\omega_0}{\sqrt{\omega_0^2 + k^2\sigma^2}}. \]  
(43)
The mode strength,
\[ S_{nk} = \exp\left[-\frac{1}{\omega_0^2 + k^2\sigma^2/2}\right], \]  
(44)
is a function that indicates the strength of each \( n \) mode,
\[ \tau_n(t) = 1 + \text{erf}\left\{Q(\omega_0 t - i\nu_i)\right\} \]  
(45)
describes the temporal buildup of the response and the decay of transient oscillations, and
\[ G_{nk}(w_z, t) = \frac{k^2}{\sqrt{\pi} \omega_0 Q} \exp\left[-Q^2w_z^2\right] \exp\left[2\nu_i w_z\right] \]  
(46)
is another rapidly decaying transient feature. In the \( \omega_0 \to \infty \) limit, both \( \tau_n(t) \) and the mode strength \( S_{nk} \) become unity, and \( G_{nk}(w_z, t) \to 0 \), such that we recover the response to impulsive perturbations given in Equation (24) as required.

It is interesting to contrast this response to an extended pulse to that in the impulsive limit. First of all, the damping factor, \( D_k(t) \), which still owes its origin to lateral mixing due to nonzero velocity dispersion, now depends not only on \( k \) and \( \sigma \) but also on the pulse frequency \( \omega_0 \). The damping time is given by
\[ \tau_D = \frac{1}{k\sigma} \sqrt{1 + \frac{k^2\sigma^2}{2\omega_0^2}}, \]  
(47)
which scales as \( \sim 1/k\sigma \) in the impulsive/short-pulse (\( \omega_0 \gg k^2\sigma^2/2 \)) limit, indicating that the response mixes away laterally, with small-scale perturbations mixing faster. In the adiabatic/long-pulse (\( \omega_0 \ll k^2\sigma^2/2 \)) limit, however, \( \tau_D \to 1/\sqrt{2\omega_0} \), i.e., the damping of the response follows the temporal behavior of the perturbing pulse itself, independent of \( k \).

The mode strength reveals several important trends: it exponentially damps away with \( n^2 \), implying that the lower-order modes are much stronger for perturbations that are slower (see also Widrow et al. 2014) and/or have larger wavelength (smaller \( k \)). Therefore, the \( n = 1 \) bending modes dominate over the \( n = 2 \) breathing modes for a sufficiently slow pulse. Note, though, that if the pulse is too slow (\( \omega_0 \to 0 \)), the mode strength is super-exponentially suppressed, especially at large scales (small \( k \)), or if the slab has a small lateral velocity dispersion, \( \sigma \), compared to that along the vertical direction, \( \sigma_z \) (recall that \( \Omega_z \sim \sigma_z/\hbar \)). This is a classic signature of adiabatic shielding of the slab due to the averaging out of the net response to zero by many oscillations of stars within the (very long) perturbation timescale (see Weinberg 1994a, 1994b; Gnedin & Ostriker 1999).

Finally, if the perturbation is not impulsive, the frequency with which the slab stars oscillate in the vertical direction is modified with respect to their natural frequency according to
\[ \Omega_z \to \frac{\omega_0^2}{\omega_0^2 + \hbar^2\sigma_z^2}, \]  
(48)
which goes to \( \Omega_z \) in the impulsive limit, as expected. For slower pulses, however, the vertical motion of the stars couples to the lateral motion (see also Binney & Schönrich 2018), resulting in a reduced oscillation frequency, especially for smaller wavelengths (larger \( k \)). In the extremely slow/adiabatic limit, \( \Omega_z \to 0 \), signaling a lack of vertical phase mixing. This is easy to understand; a forced oscillator remains in phase with the perturber if the frequency of the latter is much lower than the natural frequency. In fact, in the adiabatic limit, the response only consists of resonant stars, for which \( n\Omega_z + kv_z = 0 \) (see Appendix A), and thus no phase spiral emerges.

The above response corresponds to a temporally Gaussian pulse for a fixed wavenumber \( k \). To get the full response to a localized perturber, we substitute the expression for \( X_k \) given in Equation (31), in the \( k \)-response of Equation (39), multiply it by \( \exp[ikx] \), and integrate over all \( k \). The resultant response is an oscillating function of \( w_z \) and has a profile along \( x \) that varies with time. For the short pulse/impulsive case, we recover the expression given in Equation (32). In Figure 4 we plot the amplitude (relative to the unperturbed DF) of this oscillating response (normalized by the Fourier component of the perturber potential, \( Z_n(I) \) as a function of \( x \). The columns correspond to four different times since the time of maximum pulse strength, and the rows correspond to two different pulse times, as indicated. The solid and dashed lines indicate the bending \( (n = 1) \) and breathing \( (n = 2) \) modes, respectively. The short-pulse response shown in the top panels has a Gaussian profile centered on the point of impact at \( x = 0 \), with the initial width very similar to that of the \( \Phi_p \) profile (see Equations (32)−(35)). Over time, this response profile gets weaker and wider like a power law, as the unconstrained lateral motion of the stars causes an outward streaming, and thus decay, of the response. The long-pulse response in the bottom panels has a different, more extended profile than in the short-pulse case; it exhibits some ripples along \( x \), besides having an overall smooth behavior (see Appendix A for the response derived in the adiabatic limit). As time goes on, the response decays away and widens out owing to lateral mixing. Unlike the short-pulse case, here the response initially decays like \( \sim \exp[-\omega_0^2t^2] \) over a timescale of the pulse time, \( \tau_{\text{pulse}} = \sqrt{2}/\omega_0 \), before attaining a power-law decay at large time.

