FROM DUST TO PLANETESIMALS: CRITERIA FOR GRAVITATIONAL INSTABILITY OF SMALL PARTICLES IN GAS

JI-MING SHI¹, ² AND EUGENE CHIANG¹, ², ³

¹ Department of Astronomy, UC Berkeley, Hearst Field Annex B–20, Berkeley, CA 94720-3411, USA; jmshi@berkeley.edu
² Center for Integrative Planetary Science, UC Berkeley, Hearst Field Annex B–20, Berkeley, CA 94720-3411, USA
³ Department of Earth and Planetary Science, UC Berkeley, 307 McConic Hall, Berkeley, CA 94720-4767, USA

Received 2012 September 20; accepted 2012 December 8; published 2013 January 21

ABSTRACT

Dust particles sediment toward the midplanes of protoplanetary disks, forming dust-rich sublayers encased in gas. What densities must the particle sublayer attain before it can fragment by self-gravity? We describe various candidate threshold densities. One of these is the Roche density, which is that required for a strengthless satellite to resist tidal disruption by its primary. Another is the Toomre density, which is that required for de-stabilizing self-gravity to defeat the stabilizing influences of pressure and rotation. We show that for sublayers containing aerodynamically well-coupled dust, the Toomre density exceeds the Roche density by many (up to about four) orders of magnitude. We present three-dimensional shearing box simulations of self-gravitating, stratified, dust–gas mixtures to test which of the candidate thresholds is relevant for collapse. All of our simulations indicate that the larger Toomre density is required for collapse. This result is sensible because sublayers are readily stabilized by pressure. Sound-crossing times for thin layers are easily shorter than free-fall times, and the effective sound speed in dust–gas suspensions decreases only weakly with the dust-to-gas ratio (as the inverse square root). Our findings assume that particles are small enough that their stopping times in gas are shorter than all other timescales. Relaxing this assumption may lower the threshold for gravitational collapse back down to the Roche criterion. In particular, if the particle stopping time becomes longer than the sound-crossing time, then sublayers may lose pressure support and become gravitationally unstable.

Key words: hydrodynamics – instabilities – methods: numerical – planets and satellites: formation – protoplanetary disks

Online-only material: color figures

1. INTRODUCTION

Gravitational instability is an attractive mechanism to form planetesimals, but how it is triggered in protoplanetary disks remains unclear. In one proposed sequence of events, most of the disk’s solids first coagulate into particles 0.1–1 m in size at orbital distances of a few AU. These “boulder”-sized bodies then further concentrate by the aerodynamic streaming instability (Youdin & Goodman 2005; Johansen et al. 2007; Bai & Stone 2010 and references therein). Local densities are so strongly enhanced by the streaming instability that they can exceed the Roche density (see Section 1.3 for a definition), whereupon collections of boulders may undergo gravitational collapse into more massive, bound structures.

A weakness of this scenario is that it presumes that particle–particle sticking (i.e., chemical adhesion) can convert most of the disk’s solids into boulders, or more accurately, particles whose momentum stopping times in gas

\[ t_{\text{stop}} = \frac{m v_{\text{rel}}}{F_{\text{drag}}} \quad (1) \]

are within a factor of 10 of the local dynamical time \( \Omega^{-1} \), where \( \Omega \) is the Kepler orbital frequency, \( m \) is the particle mass, \( v_{\text{rel}} \) is the relative gas-particle velocity, and \( F_{\text{drag}} \) is the drag force whose form varies with disk environment (see, e.g., Weidenschilling 1977). Figure 1 relates \( t_{\text{stop}} \) to particle radius \( s \) as a function of disk radius \( r \) in the minimum-mass solar nebula (MMSN). For \( r \approx 1–10 \) AU, the condition \( \Omega t_{\text{stop}} = 0.1–1 \) corresponds to \( s \approx 0.1–1 \) m.

Unfortunately, particle–particle sticking might not produce boulders in sufficient numbers for the streaming instability to be significant. A comprehensive study by Zsom et al. (2011; see also Birnstiel et al. 2010) found that for realistic, experiment-based sticking models that include both bouncing and fragmentation, particles no larger than \( \sim 1 \) cm can form by sticking—even when the disk is assumed to have zero turbulence. According to Table 1 of Zsom et al. (2011), coagulation models over most of the parameter space produce \( \tau_s \sim 10^{-3}–10^{-2} \). This range is too small for the streaming instability to concentrate particles strongly—see Bai & Stone (2010), who showed that when half or more of the disk’s solid mass has \( \Omega t_{\text{stop}} < 0.1 \), densities enhanced by the streaming instability still fall short of the Roche density by more than a factor of 10. Even if particle–particle sticking could grow bodies with \( \Omega t_{\text{stop}} \sim 0.1–1 \) (e.g., Okuzumi et al. 2009, who neglected fragmentation), the disk’s solids may not be transformed into such bodies all at once. Rather, boulders may initially comprise a minority on the extreme tail of the size distribution. Unless they can multiply from a minority to a majority within the time it takes for them to drift radially inward by gas drag (\( \sim 100–1000 \) yr starting at 1 AU; Weidenschilling 1977), they threaten to be lost from the nebula by drag.

We are therefore motivated to ask whether gravitational instability is practicable for particles having realistically small sizes and concomitantly short stopping times, say \( \Omega t_{\text{stop}} \lesssim 10^{-2} \). Smaller particles suffer the disadvantage that they are harder to concentrate; since they are well-entrained in gas, turbulence in the gas can loft particles above the midplane and prevent them from collecting into regions of higher density. The streaming
Several recent studies (Chiang 2008; Lee et al. 2010a, 2010b; Sekiya 1983) have measured the maximum sublayer densities permitted by the Kelvin–Helmholtz instability. Neglecting self-gravity, they found that dust-to-gas ratios between ~2 and 30 are possible in disks that are locally enriched in metallicity by factors of 1–4 above solar. Such local enrichment can be generated by radial drifts of particles relative to gas (see Chiang & Youdin 2010 for a review). For observational evidence of radial segregation of dust from gas, see Andrews et al. (2012).

Are such enhancements in the local dust-to-gas ratio sufficient to spawn planetesimals? How high must dust + gas densities be before the effects of self-gravity manifest? Our paper addresses these questions in the limit $\Omega \to 1$, i.e., in the limit that particles are small enough to be well coupled to gas. In the next two subsections, we derive critical densities for gravitational instability in the cases of a pure gas disk (Section 1.1) and a disk composed of both gas and perfectly entrained ($\Omega \to 0$) dust (Section 1.2). The two cases give remarkably different answers for dust-rich sublayers. In Section 1.3, we add two more densities from the literature to the list of proposed criteria for gravitational collapse. Table 1 summarizes the various candidate threshold densities.

In Sections 2 and 3, we present numerical simulations of three-dimensional (3D), self-gravitating, compressible flows of thin, dense sublayers of dust. We use these simulations to try to identify which of the proposed criteria (if any) is the most relevant for gravitational instability. Section 4 summarizes our findings but also points out the limitations of our numerical simulations, which are restricted to the asymptotic limit $\Omega \to 0$. We argue in Section 4.1 how finite but still small values of $\Omega$ may lower the threshold for gravitational collapse.

