A tale of planet formation: from dust to planets
Beibei Liu$^{1,2}$ and Jianghui Ji$^{3,4}$

$^1$ Zhejiang Institute of Modern Physics, Department of Physics & Zhejiang University-Purple Mountain Observatory Joint Research Center for Astronomy, Zhejiang University, Hangzhou 310027, China
$^2$ Department of Astronomy and Theoretical Physics, Lund Observatory, Box 43, SE–22100, Sweden
$^3$ CAS Key Laboratory of Planetary Sciences, Purple Mountain Observatory, Chinese Academy of Sciences, Nanjing 210008, China
$^4$ CAS Center for Excellence in Comparative Planetology, Hefei 230026, China; bbliu@astro.lu.se, jijh@pmo.ac.cn

Received 2020 August 1; accepted 2020 September 4

Abstract The characterization of exoplanets and their birth protoplanetary disks has enormously advanced in the last decade. Benefitting from that, our global understanding of the planet formation processes has been substantially improved. In this review, we first summarize the cutting-edge states of the exoplanet and disk observations. We further present a comprehensive panoptic view of modern core accretion planet formation scenarios, including dust growth and radial drift, planetesimal formation by the streaming instability, core growth by planetesimal accretion and pebble accretion. We discuss the key concepts and physical processes in each growth stage and elaborate on the connections between theoretical studies and observational revelations. Finally, we point out the critical questions and future directions of planet formation studies.

Key words: planets and satellites: general — planets and satellites: formation — planets and satellites: dynamical evolution and stability — protoplanetary disks

1 INTRODUCTION

In this article, we review modern planet formation scenarios in the context of the core accretion paradigm. Since observation and theory are two closely-related aspects, we first recap the detection and characterization of exoplanets in Section 1.1 and protoplanetary disks in Section 1.2. The outline of general planet formation processes are given in Section 1.3, classified by the characteristic sizes of growing planetary bodies. Finally, we introduce the relevant topics that will be covered in the subsequent sections of the paper (Wang & Ip 2020).

1.1 Exoplanets

Half of the 2019 Nobel Prize in Physics was awarded to Michel Mayor and Didier Queloz, as an acknowledgement for their milestone discovery of the first exoplanet orbiting a main-sequence star. This is one of the most influential scientific breakthroughs in astronomy of the past decades. Already in 1995, the above two astronomers detected the exoplanet 51 Pegasi b around a nearby, Sun-like star in the constellation of Pegasus (Mayor & Queloz 1995). Such a discovery was extraordinary and unexpected at that time. It opened an entirely new era in astronomical observations. After that, the detection of planets beyond our Solar System has been enormously developed and grown into a rapidly evolving branch in astronomy.

One major exoplanet detection method is called radial velocity (RV, or Doppler spectroscopy). A star and its accompanying planet co-orbit their center of mass. Observers can see the periodic movement of the star induced by the planet. Due to the Doppler effect, the observed stellar spectral lines are blueshifted when the star approaches us and are red shifted when the star recedes from us. Therefore, the radial velocity of the star can be acquired by measuring the displacement of stellar spectral lines. Through this technique, the minimum mass of the planet can be obtained. Since we do not really observe the planet but infer it from the wobble of the central star, this is an indirect way to acquire information about the planet. The first exoplanet, 51 Pegasi b, was discovered by this method. Also, the radial velocity method was involved in most of the exoplanet discoveries in the early planet-hunting epoch before the launch of the Kepler satellite.

Another leading exoplanet detection method is called transit, which monitors the time variation of a star’s brightness to probe the existence of planet(s). When a
planet transits in front of its parent star, the surface of the star is partially blocked by the planet and hence the observed stellar flux drops accordingly. This periodic decrement in the stellar flux reflects the size ratio between the planet and the star. Therefore, this method can uniquely determine the radius of the planet. Compared to RV that requires high resolution spectroscopic measurements, transit is a photometric method, and is thus more efficient in detecting planets. Combining the above two methods together, we can know both the masses and radii, and therefore deduce the bulk densities and chemical compositions of the planets.

The Kepler satellite is recognized as the most successful planet hunting mission to date, which utilized transits in space to maximize the detection ability and efficiency (Borucki et al. 2010). The key to the success of the Kepler telescope is that it has both a large field of view and extremely high photometric precisions. More than 2600 confirmed exoplanets and 4000 planet candidates were detected by Kepler during its nine year operational lifetime (2000-2018). Thanks to the vastly increased number of exoplanets detected by Kepler, the analysis of planet properties from a statistical perspective has become feasible for the first time (Lissauer et al. 2011; Batalha et al. 2013; Burke et al. 2014).

The observed planets are incredibly diverse in terms of masses, sizes, compositions and orbital properties. As illustrated in Figure 1, the confirmed exoplanets span several orders of magnitude in their masses and orbital periods\(^1\). Based on the above two properties, exoplanets can be classified into the following types: hot Jupiters (Mayor et al. 1997), cold Jupiters (Zhu & Wu 2018), warm (hot) Neptunes (Dong et al. 2018), super-Earths (Borucki et al. 2011) and low-mass rocky planets. Figure 1 also marks one particular type of planets with orbital periods less than 1 day, which are called ultra-short period (USP) planets (Sanchis-Ojeda et al. 2014; Winn et al. 2018). Super-Earths strictly refer to the planets with radii between \(1\)–\(4\) \(R_{\oplus}\) in Figure 1(Borucki et al. 2011). A more general definition of super-Earths is also frequently used in literature studies, referring to planets with radii between Earth and Neptune (\(1<R_p<4\) \(R_{\oplus}\) or \(1<M_p\lesssim10\) \(M_{\oplus}\), Seager et al. 2007; Valencia et al. 2007). With this definition, super-Earths cover the ranges of terrestrial-like, rocky-dominated planets and sub-Neptune planets with non-negligible hydrogen envelopes.

We apply the latter definition of super-Earths in the following discussion of this review. Together with hot and cold Jupiters, these three major types of planets are currently the most well characterized and extensively studied samples in literature. We summarize the key observational findings and the underlying physical interpretations of these three types of exoplanets (also see reviews of Zhou et al. 2012 and Winn & Fabrycky 2015).

- Small planets are more common than large planets. When taking into account of the observational bias and for solar-type stars, the occurrence rates of hot Jupiters and cold Jupiters are \(1\%\) and \(5\%–10\%\), respectively, whereas the occurrence rate of super-Earths is \(30\%\) (Cumming et al. 2008; Howard et al. 2010; Mayor et al. 2011; Wright et al. 2012; Dong & Zhu 2013; Petigura et al. 2013; Zhu et al. 2018; Fernandes et al. 2019). Furthermore, planets are so ubiquitous that they greatly outnumber their host stars (Mulders et al. 2018; Zhu et al. 2018).

- The occurrence rate of giant planets exhibits strong dependences on both stellar mass (Johnson et al. 2007, 2010; Jones et al. 2016) and metallicity (Santos et al. 2004; Fischer & Valenti 2005; Sousa et al. 2011). The occurrence rate of super-Earths seems to be much more weakly dependent on stellar metallicity (Sousa et al. 2008; Buchhave et al. 2012, 2014; Wang & Fischer 2015; Schlaufman 2015; Zhu et al. 2016; Zhu 2019). Nevertheless, super-Earths are even more abundant around M-dwarfs compared to those around Sun-like stars (Howard et al. 2012; Bonfils et al. 2013; Dressing & Charbonneau 2015; Mulders et al. 2015; Yang et al. 2020).

- Hot Jupiters are nearly circular while the mean eccentricity of cold Jupiters is \(\sim0.25\) (Marcy et al. 2005). Obliquity defines the angle between the spin axis of the host star and the orbital angular momentum axis of the planet, which can be measured through the Rossiter-McLaughlin effect (Winn 2010). Many hot Jupiters manifest high obliquities, sometimes even polar or retrograde (Triaud et al. 2010; Albrecht et al. 2012). Theoretically, hot Jupiters were proposed to grow at further out disk locations and then migrate inward to the present-day orbits (Lin et al. 1996). Planet-disk interaction will lead the giant planets on circular and coplanar orbits (Lin & Papaloizou 1993; Artymowicz 1993; Ward 1997; Nelson et al. 2000), while the high eccentricities and inclinations of giant planets can originate from a Kozai-Lidov cycle (Kozai 1962; Lidov 1962) induced by a distant companion (Wu & Murray 2003; Fabrycky & Tremaine 2007; Naoz et al. 2011, 2013; Dong et al. 2014; Anderson et al. 2016) or planet-planet scatterings (Rasio & Ford 1996; Chatterjee et al. 2008; Juric & Tremaine 2008; Ford & Rasio 2008; Dawson & Murray-Clay 2013). The latter two planetary dynamical processes could result in the observed high obliquities. On the other hand, such misalignments could also arise

\(^1\) Data are adopted from http://exoplanet.eu/
Fig. 1 Orbital period and mass distribution of exoplanet populations, including hot Jupiters, cold Jupiters, warm-Neptunes, super-Earths and rocky-dominated terrestrial planets. A special type of planet with orbital periods less than 1 day is called an ultra-short period (USP) planet. Figure adopted from Ji & Huang (2020).

from re-orientation of the host star’s spin through internal waves (Rogers et al. 2012; Lai 2012) or tilted protoplanetary disks through binary-disk interactions (Lai 2014; Matsakos & Königl 2017; Zanazzi & Lai 2018).

– Based on Kepler data, multi-transit systems have relatively low eccentricities and inclinations while single-transit systems exhibit much higher eccentricities and inclinations (Tremaine & Dong 2012; Johansen et al. 2012; Fabrycky et al. 2014; Xie et al. 2016; Zhu et al. 2018). One hypothesis is that these single-transit planets come from multiple systems. Their orbits are further excited/disrupted by long-term planet-planet interactions or by outer companions, causing them to appear as “singles” in transit surveys (Pu & Wu 2015; Lai & Pu 2017; Mustill et al. 2017).

– Hot Jupiters/Neptunes seldom have nearby companions up to a few AU (Steffen et al. 2012; Dong et al. 2018), consistent with the Kozai-Lidov cycle and planet-planet scattering scenarios (Mustill et al. 2015). On the contrary, nearly half of warm Jupiters co-exist with low-mass planets (Huang et al. 2016). Cold Jupiters also seem to be commonly accompanied by close-in, super-Earths (Zhu & Wu 2018; Bryan et al. 2019). Besides, 25%−30% of the systems with a cold Jupiter are found to host additional giant planets (Wright et al. 2009; Wittenmyer et al. 2020).

– The period ratios of adjacent planet pairs neither show strong pile-ups at mean motion resonances (MMRs) nor uniform distribution in Kepler data. These planets exhibit an asymmetric distribution around major resonances, such as 2:1 and 3:2 MMRs (fig. 6 of Winn & Fabrycky 2015). Different scenarios are proposed to explain the above features, including tidal damping (Lithwick & Wu 2012; Batygin & Morbidelli 2013; Lee et al. 2013; Delisle & Laskar 2014; Xie 2014), retreat of the inner magnetospheric cavity (Liu et al. 2017; Liu & Ormel 2017), resonant overstability (Goldreich & Schlichting 2014), interaction with planetesimals (Chatterjee & Ford 2015), stochastic migration in highly turbulent disks (Rein 2012; Batygin & Adams 2017) or in shock-generated inviscid disks (McNally et al. 2019; Yu et al. 2010), mass growth of a planet (Petrovich et al. 2013; Wang & Ji 2017), and dynamical instability of tightly packed planetary chains (Izidoro et al. 2017, 2019; Ogihara et al. 2015, 2018).

