On dust–gas gravitational instabilities in protoplanetary discs

Henrik N. Latter* and Roxana Rosca

DAMTP, University of Cambridge, CMS, Wilberforce Road, Cambridge CB3 0WA, UK

Accepted 2016 September 26. Received 2016 September 23; in original form 2016 August 1

ABSTRACT

In protoplanetary discs the aerodynamical friction between particles and gas induces a variety of instabilities that facilitate planet formation. Of these we examine the so-called ‘secular gravitational instability’ (SGI) in the two-fluid approximation, deriving analytical expressions for its stability criteria and growth rates. Concurrently, we present a physical explanation of the instability that shows how it manifests upon an intermediate range of lengthscales exhibiting geostrophic balance in the gas component. The two-fluid SGI is completely quenched within a critical disc radius, as large as 10 au and 30 au for centimetre- and millimetre-sized particles, respectively, although establishing robust estimates is hampered by uncertainties in the parameters (especially the strength of turbulence) and deficiencies in the razor-thin disc model we employ. It is unlikely, however, that the SGI is relevant for well-coupled dust. We conclude by applying these results to the question of planetesimal formation and the provenance of large-scale dust rings.

Key words: instabilities – planets and satellites: formation – protoplanetary discs.

1 INTRODUCTION

The assembly of planets is a complex and multi-faceted phenomenon that spans a gulf of some 12 orders of magnitude in length: from micron-sized dust to 10^3 km planetary cores. It draws on an equally wide range of physical processes: collisions, dust–gas aerodynamics, gravitational collapse, instabilities, and disc structures (e.g. vortices and dust traps), to name but a few (Papaloizou & Terquem 2006; Armitage 2010; Chiang & Youdin 2010). While it is relatively straightforward to grow centimetre-sized particles from micron sizes, further growth is potentially halted by a number of ‘barriers’ (bouncing, fragmentation, and radial drift; Johansen et al. 2014). Statistically a small number of ‘lucky’ aggregates may hurdle these, but certain collective instabilities promise a more reliable route through this difficult size range. These include classical gravitational instability (GI; Safronov 1969; Goldreich & Ward 1973), streaming instability (Youdin & Goodman 2005), and the secular gravitational instability (SGI; Ward 2000; Youdin 2005). It is to the last instability that this paper is devoted.

One of the most attractive features of the single-fluid SGI is that its onset is unconditional; it should always be present. Unlike classical GI, which requires the Toomre parameter to be less than one, a single-fluid analysis presents no analogous restriction: the SGI works no matter how thin or thick the particle subdisc is (Youdin 2005). The instability attacks longer scales preferentially, which ordinarily would be stabilized by the Coriolis force; but particles can shed (or gain) angular momentum via aerodynamical drag, and hence are not obliged to undergo stabilizing epicycles. As a consequence, rings that are radially drifting towards each other continue to do so unimpeded, and the instability can proceed. On small radial scales, the SGI is suppressed by dust pressure or gas turbulence, and in fact, for well-coupled dust, turbulence decreases growth rates to potentially insignificant levels (Shariff & Cuzzi 2011; Youdin 2011). Marginally coupled particles, however, could still be subjected to respectable SGI growth rates at certain radii.

The SGI has been thoroughly explored in single-fluid models, which are applicable when the dust-to-gas density ratio is tiny (e.g. Ward 2000; Youdin 2005; Shariff & Cuzzi 2011; Youdin 2011; Michikoshi, Kokubo & Inutsuka 2012). These models assume that the angular momentum bestowed on to, or removed from, the gas disc is negligible. On sufficiently long scales, however, both sides of this momentum transaction must be included and the gas dynamics explicitly calculated. An instability criterion then appears; in a two-fluid model, the onset of SGI is no longer unconditional. Recently, Takahashi & Inutsuka (2014, hereafter TI) made a start on this problem (see also Shadmehri 2016), but there is still much to be established. Putting aside the issue of growth rates, an especially important question is: at what radii and for what particle sizes should we expect SGI to exist at all?

The first aim of this paper is to derive clean stability criteria for the SGI. In the limits of strongly coupled and weakly coupled particles these can be formulated analytically and involve a variety of parameters, including the gas’s Toomre parameter and the dust-to-gas density ratio. Because they bypass the SGI’s full sixth-order dispersion relation, these criteria make it relatively easy to assess its prevalence.

The criteria also motivate a straightforward physical picture of instability in a two-fluid system. In order for the instability to work, there must exist an intermediate range of lengthscales gas upon...
which (i) dust pressure or turbulent mass diffusion is subdominant, and (ii) the gas is prevented from executing epicycles, despite its angular momentum transactions with the dust. Going to lengthscales longer than the dust pressure (or diffusion) scale takes care of the first restriction. But the second can only be satisfied if geostrophic balance holds in the gas fluid, and so we must simultaneously find shortish scales upon which gas pressure is dominant. The existence or not of this intermediate range furnishes us with the stability criterion.

The formalism is applied to realistic disc models, where we find that it is unlikely that well-coupled dust is unstable to the SGI at any radius, unless the background turbulence is especially weak. Marginally coupled particles, however, can achieve appreciable growth rates in certain circumstances, emphasizing that the SGI could help aggregation of solids of centimetre size. We conclude, however, that SGI is probably unrelated to the dust rings recently observed by ALMA (Brogan et al. 2015).

The paper will be organized in the following way. First, in Section 2, we present the two-fluid razor-thin-disc model that we employ, alongside a critical discussion of its shortcomings. The main parameters of the analysis will also be defined. In Section 3, we revisit the single-fluid model to fix some ideas and to provide context for the subsequent analysis, while in Section 4 we briefly treat a simple two-fluid system where the gas is regarded as incompressible. The main results of the paper are given in Section 5, in which we derive analytic stability criteria in relevant limits that are then applied to realistic disc models in Section 6. We draw our conclusions in Section 7, where we discuss the relevance of the SGI in planet and structure formation in protostellar discs.

2 PRELIMINARIES

2.1 Modelling issues

The classical GI and secular GI have primarily been explored with 1D models of a vertically averaged or razor-thin disc. Recent notable exceptions are Mamatsashvili & Rice (2010) and Lin (2014), who also capture vertical convection and the magnetorotational instability, respectively. A 1D model certainly eases the analysis and it should be a reasonable approximation for unstable modes whose radial wavelength is much greater than the disc thickness; because the classical GI has minimal vertical structure (being essentially an f mode in this limit; Ogilvie 1998), it is also likely that the SGI depends on z only weakly. For wavelengths closer to the scale height, a somewhat ad hoc correction may be included (e.g. Shu 1984), which generally works against instability on these shorter lengthscales.

In a two-fluid model, however, the razor-thin assumption is complicated by the fact that the particle fluid and the gas fluid possess different thicknesses, and the former can be significantly shorter than the latter. This is a problem for the SGI because the fastest lengthscales are not far from the particle scale height (Youdin 2005), and hence of the order or less than the gas scale height. As a consequence, the approximation of a razor-thin disc is not strictly applicable, at least in the description of the gas. In a single-fluid model this issue does not crop up because it is assumed that the gas fluid is unperturbed by whatever the dust is doing; but in a two-fluid model this is not the case. It may be that the gas perturbations associated with the SGI are sufficiently small that the disc’s vertical structure plays little role. But only calculations in vertically stratified shearing boxes can decide on this point.

A second issue is the correct coupling between the two fluids. In a real system, with different gas and dust scale heights, the drag acceleration will be a function of vertical height z. Moreover, the entire column of gas will not exert drag on the dust if the dust subdisc is much thinner. For consistency, the gas external to the dust disc should be excluded from a two-fluid razor-thin treatment, with the weighting of the drag force in the momentum equation adjusted appropriately to account for the smaller surface density of the gas subdisc. Because most of the mass in both discs is near the midplane, this problem may not invalidate the main qualitative results. It should, however, be kept in mind.