The temporal behavior of the response becomes even clearer in Figure 5, where we plot the amplitude of the response as a function of time at two different positions on the slab (different rows) and for three different pulse times (different columns). As before, the solid and dashed lines indicate the \( n = 1 \) and \( n = 2 \)
modes, respectively. Initially the slab response grows nearly hand in hand with the perturbing pulse. This is captured by the $\varrho_{nk}(t)$ term in the expression for $R_k(I_z, w_z, t)$, which scales as $\exp[-\varrho^2 \sigma_z^2 t^2]$ at small $t$ but asymptotes to a constant value of 2 at late times. As the perturber strength falls off, the response decays as a Gaussian for each $k$, as described by the damping factor, $D_k(t) \propto \exp[-Q k^2 \sigma^2 t^2 / 2]$. The combined response from all $k$, however, decays at a different rate. For the
shortest pulse, for which the response asymptotes to that given by Equation (32), the damping factor $D(x, t) \propto 1/t$ at late times. In the intermediate- and long-pulse cases, the response initially tends to follow the same $\sim \exp[-\omega_I t^2]$ decay as the perturbing pulse, before finally transitioning to a power-law fall-off, which typically occurs as $\sim 1/t$, just as in the short-pulse case. Importantly, this transition sets in later for longer-lasting pulses, such that the late-time response for slower perturbations is drastically suppressed with respect to faster perturbations. From the bottom panels, it is evident that the region ($x = 10h_c$) farther away from the center of impact responds later, with a time lag of $\Delta t = 10h_c/\sigma$ (timescale of lateral streaming), which is $\sim 115$ Myr for the typical MW parameter values adopted here. The breathing mode is the dominant mode in the short-pulse case ($\tau_{\text{pulse}} = 10$ Myr), while in both the intermediate- ($\tau_{\text{pulse}} = 50$ Myr) and long-pulse ($\tau_{\text{pulse}} = 100$ Myr) scenarios the breathing mode eventually dominates. Note, though, that if the pulse becomes too long, the long-term response is adiabatically suppressed. Hence, there is only a narrow window in pulse widths for which bending modes dominate and cause a detectable response. In the next section we examine whether any of the MW satellites have encounters with the disk over timescales that fall in this regime.

The response formalism for localized, nonimpulsive perturbations developed so far can be used to model the response to transient bars and spiral arms. Encounters with such features can cause transient vertical perturbations in the potential over timescales comparable to the vertical oscillation periods of stars, thereby creating phase spirals. We discuss this in detail in Paper II for realistic disk galaxies.

6. Encounters with Satellite Galaxies

In all cases considered above we have made the simplifying assumption that the perturbing potential is separable, i.e., can be written in the form of Equation (19). However, a realistic perturber is seldom of such simple form. For example, the potential due to an impacting satellite galaxy or DM subhalo (approximated as a point perturber) cannot be written in separable form, thereby making the analysis significantly more challenging. In this section, as an astrophysical application of the perturbative formalism developed in this paper, we compute the response of the infinite slab to a satellite encounter. We relegated the far more involved computation of the response of a realistic disk to an impacting satellite to Paper II.

As shown in Appendix B, the $n = 0$ response to a satellite impacting the slab with a uniform velocity $v_p$ along a straight orbit at an angle $\theta_p$, at a distance $x$ from the point of impact, can be approximated as

$$f_l(x, v_p, x, t) = \frac{\beta}{\sqrt{2}\pi \sigma_z} \exp[-E_z/\sigma_z^2] i \frac{2GM_p}{v_p}$$

$$\times \sum_{n=-\infty}^{\infty} \frac{n\Omega_z}{\sigma_z^2} \Psi_n(x, l, t) \exp \left[ i n\Omega_z \sin \theta_p \frac{v_p}{\sigma_z} \right]$$

$$\times \exp \left[ i n(w_z - \Omega_p t) \right],$$

where

$$\Psi_n(x, l, t) = \frac{1}{2\pi} \int_0^{2\pi} dw_c \exp \left[ -i n(w_c - \Omega_z \cos \theta_p z) \right]$$

$$\times K_0 \left[ \frac{n\Omega_z (x \cos \theta_p - z \sin \theta_p)}{v_p} \right].$$

with $K_0$ the zeroth-order modified Bessel function of the second kind. This expression for the response is only valid far away from the point of impact ($x \gtrsim 1\sigma$), such that the response can be approximated as a plane wave along $x$, and at late times, after the perturber has moved far enough away from the disk, i.e., for $t \gg (x \sin \theta_p + z \cos \theta_p)/v_p$.

There are several salient features of this response that deserve special attention. The strength of the response is dictated by the $K_0$ function, whose argument depends on $\Omega_z \cos \theta_p x/v_p$ (for small $l_z$), which is basically the ratio of the encounter timescale,

$$\tau_{\text{enc}} = \frac{x \cos \theta_p}{v_p}$$

and the vertical dynamical time of the stars,

$$\tau_z = \frac{1}{\Omega_z} \sim \frac{h_z}{\sigma_z}.$$ (52)

From the asymptotic limits of $K_0$ it follows that the response scales as a power law ($\sim v_p^{-1}$) in the impulsive ($\tau_{\text{enc}} \ll \tau_z$) limit and as $\sim \exp[-n\Omega_z \cos \theta_p x/v_p]$ in the adiabatic ($\tau_{\text{enc}} \gg \tau_z$) limit. The response peaks roughly at the maximum of the $\Omega_z$ function, which occurs when the encounter timescale is comparable to the vertical dynamical time of the stars, i.e., when $\tau_{\text{enc}} \approx 0.6 \tau_z$, or in other words, when the “resonance” condition,

$$\frac{x \cos \theta_p}{v_p} \approx 0.6 \Omega_z,$$ (53)

is satisfied. For encounters faster than this, the response is suppressed like a power law, while for slower encounters it is exponentially suppressed. The $v_p^{-1}$ scaling of the response in the impulsive limit is a well-known feature of impulsive perturbations (e.g., Spitzer 1958; Aguilar & White 1985; Weinberg 1994a, 1994b; Gnedin et al. 1999; Banik & van den Bosch 2021b), and the exponential suppression is a telltale signature of adiabatic shielding, similar to the adiabatic suppression of the mode strength factor in the response to slow Gaussian pulses discussed in Section 5. While the response is heavily damped for very slow encounters, something interesting happens in the mildly slow regime, $\tau_{\text{enc}} = x \cos \theta_p / v_p \gtrsim \tau_z$. In this regime, the ratio of the $n = 2$ breathing to the $n = 1$ bending-mode response scales as

$$f_{l_1} \equiv \frac{f_{l_{n=2}}}{f_{l_{n=1}}} \sim \sqrt{2} \exp \left[ -\Omega_z \cos \theta_p x/v_p \right].$$ (54)

Thus, the bending-mode response exponentially dominates over that of the breathing mode for slower (smaller $v_p$), more distant (large $x$), and more perpendicular ($\theta_p \approx 0$) encounters. The bending mode is also more pronounced for stars with larger $\Omega_z$ or equivalently smaller $l_z$. On the other hand, for encounters with $\tau_{\text{enc}} = x \cos \theta_p / v_p < \tau_z$, the breathing modes dominate.

