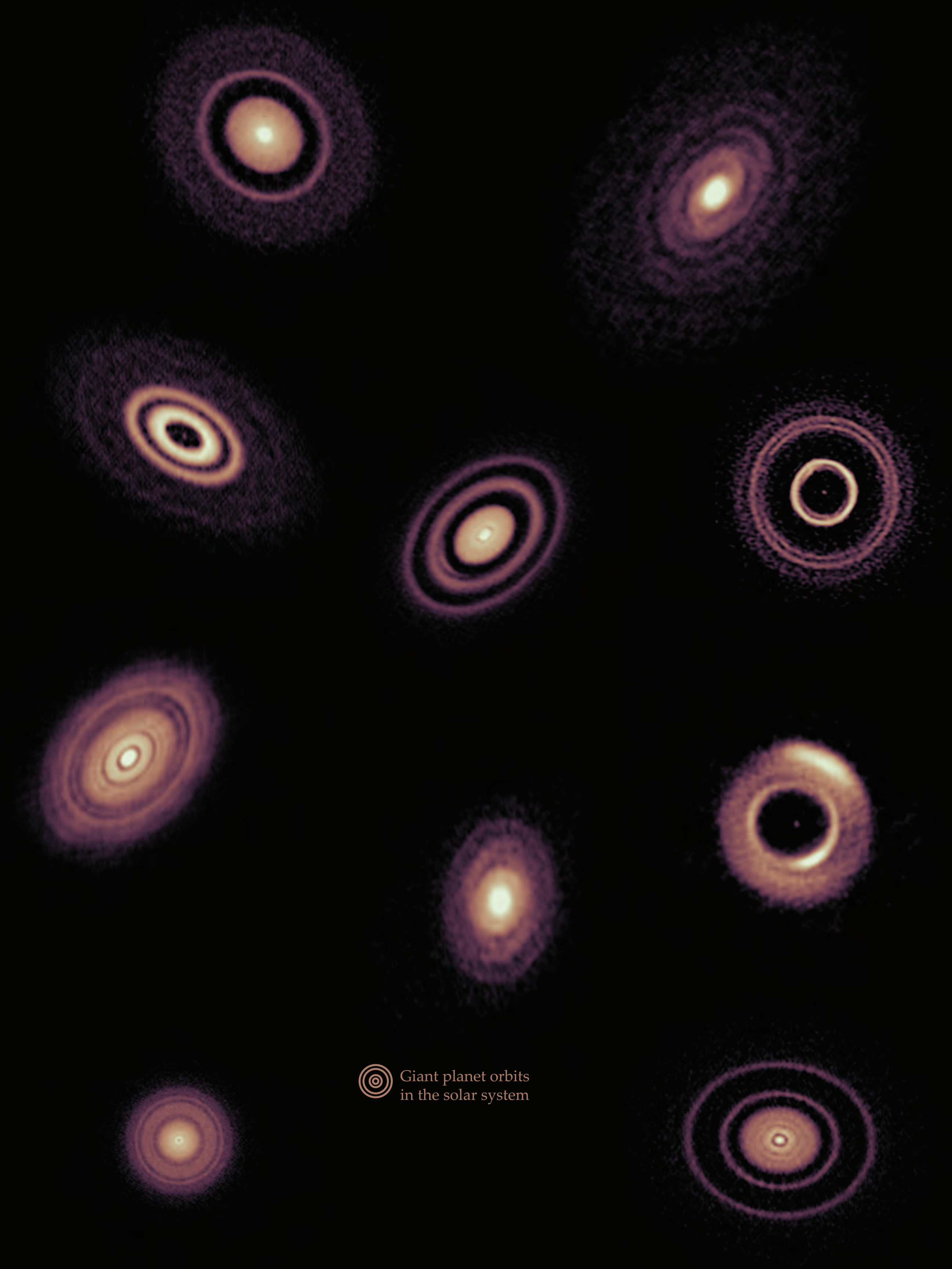
The structures of protoplanetary disks

Astronomical observations of gas, dust, and rocky material in the disks from which planets emerge help refine theoretical ideas about how they form.