A third issue concerns turbulence in the gas, its effects, and how to mathematically describe it. The gas is likely undergoing disordered motions at some (perhaps all) radii, though the underlying physics may differ between different regions (Turner et al. 2014). The disc may also support turbulence because of the settling of the dust to the midplane, and ensuing vertical Kelvin–Helmholtz instability (Cuzzi, Dobrovolskis & Champney 1993). The consequences of turbulence on particles are several. Agitation of the solids produces enhanced velocity dispersions, over and above that arising from particle collisions (Goldreich & Tremaine 1978), and thus an appreciable particle pressure that will oppose vertical settling. The precise ‘equation of state’ this pressure obeys, however, is difficult to establish. In addition, the random motions induced in the dust can potentially smooth away inhomogeneities in the dust density, and thus lead to diffusion in the continuity equation directly (Shariff & Cuzzi 2011; Youdin 2011; Michikoshi et al. 2012). The efficiency of this diffusion is pretty much unconstrained and obviously requires numerical exploration; for instance, the vertical Kelvin–Helmholtz instability will mix particles effectively in the vertical direction but not necessarily in the horizontal direction. Finally, turbulence will transport momentum, though this effect (leading to a straightforward mild damping) is not included in this paper (see TI for its treatment).

The effects of turbulence on the dust have been mathematically modelled via mean-field theories, and Langevin and Fokker–Planck treatment, with the weighting of the drag force in the momentum equation adjusted appropriately to account for the smaller surface density of the gas subdisc. Because most of the mass in both discs is near the midplane, this problem may not invalidate the main qualitative results. It should, however, be kept in mind.

2.2 Parameters

As will become clear by Section 5, the two-fluid secular GI is governed by a large number of parameters; it is hence convenient
to define them all in one place. The first key parameter is the dust
to-gas mass ratio, denoted by $\delta$ and defined via

$$\delta \equiv \frac{\sigma_a}{\sigma_g},$$

(1)

where $\sigma_a$ is the background surface density of the dust with size $a$, and $\sigma_g$ is the surface density of the gas located within the dust subdisc (see earlier). Thus $\delta$ is a function of $a$. Because the SGI is size selective, it is necessary to distinguish between particles of different sizes, and thus to use separate surface densities for each subspecies. To ease the analysis in this paper, we examine each dust species separately, though in reality different species will weakly couple via the gas phase; see Shadmehri (2016) for an attack on a system of interacting gas and two species of particle.

Generally, $\sigma_a$ is a small fraction of the total solid surface density. In fact, the surface density of millimetre- and centimetre-sized particles can be significantly less than 1 per cent of the total (e.g. Brauer, Dullemond & Henning 2008; Windmark et al. 2012a,b), though this figure varies greatly with age and as different physical processes are included or neglected (sticking, bouncing, fragmentation, and mass transfer, for instance; Garaud et al. 2013). Note that $\sigma_a$ will also be less than the total gas density, as it only includes gas situated amongst the dust subdisc. This decrease can be quantified by a factor $\sim H_g/H$, where $H_g$ and $H$ are the scale thicknesses of the gas and dust, respectively. We may then write

$$\delta \approx \frac{\sigma_a}{\sigma_g} \approx \frac{H_g}{H},$$

(2)

where $\sigma_{tot}$ is the total dust surface density (including all species), and we have assumed that the ratio of total dust to total gas density takes the standard value 0.01 (Chiang & Youdin 2010). Given the large uncertainties in the second and third factors in equation (2) (small and large, respectively), we simply set $\delta = 0.01$ for most calculations.

The second key parameter is the inverse Stokes number, which we denote by $\epsilon$ and define through

$$\epsilon = \frac{1}{\tau_s \Omega},$$

(3)

where $\tau_s$ is the stopping time of the particles, and $\Omega$ is the local orbital frequency. Large values of $\epsilon$ correspond, thus, to strongly coupled particles, and smaller values to weakly coupled particles.

The exact value of $\epsilon$ depends not only on particle size but also on radial location (e.g. Chiang & Youdin 2010). Fig. 1 gives $\epsilon$ profiles for two different minimum mass solar nebula (MMSN) models (see Section 6 for more details).

The third parameter is the ratio of dust and gas velocity dispersions, defined through

$$\eta = \frac{\epsilon_d}{\epsilon_g},$$

(4)

where $\epsilon$ and $\epsilon_g$ denote the velocity dispersions of the dust and gas, respectively. To a first approximation, the dust random velocities are controlled by ‘kicks’ delivered by the gas turbulence, rather than inter-particle collisions (Youdin & Lithwick 2007). Thus, it is possible to relate $\eta$ to properties of the turbulence, and we do so below. Collisonal agitation becomes important for larger particles, $\gtrsim cm$, and thus the $\eta$ we use in this paper may be an underestimate for these sizes.

The fourth parameter is the Toomre $Q$ of the dust, which describes the onset of classical GI. It is given by

$$Q = \frac{c \Omega}{\pi G \sigma},$$

(5)

where $G$ is the gravitational constant and for notational convenience we have dropped the subscript ‘$a$’ from the surface density. An analogous expression can be defined for the gas, which we denote by $Q_g$. The two Toomre parameters can be related via the following identity: $Q_g = (\delta/\sqrt{\pi})Q$.

Finally, the turbulent diffusion of solid particles can be quantified via the mass diffusivity $D$, and the diffusion of gas by the analogous $D_g$. As in Youdin (2011), we replace these by the dimensionless parameters $\alpha$ and $\alpha_g$, via $D = \alpha c_d^2/\Omega$ and $D_g = \alpha_g c_d^2/\Omega$, with these related via

$$\alpha = \frac{\epsilon^2 + \epsilon + 4}{(\epsilon + \epsilon^{-1})^2} \alpha_g,$$

(6)

(Youdin & Lithwick 2007; Youdin 2011). The dust’s velocity dispersion (if excited by turbulence) can also be expressed in terms of $\alpha_g$ and we find

$$\eta = \frac{\epsilon^3 + 2 \epsilon + 5/4}{(\epsilon \epsilon^{-1})^2} \alpha_g.$$

(7)

Thus $\alpha_g$ governs both $D$ and $\eta$. Unfortunately, an estimate of its magnitude is one of the great uncertainties in the theory, though it is likely to be smaller (and possibly much smaller) than the analogous dimensionless coefficient associated with angular momentum transport in the gas ($10^{-3} - 10^{-2}$ in a protoplanetary disc). For instance, if transport is controlled by magneto-centrifugal winds or strong zonal magnetic fields, then the bulk of the disc could even be laminar (e.g. Bai 2014; Lesur, Kunz & Fromang 2014; Simon et al. 2015). Having said that, vertical settling should always lead to ordered motions and some degree of radial diffusion. Putting these considerations aside, it is clear that strongly coupled particles ($\epsilon \gg 1$) have $\alpha \sim \alpha_g$ and $\eta \sim \alpha_g$, whereas weakly coupled particles...
A SINGLE-FUID MODEL

3.1 Governing equations

As we are interested in relatively short radial scales, it is convenient to employ the shearing sheet model (Goldreich & Lynden-Bell 1965), whereby a small portion of disc, centred upon a fixed radius $R_0$, is represented by a Cartesian box. In a corotating frame centred on the box, $x$ and $y$ denote the local radial and azimuthal coordinates, while $\Omega$ is the orbital frequency at $R_0$. The disc is assumed to be razor thin, so that the dust volumetric density is $\eta = \epsilon^{2/3} \alpha_0$ and are far ‘colder’, $\eta \approx \epsilon \alpha_0$.

3.2 Dispersion relation

The governing one-fluid equations admit an equilibrium characterized by a constant density $\sigma = \sigma_0$ and the perfect entrainment of the dust in the gas, both undergoing Keplerian motion $u = u_0 = -(3/2)\Omega x e_y$. To this steady state, we add small axisymmetric perturbation, $\sigma', u'$, proportional to $\exp(st + ikr)$, where $s$ is the growth rate and $k$ is the radial wavenumber. Their linearized equations are

\[ s\sigma' = -\sigma_0 i k u'_y, \]
\[ su'_x = 2\Omega u'_y - i k \Phi_\sigma - \epsilon^i \sigma_0 - \epsilon \Omega u'_x, \]
\[ su'_y = -\frac{1}{2} \Omega u'_y - \epsilon u'_x. \]

where the perturbed gravitational potential is given by $\Phi_\sigma = -(2\pi G/k^2)\sigma$ (e.g. Binney & Tremaine 1987) and where we have set $D = 0$ for the time being.