### 1.1. Critical Density for Gravitational Instability in a Pure Gas Disk

The usual criterion for gravitational instability in a razor-thin pure gas disk is expressed in terms of the dimensionless parameter

$$Q_g = \frac{c_s \Omega}{\pi G \Sigma_g}$$

where $G$ is the gravitational constant, $c_s$ is the gas sound speed, and $\Sigma_g$ is the gas surface density (Goldreich & Lynden-Bell 1965; Toomre 1964; Toomre 1981). In Equation (2), the Kepler orbital frequency $\Omega$ has been substituted for the radial epicyclic frequency. If

$$Q_g < Q_g^* = 1,$$

then the disk is gravitationally unstable to axisymmetric perturbations in the disk plane. The $Q$-criterion is a measure of the competition between stabilizing pressure, stabilizing rotation, and de-stabilizing self-gravity (see, e.g., Binney &
When $Q_g > 1$, horizontal perturbations having length scales $< 2c_g / G \Sigma_g$ are stabilized by pressure, while those having length scales $> 2c_g / G \Sigma_g$ are stabilized by rotation. When $Q_g = 1$, the first axisymmetric mode to become unstable to self-gravity has radial wavelength $2c_g / G \Sigma_g$. And as $Q_g$ approaches 1 from above, the disk is increasingly susceptible to nonaxisymmetric perturbations which swing amplify stable to self-gravity with radial wavelength $2c_g / G \Sigma_g$ (Goldreich & Lynden-Bell 1965). The criterion $Q_g \lesssim Q^*_{\text{g}}$ for gravitational instability can be translated into a criterion for the midplane density $\rho_g$ (the subscript “0” denotes the initial midplane value). We define a disk half-thickness $H_g$ using

$$\Sigma_g \equiv 2 \rho_g H_g. \quad (4)$$

We also define a half-thickness $H^{\dagger}_g$ using the usual relation from vertical hydrostatic equilibrium:

$$H^{\dagger}_g \equiv \frac{c_g}{\Omega}. \quad (5)$$

Ordinarily, $H_g \approx H^{\dagger}_g$ and we would not bother to distinguish the two; however, we will later find cases where they differ by factors of several because of the effects of dust, and thus we take care to separate the two lengths now. Upon substitution of Equations (4) and (5), the relation $Q_g \lesssim Q^*_{\text{g}}$ is shown to be equivalent to

$$\rho_g \gtrsim \rho^*_g = \left( \frac{1}{2\pi} \frac{H^{\dagger}_g}{H_g} \right)^{\frac{1}{2}}, \quad (6)$$

where we have defined a reference density

$$\rho^*_g \equiv M_\ast/r^3 \quad (7)$$

with $M_\ast$ and $r$ equal to the mass of the central star and the disk radius, respectively.

The $\rho^*_g$-criterion (6) is sometimes used (e.g., Lee et al. 2010a, 2010b) to signal gravitational instability in dusty gas disks (with $\rho_g$ replaced by the total dust + gas density $\rho_0 + \rho_g$, $Q^*_{\text{g}} = 1$, and $H^{\dagger}_g / H_g = 1$). But using $\rho^*_g$ for dust–gas mixtures is suspect because the criterion does not account explicitly for the two-phase nature of such media. In the next subsection, we make such an accounting to derive a substantially different criterion for gravitational collapse.

1.2. Critical Density for Gravitational Instability in a Dust-rich Sublayer in the Limit $\Omega \rightarrow 0$

For disks of gas and dust, gravitational instability should still be determined by the $Q$-criterion, except there is now the possibility that disk self-gravity is dominated by dust in a vertically thin sublayer at the midplane:

$$Q_d \equiv \frac{c_d \Omega}{\pi G \Sigma_d} \lesssim Q^*_{\text{d}} \quad (8)$$

In using the dust surface density $\Sigma_d$ in Equation (8), we neglect the contribution of gas to the total surface density of the sublayer. Under typical circumstances, the error accrued is small.

In the limit $\Omega \rightarrow 0$, the dust–gas mixture represents a colloidal suspension. In this suspension, dust does not contribute to the pressure $P$—which is still provided entirely by gas—but instead adds to the inertia. In other words,

$$P = \rho_g c^2_g = (\rho_g + \rho_d) c^2_d \quad (9)$$

defined by $c_d$, the speed of sound in the suspension is

$$c_d = \frac{c_g}{\sqrt{1 + \mu}}, \quad (10)$$

where $\rho_g$ is the local gas density, $\rho_d$ is the local dust density, and $\mu = \rho_d / \rho_g$ is the dust-to-gas ratio. In effect, dust increases the mean molecular weight of the gas.

Inserting Equation (10) into Equation (8) and using

$$\rho_0 = \rho^{\dagger}_g \frac{1}{2\pi Q^*_{\text{g}} H^\dagger_{\text{g}}}, \quad (11)$$

we solve for the total midplane density required for gravitational instability:

$$\rho_0 = \rho_0 + \rho_0 \gtrsim \rho^*_\text{II}, \quad (12)$$

where

$$\rho^*_\text{II} = \frac{1}{2\pi} \frac{Q^*_{\text{d}}}{Q^*_{\text{g}}} \left( \frac{\Sigma_d}{\Sigma_g} \right)^2 \frac{H^{\dagger}_d}{H_g} \rho^*_g \quad (13)$$

$$\approx 2 \times 10^4 \rho^* \left( \frac{Q^*_{\text{g}}}{30} \right) \left( \frac{1}{Q^*_{\text{g}}} \right)^2 \left( \frac{0.015}{\Sigma_d/\Sigma_g} \right)^2 \left( \frac{H^{\dagger}_d/H_g}{1} \right). \quad (14)$$

In Equation (14), our normalizations for $Q^*_\text{g}$ and the bulk (height-integrated but local to $r$) metallicity $\Sigma_d/\Sigma_g$ derive from the MMSN at $r = 1$ AU (Chiang & Youdin 2010). For these parameter choices, the critical midplane density $\rho^*_\text{II}$ is an astonishing five orders of magnitude greater than $\rho^*_g$. It is possible that real disks have masses and bulk metallicities enhanced over the MMSN by factors of a few, in which case $\rho^*_\text{II}$ would be larger than $\rho^*_g$ by about three orders of magnitude.

1.3. Other Critical Densities

Another threshold density, already alluded to at the beginning of Section 1, is the Roche density:

$$\rho^\ast_{\text{Roche}} = 3.5 \frac{M_\ast}{r^3}. \quad (15)$$

The Roche density is the density required for a strengthless, incompressible, fluid body in hydrostatic equilibrium to resist tidal disruption when in synchronous orbit at distance $r$ about a star of mass $M_\ast$ (e.g., Chandrasekhar 1987).

Yet another candidate threshold was proposed by Sekiya (1983), who found that when the midplane density exceeds

$$\rho^\ast_{\text{Sekiya}} = 0.60 \frac{M_\ast}{r^3}, \quad (16)$$

the disk becomes susceptible to an unstable, incompressible, axisymmetric mode in which in-plane motions generate out-of-plane bulges (i.e., an annulus that contracts radially becomes thicker vertically, and vice versa). The nonlinear outcome of this instability is not known, but Sekiya (1983) speculated that the dust sublayer might eventually fragment on the scale of the wavelength of the overstable mode, and that dust particles

---

Footnote: In this paper, we will superscript critical threshold densities with *; and fiducial or reference quantities with †.
might sediment toward the centers of fragments to form the first-generation planetesimals.

Table 1 summarizes the four candidate threshold densities. For realistic parameters \((Q_g \sim 10–30; \Sigma_d/\Sigma_g \sim 0.015–0.15)\), the four densities obey

\[
\rho_1^* < \rho_{\text{Sekiya}} < \rho_{\text{Roche}} \ll \rho_{\text{III}}. 
\]

The smallest three densities in this hierarchy are fixed multiples of the reference density \(\rho_1^* = M_\ast/r^3\) (with coefficients \(1/2\pi \approx 0.16, 0.6,\) and 3.5, respectively). The last density \(\rho_{\text{III}}^*\) can, in principle, be arbitrarily larger than \(\rho_1^*\); for typical, astrophysically plausible parameters, it is two to four orders of magnitude larger.

Which of the four densities in Table 1 is the most accurate predictor of gravitational collapse? In the next two sections, we describe numerical simulations performed in the \(\Omega_{\text{stop}} \to 0\) limit that attempt to answer this question. We will find, unfortunately, that the numerical expense of simulating thin sublayers of dusty gas will force us into a parameter space where the difference between \(\rho_{\text{III}}^*\) and the other densities is not as large as it is in reality; we will have to make do with what we can.

2. METHODS

2.1. Code

We simulate hydrodynamic, self-gravitating, stratified flows in disks using Athena, configured for a shearing box, with no magnetic fields (Stone et al. 2008; Stone & Gardiner 2010). Dust is assumed to be perfectly aerodynamically coupled to gas so that they share the same velocity field \(\mathbf{v}\).