Sean M. Andrews

Physics Today **74** (8), 36–41 (2021);
https://doi.org/10.1063/PT.3.4814



Millimeter-wavelength continuum emission from protoplanetary disks, captured by the Atacama Large Millimeter/Submillimeter Array. (Data adapted from refs. 6, 7, 14, and 15.)

Sean Andrews is an astrophysicist on the staff of the Smithsonian Astrophysical Observatory, part of the Center for Astrophysics|Harvard & Smithsonian, in Cambridge, Massachusetts.

THE STRUCTURES OF PROTOPLANETARY DISKS

Sean M. Andrews

Astronomical observations of gas, dust, and rocky material in the disks from which planets emerge help refine theoretical ideas about how they form.

It is striking to recognize that we have been aware of planets orbiting other stars for only the past quarter century. More than 4300 of those exoplanets have been cataloged by astronomers in that time. Statistical extrapolations from that population suggest that more than one planet orbits each star, on average.[1] The big questions in astrophysics research have always been concerned with origins. The abundance of new worlds allows us to establish the context for our solar system and to refine ideas about the creation, evolution, and biological potential of planets. Of the many exciting discoveries made about exoplanets, perhaps the most profound is their diversity. No one knows how common a planetary configuration like our solar system may be, but it's clear that a much broader range of planetary properties exists than we can find among our immediate neighbors. Much of that variety is thought to be imprinted at the epoch of planet formation.

Planetary systems are assembled in the disks of gas and solid particles that orbit young stars. Those disks are byproducts of angular-momentum conservation during star formation, and they are created when a rotating condensation in a giant cloud collapses under its own gravity. The properties and evolution of the disk material determine where planets form, how they grow, what they are made of, and how their orbital configurations can change. In turn, feedback from that early development of a planetary system influences the behavior of the disk. The epoch of planet formation is brief: Observational signatures of disks disappear in less than 10 million years. But even the relatively short burst of mutual interactions with their disk birth sites profoundly influences the properties of the exoplanets and solar-system bodies measured today, billions of years later.

It is difficult to learn more about how those processes work when astrophysicists use only direct measurements of young planets. The common techniques for finding planets are less effective when the planets are in their youth, primarily because the enhanced magnetic activity of their stellar hosts can hide planetary signals. Instead, much of the effort and progress relies on astronomical observations of disks and *in situ* measurements of planets and other bodies in the solar system. The central goal of that work is to quantify the spatial distributions of the physical and chemical conditions in disks, particularly to hunt for the telltale signatures of dynamical interactions between the planets and disk material. Measurements of the disk properties for different evolutionary states and environments are essential for refining the theoretical ideas that link young planets in their formation epoch with their descendants—the populations of mature exoplanets and the planets in the solar system.

A PLANET-FORMATION PRIMER

Even early observers of the solar system had an implicit understanding that planet formation occurs in a rotating disk, because they recognized that the planets orbit the Sun in the same direction and confined to a narrow ecliptic plane. The development of a modern planet formation theory began in the 1960s and advanced along parallel tracks in the planetary science and astronomy communities. By the 1990s, those efforts had converged into the "core-accretion" paradigm. That model considers the evolution of a trace (approximately 1% by mass) population of planetesimals—solid building blocks, roughly comparable to large asteroids—embedded in a massive gas disk.

In the original theory, the swarm of planetesimals grows slowly, over roughly a million years, by constructive collisions into a population of rocky planetary cores. If the growth is efficient enough to reach a critical mass (a few times Earth's mass), then each core can rapidly acquire an atmosphere with tens to hundreds of times Earth's mass and form a giant planet like Saturn or Jupiter. The masses, initial orbits, compositions, and subsequent evolution of such planets are primarily controlled by the spatial distributions of density and temperature in their progenitor material: the disk.

One of the major challenges for the core-accretion model is associated with initial conditions: How does a disk produce planetesimals in the first place? The solids that a disk inherits from the interstellar medium are small—perhaps a few microns across. The model's viability hinges on the ability of those small particles to grow in size by at least 10 orders of magnitude within a million years. Numerical simulations of the particle growth, grounded in results from microgravity collision experiments, indicate that millimeter- to centimeter-sized pebbles are produced easily.[2] Subsequent growth is problematic, though. Classical models assume a smooth disk structure, in which gas densities and temperatures decrease monotonically with distance from the host star. That would be expected from gravitational-collapse models, in which the primary heating mechanism is stellar irradiation.

