Observational Signatures of Young Planets in Disks

Zhaohuan Zhu

*Princeton University
E-mail: zhzhu@astro.princeton.edu

Finding young planets in protoplanetary disks is essential for understanding planet formation process and constraining the long-term evolution of planetary systems. Transitional disks, protoplanetary disks with gaps and holes, are great candidates for harboring these young planets. However, recent near-IR polarization imaging (e.g. SEEDS, VLT) and submm (e.g. SMA, ALMA) observations have posed several puzzles on transitional disks. Such as, dust and gas seems to decouple in these disks and the decoupling can occur non-axisymmetrically in disks. Spiral patterns are also discovered in these disks. In this chapter, I will first summarize observations on transitional disks. Then I will present theoretical developments on three indirect signatures of young planets in disks: gaps, large scale asymmetric structures, and spiral arms. By comparing such theories with observations, we have constrained protoplanetary disk properties and revealed the early stage of planet formation. Finally, observational strategies to directly find young planets in protoplanetary disks has been discussed and I suggest that accreting circumplanetary disks could be the key to detect young planets directly. Current direct imaging observations may have already found some circumplanetary disk candidates and there are more to come.
1. Introduction

Thanks largely to the Kepler spacecraft, more than 1800 exoplanets have been discovered. With such large number of exoplanets, the architecture of exo-planetary systems has been revealed [1], and now we know that on average every star has roughly one planet within 1 AU. However, most of these planets are billions of years old. We have limited constraints on how they formed in protoplanetary disks within their first several million years lifetime.

On the other hand, the structure of protoplanetary disks has also been revealed over the last decade by studying their spectral energy distributions (SEDs). Recently, we start to spatially resolve protoplanetary disks either through optical/near-IR telescopes (e.g. Subaru, VLT, Gemini) or submm/mm interferometer (e.g. ALMA, EVLA). In the near future, especially with ALMA, we will reveal the detailed disk structure which will provide the initial condition for planet formation.

Thus, a missing link in planet formation studies is finding young forming planets in disks. If we find young planets in disks, we know when, where, and how young planets form in disks, which will put stringent constraints on planet formation theories. Unfortunately, finding young planets that orbit around protostars through either the radial velocity or transit method is difficult mainly because their host stars are highly variable. Spitzer Space Telescope has found that more than half of the protostars show variabilities, from 0.05 to 0.2 magnitude, on a variety of timescales [2]. Such large variability dominates the variability signal induced by the planet.

The direct imaging technique is more promising to find young planets in protoplanetary disks and it has discovered several planets/planet candidates in protoplanetary disks. However, such method requires using 10 meter class telescopes to get enough photons from the young planet. Young planetary systems are far away (>100 pc away), and normally obscured by dust. For comparison, the HR 8799 system where 4 planets have been directly imaged is only 40 parsecs away.

A much easier way to find young planets is studying the imprints left by the planet on protoplanetary disks. The gravity from a young planet can perturb the protoplanetary disk, inducing spiral patterns and gaps. It can also lead to large scale disk instabilities, causing asymmetric disk structures. These large scale features in disks are much easier to be detected. Transitional disks, protoplanetary disks with gaps and holes [3], are great candidates to harbor young planets and hundreds of transitional disks have been discovered.

On the other hand, eventually we would like to confirm the existence of these young planets in disks. Directly detecting young planets in disks is still possible considering young planets may be bigger and hotter than mature planets. Furthermore, when a young planet opens a gap, it exposes itself with less extinction. The young planet can continue to grow through the accretion of the circumplanetary disk. The accretion process can release an observable amount of energy, and these circumplanetary disks can be directly found by direct imaging techniques.