The equations solved are

\[
\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{v}) = 0, 
\]

\[
\frac{\partial \rho_g}{\partial t} + \nabla \cdot (\rho_g \mathbf{v}) = 0, 
\]

\[
\frac{\partial \rho \mathbf{v}}{\partial t} + \nabla \cdot (\rho \mathbf{v} \mathbf{v} + \mathbf{P}) = -\rho \nabla \Phi - 2\rho(\Omega \hat{z} \times \mathbf{v}) + 2q\rho \Omega^2 \hat{x} - \rho \Omega^2 \hat{z}. 
\]

\[
\nabla^2 \Phi = 4\pi G \rho, 
\]

where \(\rho = \rho_g + \rho_d\) is the total density of the dust–gas suspension, \(\mathbf{P} = \rho \mathbf{I}\) is a diagonal tensor with components \(P = \rho \sigma^2\) as defined in Equation (9) with constant \(\sigma^2\) (isothermal approximation), \(\Omega\) is the mean (constant) orbital frequency, \(\hat{x}\) points in the radial direction, \(\hat{z}\) points in the vertical direction, and \(\Phi\) is the self-gravitational potential of the dust–gas mixture. We choose the shear parameter \(q = 3/2\) for Keplerian flow.

2.1.1. Algorithms and Boundary Conditions

Athena 4.0 provides several schemes for time integration and spatial reconstruction, and for solving the Riemann problem. Having experimented with various options, we adopted the van Leer algorithm for our dimensionally unsplit integrator (van Leer 2006; Stone & Gardiner 2009); a piecewise linear spatial reconstruction in the primitive variables; and the HLLC (Harten-Lax-van Leer-Contact) Riemann solver. To account for disk self-gravity, we use the routines written by Koyama & Ostriker (2009) and Kim et al. (2011) which solve Poisson’s equation using fast Fourier transforms.

Boundary conditions for our hydrodynamic flow variables (including density and velocity, but not the self-gravitational potential) are shearing-periodic in radius (\(x\)) and periodic in azimuth (\(y\)). For vertical height (\(z\)), we experimented with both periodic and outflow boundary conditions, and chose periodic boundary conditions to ensure strict mass conservation. When outflow boundary conditions were employed, mass was lost from the boundaries at early times and complicated the interpretation of our results. We verified that our results are insensitive to box height for sufficiently tall boxes; see Section 3 for explicit tests.

The Poisson solver implements shear-periodic boundary conditions in \(x\), periodic boundary conditions in \(y\), and vacuum boundary conditions in \(z\) (Koyama & Ostriker 2009; Kim et al. 2011). In our simulations, self-gravity is dominated by dust, and our boxes are tall enough to contain the entire dust layer. Both vertical and radial stellar tidal gravity are included as source terms in the van Leer integrator.

We further augmented the code to include dust in the limit of zero stopping time. In this limit, dust shares the same velocity field as gas, and contributes only to the mass density. In our modified version of Athena, two continuity equations are solved: one for the entire mixture \((\rho = \rho_g + \rho_d\), see Equation (18)), and one for the gas \((\rho_g\), see Equation (19)). The dust density is given by the difference \((\rho - \rho_g)\). The remaining hydrodynamic equations govern the dust–gas mixture \((\rho_g)\), but with gas \((\rho_g)\) contributing solely to the pressure \(P\) (see Equation (20)). For simplicity, we adopt an isothermal equation of state so that \(P\) is related to \(\rho_g\) by Equation (9) for constant \(\sigma^2\). Isothermal flows are more prone to gravitational instability than adiabatic ones (Mamatsashvili & Rice 2010).

We also modified the HLLC Riemann solver to accommodate our dust–gas mixture. Changes include the following: (1) The speeds of the left, right, and contact waves are reduced by a factor \((1 + \mu)^{-1/2}\), where \(\mu \equiv \rho_d/\rho_g\) is the local dust-to-gas ratio, to account for the added inertia from dust (see Equation (10)). (2) The pressure in the contact region is replaced by an equivalent but numerically more accurate form based on Equation (10.76) in Toro (1999). (3) When calculating left/right momentum fluxes, we ensure that only gas contributes to the pressure by using \(\rho_g\) and not \(\rho\). (4) For the flux solver to predict the pressure and wave speeds, the left/right gas densities require specification. We therefore add a reconstruction process for the gas density which interpolates cell-centered values to cell boundaries to second-order accuracy.

Previous studies of dust in the perfectly coupled limit (Chiang 2008; Lee et al. 2010a, 2010b) also introduced a static background radial pressure gradient to mimic sub-Keplerian rotation of gas in a pressure-supported disk. We could also add the appropriate source term to the van Leer integrator. However, since our goal is to determine the minimum densities required for gravitational collapse and not to study vertical shear instabilities (i.e., the Kelvin–Helmholtz instability), we omit the background pressure gradient for simplicity.

In many of our simulations, the dust layer at the midplane collapses vertically because it is gravitationally unstable. Because of our boundary conditions, “fresh” gas from outside the simulation box cannot enter into the box, and thus in the event of gravitational collapse toward the midplane, the topmost and bottommost regions of our simulation domain become evacuated.
Low-density gas in those regions becomes increasingly easy to accelerate, and the code timestep shortens by orders of magnitude, effectively halting the simulation. The dramatic reduction in timestep is not a serious limitation, as it usually occurs after the collapsing dust has attained some saturated state (see Section 3.2.1). In any case, we are more interested in the onset of gravitational instability than its nonlinear development.

2.1.2. Code Tests

The following test problems helped to validate our code.

**Linear wave propagation.** We propagated a small-amplitude one-dimensional (1D) wave in a medium with a uniform background dust-to-gas ratio, with periodic boundary conditions, no background shear, and no gravity. We found the simulated wave speed matched the reduced sound speed calculated in Equation (10). We chose our box to be one wavelength long, so that after one wave period, the wave crossed the boundaries and returned to its original position. With $N = 128$ grid cells and an initial (fractional) wave amplitude $A = 10^{-4}$, we found the deviation $\delta g \equiv (1/N)\sum_{i=1}^{N}[q_i - q_i^0] \approx 2 \times 10^{-8}$, where $q \in \{\rho, \rho_d, \rho\}$ and $q^0$ represents the initial condition.

**Dust cloud advection.** We advected a Gaussian-shaped dust cloud in a 1D domain. The cloud occupied about half the size of the box and the code was run for one box-crossing time. With $N = 256$ grid cells, the root-mean-squared deviation in the shape of the cloud was <1%.

**Hydrostatic equilibrium of a stratified but non-self-gravitating dusty disk.** Omitting self-gravity but including stellar gravity (both radial and vertical), we set up 3D dust–gas mixtures in hydrostatic equilibrium. A variety of vertical profiles for the dust-to-gas ratio were tested, ranging from uniform to linear to more complicated functional forms. All equilibria were found to be stable against small perturbations, even for dust-to-gas ratios as large as several hundred.

**Gravitational instability of 3D pure gas disks.** We simulated isothermal, gravitationally unstable disks of pure gas in 3D. The gas was initialized in hydrostatic equilibrium (computed with vertical self-gravity), and box heights spanned approximately ±4 initial gas scale heights. We found that $Q_g = 1$ did not trigger gravitational instability, whereas $Q_g = 0.5$ did. Our results are consistent with those of Goldreich & Lynden-Bell (1965), who found analytically that $Q_g^* = 0.676$ for a finite-thickness isothermal gas disk.

2.2. Initial Conditions and Run Parameters

Initial conditions for our science simulations are of a dust–gas mixture with a pre-defined vertical profile for the dust-to-gas ratio $\mu(z) = \rho_d(z)/\rho_g(z)$. We choose the form

$$ \mu(z) \equiv \mu_0 \operatorname{sech}^2 \left( \frac{z}{\zeta_d} \right), $$

where $\mu_0$ is the midplane dust-to-gas ratio. The scale height $\zeta_d$ can be thought of as the half-thickness of the dust layer insofar as $\rho_g(z)$ is constant with $z$.