In that case, the dynamical contribution of pressure support means the gas orbits slightly below the Keplerian velocity. As solids in a particular size range decouple from the gas, the sub-Keplerian flow imparts a drag force that saps the solids' angular momentum and sends them spiraling in toward the global pressure maximum near the star. That process is especially effective for pebbles: They migrate faster than they can collide and grow in the planet-forming zones of most disks.[3]

One family of possible solutions to the planetesimal growth problem invokes two modifications to the simple picture of growth by binary collisions. First, and most importantly, the solutions assume that the gas distributions in disks must not be smooth. Local pressure maxima perturb the gas flow, and thereby the aerodynamics of the solids. Even relatively low-amplitude pressure deviations can slow or stop migrating pebbles.[4] Second, the solutions presume that the concentrations of solids near such pressure traps can become high enough to trigger a growth instability or direct gravitational collapse that rapidly converts pebbles into much larger solids.[5] The key unifying hypothesis is that the production of planetesimals—and therefore the viability of the core-accretion model of planet formation—requires that localized perturbations, known as substructures, in the physical conditions are robust and fundamental aspects of protoplanetary disks.

Astrophysicists have proposed many different physical origins for the hypothesized disk substructures. Most fall into two broad categories. In one class, perturbations are associated with various fluid-dynamics processes in the gas disk—from turbulence, to hydrodynamic instabilities, to coherent magnetohydrodynamical structures, and beyond. The basic idea is that the gas disk naturally generates its own substructures, which then act as pressure traps to concentrate migrating solids and facilitate planetesimal formation. Another class of proposals considers that disk substructures are generated by interactions with already-forming young planets. Once a planet is sufficiently massive, it can clear a gap in the material around its orbit, drive spiral arms that shock the gas well away from its location, seed hydrodynamic instabilities (vortices, for instance), and potentially perturb more of the disk through its own migration. The important distinction between the two categories is that the first is a cause of planet formation, whereas

FIGURE 1. ALMA. Shown here are a few of the 66 antennas that make up the Atacama Large Millimeter/Submillimeter Array interferometer, located on the Chajnantor Plateau in the Atacama Desert in Chile. Over the past few years, ALMA's exquisite sensitivity and resolution have revealed that the disks around young stars are riddled with substructures. The details are revolutionizing our understanding of the planet formation process. (Image by D. Korden/ESO.)

the second is an effect of it. They are not mutually exclusive, but the latter implicitly presumes that an earlier generation of disk substructures was created by the former.

THE ALMA REVOLUTION

Measurements of the prevalence, locations, morphologies, sizes, and amplitudes of disk substructures are essential for assessing the roles of those mechanisms. Put simply, a deeper understanding of planet formation requires observations that find and characterize substructures. Pebbles are the optimal probes. They should show the highest-amplitude perturbations in their spatial distributions, given their concentrations near gas pressure maxima. Pebbles absorb and re-emit starlight as thermal continuum radiation, with a peak efficiency at the millimeter wavelengths detectable by radio telescopes. Although contrast is highest for the millimeter-wavelength intensity variations from substructures, resolution is a challenge. Stable perturbations to the gas disk have a characteristic size comparable to the pressure-scale height H (the ratio of the local sound speed and the Keplerian angular velocity). For typical disk properties, H is 5–10% of the roughly 10–100 AU separation from the host star. That size implies that most substructures span only a few astronomical units. (An AU, the mean Earth–Sun distance, spans roughly 1.5×10^{11} m.)

Given the distances to the nearest disks, those expected sizes suggest an angular-resolution requirement of 0.03 arcsec. That’s equivalent to resolving an airplane on the Moon. At a wavelength of 1 mm, that resolution corresponds to the diffraction limit for a telescope with a diameter of 10 km. Because a single aperture that size is an engineering impossibility, the solution is to link together many smaller apertures and form an interferometer. The necessary combination of sensitivity and resolution only recently came together with the commissioning of the revolutionary Atacama Large Millimeter/Submillimeter Array (ALMA), shown in figure 1. Constructed by an international partnership of 22 countries, ALMA is the largest ground-based astronomical observatory in the world. Located in the Atacama Desert in Chile, it is designed to measure continuum and spectral line emission at wavelengths of 0.3–7 mm with exquisite sensitivity. The 66 antennas can be arranged to probe spatial scales comparable to a telescope with a 16 km diameter.