– The occurrence rate of close-in super-Earths has a bimodal radius distribution, with a factor of two drop at $R_p\sim 1.5−2R_\oplus$ (Fulton et al. 2017; Fulton & Petigura 2018; Van Eylen et al. 2018). This so-called planetary radius valley implies a composition transition from rocky planets without H/He gaseous envelopes to planets with envelopes of a few percent in mass (Lopez & Fortney 2013; Owen & Wu 2013). The above radius gap can be explained by the gas mass loss due to stellar photoevaporation (Owen & Wu 2017; Jin & Mordasini 2018) or core-powered...
heating (Ginzburg et al. 2018; Gupta & Schlichting 2019). Besides, giant impacts may also contribute to this compositional diversity by stripping the planetary primordial atmospheres through disruptive collisions (Liu et al. 2015; Inamdar & Schlichting 2016). Based on the photoevaporation model, (Owen & Wu 2017) deduced that the compositions of these super-Earths are rocky dominated, ruling out low-density, water-world planets. Although this interpretation should be taken with caution, the water-deficit outcome may be caused by the fact that short-lived radionuclides dehydrate planetesimals during their early accretion phase (Lichtenberg et al. 2019), thermal effects take place in planetary interiors during long-term evolution phase (Vazan et al. 2018), or the planets experience a runaway greenhouse effect and lose substantial surface water through photo-dissociation (Luger & Barnes 2015; Tian & Ida 2015).

The above results are summarized based on the current demographic and orbital properties of exoplanets. For knowledge of exoplanet atmospheres, we recommend a recent review by Zhang (2020). Figure 2 exhibits the launched and future planned space missions for exoplanet detection and characterization from the National Aeronautics and Space Administration (NASA), European Space Agency (ESA) and China National Space Administration/Chinese Academy of Sciences (CNSA/CAS). For instance, the successor of the Kepler mission, Transiting Exoplanet Survey Satellite (TESS) which was launched in 2018, aims at discovering short period planets around nearby stars (Ricker et al. 2015; Huang et al. 2018a). Compared to Kepler, the advantage of TESS is that the target stars are easier for ground-based and space-based follow-up characterization observations.

On the other hand, three Chinese space missions have been initiated and approved for detection of exoplanets in the coming decades. The Chinese Space Station Telescope (CSST), scheduled for launch in 2024, will survey mature Jupiter-like planets, Neptunes and super-Earths around solar-type stars using a high-contrast imaging technique, which expects to discover tens of exoplanetary candidates and brown dwarfs. The Closeby Habitable Exoplanet Survey (CHES) mission aims at searching for terrestrial planets in habitable zones around solar-type stars within 10 pc by using astrometry in space. CHES will observe the target stars with high astrometric precisions of microarcsecond at the Sun-Earth L2 point. The mission expects to discover at least 50 Earth-like planets or super-Earths around 100 FGK stars with well-determined masses and orbital parameters. Miyin, on the other hand, is designed for detecting habitable exoplanets around nearby stars with interferometry. To achieve direct imaging of these exoplanets and assess their habitability, the mission will launch spacecrafts with groups of telescopes working in the mid-infrared wavelengths, which ensures an extremely-high spatial resolution of 0.01 arcsecond.

1.2 Protoplanetary Disk Observation

Planets form in protoplanetary disks surrounding their infant stars. Since the birth and growth of planets are tightly related with their forming environment, studying the physical and chemical conditions of protoplanetary disks becomes essential to understanding the planet formation processes. A large number of young protoplanetary disks have been observed and extensively studied in literature. Here we briefly introduce the up-to-date observations of disk substructures and provide several hints for the existence of emerging planets. Discussions on the disk solid masses and dust sizes will be presented in Section 2.2.

Thanks to the unprecedentedly high sensitivity and angular resolution of the Atacama Large Millimeter/submillimeter Array (ALMA), we now have the capability to reveal the disk structures at a spectacular level of detail. The spatial resolution of the resolved disks in nearby star forming regions is approximately 3–5 AU. Axisymmetric rings, gaps, inner cavities and spirals are commonly observed among these disks over a wide range of ages and masses of stellar hosts (Andrews et al. 2018; Huang et al. 2018b; Long et al. 2018). These disk substructures are in disagreement with the traditional picture of a smooth disk profile. For instance, Figure 3(a) shows the disk of HL Tau with a series of concentric bright rings separated by faint gaps from the dust continuum emission (ALMA Partnership et al. 2015).

There are different explanations for the formation of the above ring-like substructures. These features can be explained by grain growth (Zhang et al. 2015) or dust sintering (Okuzumi et al. 2016) at the condensation fronts of major volatile species, zonal flows in magnetized disks (Flock et al. 2015), a combined effect of the above mechanisms (Hu et al. 2019) or secular gravitational instability (Takahashi & Inutsuka 2014). Apart from those interpretations, the most widely accepted scenario is that these substructures are induced by gap-opening planets (Pinilla et al. 2012b; D’orpiro et al. 2015; Dong et al. 2015c; Jin et al. 2016; Fedele et al. 2017; Liu et al. 2018; Zhang et al. 2018; Liu et al. 2019d; Eriksson et al. 2020). One important note is that, if these ring-like features are indeed the planet origin, the young age of HL Tau (t<1 Myr) implies that planet formation may be faster than previously thought. In addition, this early formation hypothesis may also be supported by Harsono et al. (2018), who analyzed the radial distributions of disk dust and
gas around TMC1A and suggested that millimeter-sized grains have already formed around such a young Class I object at an age of $\sim 10^5$ yr.

In addition to the disk dust-component, the properties of gas, such as gas velocities, can also be obtained from molecular line emissions of CO isotopes. The kinematic deviations of gas velocities from Keplerian flows, together with the detected dust gaps at the same disk locations, strongly support the existence of embedded planets associated with gap-opening (Teague et al. 2018; Pinte et al. 2018, 2020).

The revealed spiral structures in disks from the scattered light images have been proposed to feature different origins as well (Muto et al. 2012; Grady et al. 2012; Stolker et al. 2016; Benisty et al. 2016; Muro-Arena et al. 2020). For instance, these patterns can be explained by density waves excited by the planets (Zhu et al. 2015; Dong et al. 2015b; Fung & Dong 2015; Bae & Zhu 2018). In the early phase when the disk is massive and self-gravitating, the gravitational instability can also induce large-scale spiral arms (Lodato & Rice 2005; Dong et al. 2015a). In addition, it might also be caused by the shadow from the warped disk (Montesinos et al. 2016), or a Rossby wave instability triggered vortex (Li et al. 2000; Huang et al. 2019).

Due to diverse explanations, previously mentioned disk kinematics could be treated as indirect indications of the planets. More straightforward evidence of young planets and ongoing planet formation is demonstrated as follows. The first clue comes from the RV measurements of young stars, where hot Jupiter candidates are reported around CI Tau (Johns-Krull et al. 2016) and V830 Tau (Donati et al. 2016). Furthermore, Plavchan et al. (2020) discovered a Neptune-sized planet co-existing with a debris disk around the nearby M dwarf star AU Mic by transit surveys. All above stars are in their pre-main-sequence with an age of approximately 10 Myr. In addition, two embedded planets have been detected in PDS 70’s protoplanetary disk by using the high-contrast imager Spectro-Polarimetric High-contrast Exoplanet REsearch (SPHERE) on European Southern Observatory’s Very Large Telescope (ESO’s VLT; Keppler et al. 2018; Müller et al. 2018). Figure 3(b) displays a synthetic image of the PDS 70 system, where two planets reside inside the gap of their protoplanetary disk. This is the first time that young planets have been directly imaged in their birth environment. Further analyses with submillimeter continuum and resolved Hα line emissions indicated the presence of a circumplanetary disk (Isella et al. 2019) as well as the proceeding gas accretion onto planet (Haffert et al. 2019). All these findings, together with the previous results from the disk morphologies/kinematics, provide valuable constraints on how, when and where planets can form.

1.3 Overview of Planet Formation

The study of planet formation is a highly multi-scale and multi-physics subject. The size increment of a planetary
Fig. 3  *Left* (a): protoplanetary disk of HL Tauri with multiple rings and gaps from ALMA dust continuum emission. Credit: ALMA (ESO/NAOJ/NRAO). *Right* (b): synthetic image of two young planets inside the gap of the disk around PDS 70. The data are adopted from ALMA, and SPHERE and MUSE (The Multi Unit Spectroscopic Explorer) instruments on ESO’s Very Large Telescope. Credit: ALMA (ESO/NAOJ/NRAO) A. Isella; ESO. We note that the above two images are not on the same spatial scale.

Fig. 4  Planet formation flow chart with characteristic planet bodies including μm sized dust grains, mm/cm sized pebbles, 100 km sized planetesimals, >1000 km sized protoplanets and final planets (either gas-deficit terrestrial planets, super-Earth planets or gas-rich giant planets). The physical processes are marked in red text, where the orbital migration is shown in the left while the mass growth is exhibited in the right.
body varies by more than 17 orders of magnitude, from (sub)micron-sized (dust grain) to $>10^3$ km (super-Earth/gas giant planet). It results in different physical mechanisms operating at different length scales and in different growth stages. We categorize the planetary bodies into four characteristic size objects: µm-sized dust grains, millimeter/centimeter (mm/cm)-sized pebbles, 100-km-sized planetesimals, and larger than 1000-km-sized protoplanets\(^3\)/planets. The final planets are either rocky-dominated terrestrial planets/super-Earths, or gas-dominated giant planets. Chronologically, the planet formation can be classified into the following three stages: from dust to pebbles (Sect. 2), from pebbles to planetesimals (Sect. 3) and from planetesimals to protoplanets/planets (Sects. 4, 5 and 6).

Figure 4 is a sketch of planet formation with characteristic size bodies and dominant physical processes. Small dust grains coagulate into larger particles in the beginning. Direct sticking is nevertheless stalled for pebbles of roughly mm/cm size (Güttler et al. 2010). Other mechanisms are needed to make the growth of larger bodies proceed. One leading mechanism is the streaming instability (Youdin & Goodman 2005), which clusters pebbles and directly collapses into planetesimals by the collective effect of self-gravity. The subsequent growth of planetesimals can proceed by accreting surrounding planetesimals and/or pebbles that drift inward from the outer part of the disk. When the core mass reaches a critical value ($\sim 10 M_\oplus$, Pollack et al. 1996) and there is still ample disk gas left, the protoplanets can accrete surrounding gas rapidly to form massive gas giant planets with a timescale much shorter than the disk lifetime. Otherwise, protoplanets only modestly accrete gas and form low-mass terrestrial planets or super-Earths.

Besides the mass growth, planetary bodies also interact and transfer angular momentum with disk gas, which induce orbital migration. For instance, solid particles and small planetesimals mainly feel aerodynamic gas drag, and their orbital decay is termed radial drift (Adachi et al. 1976; Weidenschilling 1977). Meanwhile, large planetesimals/planets exert gravitational forces with disk gas, and the corresponding movement is called planet migration (Goldreich & Tremaine 1979, 1980; Lin & Papaloizou 1986).

We discuss how planetary bodies grow at each stage in the following sections. Our focus is how state-of-the-art planet formation models fit into frontier observational studies. Due to the length of the paper, we do not go deep into the topics of planet migration and gas accretion in the gas-rich disk phase, and the long-term dynamical evolution of planetary systems in the gas-free phase. The key relevant questions that will be discussed in this paper are listed as follows:

- What challenges are involved in dust coagulation? What is the characteristic size of the particles that µm-sized dust grains can directly grow into (Sect. 2)?
- How can planetesimals form by the streaming instability? When and where are planetesimals likely to form (Sect. 3)?
- What are the differences and features between planetesimal accretion (Sect. 4) and pebble accretion (Sect. 5)?
- Which is the dominant accretion channel for planetesimal/protoplanet growth (Sect. 6)?

2 FROM DUST TO PEBBLES

In this section we visit the first stage of planet formation: the growth and radial drift of dust grains. We focus on the comparison between state-of-the-art theoretical and laboratory studies and the latest disk observations.

