On the Origin of Banded Structure in Dusty Protoplanetary Disks: HL Tau and TW Hya

A. C. Boley

Department of Physics and Astronomy, University of British Columbia, Vancouver BC, Canada

Received 2017 July 18; revised 2017 October 5; accepted 2017 October 12; published 2017 November 21

Abstract

Recent observations of HL Tau revealed remarkably detailed structure within the system’s circumstellar disk. A range of hypotheses have been proposed to explain the morphology, including, e.g., planet–disk interactions, condensation fronts, and secular gravitational instabilities. While embedded planets seem to be able to explain some of the major structure in the disk through interactions with gas and dust, the substructures, such as low-contrast rings and bands, are not so easily reproduced. Here, we show that dynamical interactions between three planets (only two of which are modeled) and an initial population of large planetesimals can potentially explain both the major and minor banded features within the system. In this context, the small grains, which are coupled to the gas and reveal the disk morphology, are produced by the collisional evolution of the newly formed planetesimals, which are ubiquitous in the system and are decoupled from the gas.

Key words: minor planets, asteroids: general – planet–disk interactions – protoplanetary disks

1. Introduction

The HL Tau system offers a snapshot of the planet formation process. This pre-main-sequence star is surrounded by a gaseous and dusty disk that has been observed by ALMA at unprecedented sensitivity and resolution (ALMA Partnership et al. 2015). The millimeter/submillimeter observations reveal concentric dark and bright rings that appear to be consistent, at least qualitatively, with planets clearing out material, directly reminiscent of planet formation simulations. Recent calculations and simulations have confirmed that the ring structures can indeed be matched to various degrees by embedded planets (e.g., D'pierro et al. 2015).

While the planet hypothesis is compelling and is the focus of this work, the morphology of HL Tau is not generally accepted to be the result of planets. Because the HL Tau system is thought to be < 1 Myr old, the existence of planets in the disk would challenge planet formation theory, as the age of the disk is seemingly too young to have formed planets through core nucleated instability (e.g., Pollack et al. 1996) and the azimuthal symmetry appears to be inconsistent with direct formation in a fragmenting disk (e.g., Durisen et al. 2007; Helled et al. 2014). Core nucleated instability is especially difficult at large stellocentric locations (Dodson-Robinson et al. 2009), such as those exhibited by the structure in HL Tau’s disk. These reasons, at least in part, have motivated proposals for alternative mechanisms to produce the global morphology. This includes, for example, pebble formation at the locations of condensation fronts of astrophysical ices (Zhang et al. 2015) and secular gravitational instabilities (Takahashi & Inutsuka 2016). Nevertheless, we are strongly motivated by the apparent resonant structure among the rings and gaps in the disk (ALMA Partnership et al. 2015) and further argue that the observed morphology has a dynamical, planetary origin.

Figure 1 shows the deprojected continuum image with annotations to emphasize the detailed disk structure (Figures 3(a) and (b) from ALMA Partnership et al. 2015). Dark bands (simply bands hereafter) are labeled with a “D” and are numbered sequentially by stellocentric distance, while the bright rings (rings hereafter) are denoted similarly, but with a “B.” As pointed out by the ALMA Partnership et al. (2015), multiple rings and bands appear to be near commensurabilities. While such commensurabilities should be treated with caution (Tamayo et al. 2015), they potentially provide a wealth of constraints for testing hypotheses on the disk structure origins. For example, let D1 and D2 be the locations of planets. Furthermore, let an additional planet be located at B5. The motivation for the latter assertion is that D5, B5, and D6, when taken together, appear to be a horseshoe structure, as is typical of dynamical simulations with planetesimals. For a discussion on this and other plausible arrangements, see Tamayo et al. (2015). In this proposed configuration, B5:D2 are very close to a 3:1 commensurability. D2:D1 are somewhat close to a 4:1, but are notably inside the commensurability.

Moving forward, we envisage that HL Tau does indeed have planets located at D1, D2, and B5. However, because the innermost regions are poorly resolved, we restrict our discussion to the structure exterior to B1. This restriction will not change the overall effects, and will simplify the simulations discussed in Section 2. For this discussion, we will refer to each planet as PD2 and PB5, respectively.

Table 1 shows the potential locations for mean motion resonances (MMR) with the hypothesized planets PD2 and PB5. All band and ring stellocentric distances are taken from the ALMA Partnership et al. (2015). Bands D3 and D4 correspond to a 3:2 and 2:1 with PD2. The bands D3 and D4 also approximately line up with the 2:1 and 3:2 interior to PB5, while band D7 corresponds to the exterior 3:2 with PB5.

It is well-established that planet–disk interactions can lead to observable features in disks (Bryden et al. 1999), although observing the resulting structure can be nontrivial due, in part, to the range of coupling between gas and dust grains (e.g., Birnstiel et al. 2015). The coupling itself depends on the grain sizes as well as the local gas conditions. Roughly, small grains remain coupled to the gas, large planetesimals are decoupled, and sizes within the millimeter to kilometer range are expected to migrate radially due to gas drag (e.g., Adachi et al. 1976;
Weidenschilling 1977; Haghighipour & Boss 2003), although the relevant size range that undergoes significant migration depends on the details of the disk structure. As a result, the continuum structure revealed by dust may not represent the actual gas distribution. This was highlighted by, e.g., Dipierro et al. (2015), who ran a series of simulations with different degrees of gas-dust coupling using a range of dust sizes.

Massive planets can open a gap in the gaseous disk whenever waves launched by the planet become nonlinear and dissipate angular momentum locally. The degree to which gas can be depleted around a planet depends on local disk parameters such as the viscosity and vertical scale height, as well as the planet’s mass and the duration of the interactions (Kley & Nelson 2012). This process will also open a gap in the dust disk for grains \( \lesssim \text{mm} \), depending on disk location. For the HL Tau system, a number of studies suggest that such gaps would be opened and potentially consistent with the major bands in HL Tau’s disk for planet masses in the range of approximately 0.1 to 1 \( M_{\text{Jup}} \) (Dipierro et al. 2015; Akiyama et al. 2016; Jin et al. 2016; Kanagawa et al. 2016).