The isothermal dust–gas mixture is initialized in vertical hydrostatic equilibrium, including both stellar tidal gravity and disk self-gravity:

$$ \frac{\rho_g}{\rho_g + \rho_d} \frac{d \rho_g}{dz} = -\Omega^2 z - \frac{4\pi G}{\rho_g + \rho_d} \int_0^z (\rho_g + \rho_d) dz. $$

We solve numerically the differential form of Equation (23). Taking derivatives, we find

$$ \frac{d}{dz} \left[ (1 + \mu) \frac{d \ln \rho_g}{dz} \right] = -\frac{1}{H_g^2} + \frac{4\pi G}{c_s^2} \rho_g(1 + \mu), $$

where $H_g^2 \equiv c_s/\Omega$ is a fiducial (constant) gas scale height, not to be confused with any actual disk scale height. A nondimensional form of Equation (24) is given by

$$ \frac{d}{dz} \left[ (1 + \mu) \frac{d \ln \tilde{\rho}_g}{dz} \right] = -\frac{2}{H_g Q_g} \tilde{\rho}_g(1 + \mu), $$

where we have defined the dimensionless variables $\tilde{z} \equiv z/H_g$, $\tilde{\rho}_g \equiv \rho_g/\rho_{g0}$ (where $\rho_{g0}$ is the midplane gas density), and $H_g \equiv H_g/H_g^*$, with

$$ H_g \equiv \Sigma_g/(2\rho_{g0}). $$

Upon insertion of Equation (22), Equation (25) can be solved numerically for $\rho_g(\tilde{z})$. But the solution must satisfy the following two constraints:

$$ h_g = \int_0^\infty \tilde{\rho}_g(\tilde{z}) d\tilde{z} $$

by definition of $H_g$, and

$$ \frac{\Sigma_3}{\Sigma_\infty} = \frac{1}{h_g} \int_0^\infty \tilde{\rho}_g(\tilde{z}) \mu(\tilde{z}) d\tilde{z} $$

for a fixed height-integrated (i.e., bulk) metallicity $\Sigma_3/\Sigma_\infty$.

Our procedure is as follows. We freely specify $Q_g^*, \Sigma_3/\Sigma_\infty$, and $\mu_0$ as model input parameters. We then iteratively solve Equations (25), (27), and (28) for the three unknowns $\tilde{\rho}_g(\tilde{z})$, $h_g$, and $z_\Delta$. First, we guess $z_\Delta$ and $h_g$, and integrate Equation (25) to obtain $\tilde{\rho}_g(\tilde{z})$. If $\tilde{\rho}_g$ so calculated fails Equation (27), then we revise $h_g$ and reintegrate Equation (25), repeating until Equation (27) is satisfied. Next, we check Equation (28). If $\tilde{\rho}_g(\tilde{z})$ and $h_g$ fail Equation (28), then we revise $z_\Delta$ and repeat the procedure from the beginning, reintegrating Equation (25) to obtain $\tilde{\rho}_g$, reestablishing Equation (27), and so on. Typically, ~100 iterations (~10 for $z_\Delta \times ~10$ for $h_g$) are required before all constraints are satisfied to ~1% accuracy in $z_\Delta$ and $10^{-6}$ accuracy in $h_g$.

Table 3 lists the parameters of our models. Note that these parameters do not describe plausible protoplanetary gas disks; in particular, our model metallicities $\Sigma_3/\Sigma_\infty$ are orders of magnitude above the solar value of ~0.015. Parameters are instead chosen to yield disk flows that our code can adequately resolve while still testing Equation (14). Unfortunately, more astrophysically realistic parameters correspond to dust sublayers that are too vertically thin for us to resolve numerically; the code timestep, set by the sound-crossing time across a grid cell, becomes prohibitively short as thinner dust layers are considered. This difficulty means that the difference between $\rho_{g0}^*$ and the other candidate threshold densities is much less than what it is in reality, and our ability to distinguish between the candidates degrades as a result.

Figure 2 plots the initial conditions for our standard model $(S = \text{STD32})$, for which $Q_g = 24$, $\Sigma_3/\Sigma_\infty = 8$, and $\mu_0 = 35$. For this specific case, we calculate that $h_g = 0.20 c_s/\Omega$ and $z_\Delta = 0.083 c_s/\Omega$. The top and bottom boundaries of our simulation box are indicated by dotted vertical lines; typically,
of amplitude

initial equilibrium by adding random cell-to-cell fluctuations
top and bottom of our computational box.

agreement with the exact solution is good.
computed by our 3D Poisson solver (green dashed curve); the
constant at unity. We also overplot the self-gravitational force
ratio of pressure to gravity (gray dotted line) which is practically
pressure gradient (black solid curve). It does, as evidenced by the
(blue dashed curve) and disk self-gravity (red solid curve
hydrostatic equilibrium is satisfied. The sum of stellar gravity
half of the simulation box to demonstrate how well vertical
box heights span ±4zd (see Section 3 for box height tests).

The right-hand panel of Figure 2 compares gas density profiles
encompassed by the density profiles (dashed) and disk self-gravity (red solid curve)
computed with and without self-gravity, and with and without
dust, and shows that both the weight and self-gravity of the
embedded dust layer force gas into a similarly thin layer.

Figure 3 plots the initial force densities within the upper
half of the simulation box to demonstrate how well vertical
hydrostatic equilibrium is satisfied. The sum of stellar gravity
(blue dashed curve) and disk self-gravity (red solid curve)
computed via the integral in Equation (23) should equal the
pressure gradient (black solid curve). It does, as evidenced by the
ratio of pressure to gravity (gray dotted line) which is practically
constant at unity. We also overplot the self-gravitational force
computed by our 3D Poisson solver (green dashed curve); the
agreement with the exact solution is good.

Every simulation listed in Table 3 is perturbed from its
initial equilibrium by adding random cell-to-cell fluctuations
of amplitude ∼10−3cg to the velocity field. The typical duration
of a simulation is ∼20 Ω−1. Our rationales for box size and
resolution are explained in Section 3.

3. RESULTS

Results for two-dimensional (2D) shearing sheets are
described in Section 3.1, and those for 3D shearing boxes are
in Section 3.2.

3.1. 2D Shearing Sheet

For 2D dusty disks, the criterion for gravitational instability reads

\[ Q_d = \frac{c_d Ω}{π G (Σ_d + Σ_δ)} = \frac{Q_g}{(1 + μ_0)^{3/2}} < Q^*_d, \]  

(29)

We test this criterion by constructing a series of 2D shearing sheet simulations with various values of \( Q_g \) and \( μ_0 \), thereby seeing if we can converge on a unique value for \( Q^*_d \). Although total surface densities can change during the simulation, the dust-to-gas ratio stays fixed at its initial value because of our perfect-coupling approximation. Initial conditions are as follows: for a given domain size \( L_x \) and \( L_y \), the flow velocity \( v = -(3/2)Ω x \hat{e}_y \), and the surface density \( Σ = Σ_0 + δΣ \cos(k \cdot x) \), with \( Σ_0 = Σ_{g0} + Σ_{d0} = Σ_0(1 + μ_0) \), \( δΣ/Σ_0 = 0.01 \), \( k_x = -2(2π/L_x) \), and \( k_y = 2π/L_y \). In our 2D simulations, we choose \( c_g = Ω = Σ_0 = 1 \) as our units.

Table 2 lists the parameters for our 2D runs. Our standard 2D run, labeled S2D0, has \( Q_g = 12 \) and \( μ_0 = 8.0 \), and therefore \( Q_d = 0.44 \). For this run, the domain size is chosen large enough to easily fit the critical wavelength \( \lambda_c \) for gravitational instability: \( L_x = L_y = 10c_g/Ω \gtrsim 10\lambda_c \), where

\[ \lambda_c \equiv \frac{2c^2_g}{G Σ_0} \]  

(30)

is the wavelength of the fastest growing mode according to the
WKB dispersion relation for axisymmetric waves. It is also
the wavelength of the first mode to become unstable when
\( Q_d \) just crosses \( Q^*_d \). The resolution of the standard run is \( N_x \times N_y = 256 \times 256 \), so that one critical wavelength is resolved
across ∼10 grid cells.

For S2D0, we find that the disk is indeed gravitationally
unstable: density waves steepen quickly, and dense clumps of
dusty gas form before one orbital period elapses. A simple
way to portray instability is to track the maximum dust density
max Σd versus time—this is done in Figure 4, which shows that
the maximum dust density increases by two orders of magnitude
over a few dynamical times for our standard run.
The Astrophysical Journal

Here, we simplify take (A color version of this figure is available in the online journal.)