2.1 Theoretical and Laboratory Studies

2.1.1 Radial drift of solid particles

Since the gas in protoplanetary disks is pressure-supported, it rotates around the central star at a velocity $v_{φ,g} = (1 - η)v_K$, where $φ$ refers to the azimuthal direction in cylindrical coordinates, $v_K ≡ \sqrt{GM_*/r}$ is the Keplerian velocity at the radial disk distance $r$, $G$ is the gravitational constant and $M_*$ is the mass of the central star. The headwind prefactor that measures the disk pressure gradient is given by (Nakagawa et al. 1986)

$$\eta = -\frac{c_s^2}{2v_K^2} \frac{\partial \ln P_g}{\partial \ln r} = \frac{1.5 - 0.5k_T - k_Σ}{2} \left( \frac{H_g}{r} \right)^2,$$

where $c_s$ is the gas sound speed, $P_g$ is the gas pressure, $H_g$ is the gas disk scale height, $h_g = H_g/r$ is the disk aspect ratio, and $k_Σ$ and $k_T$ are the gradients of gas surface density and temperature, respectively. The relative velocity between $v_{φ,g}$ and $v_K$ is often referred to as the headwind velocity $\eta v_K$. In most regions of the protoplanetary disk, the pressure gradient is negative, and gas orbits at a sub-Keplerian velocity. All notations utilized in this paper are listed in Table 1.

There are various disk models used for studying young protoplanetary disks, of which the Minimum Mass Solar Nebula (MMSN) model is most commonly adopted. The
MMSN model represents the minimum amount of solids necessary to build the Solar System planets (Hayashi 1981), which is given by

\[ \Sigma_g = 1700 \left( \frac{r}{1 \text{ AU}} \right)^{-3/2} \text{ g cm}^{-2}, \]
\[ T_g = 280 \left( \frac{r}{1 \text{ AU}} \right)^{-1/2} \text{ K}, \]
\[ h_g = 3.3 \times 10^{-2} \left( \frac{r}{1 \text{ AU}} \right)^{1/4}. \]

Therefore, at a 1 AU orbital distance \( \eta=1.8 \times 10^{-3} \) and the headwind velocity \( \eta v_K \approx 50 \text{ m s}^{-1} \). Here we choose the MMSN as a canonical disk model. Other disk models of surface densities include the minimum mass extrasolar nebula (MMEN) model based on mass extrapolation from short period exoplanets (Chiang & Laughlin 2013), obtained from the emission of dust disks at (sub)millimeter wavelengths (Andrews et al. 2009) or inferred from the disk accretion rate and viscous theory (Lynden-Bell & Pringle 1974; Hartmann et al. 1998). The disk heating source and dust opacity determine the gas temperature profile. The dominant heating mechanisms include viscous dissipation (Ruden & Lin 1986; Garaud & Lin 2007) and stellar irradiation (Chiang & Goldreich 1997; Bell et al. 1997; Dullemond et al. 2001; Dullemond & Dominik 2004).

Unlike gas, a solid particle does not feel the pressure gradient force and tends to move at the Keplerian velocity. The solid particle experiences a hydrodynamic gas drag when its velocity deviates from that of the gas (Whipple 1972; Weidenschilling 1977)

\[ F_{\text{drag}} = \begin{cases} \frac{4\pi}{3} \rho_s R^2 \nu_{th} \Delta v & [\text{Epstein drag law}], \\ -\frac{\pi}{2} C_D R^2 \rho_g \Delta v v & [\text{Stokes drag law}], \end{cases} \]

where \( R \) is the radius of a spherical particle, \( \rho_s \) is the gas density, \( \nu_{th}=\sqrt{8/\pi c_s} \) is the mean thermal velocity of the gas, \( \Delta v=v-v_K \) is the relative velocity between the particle and gas, and \( C_D \) is given by

\[ C_D = \begin{cases} 24 R_e^{-1} & R_e < 1, \\ 24 R_e^{-0.6} & 1 \leq R_e \leq 800, \\ 0.5 & R_e > 800. \end{cases} \]

The drag coefficient \( C_D \) depends on the particle’s Reynolds number \( R_e=2R\Delta v/\nu_{\text{mol}} \), where \( \nu_{\text{mol}}=\lambda_{\text{mfp}} \nu_{th}/2 \) is the kinematic molecular viscosity, \( \lambda_{\text{mfp}}=m_{\text{mol}}/\sigma_{\text{mol}} \rho_g \) is the gas mean free path, \( m_{\text{mol}}=\mu m_H \) and \( \sigma_{\text{mol}}=2 \times 10^{-15} \text{ cm}^{-2} \) are the mass and collisional cross section of the gas molecule respectively, and \( \mu=2.33 \) and \( m_H=1.67 \times 10^{-24} \text{ g} \) are the gas mean molecular weight and the hydrogen atom mass respectively.

The gas drag acceleration can also be written as \( \alpha_{\text{drag}}=-(v-v_K)/t_{\text{stop}} \), where \( t_{\text{stop}} \) is the stopping time of the particle, which quantifies how fast the particle adjusts its velocity toward the surrounding gas. The stopping time of the particles varies in different regimes. For instance,

\[ t_{\text{stop}} = \begin{cases} \frac{R\rho_s}{4R} & \text{when } R < 9/4\lambda_{\text{mfp}} \quad [\text{Epstein regime}], \\ 9\lambda_{\text{mfp}} \nu_{th} \rho_g & \text{when } R \geq 9/4\lambda_{\text{mfp}} \quad [\text{Stokes regime}]. \end{cases} \]

We note that the above expression in the Stokes regime holds when \( R_e \lesssim 1 \). As a particle’s size increases, \( F_{\text{drag}} \) eventually becomes quadratic in \( \Delta v \), and the stopping time is inversely proportional to \( \Delta v \).

The stopping time is also widely expressed in a dimensionless form \( \tau_s=t_{\text{stop}}/\Omega_K \), where \( \tau_s \) is termed the Stokes number, and \( \Omega_K=\Omega_K/r \) is the Keplerian angular frequency. Generally, pebbles are considered as mm/cm sized small rocks. However, from a hydrodynamical perspective, pebbles are specifically referred to as solid particles with a range of Stokes number approximately from \( 10^{-3} \) to 1. As will be demonstrated in Section 5, particles with such Stokes numbers are marginally coupled to the disk gas, and can be efficiently accreted by larger protoplanetary bodies, such as planetesimals/planets.

The radial and azimuthal velocity of the solid particle with respect to the Keplerian motion is given by Nakagawa et al. (1986),

\[ \begin{align*}
\nu_r &= -\frac{2}{1+\tau_s} \eta v_K + \frac{1}{1+\tau_s} \nu_{v,G}, \\
\nu_{v,G} &= -\frac{1}{2(1+\tau_s)} \eta v_K + \frac{\tau_s}{2(1+\tau_s)} \nu_{r,G},
\end{align*} \]

where \( \nu_{r,G} \) in the second term on the right side of the equation is the gas radial velocity due to disk accretion, much lower than the headwind velocity \( \eta v_K \) in the first term.

From Equation (6), the radial velocity of the solid particle peaks at \( \tau_s=1 \). For the MMSN, such particles are roughly meter-sized at 1 AU and cm-sized at 10 AU. They are strongly affected by the gas and drift toward the central star within a timescale of approximately 100 orbits. For pebble-sized particles of \( \tau_s \approx 10^{-2} - 1 \) and in the protoplanetary disk regions of \( r \lesssim 50 \text{ AU} \), the radial drift timescale is shorter than the gas disk lifetime (\( \sim 3 \text{ Myr} \). Haisch et al. 2001). Therefore, due to the rapid inward drift, the survival of these high Stokes number particles in protoplanetary disks is a long-standing conundrum in
planet formation (Adachi et al. 1976; Weidenschilling 1977).

2.1.2 Dust coagulation

The primordial solids in protoplanetary disks are dust grains that originate from the interstellar medium (ISM). These solid particles follow a size distribution $n(R) \propto R^{-3.5}$, in a range between nanometer and (sub)micron (Mathis et al. 1977). The microphysics of grain growth is not controlled by gravity, but relies on electromagnetic interaction, like the Van der Waals force. Such inter-particle attractive forces bring small dust grains together to form large aggregates through pairwise collisions.

The collision outcome depends on the impact velocity between dust particles. In the early phase of low-velocity gentle collision, (sub)micron-sized dust stick together to form large aggregates with porous structures. This is referred to as the “hit-and-stick” regime. As the growth proceeds, their collision velocity increases with the size of the particle. In this intermediate velocity regime, the aggregates result in restructuring through compactification (Dominik & Tielens 1997; Blum & Wurm 2000). The above process increases the mass-to-area ratio and thus the Stokes number of the particles, resulting in more energetic collisions. The further mass growth is terminated by either bouncing or fragmentation due to high velocity, catastrophic impacts (Güttler et al. 2010). This results in particles only growing up to millimeter to centimeter sized pebbles in nominal protoplanetary disk conditions (Zsom et al. 2010; Birnstiel et al. 2012).

The sticking and growth patterns of the dust aggregates depend on their material properties. In the inner protoplanetary disk regions of less than a few AU, silicates are the main constituent of dust grains. During collisions, these silicate aggregates bounce off or even fragment completely at a threshold velocity of approximately 1 m s$^{-1}$ (Blum & Wurm 2008). Meanwhile, for the regions outside the water-ice line, dust grains are dominated by water-ice. The icy or ice-coated aggregates are more porous than silicate aggregates, and thus the bouncing is less evident for them (Wada et al. 2011). In addition, the sticking still occurs at a collision of 10 m s$^{-1}$ for the icy aggregates. Due to a higher surface energy and a lower elastic modulus, these icy aggregates are more sticky compared to silicates (Gundlach & Blum 2015). Nevertheless, based on the recent laboratory experiments, Musilok & Wurm (2019) found that the surface energy of icy aggregates is comparable to that of silicates when the disk temperature is lower than 180 K. If this is true, it implies that the actual difference in the growth pattern of the above two types of aggregates might be less pronounced than anticipated in the literature (also see Gundlach et al. 2018 and Steinpilz et al. 2019).

Theoretical studies of the global dust coagulation and radial transport have shown that (sub)micron-sized dust grains succeed in growing to mm/cm-sized pebbles (Ormel et al. 2007; Brauer et al. 2008; Zsom et al. 2010; Birnstiel et al. 2010, 2012; Krijt et al. 2016; Estrada et al. 2016). The size of pebbles is either regulated by the radial drift in the outer disk region of $r \geq 10$ AU, or limited by bouncing/fragmentation in the inner disk region of $r < 10$ AU (e.g., fig. 3 of Testi et al. 2014). Nevertheless, further mass increase by incremental growth becomes problematic, due to the above mentioned bouncing and fragmentation barriers.

2.2 Observational Studies

2.2.1 Evidence for dust radial drift

The most straightforward evidence for radial drift of dust comes from the size comparison between the dust and gas components of the disks. The sizes of dust and gas disks can be separately inferred from millimeter dust continuum emission and the molecular line emission of CO isotopes (Dutrey et al. 1998; Hughes et al. 2008). By these comparisons, gas disks are generally found to be much smaller than dust disks, indicating the radial drift of millimeter dust particles has already taken place at the corresponding ages of the systems (Andrews et al. 2012; Ansdell et al. 2018; Jin et al. 2019; Trapman et al. 2020).

The other evidence is from the commonly observed substructures in protoplanetary disks, such as cavities, rings and gaps. These features resolve the drift timescale problem; otherwise, pebbles should be depleted in the disk regions of $r \lesssim 50$ AU within a few Myr (Sect. 2.1.1). The difference in spatial distributions of grains with different sizes is usually used to test the presence of pressure bumps and a strong indicator of the mobility of dust grains. The disks with large inner cavities/holes of a few tens of AU are called ‘transitional disks’ (Calvet et al. 2005; Espaillat et al. 2014; Owen 2016). In these disks, different spatial distributions are observed from the emissions of the millimeter continuum, infrared scattering light and/or molecular lines (Dong et al. 2012; van der Marel et al. 2013, 2015). One should note that the former emission probes the mm-sized pebbles, while the latter two trace $\mu$m grains and gas, respectively. The above phenomena can be explained by the dust filtration effect (Rice et al. 2006; Zhu et al. 2012; Pinilla et al. 2012a), where small grains are tightly coupled to gas flow, while large pebbles drift toward and halt at local pressure maxima, resulting in mm-sized dust cavities larger than gas/small grain cavities. Similarly, such variation in the gas and dust components
is also resolved in ring-shaped substructures (Isella et al. 2016).