While such gap-opening effects would be able to explain D1 and D2, as well as the horseshoe region D5-B5-D6 (depending on the grain size distribution), planets at these locations cannot obviously explain the additional band features such as D3, D4, and D7 through, e.g., resonances if only small grains are present in the gaseous disk. Gas drag effects on millimeter grains are too efficient, and the apparent resonant structure would not manifest. Therefore, if we consider only the dynamics of gas and millimeter grains, then additional planets must be located at the other bands (e.g., Simbulan et al. 2017), or a nonplanetary origin is required for some of the structure. On the other hand, we show that the entire detailed structure of HL Tau can be explained by assuming that (1) there are three embedded planets (we focus on only two), that (2) planetesimal formation was, for whatever reason, ubiquitous throughout the disk, and that (3) the millimeter grains observed by ALMA are direct tracers of the dynamically cold planetesimal population. We will provide physical context for the latter as we present the results. We further extend the analysis to TW Hya, in which similar substructure should be seen in more sensitive observations of the disk if the paradigm presented here is correct.

2. Methodology

We seek to explain the detailed ring/band structure with a flexible, simple process that could occur in essentially any planet-forming environment. We focus the discussion on HL Tau, but will return to TW Hya in Section 4.3. We further begin by only considering planets on initially circular orbits, although we will relax this restriction in later sections. If a planet is located at B5, as we posit, then its mass needs to be such that a large horseshoe region is produced spanning from D5 to D6. Excluding tadpole and horseshoe orbits, as well as gas drag effects, we know that a planetesimal will become significantly perturbed if it is located a distance \( \Delta < 2.4\alpha_p^{1/3} \) from the planet (Gladman 1993), where \( \mu = M_p/M_\star \) (planet to star mass) and \( \alpha_p \) is the planet’s semimajor axis. Throughout these calculations, we assume that the central mass for the HL Tau system is 1.3 \( M_\odot \). We set \( \Delta = 8 \) au for PB5 at 68.8 au based on the surface brightness profile (ALMA Partnership et al. 2015), which yields a mass \( M_{\text{PB5}} = 1.5 \times 10^{-4} M_\odot \) (\( \sim 50 M_{\text{Jup}} \)). For the inner planet at 32.3 au, using the D2 gap profile alone is not a strong constraint on D2’s true width due to the narrow angular size of the gap and the size of the beam.

Table 1

| Band/Ring | SMA (AU) | 4:3 | 3:2 | 2:1 |
|-----------|---------|-----|-----|-----|
| (P)D2     | 32.3 ± 0.1 | 39.1 | 42.3 | 51.3 |
| D3        | ∼42     | ... | ... | ... |
| D4        | ∼50     | ... | ... | ... |
| D5        | 64.2 ± 0.1 | ... | ... | ... |
| (P)B5     | 68.8 ± 0.1 | 56.8 | 52.5 | 43.3 |
| (P)B5+    | ...     | 83.3 | 90.1 | 109.2 |
| D6        | 73.7 ± 0.1 | ... | ... | ... |
| D7        | ∼91     | ... | ... | ... |

Note. Bright rings are listed with a “B” and dark bands are denoted with a “D.” Rows with a “(P)” show the locations of proposed planets. The greater and less than symbols for PB5 are used to highlight commensurabilities interior and exterior to the proposed planet. SMA is the semimajor axis.
Instead, we set the inner planet to be approximately a Neptune mass such that $M_{\text{Neptune}} = 5.2 \times 10^{-5} M_{\odot}$ ($\sim 17 M_{\oplus}$). This yields a gap size that is reasonably consistent with the width of D2. HL Tau does have significant gas present. As such, the planet masses should be expected to grow as the system evolves. Furthermore, because the planets are already well above $10 M_{\oplus}$, we envisage that growth is currently limited by gas flow into the Hill region rather than the lack of sufficient gas at the location of either planet.

We use Mercury6 (Chambers 1999) with the hybrid integrator to evolve a proposed HL Tau analog consisting of two planets and 100,000 test particles. Interactions between the planets are included. The test particles are placed uniformly in the semimajor axis between 20 and 100 au, which gives an equivalent surface density profile $\Sigma \propto r^{-1}$, assuming identical particles. The eccentricities of the test particles are set to zero, but the inclinations are drawn from a uniform random distribution between $0^\circ$ and $1^\circ$. The planets are also given $1^\circ$ inclinations, and the nodes of the planets and test particles are randomized between 0 and $2\pi$. Including inclinations was done to prevent the simulation from being strictly 2D.

The orbital time step is set to 140 days, which is used by the hybrid integrator in MVS mode. The accuracy parameter is set to $1 \times 10^{-14}$ whenever the hybrid integrator switches to the Bulirsch–Stoer method. The simulation is evolved for 100,000 years. While HL Tau could be as old as 1 Myr, the chosen integration time is long enough to allow significant resonant structure in the disk to form. It further highlights that the mechanism proposed here could, in principle, develop rapidly. Each planetesimal is envisaged to be large enough such that gas drag effects are negligible, at least relative to orbital excitation through interactions with planets. This is qualitatively consistent with results from planetesimal bow shock studies (Hood & Weidenschilling 2012; Morris et al. 2012).

We refer to the above setup as the base HL Tau, base simulation. For additional simulations, which are used to explore the effects of planet eccentricity and to explore potential signatures in TW Hya, we use the same methodology and summarize the corresponding initial conditions and other changes before discussing the results.