Finally, the slab response to the impacting satellite, given in Equation (49), consists of oscillating functions of time, lateral

4 While the adiabatic response in one-degree-of-freedom cases, e.g., the vertical phase spiral in the isothermal slab, is exponentially suppressed, that in multiple-degree-of-freedom systems such as inhomogeneous disks is usually not because of resonances (Weinberg 1994a, 1994b).
distance $x$, and the vertical oscillation amplitude, $\sqrt{2L_z}/v$ (see Equations (B9) and (B10)). This implies that the satellite induces not only temporal oscillations, which give rise to phase mixing and thus phase spirals due to different oscillation frequencies of the stars (see Section 3), but also spatial corrugations. These vertical and lateral waveforms have wavenumbers given by

$$k_z = \frac{n\Omega_z \cos \theta_p}{v_p} \quad \text{and} \quad k_x = \frac{n\Omega_z \sin \theta_p}{v_p},$$

(55)

respectively. Thus, perpendicular impacts induce only vertical corrugations while planar ones excite waves only laterally. An inclined encounter, on the other hand, spawns corrugations along both directions. Both wavelengths get longer with decreasing mode order, increasing impact velocity, and decreasing vertical frequencies, i.e., increasing actions.

6.1. Impact of Satellite Galaxies on the Milky Way Disk

The MW halo harbors many satellite galaxies. Some of these are quite massive, with DM halo mass comparable to the disk mass, and either underwent or are about to undergo an major impact. These estimates for comparison with actual data and numerical simulations. Our approach also ignores the presence of a DM halo, which can impact the disk response in several ways (see Section 6.2). Because of all these shortcomings, we caution against using the following response estimates for comparison with actual data and/or detailed numerical simulations.

We consider the MW satellites with parallax and proper-motion measurements from Gaia DR2 (Gaia Collaboration et al. 2018c) and the corresponding galactocentric coordinates and velocities computed and documented by Riley et al. (2019), their Table A.2; see also Li et al. 2020 and Vasiliev & Belokurov (2020). Of these, we only consider the satellites with known dynamical mass estimates (Simon & Geha 2007; Bekki & Stanimirović 2009; Łokas 2009; Erkal et al. 2019). Adopting the initial conditions for galactocentric positions ($R, z, \phi$) and velocities ($v_R, v_z, v_\phi$) as the median values quoted by Riley et al. (2019) and Vasiliev & Belokurov (2020), we simulate the orbits of the galaxies in the combined gravitational potential of the MW halo, disk, and bulge, which are respectively modeled by a spherical NFW (Navarro et al. 1997) profile (virial mass $M_{\text{vir}} = 9.78 \times 10^{11} M_\odot$, scale radius $r_h = 16$ kpc, and concentration $c = 15.3$), a Miyamoto–Nagai (Miyamoto & Nagai 1975) profile (mass $M_d = 9.5 \times 10^{10} M_\odot$, scale radius $a = 4$ kpc, and scale height $b = 0.3$ kpc), and a spherical Hernquist (1990) profile (mass $M_b = 6.5 \times 10^9 M_\odot$ and scale radius $r_b = 0.6$ kpc). The total mass of our fiducial MW model is thus $1.08 \times 10^{12} M_\odot$. We evolve the positions and velocities of the satellites both forward and backward in time from the present day, using a second-order leap-frog integrator. For simplicity, we ignore the effect of dynamical friction. From each individual orbit, we note the time, $t_{\text{cross}}$, when the satellite crosses the disk (i.e., crosses $z = 0$) and record the corresponding distance, $x_p$, from the Sun, which we integrate backward/forward in time using a purely circular orbit up to $t_{\text{cross}}$. We also record the velocity, $v_p = \sqrt{v^2 + v_z^2 + v_\phi^2}$, and the angle of impact with respect to the disk normal, $\theta_p = \cos^{-1}(v_z/v_p)$. Finally, we compute the disk response to the satellite encounter using Equation (49). Results are summarized in Table 1 and Figures 6 and 7.

In Figure 6, we plot the impact parameter, $x_p \cos \theta_p$ (with respect to the Sun), as a function of the encounter velocity, $v_p$, of the satellites, for the penultimate (left panel), last (middle panel), and next (right panel) disk crossings. The red (gray) symbols denote the satellites that induce a strong (weak) amplitude of the bending-mode response, $f_{1,n=1}/f_0$, for $I_s = h_s \sigma_s = 9.2$ kpc $s^{-1}$. As shown in Appendix C, we consider $f_{1,n=1}/f_0 = \delta = 10^{-4}$ as a rough estimate for the minimum detectable relative response, i.e., the boundary between strong and weak responses to satellite passage. The solid black line indicates the boundary between bending and breathing modes, i.e., where the breathing-to-bending ratio, $f_{23}/f_{21}$ (Equation (54)), is equal to unity. Hence, the blue and red shaded regions indicate where the response is dominated by bending and breathing modes, respectively. The magenta dashed line roughly denotes the boundary between a strong bending response (blue shaded region) and a response that is adiabatically suppressed (gray shaded region). The latter is defined by the condition $\exp[-\Omega_z x_p \cos \theta_p/v_p] < \delta = 10^{-4}$.

In Figure 7, we plot the amplitude of the bending-mode response, $f_{1,n=1}/f_0$ (top panel), and the breathing-to-bending ratio, $f_{23}/f_{21} = f_{1,n=2}/f_{1,n=1}$ (bottom panel), in the solar neighborhood, as a function of the time $t_{\text{cross}}$ (in Gyr) when the satellite crosses the plane of the disk, assuming the fiducial MW parameters. Negative and positive $t_{\text{cross}}$ correspond to disk crossings in the past and future, respectively, and we once again consider stars with $I_s = h_s \sigma_s = 9.2$ kpc $s^{-1}$.

Both Figure 6 and the bottom panel of Figure 7 make it clear that all the disk crossings considered here preferentially excite bending rather than breathing modes in the solar neighborhood. As shown in Section 3, these trigger one-armed phase spirals in the solar neighborhood, in qualitative agreement with the MW snail observed in the Gaia data. However, as is evident from the top panel of Figure 7, most satellites only trigger a minuscule response in the disk, with $f_{1,n=1}/f_0 < \delta = 10^{-4}$, either because the satellite has too low mass or because the encounter, from the perspective of the Sun, is too slow such that the local response is adiabatically suppressed. The strongest response by far is triggered by encounters with Sgr, for which the bending-mode response, $f_{1,n=1}/f_0$, is at least $1 - 2$ orders of magnitude larger than that for any other satellite. Based on our orbit integration, it had its penultimate disk crossing, which closely

---

5. Our MW potential is similar to GALPY MPETENTIAL2014 (Bovy 2015) except for the power-law bulge, which has been replaced by an equivalent Hernquist bulge.