Eliminating the primed variables produces a relatively neat third-order dispersion relation:

\[ s^3 + 2\Omega s^2 + (s^2 + \epsilon^2 \Omega^2) s + \epsilon \Omega (s^2 - \Omega^2) = 0. \]

Here

\[ \sigma^2 = \Omega^2 - 2\pi G\sigma_0 |k| + k^2 c_s^2 \]

is the standard expression for the squared frequency of density waves in a 1D inviscid disc. This agrees with multiple examples in the literature, notably in Ward (2000) and Youdin (2005), and also in Youdin (2011) and Shariff & Cuzzi (2011), when turbulent diffusion is omitted.

3.3 Without gas drag

It is worthwhile examining the classical case with no drag, i.e. when $\epsilon = 0$. The dispersion simplifies and one obtains $s = \pm is_0$ and $s = 0$. The first pair of solutions corresponds to 1D density waves, which can grow on a band of intermediate wavenumbers $k$ girdling $k_c = \pi G\sigma_0/c_s^2$.

The instability criterion requires the Toomre parameter $Q = \Omega/c_s/(\pi G\sigma_0)$ to be less than 1 (e.g. Safronov 1969; Goldreich & Ward 1973). Radial collapse on long wavelengths is impeded by epicyclic motion induced by the inertial forces, whereas short wavelength modes are stabilized by pressure. (Note that non-linear non-axisymmetric instability occurs for larger $Q \approx 2$.) The dust, however, must be rather thin in order to achieve $Q < 1$ (e.g. Cuzzi et al. 1993; Chiang & Youdin 2010).

The $s = 0$ ‘quasi-geostrophic’ mode neither grows nor oscillates, but instead corresponds to a steady ‘zonal flow’. From equation (12), the fundamental balance is between the Coriolis force, on one hand, and self-gravity and pressure, on the other hand. The mode corresponds to a radially varying sequence of super- and sub-Keplerian orbital motions (or ‘jets’):

\[ u'_\gamma = \frac{i k}{2 \Omega} \left(1 - 2 \frac{k_c}{|k|}\right) P', \]

where $P'$ is the associated pressure perturbation. The pressure gradient negates any type of epicyclic motion. This type of flow plays an important part in the secular GI, as will be made clear in the following sections, and indeed in the streaming instability (Jacquet, Balbus & Latter 2011).
3.4 With gas drag

Being a cubic, the dispersion relation (14) does not yield simple formulae for the various growth rates. It is straightforward, however, to obtain a general stability criterion and asymptotic expressions in the limit of strong as well as weak drag.

The last term in the cubic reveals that linear instability is assured when \(\Omega^2 - \Omega_k^2 < 0\), which is satisfied for all modes with sufficiently small wavenumbers \(k < k_c\). Hence, instability is unconditional, though in practice (putting aside the disc’s global structure and cylindrical geometric effects) an unstable dust layer must have radial extent greater than \(\approx 2\pi/k_c\), or else there will be insufficient space for the modes to manifest themselves.

In Fig. 2 we plot numerical growth rates of the secular GI for parameters characteristic of the two limits. Superimposed are the leading order estimates taken from equations (18) and (20).

3.4.1 Weak coupling limit

When \(\epsilon \ll 1\), we are in the weak drag regime corresponding to larger particles, at larger disc radii, in less massive discs. Equation (14) then yields two density waves,

\[
s = \pm i \Omega - \frac{1}{2} \epsilon \Omega \left( 1 + \frac{\Omega^2}{\Omega_k^2} \right) + O(\epsilon^3 \Omega),
\]

which are unstable, as expected, when \(Q < 1\) and thus correspond to the classical GI. Otherwise the two waves are mildly damped by drag.

The third secular mode exhibits a growth rate of

\[
s = \frac{\Omega^2 - \Omega_k^2}{\Omega_k} \epsilon \Omega + O(\epsilon^3 \Omega),
\]

Instability is assured when \(\Omega^2 < \Omega_k^2\), in accordance with the stability criterion derived above. The fastest growing mode possesses \(k = k_c\) and the maximum growth rate is

\[
s_{\text{max}} \approx \frac{\epsilon \Omega}{Q^2 - 1}.
\]

The mode always grows, but for large \(Q\) the growth rate can be small.

In the limit of small \(\epsilon\), the unstable mode is, in fact, a modified zonal flow. In equation (12), the dominant balance is quasi-geostrophic, i.e., between the pressure gradient, self-gravity, and the Coriolis force (because \(\epsilon \sim \Omega\) and \(u_r \sim \epsilon u_r\)). Using equation (16), with equations (11) and (13), gives then precisely the leading order term for \(s\).

3.4.2 Strong coupling regime

In the strong drag limit, \(\epsilon \gg 1\), the SGI grows at a rate of

\[
s = \frac{\Omega^2 - \Omega_k^2}{\Omega_k} \epsilon^{-1} + O(\epsilon^{-2}).
\]

Instability occurs for the exact same range of \(k\) as in the opposed weak drag limit, but the maximum growth rate is slightly altered:

\[
s_{\text{max}} \approx \frac{\Omega}{Q^2 \epsilon},
\]

and the mode works through a different arrangement of forces. In equation (12), instead of radial geostrophic balance, it is the last three terms that are dominant: radial contraction by self-gravity is met by the drag force and pressure. Essentially, the dust has achieved terminal velocity. Solving for \(u'_r\) and using equation (11) yield the leading term for \(s\).

In addition, there exist two density waves with \(s = \pm i \Omega + O(\epsilon^{-1} \Omega)\). As in the weak coupling limit, they grow exponentially when \(Q < 1\) but are otherwise weakly damped by drag at a rate \(\sim \Omega/\epsilon\).

3.4.3 Role of turbulent diffusion

For completeness, we next consider the impact of turbulent mass diffusivity, focusing especially on the critical wavenumber at which instability is quenched. Letting \(D \neq 0\), and taking the limit of weak drag, \(\epsilon \ll 1\), the leading order expression for the growth rate [equation (18)] picks up a term \(-k^2 D \Omega^2/\omega^2\). The critical \(k\) is then easy to compute:

\[
k_{\text{cut}} = \frac{2\pi G \sigma_0}{c^2 + D \Omega^2/\epsilon}.
\]

The importance of turbulent diffusion versus pressure is quantified by \(D \Omega^2/\epsilon c^2\). For weakly coupled particles, this parameter is \(\sim \alpha (\eta c)^{-1} \sim 1\), using the estimates on turbulent diffusion and velocity dispersion from Section 2. This is saying that turbulence is roughly as important as dust pressure, at least in setting the short scale cut-off for SGI. Nonetheless, in the remainder of the paper we often omit diffusion when dealing with weakly coupled particles, primarily in order to derive clean expressions. The neglect of diffusion will not change these results qualitatively. Instability criteria should then be regarded as necessary conditions, not sufficient, and maximum growth rates should be understood as upper bounds.

In the opposite limit, \(\epsilon \gg 1\), the situation is quite different. The growth rate [equation (20)] is modified by the term \(-k^2 D\), and the importance of turbulent diffusion on small scales is quantified instead by \(\epsilon D \Omega^2/c^2\) which is \(\sim \epsilon \gg 1\). As a consequence, the critical cut-off for small particles is completely controlled by turbulence, as shown earlier by Youdin (2011) using a heuristic (but essentially equivalent) argument.