To justify further the neglect of gas damping for the purposes of these simulations, consider the gas drag force felt by the planetesimal $F_D = -\frac{1}{2} C_D \rho \pi R^2 \nu_{\text{wind}}^2$. Here, the local gas density is $\rho$ and the planetesimal size is $R$. The gas-planetesimal relative wind speed is approximated in this case by $\nu_{\text{wind}} \sim e \nu_{\text{circ}}$ for eccentricity $e$ and Keplerian circular speed $\nu_{\text{circ}}$. We take the coefficient of drag $C_D \approx 1$ and further let $F_D = m \nu_{\text{wind}} / \nu_{\text{circ}}$ for drag time $t_D$ and planetesimal mass $m$. With these assumptions, we can write the ratio of the drag time to the local orbital period as $\frac{T}{t_D} = \frac{4}{3 \nu_{\text{circ}}} \frac{\rho}{\rho_p} \frac{R}{r}$ at a given orbital distance $r$. Setting $R = 1$ km, $\rho_p = 1$ g cm$^{-3}$ (planetesimal internal density), $r = 40$ au, $e = 0.05$, and $\rho = 5 \times 10^{-13}$ g cm$^{-3}$, we find that $\frac{T}{t_D} \approx 3000$. The gas density for the disk is based on the best-fit model from Kwon et al. (2011). These values are chosen to emphasize the effects of eccentricity damping on a small planetesimal with modest eccentricity. The distance of 40 au is used to represent roughly the location of B2. The resonant forcing timescale is approximately $\sim 100$ orbits, which is described more below. Thus, we expect dynamical heating to be effective at the locations of MMR for planetesimals that are a kilometer in size or larger, while planetesimals that reside away from regions of orbital forcing will remain on cold orbits.

Radial drift can impact large planetesimals as well as small grains. Because the drag time is much larger than the local orbital period for kilometric planetesimals, we can write that the change of the planetesimal’s orbital angular momentum is $\dot{L} = -\frac{5}{3} C_D \rho \pi R^2 \nu_{\text{wind}}^2$, assuming a circular, planar orbit. In this case, the relative wind speed $\nu_{\text{wind}}$ is the difference between the local gas orbital speed and the planetesimal’s orbital speed, which is due to the gas pressure gradient (Whipple 1973; Adachi et al. 1976; Weidenschilling 1977). The detailed value of $\nu_{\text{wind}}$ depends on the disk model, but a reasonable estimate is $\sim 5000$ cm s$^{-1}$. Taking the orbital angular momentum of the planetesimal to be $L = \frac{4}{3} \pi \rho_p R^3 \nu_{\text{wind}} r^2$ and assuming the size of the planetesimal does not change, the angular momentum can be time differentiated and set equal to the drag torque to find $\dot{r} = -\frac{3}{8 \pi} C_D \rho \nu_{\text{wind}} r^2$. Using the same conditions as above, the radial drift speed is $\dot{r} \sim 0.2$ au Myr$^{-1}$. The drift of kilometer-sized planetesimals or larger will thus be small overall. This is nonetheless dependent on the location and the conditions in the disk. During the integration time of the simulations, radial drift could cause some of the kilometer-sized planetesimals to drift into a resonance, although this effect would not be captured by the given simulations. On the other hand, planetesimals with larger eccentricities, such as those in MMR, will experience much larger relative wind speeds and could drift out of the resonant locations.

The behavior of small grains will be different. The drag time for these particles is given by $t_D \approx \rho_k s^2 / (\rho v_h)$, where $\rho_k$ is the grain’s density, $s$ is the grain radius, $\rho$ is the gas density, and $v_h$ is the mean thermal speed of the gas. If $s \sim 1$ mm, $\rho_k \sim 1$ g cm$^{-3}$, and the gas temperature is $\sim 55$ K, then again for the conditions at $r = 40$ au, $t_D \sim 0.09$ year. As such, these small grains will not respond to resonant eccentricity forcing. If the grains are suddenly released from an eccentric planetesimal, then the small grains will quickly acquire the local orbital conditions. Moreover, the difference between the gas’s orbital motion and the circular, Keplerian speed will result in a residual radial gravity term $\Delta g$ felt by the grain (Weidenschilling 1977). At 40 au $\Delta g \approx 9 \times 10^{-6}$ cm s$^{-2}$ for the assumed wind speed $\nu_{\text{wind}} = 5000$ cm s$^{-1}$. The resulting radial drift of the millimeter grains is $\dot{r} \sim \Delta g t_D \sim 50$ au Myr$^{-1}$. As such, while millimeter grains will have their eccentricities quickly damped, they will undergo substantial radial migration unless they are reaccreted onto larger planetesimals or are destroyed.

### 3. Results

#### 3.1. HL Tau, Base Simulation

The eccentricity evolution of the planetesimals in the HL Tau, base simulation is shown in Figure 2. After a few 10,000 yr, the main resonant structure is established. We do not expect the morphology to undergo significant changes for integration times longer than 100,000 yr, at least for the current $n$-body model. Moving forward, we will take 100,000 yr to be a representative snapshot for the proposed system.

Figure 3 (left) shows the radial locations of planetesimals after 100,000 yr of evolution. Binning is done according to the physical location of the planetesimal, not its semimajor axis. Horseshoe gap structures are clearly produced at the location of each planet. In addition, MMR between the planetesimals and
the planets deplete the number of planetesimals that are located near bands in HL Tau, assuming the mm grains that are observed by ALMA follow the planetesimals. However, the agreement is insufficient. The gaps are too small to be reliably imaged, and the structure as a whole is not a strong match to the observations. This is emphasized in the right panel of the figure, which shows the face-on X–Y positions of the planetesimals. The gaps are present, but are small and are azimuthally asymmetric due to variations in the orbital orientations of the high-eccentricity planetesimals. Even the horseshoe regions are not so clearly delineated in the face-on image.

We now proceed under the following framework: the small grains observed by ALMA are debris, produced by the collisions among the already-formed planetesimals. If this is the case, we might expect that the hot population of planetesimals is unlikely to collide with itself, and if a hot planetesimal does collide with an object, it does so with the cold population due to the higher density (e.g., Wyatt et al. 2010; Lawler et al. 2015). The small grain radial drift time through the cold population of planetesimals is assumed to be long compared to the reaccretion timescale of those grains onto different planetesimals due to, e.g., pebble accretion (Johansen et al. 2015). Under these assumptions (discussed in more detail below), small grains will be more strongly associated with the dynamically cold planetesimals, which will serve as a spatial filter. Throughout the analysis here, we use an eccentricity cutoff of 0.005 for the hot and cold populations. Except in regions close to the planets, planetesimals with eccentricities larger than this value are associated with MMR.