Furthermore, the drift of pebbles also leads to a radial variation of chemical compositions in disk gas. For instance, in the protoplanetary disk the CO molecule condenses into solids outside of the CO-ice line while it sublimates into vapor inside. Since some fraction of C is in refractory materials, the disk gas interior to the CO-ice line is expected to have a C/H lower than the stellar value when pebbles are assumed to be static. However, when these pebbles continuously drift inwardly and cross the CO-ice line, a substantial amount of C is loading interior to the ice line and sublimates into vapor, which naturally increases C/H interior to the CO-ice line (Krijt et al. 2018). When comparing the C/H in the stellar photosphere and in the disk gas, Zhang et al. (2020) first reported an enhanced C/H interior to the CO-ice line in the disk of HD 163296, which is 1–2 times higher than the stellar value. This C/H enrichment is in line with the large-scale radial drift of icy dust particles.

2.2.2 Constraints on pebble size

Opacity Index

Let us first look at what we can know from the dust continuum emission of young protoplanetary disks at (sub)millimeter and radio wavelengths. The observed intensity is $I_\nu=I_{\nu0}[1-\exp(-\tau_\nu)]$, where the subscript 0 refers to the value at the disk midplane, $\tau_\nu=\kappa_\nu\Sigma_d$ is the optical depth from the disk midplane to the surface layer, $\kappa_\nu$ is the disk opacity at the corresponding frequency, and $\Sigma_d$ and $T_d$ are the dust mass and temperature, respectively. When the disk is optically thin ($\tau_\nu\ll1$) at the observed wavelengths, $I_\nu\approx\tau I_{\nu0}\approx\kappa_\nu\Sigma_d I_{\nu0}$. Consequently, the observed integrated flux $F_\nu\propto\kappa_\nu\Sigma_d B_\nu(T_d)$ and $M_d$ is the dust mass. At millimeter wavelengths, the Planck function $B_\nu$ is expected to approximately follow the Rayleigh-Jeans law, $B_\nu\approx2k_B T_d\nu^2/c^2$, where $k_B$ is the Boltzmann constant and $c$ is the light speed. Therefore, $F_\nu\propto\kappa_\nu\nu^2 M_d T_d$. This means that, if $\kappa_\nu$ and $T_d$ are known, the dust mass can be estimated from the observed flux.

In addition, the spectral index of dust opacity can be obtained from disk observations. Assuming that the opacity has a power-law dependence on frequency $\kappa_\nu\propto\nu^\alpha$ and $F_\nu\propto\nu^\beta$, we have $F_\nu\propto k_\nu\nu^2\propto\nu^2+\beta$. Since the spectral index $\beta$ is measured from the spectral energy distribution, the dust opacity index can be calculated accordingly through $\beta=\alpha-2$.

The commonly accepted evidence for grain growth is based on the spectral index measurements at millimeter wavelengths (Draine 2006). The interpretation is given as follows. Based on the Mie theory calculation, the maximum grain size matters for the dust opacity. When the maximum grain size is smaller than the observed wavelength (Rayleigh regime), $\kappa_\nu$ is independent of grain sizes and $\beta$ remains high. When the maximum grain size is larger than the observed wavelength (geometric regime), $\kappa_\nu$ decreases with the grain size and $\beta$ drops to a lower value close to 0–1 (Fig. 3 of Ricci et al. 2010). As a result, grains with a maximum size $\lesssim1$ mm naturally result in a less than unity spectral index at millimeter wavelengths. In other words, the value of $\beta$ can principally reveal the size of the largest grain in disks. In a realistic situation, $\beta$ is also affected by the size distribution, composition and porosity of the dust aggregates. Nonetheless, these dependence-induced uncertainties are generally smaller compared to that due to the maximum grain size (Fig. 4 of Testi et al. 2014).

It is worth mentioning that the opacity spectral index observed in ISM gives $\beta_{\text{ISM}}\sim1.7$, while the inferred spectral index of most protoplanetary disks is $\beta_{\text{disk}}\lesssim1$, much smaller than the typical ISM value. Studies that combine multi-wavelength observations with detailed modeling suggested the ubiquitous presence of grain growth in disks with a variety of ages and around stars with different masses (Natta et al. 2004; Ricci et al. 2010, 2014; Miotello et al. 2014; Pinilla et al. 2017). Furthermore, the maximum grain size also correlates with the disk radial distance, where generally centimeter sized particles reside in the inner disk regions and millimeter sized grains are present further out (Pérez et al. 2012; Tazzari et al. 2016). Noticeably, for disks with substructures, the spectral index also varies across the rings and gaps, with lowest values in the rings and highest values in the gaps, indicating further grain growth in the high density centric ring regions (Huang et al. 2018b; Li et al. 2019c; Long et al. 2020).

The above spectral index interpretation is based on two underlying assumptions: the dust emission is optically thin, and the opacity is dominated by absorption rather than scattering at the observed wavelengths. These two assumptions may also be intrinsically correlated. For instance, if the scattering is the main source of opacity instead of absorption, the observed intensity would be $I_\nu\approx\sqrt{1-w_\nu\tau_\nu I_{\nu0}}$ (Zhu et al. 2019), where $w_\nu$ is the single-scattering albedo. The above formula reduces to the previously mentioned $I_\nu\approx\tau I_{\nu0}$ where the scattering is negligible compared to the absorption ($w_\nu\rightarrow0$). This means that the scattering causes disks to look cooler than they actually are (e.g., fig. 9 of Birnstiel et al. 2018). In this respect, the disk mass and optical depth are likely to be underestimated when the scattering is ignored in literature studies (Zhu et al. 2019; Liu 2019; Ballering & Eisner 2019). If the disks are indeed very massive and optically thick even in millimeter wavelengths, then the above approach for the grain size estimation is not valid anymore.
Recently, Carrasco-González et al. (2019) considered both scattering and absorption in dust opacity and neglected any underlying assumption on the optical depth for the study of the HL Tau disk. They still found that the grains have grown to millimeter size. Importantly, similar treatments with careful assumptions should be applied to other disks for more realistic estimates of the grain size as well.

Polarization

Several studies attempted to explain the disk polarized emission by dust scattering (Kataoka et al. 2016; Yang et al. 2016), although other interpretations such as grain alignment with the disk magnetic field could still be relevant. If the dust scattering is indeed the dominant mechanism, the maximum grain sizes can also be constrained by the observed polarization degree. For the same system, the HL Tau disk, Kataoka et al. (2017) reported a maximum grain size of 100 µm, one order of magnitude smaller than the size estimated from Carrasco-González et al. (2019). Despite these polarimetric measurements being only applied to a small number of disks, the inferred size is considerably lower than that obtained from opacity index measurements (Hull et al. 2018; Bacciotti et al. 2018; Ohashi et al. 2020).

The above polarization analysis overall supports grain growth. Nevertheless, it is still unclear whether the discrepancy in grain size estimation between these two interpretations is because of the existence of multi-specie dust grains, or due to the limitations and degeneracies in methodologies themselves. Future studies are warranted to make further claims.

Meteorites in the Solar System

There is evidence of grain growth in our Solar System. For instance, the calcium-aluminium-rich inclusions (CAIs) are sub-mm to cm-sized grains identified in the most primitive meteorites. These refractory inclusions are thought to be the earliest solids condensed from the young nebula that formed the Solar System. In cosmochemistry, the Pb–Pb isotopic dating shows that CAIs formed 4.567 Gyr ago (Amelin et al. 2010; Connelly et al. 2012). The size of CAIs supports the hypothesis of coagulation-driven growth of condensates. In addition, chondrules are igneous-textured spherules dominating in chondrites. The typical size of chondrules is 0.1 to 1 mm (Friedrich et al. 2015; Ebel et al. 2016; Simon et al. 2018). Although there are still discrepancies between Pb-Pb ages and Al-Mg ages (see Kruijer et al. 2020), chronological studies indicated that a small fraction of chondrules might be formed as early as the CAIs, while the majority formed 2–4 Myr after the formation of CAIs (Amelin et al. 2002; Kleine et al. 2009; Villeneuve et al. 2009; Connelly et al. 2012; Pape et al. 2019). The presence of chondrules and CAIs in meteorites supports grain growth in the Solar System.

In summary, the current disk observations, in line with theoretical/laboratory studies, together demonstrate that the first step of planet formation, from dust to pebbles, is robust and ubiquitous during the protoplanetary disk evolutionary stage.

3 FROM PEBBLES TO PLANETESIMALS

As stated in Section 2, the direct coagulation fails to produce aggregates much larger than pebble-sized. In this section, we emphasize one powerful planetesimal formation mechanism termed the streaming instability, which overcomes the above growth barrier by clustering and collapsing dense pebble filaments into planetesimals. The concept of the streaming instability mechanism and the operating disk conditions are presented in Section 3.1 and Section 3.2, respectively. The observational evidence that supports the planetesimal formation by streaming instability is discussed in Section 3.3.

We also note that there are other alternative scenarios to form planetesimals (see Johansen et al. 2014 for a review). For instance, incremental growth may still proceed when aggregates have very high porous structures without significant compactifications (Suyama et al. 2008). Such fluffy particles have a higher area-to-mass ratio compared to the compact grains, resulting in a higher collisional rate and a lower Stokes number. These highly porous aggregates could still overcome the above growth and radial drift barriers and form planetesimals under certain circumstances (Okuzumi et al. 2012; Kataoka et al. 2013; Homma et al. 2019). Besides, turbulent clustering is another mechanism that directly concentrates small dust grains into planetesimals. (Cuzzi et al. 2008; Hartlep & Cuzzi 2020). Different than the streaming instability, this mechanism requires a prior condition of underlying turbulence. The optimal size of operating particles is crucially related to the energy cascade models (Cuzzi et al. 2010; Pan et al. 2011; Hartlep et al. 2017; Hartlep & Cuzzi 2020).

3.1 Streaming Instability

The streaming instability is applicable for a wide size range of solid particles but most efficiently for the particles of \( \tau_s \approx 0.1–1 \). We conventionally use the word ‘pebbles’ hereafter to refer to the solid particles incorporated in the streaming instability mechanism. Basically speaking, the streaming instability includes two processes, namely the concentration of pebbles into dense clumping filaments, and gravitational collapse of pebble filaments into planetesimals. The snapshots of the above processes and the spatial distribution of solids in a streaming instability simulation are illustrated in Figure 5.
gas becomes very turbulent near the midplane due to interactions with particles

A Local Shearing-box Simulation of Particle-Gas System

Fig. 5 Illustration of 3D local shearing-box simulation of the streaming instability for the gas-particle system. Left: the box shows the simulation domain, with each side a fraction of gas scale height. Color maps the vertical gas momentum, which exhibits high turbulence near the midplane due to interactions between gas and sedimented solid particles. Right: solid particles (color refers to the surface density) in the midplane (a) form dense filaments due to the streaming instability, and (b) collapse into planetesimals through the collective effect of self-gravity once the solid density exceeds the Roche density. Courtesy of Rixin Li.