On account of the subdominance of particle pressure in the face of turbulent diffusion when \(\epsilon \gg 1\), we may dispense with it entirely.
In this case, instability is assured for sufficiently long wavelengths as earlier, with the asymptotic growth given by
\[ s = \frac{2\pi G\sigma_0}{\epsilon \Omega} k - Dk^2, \] (23)
and a maximum growth rate of
\[ s_{\text{max}} = \frac{\pi^2 G^2 \sigma_0^2}{\epsilon^2 \Omega^2 D} \approx \frac{\Omega}{\epsilon a_0 Q_x^2}, \] (24)
where we have introduced the gas Toomre parameter in the last equality, and set \( \sigma = \sigma_0 \) (cf. Section 2.2). The wavelength of maximum growth is \( \sim a_0 Q_x H_g \), where \( H_g \) is the scale height of the gas disc.

Before moving on, it should be highlighted that a turbulent mass diffusion alters the classical GI in unexpected ways. When \( \epsilon = 0 \) and \( D \neq 0 \), arbitrarily long wavelengths are rendered unstable, though they grow at negligible rates. On the other hand, mass diffusion cannot completely stabilize short scales – a pressureless turbulent fluid will be unstable for all \( k \) (in contrast to the SGI). Introducing turbulent momentum diffusion, however, does kill instability for sufficiently large \( k \). We omit details of these calculations.

### 3.5 Physical picture

As has been commented upon in the literature, the secular GI exhibits two striking features: growth for all finite \( Q \), and growth upon arbitrarily small \( k \). The latter is especially unexpected given that traditional GI prefers intermediate wavelengths, the longer scales stabilized by the dominant Coriolis force. How does secular GI overcome the epicyclic response at large scales?

Consider two dust rings located at different radii undergoing circular orbital motion. Each ring contains a quantity of angular momentum naturally associated with its home radius. Suppose the rings are displaced radially towards one another. Though their mutual self-gravitational attraction will attempt to amplify the displacement, the two rings possess an angular momentum incommensurate with their new radial location and hence undergo epicyclic oscillations that thwart any type of gravitational collapse. This is the classical picture of stabilization at long wavelengths.

Suppose, however, that there exists a drag force on both dust rings due to interactions with the background gas. Now when the two rings are radially displaced, they will exchange angular momentum with the gas, via the last term in equation (13). Suppose, in addition, that it is possible for the rings to gain or lose angular momentum via drag so that they possess an amount commensurable with their new perturbed radial location. If so, they do not undergo epicyclic motion, and self-gravity continues to amplify their radial drift towards one another. As the two rings collapse, they continuously shed or gain angular momentum as needed. In this way, the epicyclic restoring forces are negated by the drag. Motion associated with classical density waves fails to provide the correct drag-induced angular momentum exchange for this to happen, but the motion associated with the slow third mode [equation (16)] can.

Note that this scenario only works if the gas remains an infinite reservoir of angular momentum, which can be removed or added to with no ramifications. This may be a reasonable approximation in cases where the dust density is far less than the gas density, but on some sufficiently long scale even this must break down. The question then arises: on what range of scales does the secular GI operate upon, and under what circumstances may we take the single-fluid approximation? For the smooth running of the SGI, the gas fluid must resist undergoing epicyclic oscillation when perturbed by the dust drag. When and how can this be arranged? These questions will be explored in the following two sections.

### 4 TWO-FLUID MODEL: INCOMPRESSIBLE GAS

As an intermediate step between the single-fluid and fully compressible two-fluid models, we briefly analyse the case of a compressible dust disc embedded in an incompressible gas. This situation mimics the case when the dust scale height is far less than the gas scale height, and the unstable motions are very subsonic. On the vertical scale of the dust disc, the gas density is effectively constant and the problem is `vertically local’ as far as the gas is concerned. Consequently, the gas density does not contribute to the perturbed Poisson equation.

#### 4.1 Governing equations

The equations governing the evolution of the incompressible gas are

\[ \partial_t U + U \cdot \nabla U = -\nabla P_g / \sigma_g - \nabla \Phi_f - \nabla \Phi_g - 2\Omega \epsilon_x \times U + \epsilon \Omega \sigma_g \sigma_f (u - U), \] (25)

\[ \nabla \cdot U = 0, \] (26)

where \( \sigma_g \) and \( P_g \) are the gas surface density and vertically integrated pressure, respectively, within the dust layer. Because the gas is incompressible \( \epsilon_g \) is a constant. As earlier, \( U \) is the gas velocity, and \( u \) is the dust velocity. The equations governing the dust fluid are those that appear in Section 2, equations (8)–(10). To ease the analysis, \( D = 0 \).

#### 4.2 Dispersion relation

Once again, we assume the standard equilibrium state \( \sigma_g = \sigma_g^0, \sigma_f = \sigma_f^0, u = \vec{U} = -\epsilon_0 x e_x, \) where the gas pressure and dust surface density are constant. Next, axisymmetric perturbations are assumed, denoted by \( \vec{U}', \sigma_g^\prime, u', \sigma_f^\prime \), and these are taken to be \( \propto \exp(st + ikx) \).

Because of the incompressibility condition we immediately obtain \( U_y^\prime = 0 \), and \( U_x^\prime \) is computed from the y-force balance, yielding

\[ U_x^\prime = \frac{\epsilon \delta \Omega}{s + \epsilon \delta \Omega} u_y^\prime, \] (27)

where we have introduced new \( \delta = \sigma_0 / \sigma_g \) which quantifies the dust-to-gas density ratio for a given particle size. This equation states that the gas is azimuthally accelerated by dust drag. Simultaneously, a form of radial geostrophic balance holds for the gas, with the Coriolis force balanced by the radial pressure, self-gravity, and drag. Importantly, the gas perturbation cannot undergo epicyclic motion, which might impede the growth of the secular GI. The gas pressure gradient holds the fluid radially ‘in place’ and a sequence of azimuthal jets ensues, each accelerated by the dust drag. (Note that the absence of epicycles is a generic feature of incompressible flow confined to the orbital plane.)

The perturbation equations for the dust are the same as equations (11)–(13) except for the inclusion of the term \( \epsilon \Omega U_y^\prime \) on the right-hand side of the dust’s y-force balance. Eliminating the dependent variables obtains the following quartic dispersion...
We adopt the equations listed in TI in order to describe our system:

\[ s^4 + (2 + \delta)\epsilon \Omega s^3 + [\sigma^2 + (1 + 3\delta)\epsilon^2 \Omega^2] s^2 + \epsilon \Omega [\delta(\sigma^2 + 2\epsilon^2 \Omega^2) + \sigma^2 \Omega^2] s + 2\delta(\sigma^2 - \epsilon^2 \Omega^2)^2 \Omega^2 = 0, \]  

which we now briefly analyse.

4.3 Instability criterion and asymptotic growth rates

Though equation (28) may appear rather formidable, an instability criterion appears immediately. Putting aside the classical GI for now, the SGI mode is marginal when the last term is zero, yielding exactly the same instability criterion as in the single-fluid model: \( \sigma^2 < \Omega^2 \), and hence instability occurs on all \( k < 2k_c \). Though we allow for gas perturbations, gas incompressibility restricts these to a form of zonal flow which absorbs or bestows angular momentum as necessary to facilitate instability.

In the weak coupling limit, \( \epsilon \ll 1 \), the leading order term in the SGI growth rate is obtained by setting \( s = s_1 \epsilon + \cdots \) and substituting this into equation (28). One obtains the quadratic:

\[ s_1^2 + \Omega(\delta - \xi)s_1 = 2s_1(\Omega^2 - \sigma^2)\Omega^2 = 0. \]  

The resulting solution for \( s_1 \) agrees with the single-fluid expression [equation (18)] to leading order in small dust-to-gas fraction \( \delta \ll 1 \).

In the strong coupling limit, \( \epsilon \gg 1 \), assuming that \( s \sim \Omega/\epsilon \), a similar analysis reveals that the SGI growth does not depend on \( s \) at all. In fact, \( s = (\Omega^2 - \sigma^2)/(\epsilon \Omega) \), precisely the same expression as in the single-fluid case [equation (20)].