The consequences of our assumptions are highlighted in Figure 4. The ring and bands exhibit strong overlap with the morphological features in the ALMA image, although the observability of the simulated disk will ultimately depend on the detailed collisional model and subsequent dust evolution (under the debris hypothesis). We note that D2 displays an additional horseshoe region. If the planets are on circular orbits, as envisaged here, this type of structure must be a result. Because this is not seen in the current ALMA continuum images, we offer two potential explanations. (1) The horseshoe gap is present for D2, but the current image is still too low resolution/sensitivity to show this structure. If this is the case, should more sensitive and higher resolution images become available, we can test for the presence of a horseshoe gap similar to that potentially seen with the D5-B5-D6 morphology. (2) The planet PD2 has a lower mass than what is used in the base simulation (Neptune-mass) and it is on an eccentric...
orbit. We show one such possible configuration in the next subsection.

There are two additional issues. First, the location of the B6 ring is potentially at too large of a radius. The face-on image shows that planetesimals populate the necessary region, but the peak of the distribution is offset due, in part, to the use of an eccentricity cut. We interpret this to suggest that grain physics cannot be ignored completely and that a detailed collisional model is necessary to determine the exact locations of features.

The second issue is that the ALMA Partnership et al. (2015) found that the rings in the deprojected ALMA image appear to have a low \( e = 0.033 \), but non-negligible eccentricity (based on ring offsets from the 1 mm peak emission). If this is not an image artifact, then secondary interactions between the gas and the small grains (e.g., see Jin et al. 2016) would be needed to produce such structure, as an eccentric planet mainly makes a planetesimal gap wider rather than eccentric in the proposed context.

Even with these outstanding issues, which ultimately need to be addressed, the structure as a whole is interpreted to have high fidelity with the morphology of the ALMA image.

3.2. HL Tau, Eccentric Case

A planet on an eccentric orbit has several consequences under the proposed paradigm. First, the gap width would be controlled by the mass of the planet and by the planet’s orbital eccentricity (e.g., Deck et al. 2013). We illustrate this case by placing PD2 at a semimajor axis of 33.076 au, which is at the 3:1 commensurability with PB5. The mass of PD2 is reduced to 1 \( M⊕ \), and the orbital eccentricity is increased from circular to \( e = 0.005 \). The semimajor axis and mass of PB5 remain unchanged, but its eccentricity is increased to \( e = 0.001 \). The simulation setup is otherwise unchanged from the circular initial conditions.

Without any eccentricity, we would expect the resulting D2 to exhibit a width of approximately 2 au and to have a well-defined horseshoe structure. With the given eccentricity, PD2 opens D2 to a width of about 3–4 au. The horseshoe structure, while still present, is greatly reduced and could be missed entirely by observations. The full structure with the eccentricity cut is shown in Figure 5.

The circular and eccentric cases highlight the variety of conditions that could produce structure within disks, including morphologies with and without horseshoe gaps. It does not, however, give rise to limitless possibilities. Substructure must still be reconciled with the locations of resonances. Moreover, to be observable, the resonant gaps must have sufficient width. This may only be possible upon reaching a certain mass range for the planets.

4. Discussion

The agreement between the simulation results and the locations of the rings and bands suggest that the structure is consistent with planets located at D2 and B5. The model used here posits that the millimeter grains are created by planetesimals through collisions, as small grains alone would be too coupled to the gas to be affected by the dynamics of the planets, at least in a way that is consistent with the full ring and band structure. Once produced in the disk, the small grains trace the dynamically cold planetesimal population under the assumption that reaccretion of pebbles is faster than radial drift. Such a model has multiple consequences.

1. Planetesimal formation is rapid, and occurs throughout the entire disk. If correct, planetesimals do not, in general, form at special locations in the disk such as at dead zone boundaries. We do not address the formation mechanism...
Figure 4. Same as Figure 3, but using an eccentricity cutoff of $e = 0.005$, for which planetsimals with eccentricities in excess of this limit are excluded in the histogram (left panel) and the face-on positions (right panel). The locations between the rings and bands are consistent with the actual morphology in Figure 1.

4.1. Assumptions Regarding Planetesimals and Small Grains

One of the main assumptions used in this work is that the small grains trace the dynamically cold planetesimal population. This is envisaged to be due to (1) planetesimal collisions being more likely within the cold population, along with (2) recycling of those grains through reaccretion. We explore the collisional assumption in two ways, which yield comparable results.

First, let the collision rate for one planetesimal to hit another be approximated by (e.g., Safronov 1972)

$$\dot{N} = \frac{3\pi R_e^2 \Sigma_d F_G}{2 \rho_p R^3 T}, \quad (1)$$

for collision radius $R_e$, planetesimal density $\rho_p$, planetesimal size $R$, and orbital period $T$. This rate assumes that each planetesimal effectively encounters the local disk surface density of solids $\Sigma_d$ twice per orbit. The gravitational focusing parameter $F_G$ depends on the properties of the planetesimals and the speed $V_{rel}$ of a given planetesimal relative to the local population. We assume identical planetesimals, which sets $R_e = 2R$:

$$\dot{N} = \frac{6\pi \Sigma_d F_G}{\rho_p RT}, \quad (2)$$

and

$$F_G = 1 + \frac{8\pi G \rho_p R^2}{3 V_{rel}^2}. \quad (3)$$

We take the final state of the HL Tau, base simulation, and bin the particles by their semimajor axes. For our analysis, we use 0.1 au bin widths. Using smaller bins increases the noise in the calculations, while larger bin widths wash out structure. In each radial bin, we calculate a planetesimal velocity dispersion. This dispersion is determined by first finding the speed of each particle relative to a circular, planar orbit at the instantaneous cylindrical distance from the star. These relative speeds for each particle are then used in the calculation of the velocity dispersion for the given particle’s semimajor axis bin. We

here, but do argue that the mechanism must be rapid and global in extent.

2. Small grains, including millimeter sizes, do not in general reveal grain growth. HL Tau is believed to be less than 1 Myr old. If the morphology of the disk is mainly due to planetesimal dynamics, rather than small grain dynamics, then the system may be best thought of as a newly formed debris disk, which is heavily embedded in gas. This should be a common pathway of evolution for planet-forming disks.