We give a qualitative explanation here. Firstly, inward drifting pebbles get concentrated radially due to the solid-to-gas back reaction (Youdin & Goodman 2005). This is because pebbles feel the gas drag force and lose angular momentum. Similarly, the surrounding gas gains angular momentum from pebbles due to the solid-to-gas back reaction, and thus the gas velocity gets accelerated. The strength of this reaction force is determined by the volume density ratio between pebbles and gas \( \rho_{\text{peb}}/\rho_g \). We usually neglect this back reaction since the pebble density is much lower than the gas density in the nominal protoplanetary disk condition. However, when the pebble density is comparable to the gas density, the back reaction force is non-trivial. In this situation, the pebble density perturbation grows and concentrates pebbles effectively. Since the gas is accelerated towards the Keplerian velocity, and the relative velocity between gas and pebbles becomes smaller, these pebbles feel weaker gas drag and drift inward more slowly. Hence, the fast drifting pebbles from the outer part of the disk thereby catch these slower drifting pebbles and form denser filaments. Because this is a positive feedback, once an initial, radial concentration of solids is achieved \( \rho_{\text{peb}} \approx \rho_g \), the further clumping is self-amplified, and eventually the pebble density grows rapidly in a non-linear manner. Youdin & Goodman (2005) proposed the concept of streaming instability and provided the analytical linear instability solution. The robustness of the pebble clumping was later confirmed in numerical simulations (Youdin & Johansen 2007; Johansen & Youdin 2007).

Secondly, the effect of self-gravity becomes dominant when the pebble density exceeds the Roche density \( \rho_R = \frac{9 \Omega_K^2}{4 \pi G} \). In this circumstance, gravitational force overcomes tidal shear, and the pebble filaments gravitationally collapse into 100-km-sized planetesimals\(^4\). The primordial idea of the gravitational instability was proposed by Goldreich & Ward (1973). At that time these authors only considered that dust sediment into a super thin midplane layer to exceed the Roche density. However, such a vertical solid concentration generates the Kelvin-Helmholtz instability and develops vertical

\(^4\) Gerbig et al. (2020) further suggested that the collapse criterion requires gravitational force to overcome turbulent diffusion on small scales, which also regulates the sizes of forming planetesimals.
velocity stirring, which prevents the subsequent solid enrichment (Weidenschilling 1980; Cuzzi et al. 1993). On the contrary, the streaming instability induces a radial concentration of solid particles. The gravitational collapse of dense particle filaments by the above streaming effect was numerically verified by Johansen et al. (2007, 2009).

We remark a few things here. Strictly speaking, only the first step is relevant to the concept of instability – the streaming motion between solids and gas. Unlike other concentration mechanisms, (e.g., Cuzzi et al. 2008), one feature of the streaming effect does not require any underlying disk turbulence. The second step does not inherently depend on any initiated mechanisms that cluster pebbles. It broadly represents a pebble collapsing process that forms planetesimals aided by the self-gravity. These two steps have been commonly investigated together in literature studies and have been frequently recognized as a unified mechanism termed the streaming instability.

We also point out that this pebble clumping effect also triggers turbulence, even if the flow of the background gas was originally laminar. As numerically investigated by Li et al. (2018), the density fluctuation induced by the streaming instability is not sufficient to halt and concentrate pebbles. The reason for pebble clumping still arises from the previously mentioned streaming motion between pebbles and gas. It is noteworthy that most of these streaming instability simulations are conducted in a cubic box centered on a local, co-rotating coordinate frame with fixed Keplerian frequency and radial orbital distance (also called shearing-box simulations, see Fig. 5). The length scale of the box is much smaller than the orbital distance, and therefore, the motion of particles in that box is linearized with the Keplerian shear. For most streaming instability simulations, gas fluid is based on a Eulerian grid, and solid particles are treated as Lagrangian superparticles, each representing a swarm of actual pebbles. Such a particle-fluid hybrid approach substantially reduces the computational cost for investigating the non-linear pebble clumping and planetesimal formation processes.

The mass of planetesimals formed by streaming instability simulations follows a top-heavy mass distribution, which can be roughly fitted by a power-law plus exponential decay for the intermediate and high mass branch (Johansen et al. 2015; Schäfer et al. 2017; Abod et al. 2019). A turnover mass may exist in the lower mass branch (Li et al. 2019b). The planetesimals have a characteristic size of \(\sim 100\) km when they form at the asteroid belt region (Johansen et al. 2015). The characteristic mass/size increases with disk metallicity, the mass of the central star and radial distance (Johansen et al. 2012, 2015; Simon et al. 2016), modestly increases with gas pressure gradient (Abod et al. 2019), and appears to be independent of Stokes number (Simon et al. 2017). Based on the extrapolation of literature streaming instability simulations, Liu et al. (2020) derived the characteristic mass of planetesimals as

\[
M_{pl} = 5 \times 10^{-5} \left( \frac{Z}{0.02} \right)^{1/2} \left( \frac{\gamma}{\pi} \right)^{3/2} \left( \frac{b_0}{0.05} \right)^{3} \left( \frac{M_{\ast}}{M_\odot} \right) \left( \frac{\Omega}{1 \text{Myr}^{-1}} \right) \left( \frac{M}{M_\oplus} \right),
\]

where \(Z\) is the local disk metallicity and \(\gamma=4\pi G \rho_\ast / \Omega^2\) is a self-gravity parameter\(^5\), related with gas density, stellar mass and radial distance. Adopting the MMSN model, we obtain \(\gamma=0.034\) and the resultant planetesimal is \(10^{-6} M_\oplus\) in mass (100 km in radius) at \(r=2.5\) AU around a solar-mass star. Based on Equation (7), we expect that smaller planetesimals form at shorter orbital distances and around lower-mass stars.

Most of the streaming instability studies simply considered a laminar background gas, despite that the protoplanetary disk should be turbulent in nature (Lyra et al. 2019). How the streaming instability operates in a realistic turbulent condition has not been fully understood. First of all, disk turbulence induces stochastic motion and density fluctuation of the gas. Overdense pressure bumps can be produced at the region in which particles of \(\tau_g \sim 1\) are efficiently trapped. This type of solid concentration that facilitates planetesimal formation is obtained in disks where the source of turbulence is either the magnetorotational instability (MRI, Johansen et al. 2007), or the vertical shear instability (VSI, Schäfer et al. 2020), by which angular velocity depends on the disk vertical distance (Nelson et al. 2013; Stoll & Kley 2014; Lin & Youdin 2015; Flock et al. 2017). When considering the non-ideal magnetohydrodynamical (MHD) effects, Yang et al. (2018) found that dust diffusion is weak in the radial direction and strong clumping can still occur in the dead zone region where the MRI is inactive because of low ionization fraction (Gammie 1996). On the other hand, the turbulent diffusivity acts to suppress sedimentation and concentration of particles, and therefore, the birth rate of planetesimals can be slower or even quenched when the disk turbulence increases (Gole et al. 2020). These numerical studies have only been explored in a very narrow range of parameter spaces (\(\tau_g\) and \(\alpha_t\)) with various disk turbulence mechanisms.

In order to quantify the role of turbulence, Umurhan et al. (2020) and Chen & Lin (2020) have conducted linear stability analyses for the motion equations of gas and solid particles by including additional viscous forcing terms. They also found that the growth rate gets reduced even when the disk is moderately turbulent. Nevertheless,

\(^5\) We note that \(\gamma\) can be related with the Toomre \(Q_T\) parameter by \(Q_T=\sqrt{8/\pi/\gamma}\). Thus, \(\gamma=0.034\) is equivalent to \(Q_T=47\).
the adopted simplified isotropic $\alpha$ prescription cannot fully mimic the realistic non-linear pattern of the disk turbulence. For instance, the aforementioned turbulent-induced zonal flows and pressure bumps seem to promote dust concentration and subsequent planetesimal formation (Johansen et al. 2007; Schäfer et al. 2020), which is not captured in these theoretical analyses. On the whole, it is still premature to reach definitive conclusions yet, and future studies are required for building a unified and consistent picture of this topic.

3.2 How, Where and When the Streaming Instability Occurs

In order to trigger the streaming instability, the volume density of solids needs to be enhanced comparable to that of gas, $\rho_{\text{peb}} \simeq \rho_g$ (Youdin & Goodman 2005). The onset criterion can also be expressed in terms of the surface density ratio, i.e., the metallicity $Z = \Sigma_{\text{peb}} / \Sigma_g$. The dust to gas mass ratio is measured to be 0.01 in the ISM (Bohl et al. 1978), while the canonical value of the solar metallicity is 0.014 (Asplund et al. 2009). Numerical studies reported that a super-solar metallicity ($\geq 2\% - 5\%$) is required for triggering the streaming instability (Johansen et al. 2009; Carrera et al. 2015; Yang et al. 2017). Such a threshold metallicity also depends on the disk and pebble properties. The streaming instability is easier to be triggered when the disk metallicity is higher (Johansen et al. 2009), the strength of the disk pressure gradient is lower (Bai & Stone 2010), and/or the Stokes number of pebbles is higher towards unity (Carrera et al. 2015).

Nonetheless, unless some other mechanisms can operate in the first place to enhance the pebble density, the disk with solar metallicity can hardly form planetesimals by the streaming instability. Then the question is how the solid density can be enriched to satisfy this condition?

We list several scenarios that propose pebble enrichment at peculiar disk locations. For instance, the formation site can be the water-ice line (Ros & Johansen 2013; Ida & Guillot 2016; Schoonenberg & Ormel 2017; Drążkowska & Alibert 2017; Hyodo et al. 2019). This is because the water-ice in pebbles sublimes into vapor when these pebbles drift inward across the water-ice line ($T_g \simeq 170$ K). First, pebbles are locally piled-up by a "traffic jam" effect since the outer fast drifting icy pebbles catch up with the inner slow drifting silicate grains. Second, the released water vapor diffuses back to the outside of the ice line and condenses onto the continuously inwardly drifting icy pebbles. This diffusion and re-condensation process also enhances the local solid density (Stevenson & Lunine 1988; Cuzzi & Zahnle 2004). The former mechanism generates "dry" planetesimals slighter interior to the ice line while the latter one produces "wet" planetesimals with a substantial water fraction slightly exterior to the ice line. Similar processes could also be expected at the ice lines of other volatile-rich species, such as CO and NH$_3$. Besides these ice lines, other possible pebble trapping sites can be the edge of the dead zone (Drążkowska et al. 2013; Chatterjee & Tan 2014; Hu et al. 2016; Miranda et al. 2017), the vortex generated by hydrodynamical instabilities (Surville et al. 2016; Huang et al. 2018c), and the spiral arms in self-gravitating disks (Gibbons et al. 2012).

Apart from the above mentioned mechanisms that relate with local disk properties, there are other ways of increasing the disk metallicity. For instance, Drążkowska et al. (2016) showed that this enrichment can occur in the inner sub-AU disk region as a result of the global dust growth and radial drift. In addition, pebble trapping is thought to be a natural consequence of giant planet formation. The massive planet opens a gap (Lin & Papaloizou 1986) and produces a local pressure maximum in its vicinity (Lambrechts et al. 2014). Pebbles drift more and more slowly and get concentrated on their way towards this local pressure maximum. When these pebbles reach the threshold metallicity at or close to the gap edge, the streaming instability is triggered to form planetesimals (Eriksson et al. 2020). Moreover, the solid enrichment can be fulfilled in the late disk dispersal phase when stellar photoevaporation dominates. In this case, pebbles have already decoupled from gas and sedimented to the disk midplane. The photoevaporating wind blows gas away from the disk surface. Therefore, the solid-to-gas ratio increases globally (Carrera et al. 2017).