The instability is triggered appears to be between 0.5 and 1.0.

criterion (8) becomes

\[ Q \sim c \]

evaluate vertically stratified disk, there is some ambiguity as to how we

for gravitational instability in a 2D razor-thin sheet. For a 3D,

alter the evolution. Reducing the size of the box so that it can no

run (S2D1) enables higher maximum densities to be achieved

for domain size and resolution are tested in Section 3.2.2.

Our simulation box extends \( \pm 4z_d \) vertically, and \( 14z_d \) in either

directional. Each horizontal length is about twice the critical

wavelength (\( \lambda_c \approx 6.3z_d \)). The resolution is \( 32 \times 32 \times 32 \)

so that one horizontal critical wavelength spans \( \sim 16 \) cells, and

one vertical scale length \( z_d \) spans 4 cells. These choices for

domain size and resolution are tested in Section 3.2.2.

The simulation is terminated at \( \sim 10, 3\Omega^{-1} \), at which point the

timestep has become three orders of magnitude smaller than the

initial timestep (see the final paragraph of Section 2.1.1).

Figure 5 displays a time series of the volume-rendered dust density

in the bottom half of the box. Over the course of several
dynamical times, density waves shear and amplify, eventually

concentrating into a single azimuthally elongated filament. This

filament then fragments radially. The fragments gravitationally

scatter and merge; by the end of the simulation, two clumps

remain.

A simple diagnostic that we use throughout this paper is the

time evolution of the maximum dust density, shown in the left

panel of Figure 6. Comparison with Figure 5 reveals that max

\( \rho_d \) (or \( \mu \)) ceases to rise once the clumps finish

collecting. At this point, each clump is gravitationally bound,

with a maximum central density that depends on the simulation

initial conditions cannot be in perfect hydrostatic balance; however, the magnitude of the vertical motions is small and stays

An alternative is to calculate a vertically averaged, density-

weighted sound speed. We found, however, that such a procedure

made little practical difference, since dust densities are much

greater than gas densities near the midplane and drop steeply

with height.

3.2.1. Standard Run (\( S = \text{STD32} \))

To orient the reader, we present results for our standard 3D run

(\( S \), also labeled STD32 in Section 3.2.2), for which \( Q_d = 0.5 \).

The full set of model \( S \) parameters is listed in Table 3, and the

initial gas and dust density profiles are displayed in Figure 2.

Our simulation box extends \( \pm 4z_d \) vertically, and \( 14z_d \) in either

directional. Each horizontal length is about twice the critical

wavelength (\( \lambda_c \approx 6.3z_d \)). The resolution is \( 32 \times 32 \times 32 \)

so that one horizontal critical wavelength spans \( \sim 16 \) cells, and

one vertical scale length \( z_d \) spans 4 cells. These choices for

domain size and resolution are tested in Section 3.2.2.

The simulation is terminated at \( \sim 10, 3\Omega^{-1} \), at which point the

timestep has become three orders of magnitude smaller than the

initial timestep (see the final paragraph of Section 2.1.1).

Figure 5 displays a time series of the volume-rendered dust density

in the bottom half of the box. Over the course of several
dynamical times, density waves shear and amplify, eventually

concentrating into a single azimuthally elongated filament. This

filament then fragments radially. The fragments gravitationally

scatter and merge; by the end of the simulation, two clumps

remain.

A simple diagnostic that we use throughout this paper is the

time evolution of the maximum dust density, shown in the left

panel of Figure 6. Comparison with Figure 5 reveals that max

\( \rho_d \) (or \( \mu \)) ceases to rise once the clumps finish

collecting. At this point, each clump is gravitationally bound,

with a maximum central density that depends on the simulation

resolution (Section 3.2.2).

The right panel of Figure 6 shows the time evolution of various kinetic energy densities, evaluated in the three directions

and excluding the background Keplerian shear. The energy

densities are averaged horizontally and vertically over a thin slab

subtending two grid cells at the midplane (qualitatively similar

results are obtained over larger vertical averages). The horizontal

kinetic energies grow exponentially from \( t \approx 2–7 \Omega^{-1} \), with an

exponential growth rate of \( \approx 1.5 \Omega \). Radial motions dominate

azimuthal motions until the end of the simulation when they

become comparable. Vertical motions develop immediately

after the beginning of the simulation because our discretized

initial conditions cannot be in perfect hydrostatic balance; however, the magnitude of the vertical motions is small and stays

Also shown in Figure 4 are results for other runs. In S2D4,

S2D5, and S2D6, either \( Q_d \) or \( \mu \) is varied relative to our

standard run, so that \( Q_d \) varies from 0.5 to 2.0. For all of

these runs, the domain size is \( \sim 10 \lambda_c \) in each direction and

the resolution is \( \sim 10 \) cells per \( \lambda_c \), just as in the standard case.

Taken together, the results indicate that

\[ 0.5 < Q_{\text{d,3D}}^* < 1.0. \]  

(31)

Other runs explore the effects of varying resolution and domain size. Doubling both \( N_x \) and \( N_y \) relative to our standard

run (S2D1) enables higher maximum densities to be achieved

when the instability saturates, but otherwise does not seem to

alter the evolution. Reducing the size of the box so that it can no

longer accommodate even a single critical wavelength (S2D2,

S2D3) results in no instability, as expected (Gammie 2001; Johnson & Gammie 2003).

3.2. 3D Stratified Dusty Disks

Equation (8; equivalently Equation (29)) gives the criterion

for gravitational instability in a 2D razor-thin sheet. For a 3D,

vertically stratified disk, there is some ambiguity as to how we

evaluate \( c_d \) in Equation (8) because its value varies with height.

Here, we simply take \( c_d \) to be its value at the midplane, so that
criterion (8) becomes

\[ Q_d \approx Q_g \frac{1}{\Sigma_d/\Sigma_g} \frac{1}{(1 + \mu_0)^{1/2}} \lesssim Q_d^*. \]  

(32)
roughly constant for $t \lesssim 6 \Omega^{-1}$. For $t \gtrsim 6 \Omega^{-1}$, vertical motions amplify but for the most part remain smaller than horizontal motions.

The in-plane motions of the dusty clumps are illustrated in Figure 7 with a snapshot of the midplane slice of STD32 at $t = 9.6 \Omega^{-1}$. The dust clumps are seen spinning about their centers of mass as a consequence of angular momentum conservation.

### 3.2.2. Resolution and Box Size

Table 4 lists the parameters of experiments designed to test our choices for resolution, box size, and grid-cell aspect ratio.
Figure 6. Left: time evolution of maximum dust density for run STD32 (= S). Right: time evolution of kinetic energies averaged horizontally and vertically over a thin slab subtending two grid cells at the midplane (red = x-component of kinetic energy; blue = y; green = z; black = total).