5 TWO-FLUID MODEL: COMPRESSIBLE GAS

Having treated simpler models of the dust–gas system, we turn to a fully compressible two-fluid approach. The sound speed of the gas \( c_g \) and its scale height \( H_g = c_g^2/\Omega \) are assumed finite, with the dust subdisc embedded in the gas, so that \( c < c_g \) and \( H < H_g \). As in the previous section, we average over the vertical thickness of the dust disc, and thus neglect complications arising on the smaller scales \(<H \). such as shear instabilities and turbulence. These are included in an ad hoc way, using the turbulent mass diffusivity and enhanced dust pressure. Perhaps, more importantly, the gas outside the dust disc is completely neglected as far as the onset of instability is concerned. The external gas is ‘inert’ – both gravitationally and dynamically decoupled. The resulting model is workable but suffers the shortcomings discussed previously in Section 2.

5.1 Governing equations

We adopt the equations listed in TI in order to describe our system:

\[ \partial_t \sigma + \mathbf{u} \cdot \nabla \sigma = -\sigma \nabla \cdot \mathbf{u} + D \nabla^2 \sigma, \]  
\[ \partial_t \mathbf{u} + \mathbf{u} \cdot \nabla \mathbf{u} = -\frac{1}{\sigma} \nabla P - \nabla \Phi_g - \nabla \Phi_d - 2\Omega \epsilon e_z \times \mathbf{u} + \epsilon \Omega (\mathbf{U} - \mathbf{u}), \]  
\[ \partial_t \sigma_g + \mathbf{U} \cdot \nabla \sigma_g = -\sigma_g \nabla \cdot \mathbf{U}, \]  

where \( \mathbf{U} = (U_x, U_y, U_z) \) is the velocity of the gas in the laboratory frame, \( \mathbf{u} = (u_x, u_y, u_z) \) is the velocity of the gas in the dust frame, \( \sigma \) and \( \sigma_g \) are the local volume fraction of the solid and gas, respectively, \( \Phi_g \) is the gravitational potential of the gas, \( \Phi_d \) is the gravitational potential of the dust, \( P \) is the pressure, and \( D \) is the diffusion coefficient.

5.2 Dispersion relation

We assume a simple background equilibrium of homogeneous density and Keplerian shear: \( \sigma = \sigma_0, \sigma_g = \sigma_0 g, \mathbf{u} = \mathbf{U} = -(3/2)\Omega \epsilon \epsilon e_z \). To this, we add disturbances, denoted by primes, that are \( \propto \exp(\delta \pm ikx) \). The resulting linearized equations are as follows:

\[ \sigma' = -\sigma_0 \epsilon k \mathbf{u}_g', -D \epsilon k^2 \sigma', \]  
\[ \mathbf{u}_g' = -i k c_g^2 (\sigma'/\sigma_0) + 2\Omega \mathbf{u}_g' - i k \Phi' + \epsilon \Omega (\mathbf{U}_g' - \mathbf{u}_g'), \]  
\[ \sigma_0' = -\sigma_0 \epsilon k \mathbf{u}_s', \]  
\[ \mathbf{u}_s' = -i k c_s^2 (\mathbf{u}_s'/\sigma_0) + 2\Omega \mathbf{u}_s' - i k \Phi' + \epsilon \Omega (\mathbf{U}_s' - \mathbf{u}_s'), \]  
\[ \mathbf{U}_s' = -\frac{1}{2} \epsilon \Omega \mathbf{U}_s' + \epsilon \Omega (\mathbf{U}_s' - \mathbf{U}_g'), \]  

where the equilibrium of dust-to-gas density ratio is \( \delta = \sigma_0/\sigma_0 g \) for a given size \( a \). Finally, the perturbed gravitational potential is obtained from

\[ \Phi' = - (2\pi G/|k|) \left( \sigma' + \sigma_0' \right). \]  

The system of equations (35)–(40) yields a rather involved sixth-order dispersion relation:

\[ s_6 + a_5 s_5 + a_4 s_4 + a_3 s_3 + a_2 s_2 + a_1 s + a_0 = 0, \]  

with

\[ a_5 = 2(1 + \delta)\epsilon \Omega + D k^2, \]  
\[ a_4 = (1 + \delta)^2 \epsilon^2 \Omega^2 + \sigma_0^2 + 2(1 + \delta)\epsilon \Omega D k^2, \]  
\[ a_3 = \epsilon \Omega \left[ (1 + \delta)\mu^2 - \delta c_g^2 k^2 + (1 + 2\delta)\epsilon c_s^2 k^2 \right] + \left[ (1 + \delta)^2 \epsilon^2 \Omega^2 + \sigma_0^2 + \Omega^2 \right] D k^2, \]  
\[ a_2 = c_g^2 \sigma_0^2 + \Omega^2 \mu^2 - 2c_g^2 k^2 \pi G \sigma_0 |k| \left[ (1 + \delta)^2 \epsilon^2 \Omega^2 + (1 + \delta)\mu^2 + \delta c_g^2 k^2 - \delta c_s^2 k^2 \right] + [2(1 + \delta)\sigma_0^2 - c_g^2 k^2]\epsilon \Omega D k^2, \]  
\[ a_1 = \epsilon \Omega k^2 \left[ \delta c_s^2 \left( \mu^2 - c_s^2 k^2 \right) + c_s^2 \left( \mu^2 + \delta c_s^2 k^2 \right) \right] + \left( \sigma_0^2 \mu^2 + \epsilon (1 + \delta)^2 \right) - \epsilon^2 \delta (1 + \delta)\epsilon c_s^2 k^2 D k^2 \Omega^2, \]  
\[ a_0 = D c_g^2 k^4 \delta \epsilon \Omega^3. \]
In order to ease the presentation of the coefficients we have introduced the following frequencies:

\[ \omega^2 = \Omega^2 - 2\pi G \sigma_0 |k| + c^2_0 k^2, \]

\[ \mu^2 = \Omega^2 - 2\pi G (\sigma_0 + \epsilon_0) |k| + c^2_0 k^2. \]

Recall that \( \omega^2 \) is the squared frequency of density waves in the dust fluid given by equation (15).

5.3 No turbulent mass diffusion

In this subsection, we analyse the case when mass diffusion in the continuity equation is negligible, \( D \approx 0 \), but we retain the velocity dispersion of the dust particles. This situation may adequately describe a disc of weakly or marginally coupled dust and gas, \( \epsilon \ll 1 \).

5.3.1 Instability criterion

When \( D = 0 \), the dispersion relation [equation (41)] reduces to a quintic. Marginality corresponds then to \( a_i = 0 \), which we use with a condition for the onset of instability. In fact, \( a_1 = 0 \) is a quadratic equation for \( k \),

\[ c^2 k^2 - 2\pi G \sigma_0 \left( c^2 + \delta c^2_0 \right) |k| + \left( c^2 + \delta \epsilon_0 \right) / (1 + \delta) = 0, \]  

indicating that if instability occurs, it takes place on a range of wavenumbers \( k_1 < k < k_2 \), where \( k_1 \) and \( k_2 \) are solutions to the quadratic. In order for such a range to exist, the critical \( k_2 \) must be real. This is the case when the discriminant of equation (42) is positive, and the instability criterion proceeds easily:

\[ Q_\delta = \frac{\Omega^2 c_0 \sigma_0}{\pi G \sigma_0} < \sqrt{(1 + \delta) (\delta + \eta)} \approx \sqrt{1 + \frac{\delta}{\eta}}, \]

where the last scaling arises if \( \delta \ll 1 \). Alternatively, the criterion can be reworked in terms of the dust’s Toomre parameter, noting that \( Q = \sqrt{\eta \delta}^{-1} Q_\delta \). We then obtain instability when

\[ Q \lesssim \frac{\sqrt{\delta + \eta}}{\delta} \approx \delta^{-1/2}, \]

where in the last approximation we assume that \( \eta \ll \delta \) (always the case, unless turbulence is absent – see Section 2.2).

However expressed, this simple criterion encapsulates clearly the main physical effects. For instance, if we take the limit of negligible dust, \( \delta \rightarrow 0 \), the instability criterion becomes simply \( Q < \infty \), and the system is always unstable. In this limit, the dust’s drag on the gas is unimportant and the two-fluid system reproduces the same stability behaviour as the single-fluid model (Section 3); the Toomre parameter does not feature.