Even before the highest-resolution configurations of the ALMA antennas were commissioned, hints of disk substructures were emerging in images that probed scales down to 10–20 AU. They included observations of “transition” disks, a subgroup representing about 10% of the population of all disks and defined by the depletion of continuum emission in a large inner zone, tens of astronomical units from the host star. But notably, annular variations of the continuum emission were being serendipitously uncovered in a few disks that did not fall into that category.

In the fall of 2014, the ALMA staff first experimented with the most extended antenna configurations to achieve very high resolution. One of their targets was the disk around the very young star HL Tauri, previously argued to have a smooth emission distribution on 15 AU scales. The new ALMA measurements at 4 AU resolution showed a stunning morphology riddled with substructures—narrow, concentric, dark gaps and bright rings distributed to about 100 AU from the star (see the first disk[6] above the one in the bottom left in the gallery on page 36). In the next available observing season that year, I worked with a team that used a similar ALMA antenna configuration to measure another ring and gap morphology in the disk around TW Hydrae (see the bottom left image[7] on page 36). As a few more cases followed, it became clear that the goals of finding and characterizing disk substructures were readily achievable with ALMA.

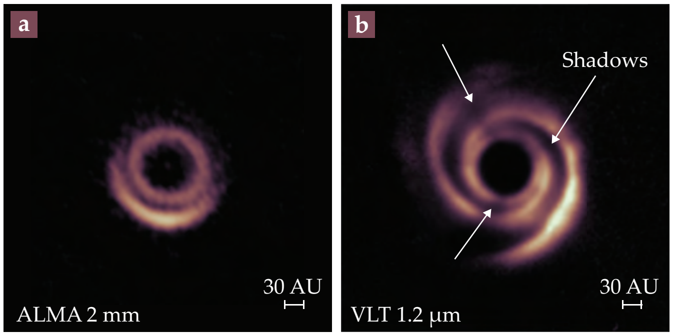


FIGURE 2. THE DISK around the young star SAO 206462, viewed in **(a)** the thermal continuum emission at a wavelength of 2 mm, from the Atacama Large Millimeter/Submillimeter Array (ALMA), and **(b)** the scattered starlight at a wavelength of 1.2 μm, from the 8.2-m-diameter Very Large Telescope (VLT). The strikingly different morphologies indicate the decoupled behaviors of the large particle in the midplane (panel a) and small solid particles in the atmosphere layers (panel b) of the disk. The narrow shadows (marked by arrows) in the scattered-light image suggest vertical perturbations at smaller separations from the central host star, not seen here. (Panel a adapted from ref. 16; panel b adapted from ref. 17.)

Over the past few years, high-resolution ALMA observations have continued to uncover examples.[8] Substructures are ubiquitous—they are found in all disks that have been explored with sufficient resolution to identify features at sizes comparable to the local scale height. The most common morphology is a set of narrow, concentric, and symmetric rings and gaps; a subset of cases shows more complex spirals or arc features. In the arc morphologies, rings and gaps are often also present. Substructures do not appear to have any preferential locations: They are found at any separation from the host star, from the inner resolution limit of a few astronomical units to the outer reaches where the millimeter continuum can be detected, more than 200 AU in some cases.

COMPLEMENTARY INSIGHTS

Although researchers have learned much about disk substructures in a short time, the samples available to study are still relatively few and biased. The focus so far has been on the brighter, larger disks that are preferentially found around stars with masses comparable to the Sun or larger. Although the samples of millimeter continuum measurements help probe systems more representative of the general disk population, the immediate focus in the field has shifted to more-detailed explorations of substructure properties in individual case studies. An important aspect of that work involves a joint analysis with complementary tracers of disk material, particularly those that are more directly sensitive to substructures in the gas phase.