To summarize, the planetesimal formation can either occur locally at peculiar disk locations such as ice lines and pressure bumps, or in a wide range of disk regions when stellar photoevaporation globally depletes the disk gas. Where and when the planetesimals form crucially depend on the detailed disk conditions and pebble concentration mechanisms. For instance, the formation location can broadly range from the most inner sub-AU disk region (dead zone edge) to the outer part of the disk extending to a few tens/hundreds of AU's (spiral arms in self-gravitating disks). The formation time is even more difficult to quantify, which might occur at an early phase when the disk is still self-gravitating ($t \lesssim 0.1 - 1$ Myr), or at a late phase when gas is significantly dispersed ($t > 3$ Myr). The other time constraint is from the meteorite chronology in our Solar System (see Krijger et al. 2020 for a review). What we learn is that these parent bodies of meteorites are apparently not all formed at once. They are more likely to form successively or undergo multiple formation phases over the entire disk lifetime. For instance, the parent bodies of iron meteorites formed within the first Myr (Krijger
shown in various systems have similar optical depths disks from the DSHARP survey. All these dusty rings shown in young protoplanetary systems have similar optical depths of the order of unity (Dullemond et al. 2018). These observed values can be interpreted as ongoing planetesimal formation regulated by the streaming instability (Stammler et al. 2019). In their explanation, drifting pebbles are concentrated in the ring where the streaming instability can be triggered. The streaming instability converts the pebbles into planetesimals when \( \rho_{\text{peb}} > \rho_g \), while it is quenched when \( \rho_{\text{peb}} < \rho_g \). Thus, such a regulated process removes the excess pebbles into planetesimals, maintaining the midplane dust-to-gas ratio to be of order unity. Since the optical depth correlates with the dust-to-gas ratio, this interpretation naturally explains the peculiar optical depths in the observed rings. \(^6\)

Evidence of the streaming instability can also be found from the minor bodies in our Solar System. Morbidelli et al. (2009) conducted collisional coagulation simulations and found that in order to reproduce the current size distribution of main-belt asteroids, the primordial planetesimals should be \( \gtrsim 100 \text{ km in size} \). Rather than incremental growth, this characteristic size and the slopes of size distributions of main-belt asteroids and Jupiter belt objects are more consistent with planetesimals obtained from streaming instability simulations (Johansen et al. 2015; Simon et al. 2016).

The most appealing evidence is from the Kuiper belt objects. Recently, a contact binary named Arrokoth (previously known as Ultima Thule or 2014 MU\(_{69}\)) was imaged by the New Horizons spacecraft during its flyby. Arrokoth, resembling other cold classical Kuiper belt objects, is thought to be well preserved in terms of the pristine properties since its formation. It consists of two equal-sized, compositionally homogeneous lobes with a narrow contact neck (Stern et al. 2019; Grundy et al. 2020; Zhao 2020). Such a peculiar shape with little distortion, and the good alignment of the two lobes strongly indicate that this type of object originated from gentle, low-speed mergers of planetesimals within a gravitationally collapsing clump of pebbles (McKinnon et al. 2020).

The prevalence of equal-sized binaries found in the Kuiper belt supports that they form by the gravitational collapsing mechanism (Nesvorný et al. 2010; Robinson et al. 2020). Most of these binaries are in the cold classical Kuiper belt which have low heliocentric orbital inclinations and eccentricities and thus remain primordial compared to other populations. Furthermore, these binaries are observed to have similar colors even though the color distribution of the binary population has a large intrinsic scatter (Benecchi et al. 2009; Marso et al. 2020). This is also expected from the gravitational collapsing, since they form from the same reservoir of solids in the pebble clumps. In addition, based on the obliquity measurements of trans-Neptunian binaries, Grundy et al. (2019) found that the prograde binaries are more common than the retrograde ones among tight binaries (22/26). Such a binary orientation distribution is consistent with the predictions of the streaming instability simulations (Nesvorný et al. 2019). In contrast, the above properties are difficult to fulfill when the binaries form by sequential coagulation and capture (Goldreich et al. 2002).

To conclude, the streaming instability seems to be the widely-accepted and leading mechanism of planetesimal formation. The success of the streaming instability is not only because the robustness of the mechanism itself is verified by numerous theoretical/numerical work, but also many key features of the planetesimal populations generated by the streaming instability are consistent with current observations, both within and beyond the Solar System. The streaming instability succeeds in bridging the gap between the pebbles and planetesimals. The growth of planetesimals after formation will be discussed in subsequent sections.

4 PLANETESIMAL ACCRETION

We review the planet formation process from planetesimals to planets from Section 4 to Section 6. In this section we focus on planetesimal accretion. The accretion cross section and accretion rates in different regimes are described in Sections 4.1 and 4.2 respectively. We further discuss the underlying physical processes in Section 4.3 and summarize the key features and applications in Sections 4.4 and 4.5 respectively.

4.1 Accretion Cross Section

The Hill radius of a planetary body orbiting a central star is defined as

\[
R_H = \left( \frac{M_p}{3M_*} \right)^{1/3} a, \tag{8}
\]

where \( M_p \) and \( a \) are the mass and semimajor axis of the body respectively. The Hill velocity is \( v_H = R_H \Omega_K \). Within the Hill sphere, the planetary body’s gravitational force is more important than that of the star.

---

\(^6\) The large-scale dust clumping at the edges of the rings are also resolved from hydrodynamic simulations (Huang et al. 2020).
We consider planetesimal accretion in the case of a few massive protoplanetary embryos embedded in a swarm of less massive planetesimals. Hereafter we call these two populations the large and small bodies, and their masses are expressed as $M$ and $m$, respectively. Only the gravitational force operates during their encounters. The collisional (accretion) cross section of the large body can be expressed as
\[ \sigma = \pi R_M^2 \left( 1 + \frac{v_{\text{esc}}^2}{\delta v^2} \right) = \pi R_M^2 \left( 1 + \frac{2GM}{R_M \delta v^2} \right), \tag{9} \]
where $R_M$ is the physical radius of the large body, and $v_{\text{esc}} = \sqrt{2GM/R_M}$ and $\delta v$ are the escape velocity of the large body and the relative velocity between the large and small bodies respectively.

\[ \text{In this circumstance, the mass ratio between the large and small bodies increases with time. Nevertheless, small planetesimals still dominate the total masses of whole populations.} \]

The above runaway phase cannot last forever. When the velocities of planetesimals are significantly stirred up by the large bodies (the latter case of Eq. (10)), the accretion cross sections of large bodies are strongly reduced compared to the runaway case. The growth of the big bodies becomes slower as $M$ increases. The accretion gradually turns into a self-regulated, oligarchic growth (Ida & Makino 1993; Kokubo & Ida 1998), featured by a decreasing mass ratio among adjacent massive bodies.

\[ \text{In this regime the massive bodies still grow faster than the small planetesimals. After the runaway and oligarchic phases, the system evolves into a bimodal population, with large protoplanets and small planetesimals (Kokubo & Ida 2000; Thommes et al. 2003; Rafikov 2003; Ida & Lin 2004a; Ormel et al. 2010). These protoplanets have a typical orbital separation of 10 mutual Hill radii (Kokubo & Ida 1998).} \]

In the late stage when there is no disk gas left, the protoplanets gradually accrete all residual planetesimals. The random velocities of the protoplanets are fully excited to their escape velocities, and their collisional cross sections reduce to the physical surface areas. In this case, the growth eventually becomes slow, and the system is chaotic in nature (Agnor et al. 1999; Chambers & Wetherill 1998; Chambers 2001; Raymond et al. 2004; Kenyon & Bromley 2006; Zhou et al. 2007).

4.2 Growth Modes: Runaway vs. Oligarchic

The collision rate of the large bodies by accreting small bodies is given by $n_m n_s \delta v$, where $n_m = \Sigma_m/mH$, $\Sigma_m$ and $H$ are the number density, surface density and vertical height of the small bodies respectively. The planetesimal accretion rate of the large body can be written as
\[ M_{\text{PIA}} = \Sigma_m n_K \pi R_M^2 \left( 1 + \frac{2GM}{R_M \delta v^2} \right) = \left\{ \begin{array}{ll}
\frac{2\pi GM \Sigma_m n_K R_M}{\delta v^2} & \propto M^{4/3} \text{when } \delta v \ll v_{\text{esc}}, \\
\frac{\pi \Sigma_m n_K R_M^2}{\delta v^2} & \propto M^{2/3} \text{ when } \delta v \gtrsim v_{\text{esc}},
\end{array} \right. \tag{10} \]
where $M \propto R_M^3$ and $H \approx 2\delta v/\Omega_K \approx \delta v/\Omega_K$ for an isotropic planetesimal velocity distribution. The growth timescale is expressed as $t_{\text{grow}} = M/(dM/dt)$, where $t_{\text{grow}}$ scales with $M^{-1/3}$ and $M^{1/3}$ in the former and latter cases, respectively.

In the early stage of planetesimal accretion (the former case of Eq. (10)), $\delta v$ is mainly excited from the mutual interactions among small planetesimals and therefore remains low compared to $v_{\text{esc}}$. Since the growth becomes faster as $M$ increases, this phase is called the runaway growth (Safronov 1972; Greenberg et al. 1978; Wetherill & Stewart 1989; Ida & Makino 1993; Kokubo & Ida 1996).

\[ t_{\text{vs}} = \frac{v}{d\delta v/dt} = \frac{\delta v^3}{4\pi G^2 n_M M^2 \ln \Lambda}, \tag{11} \]
where $\ln \Lambda \approx 3$ is the Coulomb factor, and $n_M$ is the number density of large bodies. As can be seen from
Equation (11), the effect of viscous stirring increases with the mass of the large body.

On the other hand, the random velocity of small planetesimals can also be damped by the gas friction force. The eccentricity damping timescale is given by Adachi et al. (1976),

\[ t_{e, \text{gas}} = \frac{m}{C_D \pi R_m^2 \rho_g \delta v/2} \]  

(12)

where \( C_D = 0.5 \) is the gas drag coefficient for planetesimal-sized objects in Equation (4) and \( R_m \) is the radius of the small planetesimal.

Equating \( t_{v_s} = t_{e, \text{gas}} \) and solving for the equilibrium eccentricity of small planetesimals, one can obtain

\[ e_m = \left( \frac{8 \ln \Lambda m M \Sigma M_a}{C_D \rho_g R_m^2 M^2} \right)^{1/5} \].  

(13)

Inserting this value into Equation (10), the growth timescale in the oligarchic regime can be alternatively expressed as

\[ t_{rg} = \frac{e_m^2 a^4 \Omega_K}{2\pi G \Sigma m} \frac{\Omega_K}{R_m^2} \left( \frac{8 \ln \Lambda m M \Sigma M_a}{C_D \rho_g R_m^2 M^2} \right)^{2/5} \].  

(14)

On the other hand, Ormel et al. (2010) conducted Monte Carlo simulations for the planetesimal growth and found that the runaway growth timescale can be alternatively expressed as

\[ t_{rg} = C_{rg} \frac{R_m}{\Omega_K \Sigma_m} \].  

(15)

where \( C_{rg} \sim 10 \) is a numerical prefactor. We note that \( R_m \) and \( \Sigma_m \) in Equation (15) are the initial size and surface density of the planetesimals, and the growth timescale only depends on the initial configuration. This is an important feature of the runaway growth, where the planetesimals spend most of the time doubling their initial masses.

Another important damping mechanism is called dynamical friction. It refers to the process of equipartition of the random energy between large and small bodies. The consequence of dynamical friction is that the random velocity of the body is proportional to the square root of its mass \( \delta v \propto M^{-1/2} \), in which small (large) bodies have high (low) random velocities through their mutual gravitational interactions. Besides, inelastic collisions also damp the random velocity. When two bodies collide with each other, they conserve total angular momentum but some fraction of the kinetic energy transfers into internal heat.

Now let us recall when is the transition between the runaway growth and oligarchic growth. Strictly speaking, this transition is not determined by \( \delta v/\nu_{osc} \) but relies on the relative growth rate of the massive bodies, \( (dM/dt)/M \propto M^\gamma \). The growth is in the runaway阶段 when \( \gamma > 0 \), while the growth is in the oligarchic phase when \( \gamma < 0 \). Ormel et al. (2010) proposed a physical criterion for the above transition. At the beginning of the accretion, the mass growth is faster than the excitation of random velocities, corresponding to \( t_{rg} < t_{v_s} \). The transition occurs when \( t_{rg} \sim t_{v_s} \). After that, the stirring is faster than the growth \( (t_{rg} > t_{v_s}) \), and the accretion enters the self-regulated oligarchic regime.

### 4.4 Features

We highlight a few important features of planetesimal accretion. First, the runaway accretion does not necessarily mean that the growth is rapid. The concept of “runaway” refers to a relative growth rate, \( (dM_1/M_2)/dt > 0 \) when \( M_1 > M_2 \). The absolute rate can be high or low, depending on the planetesimal surface density \( \Sigma_m \), the Keplerian orbital frequency \( \Omega_K \), and the gravitational focusing factor that relates to \( \delta v \).