For general \( \delta \), however, stability in the two-fluid disc does in fact depend on the gas’s (or dust’s) Toomre parameter: if it is too large, then instability is switched off. If \( \delta / \eta \gg 1 \), the instability criterion may be written as

\[ Q_\delta \approx \frac{c_0}{c} \left( \frac{\sigma_0}{\sigma_0} \right)^{1/2} \left( \epsilon \epsilon_0 \right)^{-1/2} \left( \frac{\epsilon}{\epsilon_0} \right)^{1/2}, \]

in which the critical \( Q_\delta \) may exceed the classical value of 1, allowing instability to occur in a finite range of conditions. In this case, we recover the secular GI.

5.3.2 Asymptotic growth rates

Explicit expressions for the growth rate are possible in the limit of \( \epsilon \ll 1 \). We set aside the classical GI and isolate the SGI mode by setting \( s = s_1 \epsilon + \cdots \). Collecting terms of order \( \epsilon \), we obtain

\[ s_1 = \left( \frac{\sigma_0}{\sigma_0} \right) \left( \frac{2\pi G \sigma_0 |k| - \Omega^2}{\sigma_0} \right) \left( \frac{k^2 c_0^2}{\Omega} \right) \left( \frac{k^2 c_0^2}{\Omega} \right) \]

leading order in small \( \delta \) and \( \eta \). The last damping term arises from the gas pressure and kicks in at short scales, of order \( H = c / \Omega \). Instability is thus restricted to scales longer than \( H \). Conversely, the Coriolis force acting on the gas (represented by the \( \Omega^2 \)) stabilizes long wavelengths, so that instability only occurs on scales shorter than \( \approx 2\pi G \sigma_0 / \Omega^2 \) (as in classical GI). The stabilizing effect of gas pressure, on the other hand, does not make an appearance. In summary, instability occurs on a range of intermediate scales. But for this range of unstable wavelengths to exist, the dust pressure must be sufficiently weak, or else the last term in equation (46) swallows the first term.

Expression (46) can be further simplified if we restrict our attention to length scales much shorter than the gas’s Jeans length \( \sim c_0^2 / (G \sigma_0) \), but not so short that the dust pressure dominates the other processes. In other words, we require \( G \sigma_0 / c_0^2 \ll k \ll G \sigma_0 / c^2 \). If this intermediate range exists, then on it we may approximate \( \omega^2 \approx k^2 c_0^2 \) in the denominator of the first term in equation (46), and obtain

\[ s \approx \frac{\Omega^2 - \omega^2}{\Omega} \epsilon = \epsilon \Omega. \]

Assuming further that \( \delta \ll 1 \), the last term may be dropped and we have an expression similar to the single-fluid one, equation (18).

The maximum growth rate is easy to compute:

\[ s_{\max} \approx \frac{\epsilon \Omega}{Q^2}, \]

using the Toomre parameter for the dust. This expression is consistent with equation (19), for sufficiently large \( Q \). But note that it only holds if there is a sufficient separation of scales between the dust and gas pressure scale heights, which is assured if \( \eta \ll 0 \). The wavelength of maximum growth is also similar to that in a single fluid \( \sim QH \).

In Fig. 3, we plot the asymptotic growth rate as a function of \( k \) alongside the full solution obtained numerically from equation (41). For the parameters chosen the asymptotic regime of small \( \delta \), \( \eta \), and \( \epsilon \). Note that the lengthscale of fastest growth is less than the gas-disc thickness \( \sim (\pi / 10) H \), though much longer than the particle-disc thickness because \( H / H_k \sim 10^{-2} \).

5.4 Pressureless turbulent dust

Having treated the non-turbulent case, more relevant for weakly coupled particles, we now turn to a pressureless dust suspended in a turbulent gas disc, the regime that best describes small dust. Thus in equations (35)–(40) we set \( D \neq 0 \), but \( c = 0 \), and it is assumed that \( \epsilon \gg 1 \). In addition, the turbulent diffusion is presumed small, \( \alpha \ll 1 \). The simplest way to capture the full effects of diffusion is to
let \( D k^2 \sim \epsilon^{-1} \Omega \), which (whatever the value of \( D \) is) will be true on some radial lengthscale. On longer wavelengths persisting with this scaling means we include harmless subdominant terms, while on shorter wavelengths we expect diffusion to quench instability in any case. The resulting equations correspond exactly to those treated in sections 2 and 3 in TI, which we now analyse in more detail and give explicit expressions for the growth rates.

5.4.1 Instability criterion

When \( D \neq 0 \), the onset of instability is difficult to calculate from the dispersion relation because unstable modes possess (small) complex frequencies. Some progress can be made, however, in the limit of dispersion relation because unstable modes possess (small) complex \( D \) when \( \epsilon \gg 1 \) and assuming that \( s = \epsilon^{-1} s_1 + \cdots \) (thus extracting only SGI modes) and \( \alpha_g = \epsilon^{-1} \alpha_1 + \cdots \). Recall from Section 2.2 that in the tight coupling limit \( \alpha \approx \alpha_g \). To leading order the dispersion relation [equation (41)] becomes the quadratic:

\[
(1 + \delta) \left[ \left( 1 + \delta \right) \mu^2 - \delta c_g^2 k^2 \right] \Omega^2 s_1^2 - \left\{ \delta c_g^2 \left( \mu^2 - c_g^2 k^2 \right) \\
\alpha_1 (1 + \delta) \left[ (1 + \delta) \mu^2 - \delta c_g^2 k^2 \right] k^2 c_g^2 \Omega s_1 \\
+ \delta c_g^2 k^4 \alpha_1 \Omega^2 \right\} = 0. \tag{49}
\]

In equation (13), TI presents an equivalent expression. For reasonable values of \( \delta \) and \( Q_g > 1 \), the coefficient of \( s_1^2 \) is positive and so the sign of the growth rates can be determined from the coefficient of \( s_1 \). After some manipulation the instability criterion is

\[
Q_g < \left[ \delta + \epsilon \alpha_g (1 + \delta) \right] \sqrt{\frac{1 + \delta}{\epsilon \alpha_g} \left[ \delta + \epsilon \alpha_g (1 + \delta) \right]}. \tag{50}
\]

which agrees with equation (18) in TI. Taking small \( \delta \), this simplifies to

\[
Q_g \lesssim \sqrt{1 + \frac{\delta}{\epsilon \alpha_g}}, \tag{51}
\]

which is remarkably similar to the diffusionless criterion [equation (43)]. The smaller the turbulent diffusion \( \alpha_g \), the greater the range of instability. Diffusion on small scales has replaced dust pressure in equation (43) in stabilizing the secular GI, but the two processes work exactly the same otherwise. Finally, note that equation (51) differs from the final criteria in TI because they make additional and unnecessary assumptions regarding the relative sizes of \( \epsilon, \delta, \) and \( \alpha_g \).

5.4.2 Asymptotic growth rates

Solutions to the quadratic [equation (49)] are ugly, but if we take the limit \( \delta \ll 1 \), we obtain

\[
\text{Re}(s) = \frac{1}{2} \left( \frac{\sigma_g}{\sigma_0} \right) \left( \frac{2 \pi G \sigma_0 |k| - \Omega^2}{\epsilon \alpha_g} \right) \left( \frac{k^2 c_g^2}{\Omega} \right) - \frac{1}{2} k^2 \alpha. \tag{52}
\]

This is almost identical to expression (46). The main difference is that the stabilizing term on short scales arises from turbulent diffusion, rather than dust pressure, once again. The long wavelength stabilization is the same – the Coriolis force. Instability can only occur if these two stabilizing scales are sufficiently well separated.

In Fig. 4, we plot the asymptotic growth rate [equation (52)] as a function of \( k \) next to the full solution obtained numerically from equation (41). For a representative set of parameters the agreement is good, though weakens the smaller \( \epsilon \) and the larger \( \delta \), as expected. Note also that at small \( k \), the approximation fails to capture the bifurcation into two monotonically growing modes with different growth rates. Finally, the wavelength of fastest growth is \( \sim H_g \), and thus the approximation of a razor-thin disc is only marginally applicable.