3. Grains are continuously recycled among the planetesimals. Collisions among planetesimals will produce grains, which then go through a pebble accretion-like mass redistribution onto the remaining planetesimals. In this way, large-scale migration of millimeter grains is avoided, as the grains spend most of their time locked up in large planetesimals (Boley & Ford 2013). It also keeps the grains largely confined to the location of the parent bodies. Grains that do form in regions of low planetesimal density may, however, experience migration into areas of higher planetesimal density, where they would be more likely to be accreted.

This mechanism should be general. The telltale signatures of this process include the following. (1) Horseshoe gaps should be present at the location of a planet, provided that the planet’s orbital eccentricity is low. High-resolution interferometry with high image fidelity may be necessary to fully see the horseshoe structure (e.g., D2 versus D5-B5-D6 in HL Tau). (2) Disks should exhibit a fine ring and band structure in which minor bands are in commensurabilities with the centers of very deep gaps or with the center of possible horseshoe structures. These features should be most noticeable when interior and exterior commensurabilities between the planets slightly overlap, widening the resulting band. (3) Unless large-scale asymmetries are present due to, e.g., gravitational instabilities (Tobin et al. 2016), all disks should show the ring-band structure while significant gas is present. Nonetheless, we expect the banded structure to fade with time. This is due, in part, to planetesimal self-stirring as the gas dissipates, which will cause the cold population to heat dynamically and wash out small radial variations.
Figure 5. Same as Figure 4, but assuming PD2 is lower mass (1M_\odot) and has e = 0.005. The mass of PB5 is kept the same, but the initial eccentricity is increased to e = 0.001. The higher eccentricity for PD2 reduces the presence of the horseshoe gap, while also allowing a wide gap with a lower-mass planet.

Figure 2

The Astrophysical Journal, 850:103 (11pp), 2017 November 20  Boley

exclude particles with an eccentricity e > 0.1. Particles with eccentricities larger than this value are almost exclusively associated with the high-eccentricity wings of the planet clearing zones (Figure 2). Particles in those wings, which extend over a large range of semimajor axes, skew the averaging used in the dispersion calculation.

Based on Equation (2), we expect the following behavior, assuming identical planetesimals. When \( F_g \sim 1 \), the collision rate is only affected by \( \Sigma_e/T \). The collision rates for a sea of small planetesimals will only show strong, local variations if there are corresponding changes in the surface density. As gravitational focusing becomes important, then collision rates will increase at semimajor axes that have dynamically cold gravitational focusing becomes important, then collision rates will increase at semimajor axes that have dynamically cold clearing zones (associated with the high-eccentricity wings of the planet orbit). This will give rise to strong local variations that depend on \( F_g \). In this way, if the picture presented here is correct, the existence of rings and bands could be a tool for evaluating the size scales of planetesimals in a disk.

The solid lines in the two upper panels of Figure 6 show the results using this method. We normalize the total mass in solids to be \( 10^{-3} M_\odot \), which is comparable to the inferred mass of the disk (Kwon et al. 2011) with an assumed solid-to-gas mass ratio of 0.01. The surface density of solids is shown in the bottom left panel. All planetesimals are assumed to be identical, with sizes \( R = 50 \) km and \( \rho_p = 1 \) g cm\(^{-3}\). The panel on the left assumes only a 2D dispersion (vertical component removed), while the right panel shows the results for the full 3D dispersion. This is done to show how the vertical extent of the disk can impact the collision rate. The particles in these simulations were given a random inclination from \( 0^\circ \) to \( 1^\circ \), which was motivated initially so that the simulations would be fully 3D. However, this should have been smaller to be consistent with the mechanism proposed here, as will be discussed momentarily. To ensure that using only a 2D dispersion is a reasonable proxy for a thinner disk, a realization of the HL Tau, base simulation was run with random initial inclinations between \( 0^\circ \) and \( 0.5^\circ \) (planets were given inclinations of \( 0.5^\circ \)). The full 3D dispersion for this thinner disk is consistent with that of the 2D dispersion in the regular HL Tau, base simulation, suggesting that the approximation is justified. The 2D and 3D dispersions for the HL Tau, base simulation are shown in the bottom-right panel of Figure 6.

For an additional determination of the collision rates, we use

\[
N = \frac{3\rho_g R_\Sigma^2}{4\rho_p R^3} V_{rel} F_g
\]

Evaluating this expression requires an estimate of the spatial volume mass density of planetesimals \( \rho_p \). To proceed, an azimuthally averaged density structure is calculated by binning particles according to their locations in spherical radius and disk altitude.\(^1\) Radial bins are logarithmically spaced between 10 and 200 au with 4096 divisions. At 45 au, this zoning corresponds to a radial grid width of about 0.033 au. The disk altitude ranges from \(-10^\circ\) to \(+10^\circ\), with 512 equally spaced divisions, which yields similar cell sizes in radial width and cell height. The density \( \rho_d \) is then found by assuming that the planetesimals are identical and that the total mass in solids is again \( 10^{-3} M_\odot \).

The collision rate further requires an estimate of the velocity dispersion in each cell. As done above, the velocity dispersion is calculated relative to a circular, Keplerian orbit at the particle’s cylindrical distance from the star. With this information in hand, we next take the current orbital elements for each particle and propagate the planetesimal along that fixed orbit for one full period. This propagation is performed using 1000 equal divisions in true anomaly. At each position along the orbit, we determine the cell in which the particle is located. The value for \( V_{rel} \) is found by adding in quadrature the cell’s velocity dispersion with the instantaneous velocity of the particle relative to a circular, Keplerian orbit. We average the values of \( V_{rel} \) and \( \rho_d \) that correspond to the start and end

\(^1\) We are using disk altitude to refer to the angle measured from the midplane of the disk toward the \( z \) axis as seen from the star. This is done to avoid potential confusion with the particle’s orbital inclination.
states for each segment of true anomaly. The resulting $N$ is then used to calculate $p = N \Delta t$, where $\Delta t$ is the time step that corresponds to the current change in true anomaly. The orbit-averaged collision rate is then found by summing $p$ over the entire orbit and dividing by the particle’s orbital period.