In this chapter, I will first summarize transitional disk observations. Then, I will discuss three indirect observational signatures left by the planet: gaps, spiral arms, asymmetric structures, and how they can explain some of the transitional disk observations. Finally, I will discuss the frontier in detecting young planets: observing accretion signatures of circumplanetary disks.
2. Transitional Disks

Transitional disks are protoplanetary disks with gaps and holes. They were first identified by their unusual SEDs at near-infrared (NIR) and mid-infrared (MIR) [4]. The transitional disks exhibit strong dust emission at wavelengths $\gtrsim 10\mu m$, while showing significantly reduced fluxes relative to typical T Tauri disks at shorter wavelengths (e.g., [5; 6; 7]). In a subgroup of transitional disks, the so-called pre-transitional disks, there is evidence for emission from warm, optically-thick dust near the star [7; 8], while in other transitional disks the emission at $\lesssim 10\mu m$ appears to be due entirely to optically-thin dust [5; 8]. The depletion of near- to mid-infrared emission is generally interpreted as being due to evacuation of dust to disk scales $\sim 5$ to $\sim 50$ AU [9; 5; 7].

Recently, thanks to high contrast NIR polarization imaging (e.g. SEEDS, [10]) and submm-interferometric techniques [11; 12; 13], these gaps/holes have been resolved spatially (Figure 1). NIR observations probe the distribution of small dust particles at the disk surface, while submm observations probe larger dust particles at the disk midplane. In many disks, NIR and submm observations reveal completely different structures, suggesting that small and big dust grains drift relative to each other in protoplanetary disks [14].

Recent ALMA observations have revealed non-axisymmetric dust distributions in some transitional disks [15; 16; 17]. In the extreme case of Oph IRS 48, there is a highly asymmetric crescent-shaped dust structure between 45 and 80 AU from the star. The peak emission from this dust structure is at least 130 times stronger than the upper limit of the opposite side of the disk. Spiral patterns have also been revealed by NIR direct imaging observations in some transitional disks (e.g. [18]).

Surprisingly, despite the wide dust gaps, most transitional disks exhibit gas accretion rates close to the averaged T Tauri disk accretion rate ($\sim 10^{-8}M_{\odot}\text{yr}^{-1}$ [20]) onto their central stars (e.g. [5; 7; 21]). Maintaining this accretion requires either a significant mass reservoir interior to the cavity, or some way of allowing mass from the outer disk to move past the cavity [22].

To summarize the observations, transitional disks normally have: 1) moderate accretion rates, 2) wide dust gaps up to several tens of AU, and in some transitional disks we see: 3) large scale...
Figure 2: Spatially resolved protoplanetary disk SAO 206462 from VLT (left panel from [19]) using polarized scattered light at NIR, and ALMA at submm (right panel from [17]). Complicated and different disk structures have been revealed at various wavelengths. Please note that the orientation of the two images are slightly different.

Figure 3: The SEDs for a full classical T Tauri star (upper panels) and a transitional disk (lower panels). At ultraviolet wavelengths (the left panels), the excess emission above the stellar photosphere is due to the magnetospherical accretion. The right panels are from [23]. The upper left panel is from [24] and the lower left panel is from [25].

asymmetric structures at submm/mm, and 4) spiral patterns at Optical/NIR.

3. Indirect Signatures of Young Planets

Theoretically, it is known that young planets can gravitationally interact with disks, launching spiral density waves [26] and inducing gaps in disks [27].

3.1 Gaps

Gap opening process has been thoroughly studied in viscous disks (e.g. [28; 29]) and MHD
turbulent disks (e.g. [30; 31]).

Observations have also suggested that gaps in some transitional disks are indeed induced by companions. For example, CoKu Tau/4 is surrounded by a transitional disk [6] with a 14 AU cavity [32], and inside the cavity it has a nearly equal mass companion with a separation of 8 AU [33]. The observed gap width and binary separation are consistent with theoretical studies on circumbinary disks [34].

However, companions have not been found inside the cavity of most transitional disks. Thus, we need to ask whether the observed gap/cavity properties are consistent with the existence of unseen planets. One serious challenge for almost all theoretical models to explain transitional disks is the fact that some transitional disks exhibit large dust clearings while still maintaining significant gas accretion rates onto the star. For example, GM Aur has an optically thin inner disk at 10 µm with an optical depth of 0.01 and an accretion rate of \( \sim 10^{-8} M_\odot \) yr\(^{-1} \). Using a viscous disk model with \( \alpha = 0.01 \), \( \Sigma_g \) is derived to be 10–100 g/cm\(^2\) at 0.1 AU. Considering that the nominal opacity of ISM dust at 10 µm is 10 cm\(^2\)/g, the optical depth at 10 µm for the inner disk is 100–1000 [22], which is 4–5 orders of magnitude larger than the optical depth (0.01) derived from observations.