One such tracer relies on the optical and IR starlight reflected by micron-sized dust particles suspended in the disk

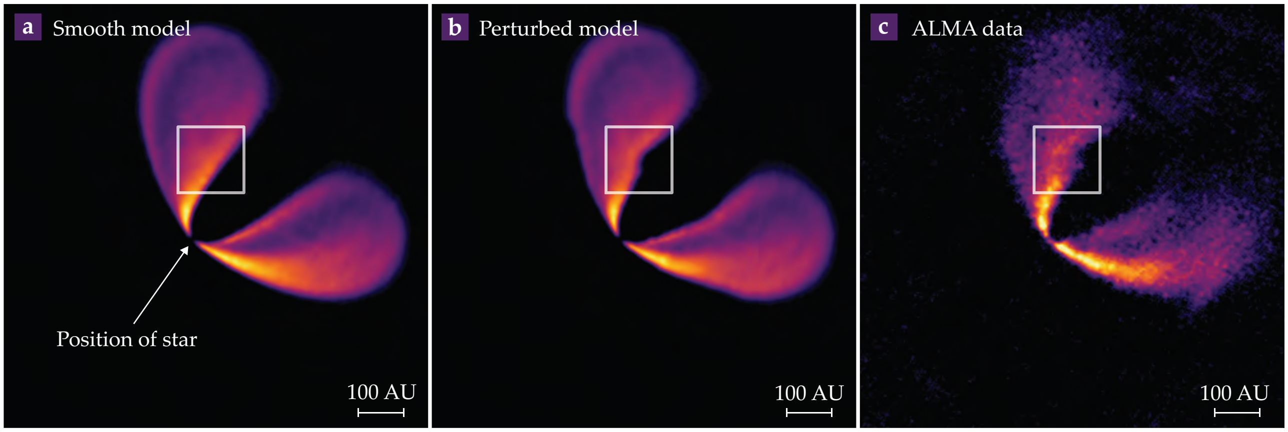


FIGURE 3. **KINEMATIC SUBSTRUCTURES** reveal themselves in subtle deviations from a smooth, Keplerian rotation model in spatially and spectrally resolved maps of molecular spectral line emission. These panels show the carbon monoxide emission maps for **(a)** a smooth model, **(b)** the same model perturbed by a planet twice as massive as Jupiter orbiting at 260 AU, and **(c)** real Atacama Large Millimeter/Submillimeter Array (ALMA) observations of the disk around a young star. The kinks in the white boxes of panels b and c are indicative of a kinematic perturbation to the gas flow. Note that the panels plot only a portion of the spectral line and don't resemble the usual morphology of a disk. (Models in panels a and b courtesy of Jaehan Bae; data in panel c adapted from ref. 14.)

atmosphere—the gas that extends vertically above the dense disk midplane.[9] With the *Hubble Space Telescope* or large ground-based observatories with adaptive optics systems to correct for the blurring effects of Earth's atmosphere, the scattered light can be measured at a resolution comparable to ALMA's. In many cases, the observed morphology is quite different from the millimeter continuum emission. Figure 2 shows a striking example. Various factors contribute to those contrasting appearances but the key insights are that each tracer probes a different vertical location in the disk and a different phase of the disk material (either the small dust grains that are coupled with the gas or the pebbles that are not). A combined analysis of those complementary tracers can be used to study the three-dimensional behavior of substructures and the associated gas–solid interactions, both of which are crucial for discriminating between models of substructure origins.

Another interesting aspect of scattered-light images is their sensitivity to vertical substructures in the inner disk, even well below the resolution limit. Vertical perturbations at small disk radii can block starlight from illuminating the material at much larger distances. The intensity variations induced by those shadows are common, and sometimes their motions can even be tracked over time. That behavior is often attributed to a warped geometry, an orbiting perturber, or stochastic upwelling of material near the inner disk edge. In any case, the shadows suggest again that substructures are 3D and persist even on very small scales.

A complementary option for tracing substructures is to directly probe the gas reservoir. Most of the disk mass is cool molecular hydrogen gas, but because of its lack of a permanent dipole moment and typical disk conditions, that reservoir of material is effectively dark. ALMA, however, can observe the spectral line emission from pure rotational transitions of other simple molecules, such as carbon monoxide, to probe the gas distribution.[10] Those measurements can be made simultaneously with the millimeter continuum and therefore achieve comparable resolution. The emission from different spectral lines probes the gas in different vertical layers. Combined measurements of the lines allow researchers to reconstruct the 3D behavior of substructures based on the variations in their line intensities and ratios. Moreover, comparing the behavior of various gas species can help disentangle any compositional changes associated with those substructures. That mapping of spectral-line-intensity variations to reveal more about substructures is really just getting started. Even so, early results are promising.