The result of this method is also shown in Figure 6. Each point represents the orbit-averaged collision rate for each particle in the 100,000 yr snapshot. As with the Safronov-based rate, the left panel shows the results assuming that the vertical components of the velocity can be neglected, while the right panel shows the full 3D result for the given simulation. Both collision rate calculations, i.e., Equations (2) and (4), are very consistent, with a shift in the rates by a factor of about 3. Importantly, locations of resonances and planetary chaotic zones show reduced collision rates, although this is modest for the 3D result. The 2D case shows a large difference between collision rates in and outside of resonances. This is driven mainly by gravitationally focusing among the dynamically cold planetesimals, which is suppressed at resonances. The differences between the 2D and 3D cases further show that the planetesimal disk needs to be vertically thin and must have a population of large planetesimals for the proposed mechanism to operate. Even in the 2D case, if the planetesimals were all about 5 km in radius, then the collision rates would look very similar to the 3D rates shown in Figure 6, just shifted upward due to the increased optical depth of the planetesimal swarm.

The above calculations establish that a difference in the production of dust is indeed expected near MMR and chaotic zones, at least for some plausible planetesimal parameters. Nonetheless, small grains produced by collisions will drift radially unless they are confined to their production region. We propose that this may be accomplished if the dust grains are reaccreted onto local planetesimals. Such accretion is only possible if the reaccretion timescale is less than the migration timescale.

Let the accretion time of a grain be estimated by
\[ t_{acc} = \left( \frac{n \pi R^2 V_{wind} \eta}{1} \right)^{-1} , \tag{5} \]
where $n$ is the number density of planetesimals in the grain’s local environment, $R$ is the planetesimal radius for all planetesimals, $\eta$ is an enhancement factor for the geometric cross section, and $V_{wind}$ is the wind speed due to the relative speed between the planetesimals and the gas, which we further
assume to be approximately the same as the relative speed between the planetesimals and the grains. For purely gravitational interactions, \( \eta = F_g \). This enhancement factor can also be represented by a pebble accretion-enhanced cross section (Hughes & Boley 2017), which can allow \( \eta > F_g \) for massive enough planetesimals. We calculate \( n \) by assuming that the midplane planetesimal number density is given by \( n = \Sigma_d/(2.5 Hm_p) \) and that \( m_p \) is the typical planetesimal mass. The value \( H \) is the scale height of the vertical planetesimal distribution, which is expected to be smaller than the scale height of the gas disk.

Using this estimate for \( n \) and assuming the typical planetesimal mass is \( m_p = 4\pi/3 \rho_p R^3 \), the reaccretion timescale becomes

\[
\tau_{acc} = \frac{10 \rho_p RH}{3 \Sigma_g V_{wind} \eta}.
\]  

(6)

This accretion time must be smaller than the inward drift timescale, which we approximate as \( \tau_{drift} \propto r/V_{drift} \) for disk radial distance \( r \) and inward grain drift speed \( V_{drift} \). As discussed, in Section 2, the radial drift speed for the millimeter grains explored here, which is \( V_{drift} = \Delta g t_d \). Recall that the residual gravity term \( \Delta g \) results from the difference between the gas orbital motion and the local circular, Keplerian orbital speed. We set \( V_{wind} = (1 - \xi) V_c \), where \( \xi \) is a parameterization. Using this definition for \( V_{wind} \), \( \Delta g = 4\pi^2(1 - \xi^2)r/T^2 \) for local orbital period \( T \) at disk radial distance \( r \). Altogether, this gives an inward drift speed of

\[
V_{drift} = \frac{T^2}{4\pi^2(1 - \xi^2)t_d}.
\]  

(7)

We seek a situation such that the ratio \( W = \frac{\tau_{wind}}{\tau_{acc}} \ll 1 \). Using Equations (7) and (6),

\[
W = \frac{40\pi^2 \rho_p RH(1 - \xi^2) t_d}{3 \Sigma_g V_{rel} \eta T^2}.
\]  

(8)

Letting the Stokes number be \( \St = 2\pi t_d/T \), Equation (8) can be rewritten as

\[
W = \frac{10 \rho_p RH(1 + \xi) \St}{3 \Sigma_g \eta} \frac{r}{T^2}.
\]  

(9)

At \( r = 40 \) au, \( T \approx 220 \) year, \( t_d \approx 0.09 \) year (see Section 2), and \( \Sigma_g \approx 0.45 \) g cm\(^{-2}\) (Figure 6), the corresponding Stokes number for the millimeter grains is \( \St \approx 0.0025 \) for these conditions. Again setting \( V_{rel} = 5000 \) cm s\(^{-1}\) as done in Section 2, \( \xi \approx 0.9906 \). We note that Equation (9) is not very sensitive to the choice of \( \xi \) (and thus \( V_{rel} \)). If we let \( R = 50 \) km, \( \eta \approx 1 \), and \( H \approx 1000 \) km (i.e., 20R), then \( W \approx 0.03 \), satisfying the reaccretion/drift condition.

Taken altogether, the results suggest that the proposed mechanism can operate for plausible planetesimal sizes, provided that the planetesimal disk is geometrically very thin. It also suggests that, should the proposed mechanism be correct, then the constraints on collision rates along with the small grain reaccretion timescale could be used to infer planetesimal properties.

4.2. Extension to the Solar System

The early debris paradigm presented here could also be reflected in the solar system’s meteoritic record. Calcium–aluminum–rich inclusions are the oldest known solids in the solar system (Bouvier & Wadhwa 2010), and are presumed to have formed during or shortly after the onset of the solar nebula. They represent the start of planet building. At least some iron meteorite parent bodies formed within 1.5 Myr of the nebula. They are thought to have been accreted from an undifferentiated solar nebula, and are presumed to have undergone some degree of differentiation.

Furthermore, planetesimal and asteroidal discs may have coexisted in the solar system, and small planetesimals may have been reaccreted onto planetesimals, thus forming larger objects. This process may have continued throughout the history of the solar system, resulting in the formation of the terrestrial planets.