Table 3

| Name  | $Q_\theta$ | $\mu_0$ | $\Sigma_d/\Sigma_0$ | $H_\theta$ ($c_s/\Omega$) | $z_d$ ($c_s/\Omega$) | $L_x \times L_y \times L_z$ ($c_s^3$) | Resolution | $\Omega t$ | $\rho_0$ ($\rho^*$) | $\rho_0^b$ ($\rho^*$) | $G^*$ |
|-------|------------|---------|----------------------|------------------------|----------------------|---------------------------------|------------|----------|-----------------|-----------------|-------|
| S     | 24         | 35.0    | 8.0                  | 0.20                   | 0.083                | 6.3 $\times$ 14 $\times$ 14 $\times$ 14 | 32 $\times$ 32 $\times$ 32 | 10.3     | 0.5             | 1.20            | 0.16             | 0.30 Y |
| R1    | 24         | 165.0   | 2.0                  | 0.91                   | 0.011                | 42.6 $\times$ 90 $\times$ 90 $\times$ 8 | 256 $\times$ 256 $\times$ 32 | 11       | 0.93            | 1.21            | 0.16             | 1.05 Y/N |
| R2    | 12         | 93.0    | 0.67                 | 1.04                   | 0.008                | 150.3 $\times$ 256 $\times$ 256 $\times$ 8 | 256 $\times$ 256 $\times$ 32 | 30       | 1.86            | 1.20            | 0.16             | 4.09 N  |
| R3    | 12         | 143.0   | 0.54                 | 1.07                   | 0.004                | 255.4 $\times$ 400 $\times$ 400 $\times$ 8 | 400 $\times$ 400 $\times$ 32 | 30       | 1.86            | 1.77            | 0.16             | 6.12 N  |
| R4    | 12         | 322.0   | 0.56                 | 1.07                   | 0.002                | 220.7 $\times$ 400 $\times$ 400 $\times$ 8 | 400 $\times$ 400 $\times$ 32 | 30       | 1.22            | 4.0             | 0.16             | 5.16 N  |
| R5    | 12         | 322.0   | 1.33                 | 0.87                   | 0.004                | 44.7 $\times$ 400 $\times$ 400 $\times$ 8 | 400 $\times$ 400 $\times$ 32 | 3.6      | 0.5             | 4.9             | 0.16             | 1.24 Y  |
| SR    | 24         | 35.0    | 8.0                  | 0.20                   | 0.083                | 6.3 $\times$ 400 $\times$ 400 $\times$ 8 | 400 $\times$ 400 $\times$ 32 | 10.0     | 0.5             | 1.20            | 0.16             | 0.30 Y  |
| Z     | 24         | 35.0    | 4.0                  | 0.52                   | 0.077                | 13.1 $\times$ 30 $\times$ 30 $\times$ 8 | 32 $\times$ 32 $\times$ 32 | 30       | 1.0             | 0.46            | 0.16             | 0.46 N  |
| Q     | 48         | 35.0    | 8.0                  | 0.35                   | 0.12                 | 8.6 $\times$ 20 $\times$ 20 $\times$ 8 | 32 $\times$ 32 $\times$ 32 | 30       | 1.0             | 0.34            | 0.16             | 0.34 N  |
| M     | 24         | 8.0     | 8.0                  | 0.24                   | 0.71                 | 3.2 $\times$ 8 $\times$ 8 $\times$ 8 | 32 $\times$ 32 $\times$ 32 | 30       | 1.0             | 0.25            | 0.16             | 0.25 N  |

Notes.

a Values are derived using $Q_z^* = 1$.

b Values are derived using $Q_{\perp} = 1$.

c GI = Gravitational Instability. Y means max $\rho_d$ increases by orders of magnitude over a few dynamical times, and N means it does not.

d See Figure 11.

Figure 7. Snapshot of the midplane for run S = STD32 at $t = 9.6\Omega^{-1}$. The largest in-plane velocity shown is 2.16 $c_s$.

(A color version of this figure is available in the online journal.)

Figure 8 shows how varying the resolution changes the evolution of our standard, gravitationally unstable run (STD32—also labeled S in Table 3). Again, we use the simple metric of max $\rho_d$ versus $t$. Broadly speaking, the runs STD16, STD32, and STD64 are all “acceptable” insofar as they all yield increases in max $\rho_d$ by orders of magnitude within several dynamical times ($t \ll 8\Omega^{-1}$). By contrast, the lowest resolution run, STD8, is unacceptable. Thus, the minimum acceptable resolution appears

Table 4

| Name        | $L_x \times L_y \times L_z$ ($c_s^3$) | Resolution | GI a | Duration ($\Omega^{-1}$) |
|-------------|-------------------------------------|------------|------|-------------------------|
| STD32(S)    | 14 $\times$ 14 $\times$ 8           | 32 $\times$ 32 $\times$ 32 | Y    | 9.8                     |
| STD8        | 14 $\times$ 14 $\times$ 8           | 8 $\times$ 8 $\times$ 8    | N    | 30.0                    |
| STD16       | 14 $\times$ 14 $\times$ 8           | 16 $\times$ 16 $\times$ 16 | Y    | 11.0                    |
| STD64       | 14 $\times$ 14 $\times$ 8           | 64 $\times$ 64 $\times$ 64 | Y    | 11.0                    |
| U32         | 14 $\times$ 14 $\times$ 8           | 56 $\times$ 56 $\times$ 32 | Y    | 10.0                    |
| LZ2         | 14 $\times$ 14 $\times$ 2           | 32 $\times$ 32 $\times$ 8  | N    | 30.0                    |
| LZ4         | 14 $\times$ 14 $\times$ 4           | 32 $\times$ 32 $\times$ 16 | Y    | 8.5                     |
| LZ6         | 14 $\times$ 14 $\times$ 6           | 32 $\times$ 32 $\times$ 24 | Y    | 8.0                     |
| LZ10        | 14 $\times$ 14 $\times$ 10          | 32 $\times$ 32 $\times$ 40 | Y    | 9.0                     |
| LZ14        | 14 $\times$ 14 $\times$ 14          | 32 $\times$ 32 $\times$ 56 | Y    | 10.5                    |
| LXY6        | 6 $\times$ 6 $\times$ 8             | 16 $\times$ 16 $\times$ 32 | N    | 30.0                    |
| LXY10       | 10 $\times$ 10 $\times$ 8           | 24 $\times$ 24 $\times$ 32 | Y    | 8.6                     |
| LXY20       | 20 $\times$ 20 $\times$ 8           | 48 $\times$ 48 $\times$ 32 | Y    | 8.7                     |

Notes. a GI = Gravitational Instability. Y means max $\rho_d$ increases by orders of magnitude over a few dynamical times, and N means it does not.
to be \( \sim 2 \) cells per scale length \( z_d \) in the vertical direction (see Nelson 2006, who found that a minimum of four smoothing lengths per scale height is required for smoothed particle hydrodynamics simulations), and \( \sim 8 \) cells per critical wavelength \( \lambda_c \) in the horizontal directions. Our standard choices for resolution—as well as the resolutions characterizing all our “science” runs, listed in Table 3 and discussed in Section 3.2.3—satisfy these minimum requirements by a safety factor of two.

Examining Figure 8 more critically, we see that the maximum value attained by \( \max \rho_d \) has not converged with resolution. Increasing the resolution enables us to resolve ever higher densities in the collapsing clumps. Another point of concern is the non-uniform aspect ratios of individual grid cells, which range from \( x:y:z \approx 2:2:1 \) to 4:4:1 over our set of science simulations (Table 3). The run U32 is characterized by perfectly cubical grid cells (1:1:1); the evolution is similar to STD32, but is characterized by an earlier onset of gravitational instability and stronger density fluctuations. This comparison suggests that our science runs with non-cubical grid cells are biased slightly against gravitational instability.

We next investigate how box size affects our results. For all box size experiments, the spatial resolution is kept at its standard value (32 grid cells per \( 14 z_d \) in either horizontal direction, and 4 grid cells per \( z_d \) in the vertical direction). Runs LZ2 through LZ14 vary box height \( L_z \), while keeping \( L_x \) and \( L_y \) fixed at their standard (STD32 = S) values. As Figure 9 reveals, box heights of \( 4-14 z_d \) yield comparable results, while a box height of \( 2 z_d \) is unacceptable. For the most part, increasing the box height seems to delay the onset of gravitational instability, with LZ4 being the exception to this rule.

Our 2D simulations indicated that \( L_x \) and \( L_y \) must be large enough to encompass at least one critical wavelength \( \lambda_c \). Our 3D simulations bear out the same requirement. Figure 9 shows that run LXY6, for which the box size is just under one critical wavelength, does not exhibit gravitational instability, unlike its bigger box counterparts.

To summarize our findings in this subsection: (1) The simulation box should be at least \( 4 z_d \) tall (\( 2 z_d \) above and below the midplane). (2) Each horizontal dimension must be longer than one critical wavelength \( \lambda_c \) as given by Equation (30). (3) Simulations require a vertical resolution of \( \geq 2 \) grid cells per scale length \( z_d \), and a horizontal resolution of \( \geq 8 \) grid cells per critical wavelength. (4) Individual grid cells that have increasingly non-uniform aspect ratios (squatter vertically than horizontally) tend to suppress gravitational instability, but the bias is minor and aspect ratios up to 4:4:1 appear acceptable. All of our science simulations (Table 3; Section 3.2.3) satisfy these requirements, in some cases by factors of two.