Second, the size of planetesimals matters for the growth. As introduced in Sect. 3, when planetesimals are assumed to form by the streaming instability, they are primarily large (e.g., 100-km in size). These planetesimals are well-decoupled from gas and their orbits remain in-situ during the disk lifetime. In this circumstance, the massive protoplanets only accrete nearby planetesimals within their feeding zones (~10 mutual Hill radii). The final masses of protoplanets are only related to local disk properties (e.g., the planetesimal surface density \( \Sigma_m \)).

However, if the initial sizes of planetesimals are kilometer or smaller, the above accretion paradigm changes. Planetesimals of smaller sizes undergo non-negligible radial drift during the disk lifetime. The accreting materials are not limited to local planetesimals in the vicinity of the protoplanets anymore. Therefore, the planets can finally reach higher masses. On the other hand, for smaller planetesimals, their random velocities remain lower due to the weaker viscous stirring and stronger gas damping (Eqs. (11) and (12)). As a result, the growth of <1-km-sized boulders is faster compared to the case of 100-km-sized planetesimals (Coleman & Nelson 2016). Planet growth has several advantages when the initial size of planetesimals is small. However, the major issue is whether the disk can form planetesimals with dominant size of \( \lesssim 1 \) km.

Third, planetesimal accretion also depends on their disk locations. As shown in Equation (10), when the planetesimals are further out, the accretion is slower due to a longer orbital time. In addition, the encounters tend to result in ejections rather than collisions when the orbital distance is larger. The outcome of the planet-planet encounters can be quantified as (Goldreich et al. 2004; Ida
Hf isotopes, among which the hafnium-tungsten (Hf-W) ratio in some planetary materials have been measured by using radioactive decay of short-lived Hf isotopes. The Hf-W ratio is widely adopted for radiometric dating. This isotope analysis is also widely used to trace the source of materials and to constrain the timing of major events in the solar system.

### 4.5 Applications

#### Solar System

The formation timescales of terrestrial planets can be measured by using radioactive decay of short-lived isotopes, among which the hafnium-tungsten (Hf-W) isotope is widely adopted for radiometric dating. This is not only due to its applicable radioactive decay half-life time of 9 Myr, but is also related to the chemical properties of these two elements: Hf is lithophile (“rock loving”) and W is siderophile (“iron loving”). W preferentially settles into the metal core and Hf remains in the silicate mantle before the protoplanet becomes massive enough to segregate. For instance, if the core forms early (early mantle-core segregation), the measured W abundance in the mantle would be high. On the other hand, if the core forms late, W would be high in the core, leaving the W-deficient mantle. Therefore, measuring the Hf/W ratio in the current Earth mantle can be utilized to constrain its core formation and differentiation. Such isotopic analyses indicate that core formation of Earth occurred 30–100 Myr after formation of the Solar System (Kleine et al. 2002; Yin et al. 2002; Jacobsen 2005; Touboul et al. 2007; Kleine et al. 2009; Rudge et al. 2010). Similarly, based on Martian meteorites, the formation time of Mars is inferred to be within a few Myr (Kleine et al. 2004; Foley et al. 2005; Dauphas & Pourmand 2011), comparable to the lifetime of the proto-solar nebula.

The final accretion stage of Solar System terrestrial planets, or called late giant-impact stage, is typically modeled by a numerical N-body approach (see Izidoro & Raymond 2018 for a summary of numerical methods). Before this stage, protoplanetary embryos of lunar to Martian masses were already formed in the terrestrial planet forming region by accreting planetesimals (see Sect. 4.2). The random velocities of the embryos will be stirred without efficient damping, and planet-planet collisions frequently occur after dispersal of the disk gas. On the other hand, giant planets have already grown to their current masses. The influence of gas giants (in particular Jupiter and Saturn) is crucial for sculpting the architecture and water delivery of the asteroid belt and the inner terrestrial planets (Raymond & Izidoro 2017; Zheng et al. 2017). Numerous numerical simulations have attempted to reproduce the terrestrial planets, both in terms of the dynamical properties and geochemical accretion timescale (Chambers 2001; Raymond et al. 2004, 2006, 2009; O’Brien et al. 2006; Thommes et al. 2008; Morishima et al. 2010; Jacobson et al. 2014). The above models are either assumed to have giant planets on their current positions with nearly circular or slightly eccentric orbits. However, these models still have shortcomings. For instance, the planets emerging out of the Mars forming region have masses comparable to Earth and Venus. In other words, the Mars analogs are difficult to produce in the above numerical simulations, known as the “small Mars problem”. Hansen (2009) first proposed that the proper Mars-size objects can be yielded when the planetesimal disk is truncated at around 1 AU. As a further step, Walsh et al. (2011) proposed that such a truncation can be physically caused by the inward-then-outward migration of Jupiter and Saturn through planet-disk interactions, which is known as the Grand Tack model. The key ingredient of this model, the migration of two giant planets, is consistent with our current understanding of the disk migration theory (Kley & Nelson 2012; Baruteau et al. 2014) and has been validated by hydrodynamic simulations (Masset & Snellgrove 2001; Morbidelli & Crida 2007; Pierens & Nelson 2008a; Zhang & Zhou 2010). The appealing point of the Grand Tack model is that it satisfactorily reproduces key observational features of the inner Solar System, such as the orbital and mass distributions of the terrestrial planets, mass depletion and the chemical compositions of planetesimals in the asteroid belt, and water content of Earth (O’Brien et al. 2014). On the other hand, Clement et al. (2018) proposed that Mars’ growth can also be stunted if the giant planet instability occurs relatively early during the terrestrial planet formation (e.g., $\lesssim$10 Myr after the gas disk dispersal).

It is also worth mentioning that most of the current N-body simulations only consider perfect mergers when two bodies collide with each other. However, the random velocities of planetesimals would be excited when the masses of the protoplanets increase. The collisional outcome sensitively relates to the masses and velocities of the colliding bodies, which could be catastrophic disruption, grinding or fragmentation (Leinhardt & Stewart 2012; Agnor & Asphaug 2004; Genda et al. 2012; Liu et al. 2015). Dedicated N-body simulations including realistic collision recipes showed that the final masses and numbers of surviving planets are comparable to the case when only
perfect mergers are considered, but differences still exist in planet spins, eccentricities, core mass fractions and formation timescales (Kokubo & Genda 2010; Chambers 2013; Carter et al. 2015; Clement et al. 2019).

Exoplanetary systems

The observed exoplanets are more diverse than Solar System planets. One important question is how to form different planet populations, such as gas giants and super-Earths. The growth of giant planets first requires the assembly of massive cores ($\sim 10 \, M_\oplus$) that can initiate rapid gas accretion (Pollack et al. 1996; Ikoma et al. 2000; Hubickyj et al. 2005; Movshovitz et al. 2010) before the dispersal of the disk gas. For a nominal disk model, the planetesimal isolation mass at a distance of a few AU (the formation zone of Jupiter and Saturn) is generally lower than this critical mass (Ida & Lin 2004a). Therefore, giant planet formation was thought to be challenging, unless the disks are extremely massive.

Protoplanets approaching such isolation masses can undergo substantial orbital migration. These protoplanets migrate towards and get trapped into MMRs at some zero-torque locations (or called trapping locations, Lyra et al. 2010; Horn et al. 2012; Kretke & Lin 2012). Depending on the number of planets and the gas disk mass, the resonant configurations can be disrupted with rapid migration, resulting in frequent planet-planet collisions even in the gas-rich disk phase (Hellary & Nelson 2012; Pierens et al. 2013; Zhang et al. 2014; Liu et al. 2015). In this circumstance, massive cores can be attained early enough and subsequently grow into gas giants. This interpretation prefers that the cores of gas giants only form early when the disk is massive, e.g., the disk accretion rate is higher than $\sim 10^{-7} \, M_\odot \, yr^{-1}$. Comparing that with the typical observed disk accretion rates and the $M_k - M_*$ dependence (Hartmann et al. 1998; Natta et al. 2006; Manara et al. 2016), it explains why only a minor fraction of stars harbor gas giant planets (Liu et al. 2015) and why giant gas fraction increases with the mass of the central star (Liu et al. 2016). On the other hand, if planet-planet collisions do not occur early enough, planets remain low-mass and they evolve into super-Earth systems with compact, (near) resonant configurations (Terquem & Papaloizou 2007; Ogihara & Ida 2009; Cossou et al. 2014; Ogihara et al. 2015; Izidoro et al. 2017). Numerous works also explored the detailed resonant trapping and stability of multiple super-Earth systems both numerically (Wang et al. 2012; Wang & Ji 2014; Sun et al. 2017; Pan et al. 2020) and analytically (Quillen 2011; Hadden & Lithwick 2018; Petit et al. 2020).

When examining the giant planets in the Solar System and based on the results of the Juno mission, the core of Jupiter is inferred to be diluted with an extended layer of heavy elements (Wahl et al. 2017; Helled & Stevenson 2017). The growth of proto-Jupiter has a strong influence on nearby planetesimals and protoplanets (Zhou & Lin 2007; Ida et al. 2013). Such a diluted internal structure is likely to result from the giant impacts among protoplanets during their assembling phase (Liu et al. 2019c).

The other possible explanation of super-Earths is that they only form late and in-situ when disk gas dissipates significantly. In that case, gas damping is weak and the super-Earth cores can grow quickly through the mergers of protoplanets (Lee & Chiang 2016). This could also explain why observed massive super-Earths ($\sim 10 \, M_\oplus$) do not undergo runaway gas accretion to become gas giants (Lee et al. 2014). The exact compositions and orbital properties of the formed super-Earths are determined by the combined effects of gas damping and solid accretion (Dawson et al. 2016), and potential further out giant planets (Ji et al. 2011; Jin & Ji 2011). On the other hand, recent hydrodynamic studies found that because the gas recycles efficiently between the planetary envelopes and surrounding disk, the rapid gas accretion onto super-Earths can be actually protracted (Ormel et al. 2015; Cimerman et al. 2017; Lambrechts & Lega 2017; Kuwahara et al. 2019). This also provides an interpretation for the ubiquitous presence of super-Earths but not gas giant planets.

The population synthesis model is an ideal tool to explore the influence of key physical processes on planet formation and evolution. In this approach, different physical processes are simplified into specialized recipes and combined into a unified deterministic model. By Monte Carlo sampling the initial conditions over appropriate distributions, the synthetic planetary populations can be generated and thus compared to the observed exoplanet sample in a statistical manner (see Benz et al. (2014) for a review). Ida & Lin (2004a) first applied such a calculation to investigate planet formation around Sun-like stars, and further extended their study to systems around stars of various masses and metallicities (Ida & Lin 2004b, 2005). The predicted correlation between gas giant planets and their stellar hosts exhibited good agreements with RV measurements (Fischer & Valenti 2005; Johnson et al. 2007). Sophisticated population synthesis models based on planetesimal accretion are further developed to make testable observational comparisons and to study how forming planets are related to the initial disk and stellar properties (Mordasini et al. 2009, 2012a,b; Ida et al. 2013;
Jin et al. 2014; Coleman & Nelson 2014, 2016; Alibert & Benz 2017; Mulders et al. 2019; Miguel et al. 2020).

**Stellar Binary System**

The previously mentioned studies focus on planet formation around single stars. However, nearly half of Sun-like stars are in binaries (Duquennoy & Mayor 1991; Raghavan et al. 2010), and this fraction is even higher for higher-mass stars (Kouwenhoven et al. 2007). Thus, studying how planets form in stellar binary systems is of crucial importance. Up to now, more than 150 exoplanets have been discovered in stellar binary systems, including both S-type (satellite-like orbits around one of the single stars) and P-type (planet-like orbits around binary stars). P-type planets are also referred to as circumbinary planets.