One possible solution is that multiple planets are present in transitional disks. If multiple planets from 0.1 AU to tens of AU can open a mutual gap, the gas flow can be continuously accelerated and passed from one planet to another so that a low disk surface density can sustain a substantial disk accretion rate onto the star [22; 35]. However, considering giant planets tend to be trapped at 2:1 mean-motion resonance, a large number of giant planets are needed in this model to pass gas from several tens of AU all the way down to 0.1 AU. Furthermore, when gas passes through each planet, a fraction of the gas can be intercepted and accreted by the planet so that the final accretion rate onto the star could be smaller than observed rates.

Another possibility is that the small dust grains in the inner disk are highly depleted by physical removal or grain growth, to the point where the dust opacity in the NIR is far smaller than the ISM opacity. Dust filtration [36] is one possible mechanism to differentiate dust from gas. Due to the gas drag, dust particles always drift to the pressure maximum in disks [37]. At the outer edge of a planet-induced gap, where the pressure reaches a maximum, dust particles drift outwards, possibly overcoming their coupling to the inward accreting gas. Dust particles will then remain at the gap’s outer edge while the gas flows through the gap. This dust trapping at the gap edge was first simulated by [38; 39]. Particle diffusion due to disk turbulence is later included [40] in 2-D simulations evolving over viscous timescales, where a quasi-steady state for both gas and dust has been achieved. [40] found that micron sized particles are difficult to filter by a Jupiter mass planet in a \( \dot{M} = 10^{-8} M_\odot \) yr\(^{-1} \) disk. Thus, cavities revealed by submm observations may not be present in NIR observations.

This observational signature is supported by recent observations that cavities revealed by submm/mm observations are missing at NIR [41; 19; 42] or are larger than cavities revealed by NIR observations of the same objects. The dust filtration picture is also supported by the metal depletion in Herbig Ae/Be stars surrounded by transitional disks compared with metal abundance of stars surrounded by full disks [43].

On the other hand, how to deplete micron sized dust within the hole to explain the NIR deficit of transitional disks is still unclear. [44] has done 1-D calculations considering dust growth and dust fragmentation at the gap edge and suggested that micron-sized particles may also be filtered.
Also by constructing a 1-D model, [45] has suggested that radiation pressure from the accreting planet can blow small dust away during the filtration process. Considering the flow pattern is highly asymmetric within the gap, 2-D simulations including dust growth, fragmentation, and radiation pressure may be needed to better understand the dust filtration process.

Figure 4: The disk surface density for the gas (leftmost panels) and dust in viscous (\(\alpha=0.01\), upper panels) and inviscid (lower panels) simulations. In the viscously accreting disk, the accretion flow carries small particles from the outer disk to the inner disk while leaves big particles at the gap edge. In the inviscid disk, particles are trapped in the gap edge vortex and planet co-orbital region. ([3])

3.2 Large Scale Asymmetric Structures

To explain the large scale lopsided dust structure observed in transitional disks, particle trapping in the azimuthal direction by vortices is proposed. It is known that anticyclonic vortices are long lived and can efficiently trap dust particles [46; 47]. Although vortices can be generated by various hydrodynamical instabilities in disks [48], in the scenario of gap opening by planets, the gap edge is subject to the Rossby wave instability (RWI) which can naturally lead to vortex formation [49; 50]. Spiral shocks excited by a planet push the disk material away from the planet, leading to gap opening. This process also piles up material at the gap edge, leading to a density bump which has a vortensity minimum and is subject to the RWI [51; 52; 53; 54; 55]. Particle trapping in 3-D vortices has also been studied using 3-D hydrodynamical global simulations including dust dynamics [56; 57], which suggests that particles at certain sizes could have a factor of more than 100 increase in the dust surface density within the vortex.