The proposed model may also be connected with chondrules, which are mm-sized igneous spherules found in abundance in undifferentiated meteorites (Scott & Krot 2005). Radiogenic dating presents evidence that the majority of chondrules formed between 2 and 3 Myr after CAI formation, although some may have formed contemporaneously with CAIs (Villeneuve et al. 2009). While chondrule formation is hotly debated, their formation mechanism may be linked to an initial population of planetesimals (Hood 1998; Asphaug et al. 2011; Morris et al. 2012; Johnson et al. 2015; Mann et al. 2016). Furthermore, chondrules and planetesimals must have coexisted in the solar nebula regardless of their formation mechanism (e.g., certain classes of asteroids are made mostly of chondrites). This itself has several consequences.
Any planetesimal that exceeds about 100 km in radius would begin to accrete small solids efficiently through combined gravity and gas drag effects (Lambrechts & Johansen 2012). This pebble accretion would be most efficient for chondrule-sized objects in the regime of the asteroid belt, with some dependence on the disk model (Johansen et al. 2015). Collisions between chondritic planetesimals would cause mixing and mass migration among the chondrule population and their parent bodies. This could, in principle, reduce the amount of chondrules that would be lost due to radial drift and could help to explain why chondrites contain chondrules with a range of formation ages (e.g., Villeneuve et al. 2009). It may also be consistent with recent proposals regarding the types and internal structures of parent bodies (Weiss & Elkins-Tanton 2013).

4.3. Extension to TW Hya

As an additional example, we extend our model and computational methods to TW Hya. Observations by Andrews et al. (2016) show that this system harbors wide gaps with some potential band and ring substructure. The most obvious dark bands in the Andrews et al. observations are at ~22, 37, and 43 au. A bright ring separates the 37 and 43 au bands, again reminiscent of a horseshoe structure. However, the disk brightness drops off rapidly after the 43 au band, which is different from that seen in HL Tau. Minor bands in TW Hya appear to be present at 28 and 31 au, although these are much harder to discern in the current observations. There is additional structure at disk radii smaller than 22 au, but as with HL Tau, we ignore this region for computational ease. It is straightforward to extend our analysis to these inner regions. The radial structure of TW Hya as presented by Andrews et al. (2016) is shown in Figure 7.

The 22 au band is broad, and appears to extend from roughly ~19 to 25 au. If a planet is placed at 22 au, then the 3:2 commensurability will be located at roughly 28.8 au, overlapping the minor band near that radial distance. If an additional planet is located at about 40 au, then the 2:3 commensurability would be at about 31 au, coinciding with the other minor band. A planet at 40.5 au would further yield a ring at ~40 au with gaps at about 37 and 43 au.

Using the same methods as given in Section 2, we place 100,000 test particles on circular orbits between 10 and 60 au. A 10 $M_j$ planet is set at 22 au and a ~19 $M_j$ planet is placed at 40.52 au, which is at the 5:2 commensurability with the inner planet. The strict commensurability may not be necessary. We place it at this location because the bright ring between the 37 and 43 au bands is slightly offset from 40 au. The masses are chosen to be consistent with the gap widths, assuming a horseshoe structure for the outer gap-ring-gap morphology. The outer planet is on a circular orbit, and the inner planet has an eccentricity of 0.005.

The results are shown in Figure 8. The mild eccentricity of the inner planet produces a wide 22 au gap, while also reducing the amount of low-eccentricity, co-orbital material. The gap in this particular simulation is slightly too wide to reproduce the TW Hya 22 au band, but this could be altered with the planet’s eccentricity. The 28 au minor band is also produced. The feature is de-emphasized in the face-on plot compared with the radial eccentricity. The 28 au minor band is also produced. The feature is de-emphasized in the face-on plot compared with the radial binning, but this is also true in the observations (e.g., Andrews et al. 2016). The 40.5 au planet produces the co-orbital region as expected, and also produces a minor band at about 31 au.

The surface brightness profile of the TW Hya millimeter observations (Andrews et al. 2016) shows that the disk brightness never strongly recovers following the 37 au band, which would be inconsistent with the simulation presented here. However, this may be due to an intrinsic structure of the disk. Regardless of the reason, additional bands at larger disk radii will be present and coincide with commensurabilities should the planet hypothesis be correct, unless the entire disk exterior to 40 au has simply become highly stirred. Likewise, the commensurability structure should extend to smaller disk radii, interior to the 22 au planet. Hints of such structure are present in the existing millimeter data.

A potential issue with the proposed paradigm is that the ~40 au ring in the ALMA continuum image of TW Hya is symmetric, within the fidelity of the observations. In contrast, the simulations suggest that a distinct, azimuthal horseshoe gap should be present. We do not have a full explanation for this...
apparent discrepancy. Dust within the planet’s Hill sphere will partly fill in such gaps, but this alone appears to be insufficient. For example, the Hill sphere for the proposed 40 au planet is about 1 au, while the horseshoe corotation gap (azimuthal gap) is about 12–13 au. This leaves approximately 5 au on either side of the planet that must be accounted for. This gap could be closed, in part, through gas-drag effects on the millimeter grains, but this must be tested in simulations that take into account the recycling of millimeter grains (as proposed here) along with the drag. Should large azimuthal gaps prove to be required under the proposed mechanism, then high-fidelity millimeter imaging can test for this signature.

5. Conclusions

We have presented a model to explain the presence of gaps in disks. The points central to the argument are that (1) planetesimals form early and form everywhere in the disk, (2) the millimeter grains are byproducts of planetesimal collisions and evolution, and (3) the small grains are recycled among the dynamically cold population of the planetesimals, keeping those grains as tracers of that population. The model will slow the inward migration of small grains, as grains will spend most of their time in larger bodies. The recycling of material will give rise to pebble-like mass transfer/migration from smaller planetesimals to larger planetesimals and embryos. Any small solids that become thermally processed while in the nebula (such as chondrules) will have a range of ages, such as seen in meteorites. When the model is applied to HL Tau, the complex banded structure can be reproduced with three planets (only two are directly modeled), provided the planetesimal disk is geometrically thin.