6. Dynamical friction might play an important role in the orbital evolution of massive satellites like the Large Magellanic Cloud (LMC) and Sgr, pushing their orbital radius further out in the past.
Table 1

MW Disk Response to Satellites for Stars with $L_*=h_z\sigma_z$ in the Solar Neighborhood

| MW Satellite Name | Mass ($M_\odot$) | $f_{1,n=1}/f_0$ Penultimate (3) | $\tau_{\text{enc}}$ (Gyr) Penultimate (4) | $f_{1,n=1}/f_0$ Last (5) | $\tau_{\text{enc}}$ (Gyr) Last (6) | $f_{1,n=1}/f_0$ Next (7) | $\tau_{\text{enc}}$ (Gyr) Next (8) |
|-------------------|-----------------|---------------------------------|------------------------------------------|--------------------------|---------------------------------|--------------------------|---------------------------------|
| Sagittarius       | $10^7$          | $4.3 \times 10^{-2}$           | $-0.92$                                   | $1.4 \times 10^{-10}$    | $-0.3$                          | $8.3 \times 10^{-4}$    | 0.03                            |
| Hercules          | $7.1 \times 10^6$| $\cdots$                       | $-3.57$                                   | $1.2 \times 10^{-4}$     | $-0.51$                         | $6.4 \times 10^{-5}$    | 3.16                            |
| Leo II            | $8.2 \times 10^6$| $\cdots$                       | $-3.61$                                   | $3.5 \times 10^{-5}$     | $-1.81$                         | $9.3 \times 10^{-5}$    | 2.34                            |
| Segue 2           | $5.5 \times 10^5$| $5 \times 10^{-5}$             | $-0.84$                                   | $3.4 \times 10^{-5}$     | $0.25$                          | $1.8 \times 10^{-6}$    | 0.27                            |
| LMC               | $1.4 \times 10^{11}$ | $1.4 \times 10^{-4}$ | $-6.97$                                   | $\cdots$                | $-2.37$                         | $7.2 \times 10^{-5}$    | 0.12                            |
| SMC               | $6.5 \times 10^9$  | $3.6 \times 10^{-8}$          | $-3.22$                                   | $\cdots$                | $-1.39$                         | $1.2 \times 10^{-9}$    | 0.22                            |
| Draco I           | $2.2 \times 10^7$  | $\cdots$                       | $-2.43$                                   | $5 \times 10^{-7}$       | $-1.23$                         | $1 \times 10^{-7}$     | 0.24                            |
| Bootes I          | $10^7$           | $\cdots$                       | $-1.65$                                   | $4.1 \times 10^{-7}$     | $-0.35$                         | $\cdots$                | 0.87                            |
| Willman I         | $4 \times 10^6$   | $\cdots$                       | $-0.63$                                   | $1.4 \times 10^{-7}$     | $-0.21$                         | $2.5 \times 10^{-8}$    | 0.4                             |
| Ursa Minor        | $2 \times 10^7$   | $\cdots$                       | $-2.26$                                   | $5.5 \times 10^{-8}$     | $-1.16$                         | $8.6 \times 10^{-9}$    | 0.29                            |
| Ursa Major II     | $4.9 \times 10^6$ | $4.5 \times 10^{-8}$          | $-2$                                      | $6.2 \times 10^{-10}$    | $-0.1$                          | $\cdots$                | 0.9                             |
| Coma Berenices I  | $1.2 \times 10^6$ | $7 \times 10^{-10}$           | $-2.47$                                   | $\cdots$                | $-0.25$                         | $\cdots$                | 0.69                            |
| Sculptor          | $3.1 \times 10^7$ | $\cdots$                       | $-2.7$                                    | $2 \times 10^{-10}$      | $-0.46$                         | $\cdots$                | 1.47                            |

Note. Column (1): the name of the MW satellite. Column (2): its dynamical mass estimate from literature (Simon & Geha 2007; Bekki & Stanimirović 2009; Lokas 2009; Erkal et al. 2019; Vasiliev & Belokurov 2020). We assume $10^6 M_\odot$ for the Sagittarius mass; note that there is a discrepancy between its measured mass of $\sim 4 \times 10^4 M_\odot$ (Vasiliev & Belokurov 2020) and the required mass of $10^7 – 10^{11} M_\odot$ for observable phase spiral signatures in N-body simulations (see, e.g., Bennett et al. 2021). Columns (3) and (4): the bending-mode response assuming fiducial MW parameters and the crossing time for the penultimate disk crossing, respectively. Columns (5) and (6): the same for the last disk crossing. Columns (7) and (8): the same for the next one. Only the satellites that trigger a bending-mode response, $f_{1,n=1}/f_0 \gtrsim 10^{-10}$, in at least one of the three cases are shown. The responses smaller than $10^{-10}$ are considered far too adiabatic and negligible and are marked by dashes. The case most relevant for the Gaia phase spiral is highlighted in red.

...
The MW disk is modeled as an isothermal slab, which lacks the axisymmetric density profile and velocity structure that characterize a realistic disk. In particular, whereas the lateral motion in our slab is uninhibited, the in-plane motion in a realistic disk consists of an azimuthal rotation combined with a radial epicyclic motion. Among others, this will have important implications for the global disk response and the rate at which phase spirals damp out owing to lateral mixing. In Paper II we apply our perturbative formalism to a realistic self-gravitating disk galaxy with a pseudo-isothermal DF (Binney 2010) and consider both external perturbations (encounters with satellites) and internal perturbations (bars and spiral arms).