Alternatively, the same spectral line data can be used to probe substructures through their kinematic perturbations of the gas velocities.[11] ALMA observations provide spatially resolved maps of the spectral line emission at resolutions of 100 m/s or better in projected velocity. That's sufficient to measure deviations as small as a few percent from a smooth Keplerian flow. The larger-amplitude deviations can be seen directly in spectral line images, as demonstrated in figure 3. Observations of those deviations in multiple spectral lines could be used to reconstruct the 3D behavior of the gas flows near substructures and thereby determine the associated pressure gradients around the local maximum. Linking the dynamical information with the complementary, intensity-based constraints from spectral lines, the millimeter continuum, and scattered light promises to provide detailed quantitative constraints on the substructure properties. That information, coupled with hydrodynamics simulations, will be essential for determining the mechanisms responsible for generating the foundational perturbations in disks.

CURRENT LANDSCAPE, NEW DIRECTIONS

Planet formation research has experienced a profound shift in the past few years, as observations have forced astronomers to appreciate that small-scale substructures are an essential aspect of the disks around young stars. As I've mentioned, those perturbations to the conditions of a disk appear to be ubiquitous and are usually symmetric and small, with sizes comparable to the local pressure-scale height. The available samples show that they do not seem to have a preferential location in the disk or any association with their host stars. Concentric

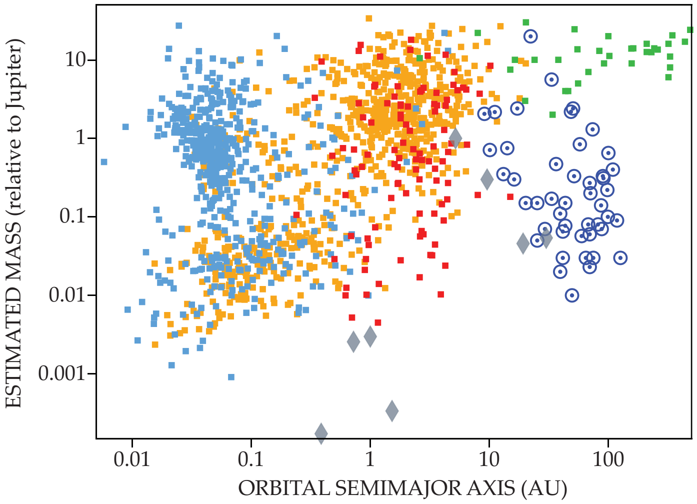


FIGURE 4. **ESTIMATED MASSES** and orbital locations for hypothesized planets (blue) that would clear the gaps in young disks observed with the Atacama Large Millimeter/Submillimeter Array,[12,18] compared with the known population of exoplanets (squares) and planets (diamonds) in our solar system. The exoplanet data are colored according to how they were detected: light blue for transits across a star, orange for radial velocity variations, red for gravitational microlensing, and green for direct imaging. (Data collated from the NASA Exoplanet Archive.) The masses and orbital locations inferred from disk substructures mostly occupy a swath of parameter space that has not yet been well searched for mature exoplanets. New facilities, however, will provide access to the older exoplanets with those masses and orbital locations and thereby enable tests of planetary systems' evolutionary pathways.

rings and gaps seem to be the most common morphology. Although researchers have learned much about substructures, they are still in the early stages of interpreting high-resolution data to determine the substructures' properties.

In general, the mechanisms that form those substructures remain unclear. Researchers do not yet know whether they are sites of planetesimal formation or signs that planetary systems already exist and are interacting with material in their birth sites. For some of the most striking examples, support is growing for the latter possibility. Much of the confidence in that hypothesis comes from hydrodynamic simulations of planet–disk interactions, which predict the gaps and rings observed in the millimeter continuum emission, and from the few cases that show pronounced and spatially localized perturbations in the spectral line emission.[12] Moreover, at least one case exists—the remarkable PDS 70 system—in which young giant planets still accreting their own circumplanetary material have been directly imaged and are clearly associated with disk substructures.[13]

If support for that possibility firms up and can be generalized, the simulations suggest that giant planet formation is well underway at 10–100 AU separations at ages less than a million years old. That's much faster than would naively be expected in the classical core-accretion model. Figure 4 shows the best estimates of young planet masses and orbits inferred from disk substructures, relative to the known planet population. The efficiency implied by those inferred planetary properties suggests that the epochs of planet and star formation overlap. Planetesimal and planetary core formation are therefore more robust than one would traditionally expect. An important test of that idea is to search for substructures in even younger disks, still embedded in their star-forming envelopes.