3.2.3. Criteria for Gravitational Collapse

Table 3 lists the simulations designed to test which of the various proposed criteria for gravitational instability is the best predictor of collapse. Figures 2 and 10 describe the initial dust and gas profiles, while Figure 11 displays the results using our simple diagnostic of \( \max \rho_d \) versus time.

First, consider runs S and R1–R5, and ask whether these runs favor \( \rho^* \) or \( \rho^*_II \) for the density required for gravitational collapse. Because dust is a major component of our disks, we do not expect \( \rho^* \)—which is strictly valid only for pure gas disks—to be a good predictor. Indeed, in all six of these runs, the midplane density \( \rho_0 \) exceeds \( \rho^*_I \), by factors of 7.5–30, yet only runs S and R5, and to a much lesser extent R1, exhibit collapse. All six runs indicate instead that \( \rho^*_II \)—equivalently, \( Q_d \)—is the better predictor, with the critical value

\[
0.5 < Q_d^* < 0.9.
\]  
(33)

There is some concern that the comparison between runs R2–R5 and run S may not be fair because runs R2–R5 have a factor of \( \sim 2 \) poorer spatial resolution in \( x \) and \( y \) compared to run S. This concern is allayed by run SR, which has the same physical parameters as S but is run with the box size and resolution of R3, and which turns out to behave qualitatively similarly to S (see Figure 11).

Our conclusion that \( \rho^*_II \) is relevant and that \( Q_d^* \) obeys Equation (33) is supported further by runs Z, Q, and M, each of which varies one of the three input parameters \( \Sigma_d/\Sigma_g \), \( Q_d \), and \( \mu_0 \).

Although runs R2–R4 do not exhibit the dramatic growth in \( \rho_d \) shown by runs S, SR, and R5—a result that we interpret to mean that \( \rho^*_II \) gives the correct criterion for gravitational collapse—runs R2–R4 do show some clumping. Figure 12 compares snapshots of runs R3 and SR (performed with the same box size and resolution), taken at the same time \( t = 10 \Omega^{-1} \).
Filaments do form in R3, although they are much weaker in density contrast compared to the filaments in SR. The mild growth shown in runs R2 and R3 might simply reflect the fact that their values for $Q_d = 1.86$ are still too close to $Q^*_d$ to suppress instability entirely. An alternative (and not mutually exclusive) possibility is that because $\rho_0 > \rho^*_{\text{Sekiya}} = 0.60\rho_0^i$ for runs R2–R4, the disk might be exhibiting the unstable (and formally incompressible) mode found by Sekiya (1983). Whatever the interpretation, the modest growth factors exhibited by R2–R4 seem unlikely to lead to planetesimal formation. In particular, the density concentrations in runs R2–R4 eventually disperse, unlike the density concentrations in runs S, SR, and
R5, for which \( \rho_0 > \rho^*_\text{Roche} \). What evidence we have suggests that Sekiya’s mode is not important for planetesimal formation, but higher resolution simulations that better separate \( \rho^*_{\text{Sekiya}} \) from \( \rho^*_\text{Roche} \) are needed for a more definitive assessment.

Finally, what about \( \rho^*_\text{Roche} \) versus \( \rho^*_\text{II} \)? Here, runs R4 and R5 are the most telling. Both runs are characterized by the largest midplane densities \( \rho_0 > \rho^*_\text{Roche} \), but only R5, for which \( \rho_0 > \rho^*_\text{II} \), undergoes gravitational collapse (see Figure 11).

Table 5 summarizes how the various candidate critical densities relate to one another and to the midplane density for our science simulations. From Table 5, \( \rho^*_\text{II} \) emerges as the best predictor of collapse.

4. SUMMARY AND DISCUSSION

Dust grains settle toward the midplanes of protoplanetary disks, forming a sublayer of solid particles sandwiched from above and below by gas. Whether this sublayer can become thin enough and dense enough to undergo gravitational instability and fragment into planetesimals is an outstanding question. We have found in this work that the density threshold for gravitational collapse can be extraordinarily high—much higher even than the Roche density \( \rho^*_\text{Roche} = 3.5 M_\ast / r^3 \), where \( M_\ast \) is the mass of the central star and \( r \) is the orbital radius. To trigger collapse in the limit that dust particles are small enough to be tightly coupled to gas, the density \( \rho_0 \) in the sublayer must be such that the Toomre stability parameter

\[
Q_d \approx \left( \frac{\rho^*_\text{II}}{\rho_0} \right)^{1/2} \lesssim 1 \tag{34}
\]

where

\[
\rho^*_\text{II} \approx \frac{M_\ast}{2 \pi r^3} \frac{Q_g}{Q^2_d} \left( \frac{\Sigma_g}{\Sigma_d} \right)^2. \tag{35}
\]

(For more precise relations, see Equations (8), (13), and (33).) Here, \( Q_g \) is the Toomre parameter for the ambient (and much thicker) gas disk, \( \Sigma_d/\Sigma_g \) is the ratio of surface densities of dust and gas (i.e., the height-integrated metallicity), and \( 0.5 < Q^2_d < 0.9 \) as measured from our simulations. For an astrophysiologically plausible disk having three times the mass of the minimum-mass solar nebula (\( Q_g \approx 10 \)) and a bulk metallicity enriched over solar by a factor of 4 (\( \Sigma_d/\Sigma_g \approx 0.06 \)), the critical density

\[
\rho^*_\text{II} \approx 1.3 \times 10^2 Q_d^{-2} \rho^*_\text{Roche}. \tag{36}
\]

Figure 13 portrays two sublayers—one for which \( \rho_0 = \rho^*_\text{Roche} \) and another, much thinner sublayer for which \( \rho_0 \approx \rho^*_\text{II} \approx 10^2 \rho^*_\text{Roche} \) \( (Q_d = 1) \). The results of our simulations, performed in the limit of perfect aerodynamic coupling between particles and gas, indicate that only the latter, much denser disk is on the verge of fragmenting.
Qualitatively, such extraordinary densities are required for gravitational instability because gas pressure renders the sublayer extremely stiff. Sound-crossing times for thin layers are easily shorter than free-fall times. We can examine the competition between stabilizing pressure, stabilizing rotation, and de-stabilizing self-gravity in both the horizontal (in-plane) and vertical directions. Horizontal stability is controlled by $Q_d$: when $Q_d > Q_d^* \sim 1$, all horizontal lengthscales $\lambda \gtrsim 2c_d^2/G\Sigma_d$ are stabilized by pressure, and all scales $\lambda \gtrsim 2c_d^2/G\Sigma_d$ are stabilized by rotation, where $c_d$ is the effective sound speed in the dust–gas mixture. At the same time, vertical stability is assured whenever the sound-crossing time across the vertical thickness of the sublayer $2H_d$ is shorter than the free-fall time:

$$\frac{2H_d}{c_d} < \frac{1}{\sqrt{G\rho_d}} \tag{37}$$

which, after substituting $H_d \approx \Sigma_d/2\rho_d$ and $c_d \approx c_g\sqrt{\rho_g/\rho_d}$, translates to

$$\left(\frac{\Sigma_d}{\Sigma_g}\right)^2 \frac{1}{Q_g} < \frac{\pi}{2} \tag{38}$$

which is easily satisfied for reasonable disk parameters.

The severe obstacle that gas pressure presents to gravitational collapse of aerodynamically well-coupled particles is discussed by Cuzzi et al. (2008, see their Section 3.1). Our 3D disk simulations support their 1D considerations.

### 4.1. Directions for Future Research

Taken at face value, the higher density threshold $\rho_d^*$ established by our work argues against using aerodynamically well-coupled particles to form planetesimals. The Kelvin–Helmholtz instability (KHI) may prevent dust from settling into the extraordinarily thin sublayers needed to cross the density threshold. One potential loophole is provided by Sekiya (1998) and Youdin & Shu (2002), who found in 1D that self-gravitating, non-rotating sublayers having a constant Richardson number $Ri$ could develop cusps of infinite density at the midplane. The presumption of these studies is that dust settles into a state that is marginally KH-stable and that this state is characterized by a constant $Ri$. Some evidence for a spatially constant $Ri$ was found in the settling experiments of Lee et al. (2010b), but only near the top and bottom faces of the dust sublayer and not at the midplane. These numerical experiments suffered, however, from lack of spatial resolution toward the midplane, and moreover neglected self-gravity. Future simulations of cuspy dust profiles including self-gravity would be welcome.