All responses calculated in this paper only account for the direct response to a perturbing potential. In general, though, the response also has an indirect component that arises from the fact that neighboring regions in the disk interact with each other gravitationally. This self-gravity of the response, which we have ignored, triggers long-lived normal-mode oscillations of the slab that are not accounted for in our treatment. Several simulation-based studies have argued that including self-gravity is important for a realistic treatment of phase spirals (e.g., Darling & Widrow 2019a; Mathur 2019). Using the Kalnajs matrix method (Kalnajs 1977; Binney & Tremaine 2008), we have made some initial attempts to include the self-gravity of the response in our perturbative analysis, along the lines of Weinberg (1991). Our preliminary analysis shows that the self-gravitating response is a linear superposition of two terms: (i) a continuum of modes given in Equation (12), dressed by self-gravity, that undergo phase mixing and give rise to the phase spiral; and (ii) a discrete set of modes called point modes or normal modes (see Mathur 1990; Weinberg 1991) that follow a dispersion relation. The continuum response can be amplified by self-gravity when the continuum frequencies, $n\Omega_c + kv_c$, are close to the point-mode frequencies, $\nu_i$. Depending on the value of $k$, the normal modes can be either stable or unstable. Araki (1985) finds that in an isothermal slab the bending normal mode undergoes finite instability below a certain critical wavelength if $\sigma_i/\sigma \lesssim 0.3$, while the breathing normal mode becomes unstable above the Jeans scale. In the regime of stability, the normal modes are undamped oscillation modes in absence of lateral streaming (Mathur 1990) but are Landau damped otherwise (Weinberg 1991). For an isothermal slab with typical MW-like parameter values, the point modes are strongly damped since their damping timescale (inverse of the imaginary part of $\nu$) is of order their oscillation period (inverse of the real part of $\nu$), which turns out to be of order the vertical dynamical time, $\tau_h/\sigma$. Moreover, the normal-mode oscillations are coherent oscillations of the entire system, independent of the vertical actions of the stars, and are decoupled from the phase spiral in linear theory since the full response is a linear superposition of the two. Based on the above arguments, we conclude that self-gravity has little impact on the evolution of phase spirals in the isothermal slab, at least in the linear regime. We emphasize that Darling & Widrow (2019a), who found their phase spirals to be significantly affected by the inclusion of self-gravity, assumed a perturber-induced velocity impulse with magnitude comparable to the local velocity dispersion in the solar neighborhood; hence, their results are likely to have been impacted by nonlinear effects. Moreover, the self-gravitating response of an inhomogeneous disk embedded in a DM halo, as in the simulations of Darling & Widrow (2019a), can be substantially different from that of the isothermal slab. We intend to include a formal treatment of self-gravity along the lines of Weinberg (1991) in future work.

The disk of our MW is believed to be embedded in an extensive DM halo, something we have not taken into account. The presence of such a halo has several effects. First of all, the satellite perturbs not only the disk but also the halo. In particular, it induces both a local wake and a global modal response (e.g., Weinberg 1989; Tamfal et al. 2021). The former typically trails the satellite galaxy and boosts its effective mass by about a factor of two (Binney & Tremaine 2008), which might boost the (direct) disk response by about the same factor. The global halo response is typically dominated by a strong $l = 1$ dipolar mode followed by an $l = 2$ quadrupolar mode (Tamfal et al. 2021), which might have a significant impact on the disk. The presence of a halo also modifies the total potential. At large disk radii and vertical heights, the halo dominates the potential and will therefore significantly modify the actions and frequencies of the stars, and consequently the shape of the phase spirals. Finally, since
the disk experiences the gravitational force of the halo, a (sufficiently massive) satellite galaxy can excite normal-mode oscillations of the disk in the halo (see, e.g., Hunt et al. 2021). We intend to incorporate some of these effects of the MW halo in Paper II.

7. Conclusion

In this paper we have used linear perturbation theory to compute the response of an infinite, isothermal slab to various kinds of external perturbations with diverse spatiotemporal characteristics. Although a poor description of a realistic disk galaxy, the infinite, isothermal slab model captures the essential physics of perturbative response and collisionless equilibration via phase mixing in the disk and thus serves as a simple yet insightful case for investigation.

We use a hybrid (action-angle variables in the vertical direction and position-momentum variables in the lateral direction) linear perturbative formalism to perturb and linearize the CBE and compute the response in the DF of the disk to a gravitational perturbation. We have considered external perturbations of increasing complexity, ranging from an instantaneous (laterally) plane-wave perturbation (Section 3); to an instantaneous localized perturbation, represented as a wave packet (Section 4); to a nonimpulsive, temporally extended, localized perturbation (Section 5); and ultimately to an encounter with a satellite galaxy moving along a straight-line orbit (Section 6). This multilayered approach is ideal for developing the necessary insight into the complicated response that is expected from a realistic disk galaxy exposed to a realistic perturbation. We summarize our conclusions below.

1. The two primary Fourier modes of slab oscillation are the $n=1$ bending mode and the $n=2$ breathing mode, which correspond to antisymmetric and symmetric oscillations about the midplane, respectively. For a sufficiently impulsive perturbation, the dominant mode is the breathing mode, which initially causes a quadrupolar distortion in the $(z, vz)$-phase space, which evolves into a two-armed phase spiral as the stars with different vertical actions oscillate with different vertical frequencies. If the perturbation is temporally more extended (less impulsive), the dominant mode is the bending mode. This causes a dipolar distortion in the $(z, vz)$-phase space that evolves into a one-armed phase spiral (see also Widrow et al. 2014; Hunt et al. 2021). Due to vertical phase mixing, the phase spiral wraps up tighter and tighter until it becomes indistinguishable from an equilibrium distribution in the coarse-grained sense.

2. Besides vertical phase mixing, the survivability of the phase spiral is also dictated by the lateral streaming motion of stars. The initial lateral velocity impulse toward the minima of $\Phi_P$ tends to linearly boost the contrast of the phase spiral. This is, however, quickly taken over by lateral streaming (with velocity dispersion $\sigma$), which causes mixing between the over- and underdensities and damps out the phase spiral amplitude. For an impulsive, laterally sinusoidal perturbation, the disk response is also sinusoidal and damps out...
like a Gaussian (due to the Maxwellian/Gaussian distribution of the unconstrained lateral velocities) over a timescale of $\tau_D \sim 1/\kappa \sigma$, i.e., small-scale perturbations damp out faster, as expected.

3. Lateral mixing operates differently for a spatially localized perturbation that can be expressed as a superposition of many plane waves. The response to each of them damps out like a Gaussian (if the perturber is impulsive). Since the power spectrum of a spatially localized perturber with a lateral Gaussian profile is dominated by its largest scales (small $k$) that mix and damp out slower, the net response from all $k$ damps away as $\sim t^{-1}$ (the response profile spreads out as $\sim t$), much slower than the Gaussian damping in case of a sinusoidal perturber.

4. The disk response to a nonimpulsive perturbation is substantially different from that to an impulsive one. If the temporal strength of the perturber follows a Gaussian pulse with pulse frequency, $\omega_0$ (e.g., a transient bar or spiral arm), the response grows and decays following the temporal profile of the pulse before eventually attaining a $\sim 1/t$ power-law fall-off. The response peaks when the pulse frequency, $\omega_0$, is comparable to the vertical oscillation frequency, $\Omega_z$. The response to more impulsive perturbations ($\omega_0 \gg \Omega_z$) is suppressed as $\sim 1/\omega_0$, whereas much slower ($\omega_0 \ll \Omega_z$) perturbations trigger a superexponentially ($\sim \exp[-n^2 \Omega_z^2/4 \omega_0^2]$ at small $k$) suppressed response. In this adiabatic limit, the stars tend to remain in phase with the perturber, oscillating at frequencies much smaller than $\Omega_z$, which inhibits the formation of a phase spiral.