On the other hand, the vortex can dissipate quickly in a disk even with a small viscosity (\(\nu \sim 10^{-5}\)) [58; 59]. The vortex is also unstable in a disk with magnetic fields [60]. By considering non-ideal MHD effects which significantly suppress the turbulence, [61] found that the vortex can be present for thousands of orbits in a disk with equivalent \(\alpha \lesssim 10^{-3}\), and dust concentration in such vortices can explain ALMA observations.
Recent theoretical works start to include dust feedback to vortices (e.g. [62; 63; 64; 65]) and consider the effect of gas gravity on particle concentration [66; 67]. Observationally, multi-band ALMA and EVLA observations start to reveal particle size distributions within the lopsided structure [68; 69]. Eventually, with ALMA’s high spatial and spectral resolution, we may be able to probe the anticyclonic gas motion directly through velocity channel maps, which could confirm or disprove the scenario that particle concentration by vortices is responsible to these large scale asymmetric structures.

### 3.3 Spiral Structures

Recent high-resolution direct imaging observations have revealed spiral structure in protoplanetary disks around Herbig Ae/Be stars (e.g. SAO 206462, [18; 19], MWC 758, [70; 71]). In the NIR polarized intensity images, two spiral arms with roughly 180° rotational symmetry are present in both SAO 206462 and MWC 758, similar to the grand design in a spiral galaxy (e.g. the Whirlpool Galaxy M51). The spiral arms also exhibit a high contrast against the background disk.

Since these disks also have holes at the center, one scenario to explain both spiral patterns and gaps/holes is that these disks harbor low-mass companions (e.g. young planets) which can open gaps and excite spiral waves at the same time.

However, there are two difficulties in explaining the observed spiral patterns using spiral wakes induced by a planet inside the gaps. First, the large pitch angle of all the observed spiral arms suggests that the disk has a relative high temperature, e.g. \( \sim 200 \, \text{K} \) at \( R \sim 100 \, \text{AU} \) which is too high for any realistic disk structure. Second, the observed spiral arms exhibit much higher brightness contrasts than suggested by the synthetic observations based on two dimensional (2-D) planet-disk simulations with the assumption that the disk is in vertical hydrostatic equilibrium [72].

Recent 3-D simulations have suggested that the pitch angle formula derived from the linear theory does not apply to the high planet mass cases [73; 74]. Spiral wakes that are excited by high mass planets (e.g. 1 \( M_J \)) become spiral shocks which propagate at speeds faster than the local sound speed [75; 76]. Thus, the pitch angle difficulty above can be alleviated by considering the non-linear extension of the spiral shock theory. The spiral arms (especially the inner arms) have complicated non-hydrostatic 3-D structure which can lead to strong density perturbation at the disk surface. Since NIR observations is sensitive to the density perturbation at the disk surface, this effect alleviates the second difficulty mentioned above. Thermodynamics can also play an important role on the spiral structure [77; 78]. Furthermore, a secondary inner arm is also excited by the planet (see also [79]). Thus only one planet is necessary to explain two spiral arms. Detailed modeling [74] has shown that planet-induced inner spiral arms can explain recent NIR direct imaging observations of SAO 206462 and MWC 758. However, since the planets are outside the spiral arms in this model, it is difficult to explain the gaps discovered at small radii in these disks. Other mechanisms, e.g. another planet or photoevaporation, are needed to explain these gaps.

Overall, companion-induced spiral arms not only pinpoint the companion’s position but also provide three independent ways (pitch angle, separation between two arms, and contrast of arms) to constrain the companion’s mass.
4. Direct Signatures of Young Planets

Accreting circumplanetary disks can release a large amount of thermal energy due to the small size and deep potential of the planet [73]. A disk around a 1 $M_J$ planet accreting at $\dot{M} = 10^{-8}M_\odot\text{yr}^{-1}$ has an accretion luminosity of

$$L_{\text{disk}} = \frac{GM_J\dot{M}}{2R_J} = 1.5 \times 10^{-3}L_\odot,$$

which is as bright as a late M-type/early L-type brown dwarf.