The approximate maximum growth rate can be obtained by a similar argument to that given in Section 5.3.2. We find to leading order on the intermediate range \( G \sigma_0/c_g^2 \ll k \ll G \sigma_0/(\epsilon \alpha \Omega) \) and for \( \delta \ll 1 \) that

\[
s \approx \frac{1}{2} \frac{\pi G \sigma_0}{\epsilon \Omega} |k| - k^2 \alpha. \tag{53}
\]

This expression is similar to the growth rate in the single-fluid pressureless case, equation (23), differing only by a half (on account of the unstable mode appearing as a complex conjugate pair). The
The disc model we use is a variant of the MMSN (see Youdin 2011). The background gas disc surface density is given as a function of disc radius by

$$\sigma_0 = 2 \times 10^4 F R_{\text{au}}^{-3/2} \text{ g cm}^{-2}. \quad (55)$$

Here $F$ is a free dimensionless parameter, and $R_{\text{au}}$ is the disc radius in units of au. The temperature of the nebula is given by

$$T = 200 R_{\text{au}}^{-1/2} \text{ K}. \quad (56)$$

As a consequence, the gas's Toomre parameter is

$$Q_g \approx \frac{40}{F} R_{\text{au}}^{-1/4}. \quad (58)$$

Thus, at 1 au, $Q_g$ is roughly 40 and this falls to about 10 as we approach 100 au. The inverse Stokes number can be approximated by

$$\epsilon = 4 \times 10^3 \frac{F}{a_{\text{min}}} R_{\text{au}}^{-3/2}, \quad (59)$$

where $a_{\text{min}}$ is particle size in units of millimetre. We have assumed that the particles always lie in the Epstein regime. Only the largest particles at the smallest radii enter the Stokes drag regime, so to make life simple, we omit it. As a result, stability can be determined once $\delta$, $F$, $\alpha_g$, particle size, and the disc radius are specified.

6.1 Weakly coupled particles

First consider larger particles with a largish Stokes number $\epsilon \lesssim 1$. From Fig. 1, in a standard MMSN these might correspond to $\sim 10$ cm particles at 30 au or more or $\sim 1$ cm particles at 100 au. In an older less massive disc, this class may also include millimetre particles but only at 100 au. Thus most particles do not fall into this regime.

Secular GI arises when criterion (45) is fulfilled. Taking the standard value for the dust-to-gas ratio $\delta = 10^{-2}$, we only need to estimate the ratio of velocity dispersions. Assuming that the particles' random motions are controlled by the background turbulence, as in Section 2, we have $c/c_g \sim \sqrt{\epsilon \alpha_g}$ and thus instability occurs when

$$Q_g < Q_{\text{crit}} \approx 10^{-1} \epsilon^{-1/2} \alpha_g^{-1/2}. \quad (59)$$

This criterion includes both the classical GI of a dust layer and the secular GI, in which we are more interested. As discussed earlier, the properties of the turbulence are poorly constrained. We thus allow $\alpha_g$ to vary between $10^{-4}$ (a perhaps unrealistically low level) and $10^{-3}$. Next, to fix ideas, we set $\epsilon \sim 0.1$ and find that the critical Toomre parameter for the gas is

$$Q_{\text{crit}} \approx 10^{-10},$$

with the lower value corresponding to a thick disc of relatively 'hot' particles ($\alpha_g = 10^{-3}$), and the higher value to a thin disc of colder particles ($\alpha_g = 10^{-7}$).

Our standard MMSN with $F = 1$ yields $Q_g$ that falls directly in this range. If $\alpha_g = 10^{-3}$ instability is not possible, but if $\alpha_g \gtrsim 10^{-4}$, then instability can occur on most radii. Less massive discs exhibit larger $Q_g$; for instance, with $F = 0.1$, we have $Q_g > 100$ at all radii, and thus the existence of SGI is very much conditional on the efficiency of the turbulence. If $\alpha_g \sim 10^{-7}$, then centimetre-sized particles, or even smaller, could be unstable in the outer regions of low-mass discs.
What are the growth rates of the unstable modes? Equation (48) gives an upper bound on $s$, in the regime of larger particles. This can be reworked into

$$s_{\text{max}} \approx \delta^2 Q_g^{-\frac{1}{2}} \alpha_g^{-2} \Omega_s \sim 10^{-5} \alpha_g^{-1} \Omega_s,$$

where the last equality comes by setting $\delta = 10^{-2}$ and $Q_g \sim 10$. For relatively strong turbulence $\alpha_g \sim 10^{-4}$, the $e$-folding time of an unstable mode is $10^4$ orbits, too long to be relevant at 100 au, but possibly significant at smaller radii, for example 10 cm size particles at $\sim 10$ au. Smaller $\alpha_g$, of course, yields faster growth on relevant time-scales.

6.2 Tightly coupled particles

We next consider well-coupled particles, a class that covers most solids of interest (see Fig. 1). The relevant stability criterion for this dust is given by equation (51). The dimensionless diffusion coefficient for tightly coupled particles is $\alpha \approx \alpha_g$. Setting $\delta = 0.01$ and $\epsilon = 10$ yields the following criterion:

$$Q_g \lesssim \sqrt{1 + 10^3 \alpha_g^{-1}},$$

which is more difficult to satisfy than in the weakly coupled case. If we assume that $\alpha_g = 10^{-7} - 10^{-5}$, the condition becomes $Q_g < Q_{\text{crit}} \approx 1 - 100$, where the larger value corresponds to inefficient turbulent diffusion ($\alpha_g = 10^{-7}$) and the lower bound corresponds to efficient diffusion ($\alpha_g = 10^{-5}$). Once again, this suggests that instability occurs when the turbulence is sufficiently weak. In fact, given $F = 1$, the SGI is suppressed if $\alpha_g \geq 10^{-5}$. The situation worsens when $\epsilon > 10$. The conclusion is that the instability may not be widespread in smaller dust.

Let us next turn to growth rates, in particular expression (54). To fix some numbers, we generously set $\epsilon = 10$ and $\delta = 0.01$, and after some manipulation obtain

$$\text{Re}(s_{\text{max}}) \sim 10^{-6} Q_g^{-\frac{1}{2}} \alpha_g^{-1} \Omega_s.$$  

If $Q_g \sim 10$, then we have

$$\text{Re}(s_{\text{max}}) \sim 10^{-8} \alpha_g^{-1} \Omega_s.$$  

While $\alpha_g \lesssim 10^{-5}$ yields appreciable growth at all radii, $\alpha_g \sim 10^{-6}$ does so only for $R \lesssim 10$ au. For larger $\epsilon$, growth times lengthen. This further reinforces the conclusion that the SGI is only relevant to the dynamics of tightly coupled particles when turbulence is very low indeed.

6.3 Marginally coupled particles in realistic disc models

In this subsection, we solve the full dispersion relation at each radius of our disc model. Fig. 1 indicates that particle sizes of a millimetre and above couple to the gas differently at different radii, potentially passing from the well-coupled regime to the weakly coupled regime as we go further out radially. At certain radii, $\epsilon \sim 1$ and our analytic results are no longer strictly valid, meaning that the dispersion relation (41) must be solved numerically.

Some stability curves are plotted in Fig. 5 for two different disc models, $F = 1$ (left panels) and $F = 0.1$ (right panels), for several values of the turbulence parameter $\alpha_g$, and for two values of the dust-to-gas density ratio, $\delta = 0.01$ (top row) and 0.1 (bottom row). Parameter regions above a given curve are subject to instability, and thus for a given $\alpha_g$ and particle size there exists a critical radius within which the SGI is completely suppressed. In fact, for a relatively turbulent disc with $\alpha_g = 10^{-5}$, $F = 1$, and $\delta = 0.01$, all particles smaller than $\sim 3$ cm are stable, and all particles smaller than 1 mm are stable, when $\alpha_g = 10^{-5}$. Weaker levels of turbulence, of course, permit instability upon smaller particle sizes and for larger swathes of the disc. But one must drive $\alpha_g$ to levels $\sim 10^{-7}$ to obtain SGI at radii $< 10$ au for particles larger than a millimetre.