5. The timescale of perturbation dictates the excitability of different modes, with slower (faster) pulses triggering stronger bending (breathing) modes. An encounter with a satellite galaxy that hits the disk with a uniform velocity $v_P$ and an angle $\theta_P$ with respect to the normal at a distance $x_P$ away from an observer in the disk perturbs the potential at an observer’s location with a characteristic timescale $\tau_{\text{enc}} \sim x_P \cos \theta_P/v_P$. If $\tau_{\text{enc}}$ is long (short) compared to the typical vertical oscillation time, $\tau_\zeta \sim h_z/\sigma_z$, at the observer’s location, the dominant perturbation mode experienced is a bending (breathing) mode. Thus, bending modes are preferentially excited not only by low-velocity encounters but also by more distant and more perpendicular ones. Since the velocities of all MW satellites are much larger than $\sigma_z$, the decisive factor for bending versus breathing modes is the distance from the point of impact. This is in qualitative agreement with the results from $N$-body simulations of the MW–Sgr encounter performed by Hunt et al. (2021), which show more pronounced bending (breathing) modes farther from (closer to) the location where Sgr impacts the disk. Moreover, for a given encounter, stars with larger actions undergo stronger breathing-mode oscillations since they oscillate slower.

6. Besides phase spirals, satellite encounters also induce spatial corrugations in the disk response, with vertical and lateral wavenumbers given by $k_z = n \ell z \cos \theta_P/v_P$ and $k_x = n \ell z \sin \theta_P/v_P$, respectively.

As an astrophysical application of our formalism, we have investigated the direct response of the MW disk (approximated as an isothermal slab) to several of the satellite galaxies in the halo for which dynamical mass estimates and galactocentric phase-space coordinates from Gaia parallax and proper-motion measurements are available. We integrate the orbits of these satellites in the MW potential and note the impact velocity $v_P$, angle of impact $\theta_P$ with respect to the normal, and the impact distance from the solar neighborhood $x_P$ during their penultimate, last, and next disk crossings. We use these parameters to compute the direct response to the MW satellites and find that all of them excite bending modes and thus one-armed phase spirals in the solar neighborhood, similar to that discovered in the Gaia data by Antonio et al. (2018). In the solar vicinity, the largest direct response, by far, is due to the encounter with Sgr. The direct responses triggered by other satellites, most notably Hercules and the LMC, are at least 1–2 orders of magnitude smaller. Hence, we conclude that if the Gaia phase spiral was triggered by an encounter with an MW satellite, the strongest contender is Sgr. Although Sgr has been considered as the agent responsible for the Gaia phase spiral and other local asymmetries and corrugations, several studies have pointed out that it cannot be the sole cause of all these perturbations (see, e.g., Bennett & Bovy 2021; Bennett et al. 2021). Our work argues, though, that the direct response in the solar neighborhood from the other MW satellites, including the LMC, is not significant enough, at least in the range of actions covered by the Gaia survey. Of course, as discussed in Section 6.2, the indirect response from the DM halo of the MW might play an important role especially for the more massive satellites such as Sgr and the LMC. Moreover, the global response of a realistic disk will be different from that of the isothermal slab model considered here. We investigate the realistic disk response in Paper II and leave a sophisticated analysis incorporating self-gravity and halo response for future work. It remains to be seen whether a combination of Sgr plus other (internal) perturbations due to, for example, spiral arms (Faure et al. 2014) or the (buckling) bar (e.g., Khoperskov et al. 2019) can explain the fine structure in the solar neighborhood, or whether perhaps a solution requires modifying the detailed MW potential. It is imperative, though, to investigate the structure of phase spirals at other locations in the MW disk, in particular whether they are one- or two-armed. This would help to constrain both the timescale and location of the perturbation responsible for the various out-of-equilibrium features uncovered in the disk of our MW.

The authors are grateful to the anonymous referee for thoughtful comments and to Kathryn Johnston, Jason Hunt, Adrian Price-Whelan, Kaustav Mitra, Elena D’Onghia, Chris Hamilton, and Dhruba Dutta-Chowdhury for insightful discussions and valuable suggestions. M.D.W. is supported by the National Science Foundation through grant No. AST-1812689. F.C.v.d.B. is supported by the National Aeronautics and Space Administration through grant No. 19-ATP19-0059 issued as part of the Astrophysics Theory Program.

Appendix A

Adiabatic Limit of Slab Response

In the adiabatic/slow limit, the slab response can be computed by taking the $\omega_0 \to 0$ limit and performing the $\tau$
integral in Equation (18) to obtain
\[
\begin{align*}
    f_{nk} = -i \pi \Phi_N Z_n(L, \mathcal{X}) n \Omega_{\zeta} \sum_{n=-\infty}^{\infty} Z_n(L) \exp \left[ \frac{n^2 \Omega_{\zeta}^2}{2k^2 \sigma^2} \right] \\
    \times n \Omega_{\zeta} \left( \frac{1}{\sigma_{\zeta}^2} - \frac{1}{\sigma^2} \right) \exp [imw_i].
\end{align*}
\] (A2)

Substituting the Gaussian form for $\mathcal{X}_K$ given in Equation (31) in the above expression, multiplying it by $\exp [ikx]$, and integrating over all $k$, we obtain the following final expression for the slab response in the slow limit:
\[
\begin{align*}
    f_l(I, w_z, x) = -i \pi \Phi_N \mathcal{X}_{K} \sum_{k=-\infty}^{\infty} Z_n(I, \mathcal{X}) n \Omega_{\zeta} \left( \frac{1}{\sigma_{\zeta}^2} - \frac{1}{\sigma^2} \right) \exp [imw_i].
\end{align*}
\] (A3)

where
\[
J_k(x) = \int_{-\infty}^{\infty} dk \frac{\exp [ikx]}{|k|} \times \exp [-k^2 \Delta_x^2 / 2] \exp \left[ -n^2 \Omega_{\zeta}^2 / 2k^2 \sigma^2 \right].
\] (A4)

The above integral can be approximately evaluated in the small and large $x$ limits by the saddle point method to obtain the following asymptotic behavior of $J_k(x)$:
\[
\begin{align*}
    J_k(x) &\sim \left\{ \begin{array}{ll}
        \left[ \frac{\pi \sigma}{2n} \right] n \Omega_{\zeta} \Delta_x \exp [-n \Omega_{\zeta} \Delta_x / \sigma] & \text{small } x,
        \\
        \left( \frac{\pi}{2} \right)^{1/2} \Delta_x \exp [-x^2 / 2 \Delta_x^2] & \text{large } x.
    \end{array} \right.
\end{align*}
\] (A5)