Disk structures are produced through resonances with existing planets. Planets that are on circular orbits will produce a horseshoe structure with a band–ring–band morphology. As the planet gains eccentricity, this structure will develop into a single broad band. The model predicts that minor bands should be associated with the locations of commensurabilities with proposed planets, provided that the planet is massive enough to cause significant excitation. This creates a strong test of the model, although there is some flexibility. For example, eccentric planetary embryos may be able to create deep bands regardless of whether they are in or out of commensurability, although a mean motion resonance may be the most plausible mechanism for forcing an embryo’s orbital eccentricity.

While the model appears to have success in reproducing the morphology of HL Tau and TW Hya, at least at the level of detail presented here, the model must be developed further. The coevolution of millimeter grains and planetesimals must ultimately be simulated to establish fidelity between models and the corresponding millimeter continuum imaging.

A.C.B. thanks S. Andrews, B. Gladman, P. Granados, M. Payne, D. Tamayo, C. van Laerhoven, J. White, and D. Wilner for discussions that helped to improve this manuscript. A.C.B. also thanks the referee, Stuart Weidenschilling, for very helpful comments that improved this manuscript. This work was supported by an NSERC Discovery grant, The University of British Columbia, the Canadian Foundation for Innovation, and the BC Knowledge Development Fund.

ORCID iDs

A. C. Boley © https://orcid.org/0000-0002-0574-4418

References

Adachi, I., Hayashi, C., & Nakazawa, K. 1976, PThPh, 56, 1756
Akiyama, E., Hasegawa, Y., Hayashi, M., & Iguchi, S. 2016, ApJ, 818, 158
ALMA Partnership, Brogan, C. L., Pérez, L. M., et al. 2015, ApJL, 808, L3
Andrews, S. M., Wilner, D. J., Zhu, Z., et al. 2016, ApJL, 820, L40
Asphaug, E., Jutzi, M., & Mollov, N. 2011, E&PSL, 308, 369
Birnstiel, T., Andrews, S. M., Pinilla, P., & Kama, M. 2015, ApJL, 813, L14
Boley, A. C., & Ford, E. B. 2013, arXiv:1306.0566
Bouvier, A., & Wadiwa, M. 2010, NatGe, 3, 637
Bryden, G., Chen, X., Lin, D. N. C., Nelson, R. P., & Papaloizou, J. C. B. 1999, ApJL, 514, 344
Chambers, J. E. 1999, MNRRS, 304, 793
Deck, K. M., Payne, M., & Holman, M. J. 2013, ApJ, 774, 129
Dipierro, G., Price, D., Laibe, G., et al. 2015, MNRRS, 453, L73
Dodson-Robinson, S. E., Veras, D., Ford, E. B., & Beichman, C. A. 2009, ApJ, 707, 79
Durisen, R. H., Boss, A. P., Mayer, L., et al. 2007, in Protostars and Planets V, ed. B. Reipurth, D. Jewitt, & K. Keil (Tucson, AZ: Univ. Arizona Press), 607
Gladmans, B. 1993, Icar, 106, 247
Haghighipour, N., & Boss, A. P. 2003, ApJ, 583, 996
Hellé, R., Bodenheimer, P., Podolak, M., et al. 2014, in Protostars and Planets VI, ed. H. Beuther et al. (Tucson, AZ: Univ. Arizona Press), 643
Hood, L. L. 1998, M&PS, 33, 97
Hood, L. L., & Weidenschilling, S. J. 2012, M&PS, 47, 1715
Hughes, A., & Boley, A. C. 2017, MNRRS, 472, 3543
Jin, S., Li, S., Isella, A., Li, H., & Ji, J. 2016, ApJ, 818, 76
Johansen, A., Mac Low, M.-M., Lacerda, P., & Bizzarro, M. 2015, SciA, 1, 1500109
Johnson, B. C., Minton, D. A., Melosh, H. J., & Zuber, M. T. 2015, Natur, 517, 339
Kanagawa, K. D., Muto, T., Tanaka, H., et al. 2016, PASJ, 68, 43
Kley, W., & Nelson, R. P. 2012, ARA&A, 50, 211
Kwon, W., Looney, L. W., & Mundy, L. G. 2011, ApJ, 741, 3
Lambrechts, M., & Johansen, A. 2012, A&A, 544, A32
Lawler, S. M., Greenstreet, S., & Gladman, B. 2015, ApJL, 802, L20
Mann, C. R., Boley, A. C., & Morris, M. A. 2016, ApJ, 818, 103
Morris, M. A., Boley, A. C., Desch, S. J., & Athanassiadou, T. 2012, ApJ, 752, 27
Pollack, J. B., Hubickyj, O., Bodenheimer, P., et al. 1996, Icar, 124, 62
Saito, Y., Sato, T., & Yamashita, H. 2017, ExoPlanets, ed. B. Reipurth, D. Jewitt, & K. Keil (Tucson, AZ: Univ. Arizona Press), 643
Schersten, A., Elliott, T., Hawkesworth, C., Russell, S., & Masarik, J. 2006, E&PSL, 241, 530
Scott, E. R. D., & Krot, A. N. 2005, in ASP Conf. Ser. 341, Chondrites and the Earth and Planets (San Francisco, CA: ASP), 15
Simon, S. M., Tamayo, D., Petrovich, C., Reiners, H., & Murray, N. 2017, MNRRS, 469, 3337
Takahashi, Z. S., & Inutsuka, S.-i. 2016, ApJ, 152, 78
Tamayo, D., Triaud, A. H. M. J., Menou, K., & Reiners, H. 2015, ApJ, 805, 100
Tobin, J. J., Kratter, K. M., Persson, M. V., et al. 2016, Natur, 538, 483
Tobin, J. J., Kratter, K. M., Persson, M. V., et al. 2016, Natur, 538, 483
Villeneuve, J., Chaussidon, M., & Libourel, G. 2009, Sci, 325, 985
Weidenschilling, S. J. 1977, MNRRS, 180, 7
Weiss, B. P., & Elkins-Tanton, L. T. 2013, AREPS, 41, 529
Whipple, F. L. 1973, NASSP, 319, 355
Wyatt, M. C., Booth, M., Payne, M. J., & Churich, L. J. 2010, MNRRS, 402, 657
Zhang, K., Blake, G. A., & Bergin, E. A. 2015, ApJL, 806, L7