Ultimately, the final assessment of the viability of the planet–disk interactions hypothesis will come through direct imaging searches for young planets in the gaps in those disks. Some of that work is already underway and takes advantage of the fact that those systems shine substantially brighter than just their photospheres, because of their own compact circumplanetary dust disks and associated accretion flows. But the near future promises a considerable improvement in the sensitivity to lower-mass planets at smaller separations from their host stars. That improvement will come from the *James Webb Space Telescope*, planned to launch this autumn, and a new suite of large, 30-m-diameter ground-based observatories planned to be constructed over the next decade. Eventually, statistical comparisons of the distributions in mass and orbital semimajor axes for the young planets and the older exoplanet populations will provide a definitive constraint on planetary migration histories.

Studying disk substructures in increasing detail is also exciting. Astronomers are looking forward to improvements in existing facilities, such as ALMA, and to the development of new observatories that will push the resolution frontier. For example, the National Radio Astronomy Observatory's proposed next-generation Very Large Array, a radio interferometer that will specialize in longer, centimeter-scale wavelengths, will be able to resolve substructures at sub-astronomical-unit scales in the innermost parts of nearby protoplanetary disks. The combined access to such features from new and existing facilities promises that protoplanetary disk observations will continue to help refine our understanding of planet formation in the years to come.

I thank David Wilner and Rich Teague for their comments and suggestions, and Jaehan Bae, Myriam Benisty, and John Carpenter for kindly sharing their results to help with the figures.

REFERENCES

1. J. N. Winn, D. C. Fabrycky, *Annu. Rev. Astron. Astrophys.* **53**, 409 (2015).
2. T. Birnstiel, M. Fang, A. Johansen, *Space Sci. Rev.* **205**, 41 (2016).
3. S. J. Weidenschilling, *Mon. Not. R. Astron. Soc.* **180**, 57 (1977).
4. P. Pinilla et al., *Astron. Astrophys*. **538**, A114 (2012).
5. E. Chiang, A. N. Youdin, *Annu. Rev. Earth Planet. Sci.* **38**, 493 (2010).
6. ALMA Partnership et al., *Astrophys. J. Lett*. **808**, L3 (2015).
7. S. M. Andrews et al., *Astrophys. J. Lett*. **820**, L40 (2016).
8. S. M. Andrews, *Annu. Rev. Astron. Astrophys*. **58**, 483 (2020).
9. A. Garufi et al., *Astron. Astrophys*. **620**, A94 (2018).
10. A. Isella et al., *Phys. Rev. Lett*. **117**, 251101 (2016).
11. C. Pinte et al., *Astrophys. J. Lett*. **860**, L13 (2018).
12. S. Zhang et al., *Astrophys. J. Lett*. **869**, L47 (2018); J. Bae, Z. Zhu, L. Hartmann, *Astrophys. J*. **850**, 201 (2017).
13. M. Keppler et al., *Astron. Astrophys*. **617**, A44 (2018); S. Y. Haffert et al., *Nat. Astron*. **3**, 749 (2019).
14. S. M. Andrews et al., *Astrophys. J. Lett*. **869**, L41 (2018).
15. J. Huang et al., *Astrophys. J*. **891**, 48 (2020); S. Pérez et al., *Astron. J*. **158**, 15 (2019); R. Dong et al., *Astrophys. J*. **860**, 124 (2018).
16. P. Cazzoletti et al., *Astron. Astrophys*. **619**, A161 (2018).
17. T. Stolker et al., *Astrophys. J*. **849**, 143 (2017).
18. J. Bae, P. Pinilla, T. Birnstiel, *Astrophys. J. Lett*. **864**, L26 (2018); G. Lodato et al., *Mon Not. R. Astron. Soc*. **486**, 453 (2019). PT