Appendix B

Slab Response to Satellite Encounters

The perturbing potential, $\Phi_p$, at $(x, z)$ due to a satellite galaxy impacting the disk along a straight orbit with uniform velocity $v_p$ at an angle $\theta_p$ with respect to the normal is given by Equation (23). Computing the Fourier transform, $\Phi_{nk}$, of $\Phi_p$ and substituting this in Equation (18) yields
\[
\begin{align*}
    f_{nk}(L, \mathcal{X}) = \int \frac{GM_p}{v_p} f_0(v_x, v_y, E_z) \left( \frac{n \Omega_{\zeta}}{\sigma_{\zeta}^2} + k v_z \right) \\
    \times \exp [-i(n \Omega_{\zeta} + kv_z)t] \mathcal{F}_{nk}(t).
\end{align*}
\] (B1)

where
\[
\begin{align*}
    \mathcal{F}_{nk}(t) &= \frac{1}{2\pi} \int_0^{2\pi} dv_z \exp [-imw_i] \int_{-\infty}^{\infty} dx' \exp [-ikx'] \\
    &\times \int_{-\infty}^{t} dt \exp [i(n \Omega_{\zeta} + kv_z)\tau] \\
    &\times \sqrt{\left( \tau - \frac{v'}{v_p} \cos \theta_p + x' \sin \theta_p \right)^2 + \left( \tau - \frac{v'}{v_p} \cos \theta_p - z' \sin \theta_p \right)^2}.
\end{align*}
\] (B2)

The $\tau$ integral can be computed in the large $t$ limit to yield
\[
\begin{align*}
    \mathcal{F}_{nk}(t \to \infty) &= \frac{1}{2\pi^2} \int_0^{2\pi} dv_z \exp [-imw_i] \int_{-\infty}^{\infty} dx' \exp [-ikx'] \\
    &\times \exp \left[ i(n \Omega_{\zeta} + kv_z) \cos \theta_p z' \right] \\
    &\times \exp \left[ i(n \Omega_{\zeta} + kv_z) \sin \theta_p x' \right] \\
    &\times \frac{K_0(n \Omega_{\zeta} + kv_z)}{v_p} \left( \frac{x' \cos \theta_p - z' \sin \theta_p}{v_p} \right).
\end{align*}
\] (B3)

where $K_0$ denotes the zeroth-order modified Bessel function of the second kind. Recalling that the unperturbed DF is isothermal, given by Equation (13), we integrate Equation (B1) over $v_x$ and $v_y$ to obtain
\[
\begin{align*}
    \int_{-\infty}^{\infty} dv_x \int_{-\infty}^{\infty} dv_y f_{nk}(L, v_x, v_y, t) \\
    \approx \frac{\rho_i}{\sqrt{2\pi} \sigma_{\zeta}} \exp \left[ -E_z / \sigma_{\zeta}^2 \right] \frac{GM_p}{v_p} \\
    &\times \int_{-\infty}^{2\pi} dv_z \exp [-imw_i] \exp \left[ i(n \Omega_{\zeta} \cos \theta_p z') \right] \\
    &\times \int_{-\infty}^{\infty} dx' \exp [-ikx'] \exp \left[ i(n \Omega_{\zeta} \sin \theta_p x') \right] \\
    &\times \exp \left[ -\frac{1}{2} k^2 \sigma_{\zeta}^2 \left( t - S/v_p \right)^2 \right] \\
    &\times \frac{K_0(n \Omega_{\zeta} - ik^2 \sigma_{\zeta}^2 (t - S/v_p))}{v_p} \left( \frac{x' \cos \theta_p - z' \sin \theta_p}{v_p} \right).
\end{align*}
\] (B4)

where we have defined
\[
S = z' \cos \theta_p + x' \sin \theta_p.
\] (B5)

Multiplying Equation (B4) by $\exp [ikx]$ and integrating over $k$ yields
\[
\begin{align*}
    \int_{-\infty}^{\infty} dk \exp [ikx] \int_{-\infty}^{\infty} dv_x \int_{-\infty}^{\infty} dv_y f_{nk}(L, v_x, v_y, t) \\
    \approx \frac{\rho_i}{\sqrt{2\pi} \sigma_{\zeta}} \frac{GM_p}{v_p} \\
    &\times \int_{-\infty}^{2\pi} dv_z \exp [-imw_i] \exp \left[ i(n \Omega_{\zeta} \cos \theta_p z') \right] \\
    &\times \sqrt{2\pi} \int_{-\infty}^{\infty} d\Delta_x \frac{1}{2} \exp \left[ -\frac{1}{2} \left( \frac{\Delta x}{\sigma_{\zeta}} \right)^2 \right] \\
    &\times \left[ \frac{1}{\sigma_{\zeta}^2 \theta_p} \left( 1 + \frac{\Delta x^2}{\sigma_{\zeta}^2 \theta_p^2} \right) + \frac{n \Omega_{\zeta}}{\sigma_{\zeta}^2} \right] \\
    &\times \exp \left[ i\frac{n \Omega_{\zeta} \sin \theta_p x'}{v_p} \right] \\
    &\times \frac{K_0(n \Omega_{\zeta} + i(\Delta x)^2 / \sigma_{\zeta}^2 \theta_p^2)}{v_p} \left( \frac{x' \cos \theta_p - z' \sin \theta_p}{v_p} \right).
\end{align*}
\] (B6)
where $\Delta x = x - x'$ and $t' = t - S/v_p$. In the large time limit, using the identity that $\lim_{\sigma \to \infty} \exp\left[-(\Delta x)^2/2\sigma^2\right]/\sigma^3 = \sqrt{2\pi} \delta(\Delta x)$, the integration over $\Delta x$ is simplified. Upon performing this integral, multiplying the result by $\exp\left[i m w_z\right]$, and summing over all $n$, we obtain the following response:

$$f(t, w_z, x, t) \approx \frac{\rho_0}{\sqrt{2\pi} \sigma_z} \exp\left[-E_z/\sigma_z^2\right] \times \frac{2GM_p}{v_p}$$

$$\times \sum_{n=-\infty}^{\infty} \left[ \frac{n\Omega_z}{\sigma_z^2} \Psi_n(x, I_z) \exp\left[i n\Omega_z \sin \theta_p z/v_p\right] \right] \times \exp\left[i m(w_z - \Omega_z t)\right],$$