We have worked with the limit that the stopping times $t_{\text{stop}}$ of particles in gas are small compared to all other timescales. But in reality, finite particle sizes imply finite $t_{\text{stop}}$ (see Figure 1). When the assumption of infinitesimal stopping time breaks down, new effects may appear that might lower the threshold for gravitational instability.

One such effect is as follows. Consider again the competition between stabilizing pressure and de-stabilizing self-gravity (in either the vertical or horizontal directions). A major reason why the sublayer so strongly resists collapse is that sound waves travel quickly across it. We have taken the sound speed for our dust–gas suspension to be $c_d = c_g\sqrt{1 + \rho_d/\rho_g} \approx c_g\sqrt{\rho_g/\rho_d}$ (Equations (9) and (10)). But this presumes that particles are perfectly coupled to gas. If the sound-crossing time across some scale $\lambda$ were to become shorter than the particle stopping time, i.e., if

$$\frac{\lambda}{c_d} < \frac{\sqrt{\rho_d}}{c_g} \approx \frac{t_{\text{stop}}}{t_{\text{stop}}} \tag{39}$$

then our use of $c_d \approx c_g\sqrt{\rho_g/\rho_d}$ would be invalid. Particles on scales $\lambda$ would lose support from gas pressure and become susceptible to gravitational instability.

To get a sense of where in parameter space this instability may lie, we normalize $\lambda$ to the full vertical thickness of the sublayer:

$$\lambda = 2H_d\hat{\lambda} = \frac{\hat{\lambda}\Sigma_d}{\rho_d} \tag{40}$$

where $\hat{\lambda}$ can take any value (larger than or smaller than unity). Then, Equation (39) for the loss of pressure support translates to a midplane density (dominated by dust) of

$$\rho_d \approx \rho_d^* \gtrsim \frac{2}{\pi} \frac{M_s}{r^3} \left(\frac{\Sigma_d}{\Sigma_g}\right)^2 \frac{\hat{\lambda}^2}{Q_g} \frac{1}{(\Omega t_{\text{stop}})^2} \tag{41}$$

where $\Omega$ is the Kepler orbital frequency. For self-gravity to resist tidal disruption, $\rho_d = \rho_d^{*\text{Roche}} = 3.5M_s/r^3$. Substituting
this requirement into Equation (41), we find that
\[
\Omega_{\text{stop}} \gtrsim \left( \frac{2}{3.5\pi} \right)^{1/2} \left( \frac{\Sigma_d}{\Sigma_g} \right) \left( \frac{\lambda}{Q_g} \right)^{1/2} 
\]
\[
\gtrsim 8 \times 10^{-3} \left( \frac{\Sigma_d}{\Sigma_g} \right) \left( \frac{\lambda}{Q_g} \right)^{1/2} (42)
\]
for particles on scales $\lambda$ to decouple from sound waves. For $\lambda = 1$, requirement Equation (42) could be fulfilled by particles having sizes of a few millimeters to a few centimeters at distances of 1–10 AU (Figure 1—but note that the curves in the figure need to be adjusted by factors of a few for mass-enriched nebulae). For $\lambda < 1$, even smaller particles could lose pressure support and collapse gravitationally.

Future simulations that include finite particle stopping times could try to find such an instability. A complication would be that accounting for finite $t_{\text{stop}}$ would introduce the streaming instability, which could prevent the dust density from attaining the Roche value—see, e.g., runs R21-3D and R41-3D in Figure 5 of Bai & Stone (2010), for which $\Omega t_{\text{stop}} \leq 0.1$ and $\rho_d < \rho^*_{\text{Roche}}$. To find the instability that we are envisioning, one would have to restrict $\Omega t_{\text{stop}}$ to small enough values to suppress the streaming instability—thereby permitting the setting of grains into sublayers for which $\rho_d = \rho^*_{\text{Roche}}$—while at the same time keeping $\Omega t_{\text{stop}}$ large enough to satisfy Equation (42) and nullify pressure support.

We are grateful to Eve Ostriker and Andrew Youdin for discussions. Section 4.1 was inspired by discussions with Eve that clarified the limitations of our study and pointed the way to a possible new route to gravitational instability. We thank Xuening Bai, Chang-Goo Kim, Eve Ostriker, Ian Parrish, and Jim Stone for help in augmenting Athena. The simulations were performed with the Berkeley cluster Heneyy, which was made possible by a National Science Foundation Major Research Instrumentation (NSF MRI) grant. Financial support for the authors was provided by the Berkeley Center for Integrative Planetary Science, the Berkeley Theoretical Astrophysics Center, and grants from NSF (AST-0909210) and NASA Origins.

REFERENCES

Andrews, S. M., Wilner, D. J., Hughes, A. M., et al. 2012, ApJ, 744, 162

Bai, X.-N., & Stone, J. M. 2010, ApJ, 722, 1437

Binney, J., & Tremaine, S. 2008, Galactic Dynamics (2nd ed.; Princeton, NJ: Princeton Univ. Press)

Birnstiel, T., Ricci, L., Trotta, F., et al. 2010, A&A, 516, L14

Chandrasekhar, S. 1987, Ellipsoidal Figures of Equilibrium (1st ed.; New York: Dover)

Chiang, E. 2008, ApJ, 675, 1549

Chiang, E., & Youdin, A. 2010, AREPS, 38, 493

Curzi, J. N., Hogan, R. C., & Shariff, K. 2008, ApJ, 687, 1432

Gammie, C. F. 2001, ApJ, 553, 174

Goldreich, P., & Lynden-Bell, D. 1965, MNRAS, 130, 125

Johansen, A., Oishi, J. S., Low, M., et al. 2007, Natur, 448, 1022

Johnson, B. M., & Gammie, C. F. 2003, ApJ, 597, 131

Kim, C.-G., Kim, W.-T., & Ostriker, E. C. 2011, ApJ, 743, 25

Koyama, H., & Ostriker, E. C. 2009, ApJ, 693, 1316

Lee, A. T., Chiang, E., Asay-Davis, X., & Barranco, J. 2010a, ApJ, 718, 1367

Lee, A. T., Chiang, E., Asay-Davis, X., & Barranco, J. 2010b, ApJ, 725, 1938

Mamatsashvili, G. R., & Rice, W. K. M. 2010, MNRAS, 406, 2050

Nelson, A. F. 2006, MNRAS, 373, 1039

Okuzumi, S., Tanaka, H., & Sakagami, M.-a. 2009, ApJ, 707, 1247

Sekiya, M. 1998, Icar, 133, 298

Stone, J. M., & Gardiner, T. 2009, NewA, 14, 139

Stone, J. M., & Gardiner, T. A. 2010, ApJS, 189, 142

Stone, J. M., Gardiner, T. A., Teuben, P., Hawley, J. F., & Simon, J. B. 2008, ApJS, 178, 137

Toomre, A. 1964, ApJ, 139, 1217

Toomre, A. 1981, in Structure and Evolution of Normal Galaxies, ed. S. M. Fall & D. Lynden-Bell (Cambridge: Cambridge Univ. Press), 111

Toro, E. 1999, Riemann Solvers and Numerical Methods for Fluid Dynamics (Berlin: Springer)

van Leer, B. 2006, CCoPh, 1, 192

Weidenschilling, S. J. 1977, MNRAS, 180, 57

Weidenschilling, S. J. 1980, Icar, 44, 172

Weidenschilling, S. J. 2006, Icar, 181, 572

Weidenschilling, S. J. 2010, M&PS, 45, 276

Youdin, A. N., & Goodman, J. 2005, ApJ, 620, 459

Youdin, A. N., & Shu, F. H. 2002, ApJ, 580, 494

Zsom, A., Ormel, C. W., Dullemond, C. P., & Henning, T. 2011, A&A, 534, A73