The SGI’s struggles worsen when the disc is less massive ($F = 0.1$), with millimetre-sized particles unstable only for very low values of $\alpha_g$. But moving to the lower row of plots, it is immediately clear that increasing $\delta$ improves its range. For example, if $F = 1$ and $\alpha_g = 10^{-4}$ then all particles smaller than millimetre sizes are stable. If $\alpha_g = 10^{-5}$ then all particles smaller than $\sim 0.1$ mm are stable. On the other hand, when $\delta < 0.01$ the prospects for instability become increasingly bleak. We also compute the minimum $e$-folding times for the SGI, for $F = 1$ and $\delta = 0.01$. These are plotted in Fig. 6. Solid curves correspond to centimetre-sized particles and dashed curves correspond to millimetre-sized particles. We omit smaller particles because they are always in the well-coupled regime that treated in Section 6.2. The different colours represent different values of $\alpha_g$.

As is clear, a turbulence level of $\alpha_g > 10^{-5}$ yields a growth time too long to be important for both particle classes, while a value of $\alpha_g = 10^{-6}$ yields a growth time of a few $10^4$ years for centimetre-sized particles, at a large range of radii $\geq 20$ au. Millimetre-sized particles exhibit appreciable growth only for low turbulence levels $\alpha_g \sim 10^{-7}$ and at larger radii $\geq 50$ au. Less massive discs (such as with $F = 0.1$) yield even weaker growth. Increasing $\delta$ to 0.1 exacerbates growth by up to an order of magnitude, whereas decreasing $F$ to 0.1 reduces the growth rate by roughly an order of magnitude.

These growth rates are, on the whole, consistent with Youdin (2011) at radii $\geq 20$ au, but closer to the disc, two-fluid effects lead to noticeable departures. The growth times diverge at certain critical radii. These occur, of course, when the SGI is stabilized, as indicated by Fig. 5. For a given $\alpha_g$, any given particle size has a critical radius within which the SGI is stable. The takeaway message is that millimetre-sized particles require very low levels of turbulence to be SGI unstable, and then this is localized to the outer parts of the disc. Centimetre-sized particles do better, and the SGI may play some role in their dynamics across a range of disc radii and disc properties. Finally, older, less massive discs struggle to host the SGI in any form, though a larger dust-to-gas ratio can mitigate this to some degree. Unfortunately, like $\alpha_g$, the parameter $\delta$ is difficult to constrain.

7 CONCLUSION

In this paper, we have explored the SGI using a simple two-fluid model. Despite the complexity of its associated sixth-order dispersion relation, analytic stability criteria and growth rates can be obtained in the two limits of weakly and strongly coupled particles. We find that on sufficiently long and short radial scales the SGI is stabilized; the existence of an unstable range of intermediate scales leads to an explicit instability condition involving the gas’s Toomre parameter, a distinctive feature of the two-fluid SGI, as opposed to the single-fluid version.

The mathematical analysis suggests a straightforward way to understand the instability mechanism. The SGI favours intermediate scales upon which stabilizing dust pressure or turbulence is weak, but upon which the gas pressure is strong. The latter condition permits the gas to fall into geostrophic balance; hence, when the gas is azimuthally accelerated by the dust drag, it will form a zonal

---

flow rather than undergoing epicycles that would disrupt the radially collapsing dust.

An assessment of the prevalence of SGI in real disc models is handicapped by uncertainties in two parameters, the strength of the turbulence, $\alpha_g$, and the mass ratio of a certain species of dust to the gas within the dust subdisc, $\delta$. Starting with a fiducial value of $\delta = 0.01$, we find that a moderate level of turbulence $\alpha_g = 10^{-5}$ prohibits the SGI on most radii, and when it does occur, it grows too slowly $\sim 10^5$ yr compared to the time-scale of the large-scale evolution of the disc, and of appreciable radial drift. Weaker turbulence $\alpha_g = 10^{-6}$ permits growth for centimetre-sized particles on radii $\gtrsim 10$ au, with e-folding times of a few $10^4$ yr. Smaller-sized particles may be subjected to SGI but grow too slowly. It is only for $\alpha_g = 10^{-7}$ that millimetre-sized particles sustain growth at reasonable levels, and then only for $R > 10$ au. Increasing $\delta$ improves the situation, of course, and $\delta > 0.01$ might be the case for particularly well-settled and populous subclasses of particles, though further work is needed to better constrain this parameter. Even so, if $\alpha_g > 10^{-6}$ it may prove to be difficult for the SGI to meaningfully impose itself on the disc dynamics.

We also discuss the various shortcomings of the razor-thin model we employ, which is especially a problem when the dust and gas discs exhibit different scale thicknesses. These issues no doubt impact quantitatively on our results, but the main qualitative conclusions and our picture of instability should be robust. They can be checked with a suitable vertically stratified analysis akin to Mamatashvili & Rice (2010) and Lin (2014), which will also provide more reliable quantitative estimates on the stability curves and growth rates.

Our results extend previous analyses of the SGI, and for larger radii are in relative agreement with Youdin (2011) and Shariff & Cuzzi (2011). A notable difference is that the two-fluid model prohibits SGI on radii less than a critical radius. As a result, the SGI is certainly unviable on radii $< 1$ au, and possibly absent on radii $< 10$ au, the expected regions of planet formation. The prospects for SGI in the centimetre class of particles on disc radii $\sim 10$ au are reasonable as long as gas turbulence is not too efficient $\alpha_g \ll 10^{-6}$. The instability could then be an important route by which large aggregates could form further out, leapfrogging the entire range of difficult centimetre to kilometre sizes. Note that our results are only for axisymmetric instability. It is likely, via analogy with classical GI, that non-axisymmetric SGI occurs for larger $Q_g$, in which case our stability curves may need some revision.

It has been hypothesized that the SGI generates observed dust ring structures at larger radii in protoplanetary discs (TI). As discussed in Section 6, however, the SGI has great difficulty on radii $\gtrsim 10$ au for small particle less than a centimetre in size. The dust-to-gas ratio $\delta$ needs to be increased, and $Q_g$ taken to levels approaching 1 in order

Figure 5. Stability curves in a plane of disc radius and particle size. The pink, black, red, and blue curves correspond to $\alpha_g = 10^{-7}, 10^{-6}, 10^{-5},$ and $10^{-4}$, respectively. The regions above the respective curves are unstable. Two disc models are employed, a more massive disc with $F = 1$ (left panels) and a less massive disc with $F = 0.1$ (right panels). The upper row corresponds to a dust-to-gas mass ratio of $\delta = 0.01$, while the lower row to $\delta = 0.1$. 

Figure 6. Maximum e-folding times for the SGI at different radii in a standard disc model with $F = 1$ and dust-to-gas ratio $\delta = 0.01$. The solid curves correspond to centimetre-sized particles, and the dashed curves to millimetre-sized particles. The blue colour represents a turbulent mass diffusivity of $\alpha_g = 10^{-5}$, red represents $\alpha_g = 10^{-6}$, and green represents $\alpha_g = 10^{-7}$.

to obtain instability. While it may be possible to justify increasing $\delta$, such a low $Q_g$ would mean that the gas disc is marginally unstable to classical GI. Perhaps a more important point is that, while the linear phase of the SGI evolution is axisymmetric, its non-linear phase will most likely involve a non-axisymmetric breakdown into disordered flow, as in classical GI, not the formation of large-scale quasi-steady rings. Dedicated non-linear simulations are required to test what dynamics the SGI exhibits once it reaches non-linear amplitudes, and how readily it forms planetesimal clumps. This forms the basis of future work.

ACKNOWLEDGEMENTS

The authors thank the reviewer, Dick Durisen, for a set of comments that led to a much improved manuscript. HNL acknowledges partial funding from STFC grant ST/L000636/1, and RR from a Bridgewater summer internship and from Newnham college.

REFERENCES


Chiang E., Youdin A. N., 2010, Annu. Rev. Earth Planet. Sci., 38, 493

This paper has been typeset from a \TeX/L\TeX file prepared by the author.