(B7)

where

$$\Psi_n(x, I_z) = \frac{1}{2\pi} \int_0^{2\pi} \, dw_z \exp\left[-in\Omega_z \cos \theta_p z/v_p\right]$$

$$\times K_0\left[\left| \frac{n\Omega_z \cos \theta_p z/v_p}{v_p} \right| \right].$$

(B8)

The above expression for $\Psi_n$, can be simplified by evaluating the $w_z$ integral under the epicyclic approximation (small $I_z$ limit), to yield the following approximate form:

$$\Psi_n(x, I_z) \approx K_0\left[\left| \frac{n\Omega_z \cos \theta_p z/v_p}{v_p} \right| \right] \Phi_n^{(0)}(I_z) - i \frac{n\Omega_z \sin \theta_p}{v_p}$$

$$\times K_0\left[\left| \frac{n\Omega_z \cos \theta_p z/v_p}{v_p} \right| \right] \Phi_n^{(1)}(I_z)$$

$$- \frac{1}{2} \left( \frac{n\Omega_z \sin \theta_p}{v_p} \right)^2 K_0\left[\left| \frac{n\Omega_z \cos \theta_p z/v_p}{v_p} \right| \right] \Phi_n^{(2)}(I_z) + ...$$

(B9)

Here each prime denotes a derivative with respect to the argument of the function. $\Phi_n^{(j)}(I_z)$, for $j = 0, 1, 2, \ldots$, is given by

$$\Phi_n^{(j)}(I_z) = \frac{1}{2\pi} \int_0^{2\pi} \, dw_z \, z^j \exp\left[-im\Omega_z \cos \theta_p z/v_p\right]$$

$$\approx \frac{2l_z}{\nu}^{j/2} J_{n+j}(2l_z/v_p) \left( \frac{n\Omega_z \cos \theta_p z/v_p}{v_p} \right)^{j/2}.$$  

(B10)

Here the implicit relation between $z$, $w_z$, and $I_z$ given in Equation (9), which yields $z = \sqrt{2} l_z/v \nu$ for small $I_z$, has been used. $J_n$ denotes the $n$th derivative of the $n$th-order Bessel function of the first kind, and $\nu = \sqrt{2} \sigma_z/h_0$ is the vertical epicyclic frequency. In Equation (B7), well after the encounter (large $t$), the term, $1/\sigma_z^2$, can be neglected relative to $in\Omega_z/v_p$ for $n = 0$, thus yielding the expression for the disk response to satellite encounters given in Equation (49).

**Appendix C**

**Detectability Criterion for the Phase Spiral**

The demarcation between strong and weak amplitudes of a phase spiral is dictated by the minimum detectable relative response, $\delta$, which can be determined in the following way. Let there be a phase spiral that we want to detect with a total number, $N_a$, of stars by binning the phase-space distribution in the $\sqrt{2} l_z \cos w_z - \sqrt{2} l_z \sin w_z$ plane. Let us define the unperturbed DF, $f_0$, and the normalized unperturbed DF, $\tilde{f}_0$, such that

$$N_a = \iint f_0 \, dl_z \, dw_z, \quad \tilde{f}_0 = \frac{f_0}{N_a}.$$  

(C1)

The perturber introduces a perturbation in the (normalized) DF, $\tilde{f}_p$, which manifests as a spiral feature in the phase space due to phase mixing. To recover $\tilde{f}_p$, we bin the data in $I_z$ and $w_z$, such that the perturbation in the number of stars in each bin $(\Delta I_z, \Delta w_z)$ is given by

$$N(\Delta I_z, \Delta w_z) = N_a \tilde{f}_p \Delta I_z \Delta w_z.$$  

(C2)

The optimum binning strategy can be determined as follows. The phase spiral is a periodic feature in both $I_z$ and $w_z$. Therefore, to pull out the periodicity in $I_z$, we need to sample with a frequency exceeding the Nyquist frequency, i.e., the bin size, $\Delta I_z$, should be less than $I_{z,\text{max}}/N_{\text{wind}}$, where $I_{z,\text{max}}$ is the maximum $I_z$ in the sample and $N_{\text{wind}}$ is the number of winds of the spiral. Moreover, $\Delta I_z$ is required to exceed the Gaia measurement error so that the error is dominated by Poisson noise, i.e., we require $\Delta I_z/I_z > \Delta G_{\text{Gaia}} \sim 10^{-2}$ (see Luri et al. 2018; Katz et al. 2019, for parallax and radial velocity errors, the two dominant sources of measurement errors in Gaia). Within each $I_z$ bin, the data are further divided into $N_a$ azimuthal bins, each of size $\Delta w_z = 2\pi/N_a$.

For optimum sampling in $w_z$, $N_a$ should be greater than $2n$ (for spiral mode $n$) and less than $2\pi/\Delta G_{\text{Gaia}}$. After binning the data as discussed above, a reliable detection of the phase spiral can be made with a given signal-to-noise ratio ($S/N$) when the perturbation in the number of stars in each bin

$$N(\Delta I_z, \Delta w_z) = N_a \times \tilde{f}_p \Delta I_z \Delta w_z \geq (S/N)^2.$$  

(C3)

Here we have assumed that the error in recovering the spiral feature is dominated by Poisson noise. This yields the following estimate for the minimum detectable relative response for an isothermal slab:

$$\frac{\tilde{f}_p}{f_0} \geq \delta = 3.6 \times 10^{-4} \times \left( \frac{S/N}{3} \right)^2 \left( \frac{10^6}{N_a} \right) \left( \frac{N_a}{10} \right)$$

$$\times \left( \frac{0.1}{\Delta I_z/I_z} \right) h_0 \frac{\sigma_z}{I_z} \exp\left( \frac{E_z(I_z)}{\sigma_z^2} \right).$$  

(C4)

Provided that there are about a million stars in the Gaia data of the solar neighborhood (Antoja et al. 2018), we consider $\delta = 10^{-4}$ to be a rough estimate for the minimum detectable relative response.

**ORCID iDs**

Uddipan Banik @ https://orcid.org/0000-0002-9059-381X
Martin D. Weinberg @ https://orcid.org/0000-0003-2660-2889
Frank C. van den Bosch @ https://orcid.org/0000-0003-3236-2068

**References**

Aguilar, L. A., & White, S. D. M. 1985, ApJ, 295, 374
Antoja, T., Helmi, A., Romero-Gómez, M., et al. 2018, Natur, 561, 360
Araki, S. 1985, PhD thesis, Massachusetts Institute of Technology
Banik, U., & van den Bosch, F. C. 2021a, ApJ, 912, 43