Detailed SEDs for CPDs accreting at different accretion rates are shown in Figure 5 [73]. Circumplanetary disks only accreting at $10^{-10}M_\odot\text{yr}^{-1}$ around a 1 $M_J$ planet can be brighter than the planet itself. A moderately accreting circumplanetary disk ($\dot{M} \sim 10^{-8}M_\odot\text{yr}^{-1}$; enough to form a 10 $M_J$ planet within 1 Myr) around a 1 $M_J$ planet has a maximum temperature of $\sim$2000 K, and at NIR wavelengths ($J$, $H$, $K$ bands), this disk is as bright as a late M-type brown dwarf or a 10 $M_J$ planet with a “hot start”. To use direct imaging to find the accretion disks around low mass planets (e.g., 1 $M_J$) and distinguish them from brown dwarfs or hot high mass planets, it is crucial to obtain photometry at mid-infrared bands ($L'$, $M$, $N$ bands) because the emission from circumplanetary disks falls off more slowly towards longer wavelengths than those of brown dwarfs or planets. When the giant planet has strong magnetic fields (>100 Gauss), they can truncate the CPD, leading to magnetospheric accretion [80; 73]. Besides the emission from the disk, an additional blue component from the hot spot on the star is also present (the right panel of Figure 5). Magnetospheric accretion around T-Tauri stars leads to strong line emission (e.g. H$\alpha$). The recently discovered H$\alpha$ source within LkCa 15 [81] could be a CPD undergoing magnetospheric accretion. Table 1 summarizes recent detections of point sources in protoplanetary disks, and model predictions based on [73]. For the theoretical predictions, the disk inner radius is assumed to be 1 Jupiter radii. To derive the absolute magnitudes, the distance and disk inclination are from the given references for each source. We fit the model to the observed $L'$ band magnitudes, and predict magnitudes at other wavelength bands. Generally, the model agrees with observations reasonably well.

Besides thermal emission, CPD can also be probed by dust continuum [86], velocity channel maps [87], or disk chemical abundances [88]. With ALMA’s high sensitivity and spatial resolution, we could probe all these three signatures of CPDs.

5. Conclusion

Features in protoplanetary disks (e.g. gaps, large scale asymmetry, spiral arms) may have already implied young planets in disks. With ALMA and Extreme-AO, we start to see narrower gaps and finer features, which can limit the potential parameter space of young planets. At the same time, we also have great hope to detect young planets directly through accreting circumplanetary disks. In the future, both indirect and direct methods will allow us to determine the distribution and occurrence of young planets, which will shed light on the decades-old problem of planet formation, and reveal how young planetary systems can evolve into older ones such as our Solar System, billions of years after they were born.
Table 1: Magnitudes of Accreting Circumplanetary Disks compared with Observations

<table>
<thead>
<tr>
<th></th>
<th>J</th>
<th>H</th>
<th>K</th>
<th>L'</th>
<th>M</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Full Disk</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Obs. HD169142 from [82]</td>
<td>&gt;13.8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>The. 10^{-5}</td>
<td>14.8</td>
<td>14.66</td>
<td>13.82</td>
<td>12.2</td>
<td>11.62</td>
<td>10.12</td>
</tr>
<tr>
<td>Obs. HD100546 from [83; 84]</td>
<td>19.4±0.32</td>
<td>&gt;15.43±0.11</td>
<td>13.92±0.1</td>
<td>13.33±0.16</td>
<td></td>
<td></td>
</tr>
<tr>
<td>The. 2×10^{-6}</td>
<td>20.66</td>
<td>18.41</td>
<td>16.50</td>
<td>13.9</td>
<td>13.05</td>
<td>11.37</td>
</tr>
<tr>
<td>Obs. LkCa 15 from [81]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>The. 7×10^{-6}</td>
<td>16.09</td>
<td>15.92</td>
<td>14.96</td>
<td>13.2</td>
<td>12.54</td>
<td>11.04</td>
</tr>
</tbody>
</table>

Figure 5: The SEDs of accreting circumplanetary disks (black curves) and young planets (red and blue curves from [85]). The red curves are the SEDs based on the “hot start” models, while the blue curves are from the “cold start” models. For another comparison, the green curve is the SED of the protostar GM Aur scaled to 100 pc. At the top of each panel, the black curves indicate the transmission functions of J, H, K, L', M, and N bands. Left: the circumplanetary disk extends all the way to the planet surface. Right: the disk is truncated by the magnetic field from the young planet and the disk is undergoing magnetospheric accretion. For comparison, the SED of GM Aur is shown as the green curve. The blue and green dots/upper limit are detections for HD 169142b and HD100546b.
References