Planets are less common in binary systems compared to single hosts (Wang et al. 2014a,b). The S-type planets have not yet been found in binaries with a period less than 1000 days. Close binary companions play a destructive role in forming S-type planets (Thebault & Haghighipour 2015). First, the protoplanetary disk would be tidally truncated by the secondary companion, reducing the disk mass and lifetime (Artymowicz & Lubow 1994; Miranda & Lai 2015). Second, secular perturbations induced by the companion excite high relative velocities among planetesimals, leading to catastrophic collisions (Heppenheimer 1978; Thebault et al. 2006, but also see Xie et al. 2010b). Third, hydrodynamic simulations showed that the dynamics of disk gas in the binary system is much more complicated than typical assumed static, axisymmetric configurations in theoretical analyses (Paardekooper et al. 2008; Kley et al. 2008; Marzari et al. 2009; Müller & Kley 2012). All above effects are detrimental for the growth of planetesimals. A noteworthy point is that disruptive collisions among planetesimals would produce a reservoir of dust debris. The sweeping of the dust debris can nevertheless boost further mass growth of surviving, leftover planetesimals (Paardekooper & Leinhardt 2010; Xie et al. 2010a).

Although in-situ formation of S-type planets in close binaries seems to be very challenging, it is comparatively easy to form P-type planets. Disk-driven migration needs to be considered when the planets grow massive (Pierens & Nelson 2008b; Kley & Haghighipour 2015). The subsequent evolution of planetary systems includes the MMR capture, excitation of eccentricities along with the gas disk dissipation and planet-planet scattering. Gong & Ji (2018) suggested that S-type planets can form through planet-planet scattering from P-type planets and then tidally capture in various binary configurations. A smaller eccentricity or a lower mass ratio of the binary leads to a higher capture probability up to 10% and produces S-type planets with retrograde orbits, consistent with the result of the two unequal-mass planet ejection model (Gong & Ji 2017).

Another important topic is the orbital configuration of circumbinary planets (P-type orbits). More than 20 circumbinary planets have been detected so far (Schwarz et al. 2016). For those discovered by the Kepler mission, the planets are inferred to be inclined by less than a few degrees relative to the binary plane (Kostov et al. 2014). However, the above coplanarity might be caused by the observational bias: planets orbiting in the binary plane are easier to detect. On the other hand, circumbinary disks with high inclinations have been subsequently discovered (Kennedy et al. 2012; Brinch et al. 2016; Czekala et al. 2019), even on polar orbits (Kennedy et al. 2019). Such misaligned disks may strongly indicate the existence of high-inclination circumbinary planets. In addition, since most stars originate from star clusters and stellar associations (Lada & Lada 2003), the planetary orbits can be influenced by perturbations from stellar fly-bys. This idea has been explored mainly for planetary systems around Sun-like single stellar hosts, and the corresponding studies generally found that close encounters from the passing-by stars are prone to be destructive for planetary systems, reducing their multiplicities with surviving planets on more eccentric and inclined orbits. (Spurzem et al. 2009; Malmberg et al. 2011; Pfalzner 2013; Liu et al. 2013; Hao et al. 2013; Li & Adams 2015; Zheng et al. 2015; Cai et al. 2017; Li et al. 2019a). Similarly, the cluster environment can also significantly impact the planets in binary systems. Ma et al. (2020) showed that these stellar fly-bys can affect the inclination distribution of circumbinary planets. For instance, for close binary systems originating in open clusters with a spacing of >1 AU, a few repeated fly-bys can already induce their highly-inclined orbits.

5 **PEBBLE ACCRETION**

In this section, we recapitulate pebble accretion, including the physical mechanism (Sect. 5.1), the accretion rates and efficiencies in different regimes (Sect. 5.2), the properties (Sect. 5.3) and applications (Sect. 5.4).

### 5.1 Onset and Terminal Condition

In contrast to planetesimal accretion, pebble accretion refers to pebble-sized (~ mm-cm) solid particles that get accreted by planetary bodies (Ormel & Klahr 2010; Lambrechts & Johansen 2012). Since these small pebbles are strongly influenced by the surrounding disk gas, both gas drag and gravitational force play decisive roles during the pebble-planet encounter (also see reviews of Johansen & Lambrechts 2017 and Ormel 2017). Here we provide a
impact parameter during the encounter and $t_{\text{stop}}$ is the stopping time of the pebble, which quantifies how fast the orbit of the pebble is adapted to the gas due to the gas friction force. The pebble reaches a terminal velocity to sediment onto the planet such that $v_{\text{set}} = GM_p t_{\text{stop}}/b^2$. The settling time and encounter time are given by $t_{\text{set}} = b/v_{\text{set}}$ and $t_{\text{enc}} = b/\Delta v$ respectively (see Fig. 6), where $\Delta v$ is the unperturbed relative velocity between the planet and pebble. When the above two terms are equal, it gives the largest pebble accretion radius in the settling regime (Ormel & Klahr 2010; Lambrechts & Johansen 2012; Ida et al. 2016)

$$b_{\text{set}} \approx \sqrt{GM_p t_{\text{stop}} / \Delta v}.$$  \hspace{1cm} (17)

As can be seen from the above equation, $b_{\text{set}}$ is smaller as $t_{\text{stop}}$ becomes shorter. Physically, this is because smaller pebbles with a shorter $t_{\text{stop}}$ are more tightly coupled to gas. The trajectories of these pebbles become strongly affected by the planet only when they sediment deeply enough to the planet. The accretion radius is therefore smaller.

Equivalently, criterion (1) can also be derived from the concept of gravitational deflection (Lambrechts & Johansen 2012). Pebble accretion occurs when the gravitational deflation time $t_{\Delta v} = \Delta v/v_s = \Delta v/(GM_p/b^2)$ is shorter than the pebble stopping time $t_{\text{stop}}$. Equation (17) can be obtained by equating $t_{\Delta v}$ with $t_{\text{stop}}$.

Criterion (2) breaks down when the unperturbed encounter velocity $\Delta v$ is so high that $t_{\text{stop}} > t_{\text{enc}}$. In this limit, gas drag is too weak to influence the orbit of the pebble during this short timespan. By equating the above two timescales, we obtain a threshold velocity $v_s = \sqrt{GM_p/b \tau_s \eta v_K}$, where $v_s \sim v_{\text{enc}},$ pebble accretion is in the settling regime (Ormel & Klahr 2010) (equivalent to the strongly-coupled regime in Lambrechts & Johansen 2012). When $\Delta v \ll v_s$, the accretion enters the inefficient ballistic regime (weakly-coupled regime), where gas drag is unimportant. An exponential decay function is applied to fit such a transition (Ormel & Klahr 2010; Liu & Ormel 2018),

$$f_{\text{set}} = \exp \left[ -0.5 \left( \frac{\Delta v}{v_s} \right)^2 \right].$$  \hspace{1cm} (18)

and the accretion radius is expressed as $b_{\text{PA}} = b_{\text{set}} f_{\text{set}}$.

In order to distinguish from planetesimal accretion, conventionally, pebble accretion refers to the settling regime when the growth is efficiently assisted by the gas drag ($f_{\text{set}} \equiv 1$). However, when the mass of the planet/planetesimal is low or $\tau_s$ is high, the settling condition is not always satisfied and the accretion can be inefficient. We can obtain the critical mass for the onset of efficient pebble accretion ($f_{\text{set}} \equiv 1$) by equating $v_s$ with $\Delta v$ (adopted to be the headwind velocity $n v_K$, see

---

**Fig. 6** Sketch of the pebble-planet encounter, viewed in a co-moving frame with a central planet. The approach velocity is given by the unperturbed relative velocity between these two bodies ($\Delta v$). The perturbed and unperturbed trajectories are represented in the blue solid and dashed lines respectively with an impact distance parameter $b$. The important timescales are: the stopping time of the pebble $t_{\text{stop}}$, the pebble-planet encounter time $t_{\text{enc}} = b/\Delta v$, and the settling time $t_{\text{set}} = b/v_{\text{set}}$, where $v_{\text{set}}$ is the sedimentation velocity when the planet’s gravity equals the gas drag. Figure adopted from Liu & Ormel (2018).

Physical picture of pebble accretion based on the order-of-magnitude timescale analysis. How and in which condition pebble accretion commences are discussed as follows.

During a pebble-planet encounter, the operation of pebble accretion needs to satisfy the following two conditions:

1. The time a pebble settled onto the planet $t_{\text{set}}$ is shorter than the pebble-planet encounter time $t_{\text{enc}}$; otherwise, the pebble cannot be accreted onto the planet.
2. The stopping time of the pebble $t_{\text{stop}}$ is shorter than the pebble-planet encounter time $t_{\text{enc}}$, which means that gas drag matters during pebble-planet interaction; otherwise, the pebble-planet interaction is similar to the planetesimal-planet interaction, where only gravitational force plays a role.

An illustration of the pebble-planet encounter is displayed in Figure 6. When the pebble closely interacts with the planet, the gas drag acceleration $a_{\text{drag}} = (v_{\text{peb}} - v_g)/t_{\text{stop}}$ is adjusted quasi-statically to balance the gravitational acceleration $a_g = GM_p/b^2$, where $b$ is the
Sect. 5.2), which gives

\[ M_{\text{onset}} = \tau_s \eta^3 M_\star \]
\[ = 2.5 \times 10^{-4} \left( \frac{\tau_s}{0.1} \right) \left( \frac{\eta}{2 \times 10^{-3}} \right)^3 \left( \frac{M_\star}{M_\odot} \right) M_\odot, \]

(19)
corresponding to a 500 km radius planetesimal\(^8\). Since the typical planetary bodies we consider are much more massive than \(M_{\text{onset}}\), for simplicity, the referred pebble accretion hereafter only corresponds to the accretion in the settling regime (\(b_{PA} = b_{\text{set}}\)).

The above analysis holds when the planet feedback onto the disk gas is negligible. However, the planet perturbation increases with its growing mass. When the planet is massive enough to open a gap and reverse the local gas pressure gradient, the inward drifting pebbles stop at the local pressure maximum. In this case, pebble accretion cannot proceed, quenching the core mass growth. Such a process leads to a ‘pebble isolation’ from the planet, and the onset mass of the planet is called pebble isolation mass (Lambrechts et al. 2014; Bitsch et al. 2018; Ataiee et al. 2018). For instance, Bitsch et al. (2018) obtained a fitting formula for the pebble isolation mass from hydrodynamical simulations, which reads

\[ M_{\text{iso}} = 25 \left( \frac{h_g}{0.05} \right)^3 \left( \frac{M_\star}{M_\odot} \right) \left[ 0.34 \left( \frac{\eta}{\log_{10} \eta} \right)^4 + 0.66 \right] \left[ 1 - \partial \ln P_g / \partial \ln \nu + 2.5 \right] M_\odot, \]

(20)

where \(\alpha_t\) is the efficiency factor of the turbulent viscosity where \(\nu_t = \alpha_t c_s H_g\) (Shakura & Sunyaev 1973). The termination of pebble accretion corresponds to 10–20 \(M_\odot\) super-Earth planets around solar-mass stars and Earth-mass planets around 0.1\(M_\odot\) late M-dwarf stars.

### 5.2 Accretion Rate and Efficiency

As shown in Equation (17), the accretion radius depends on the relative velocity between the pebble and planet \(\Delta v\). For a planet on a circular orbit, \(\Delta v\) is a sum of the headwind velocity \(\eta \nu_K\) and the Keplerian shear velocity \(\Omega_K b_{\text{set}}\). The accretion is in the headwind regime (Bondi regime) when \(\Delta v\) is dominated by \(\eta \nu_K\), while it is in the shear regime (Hill regime) when \(\Delta v\) is dominated by the Keplerian shear. The transition planet mass between the headwind and shear regimes can be calculated by equating the above two velocities,

\[ M_{\text{hw/sh}} = \eta^3 M_\star / \tau_s \]
\[ \approx 0.025 \left( \frac{\eta}{2 \times 10^{-3}} \right)^3 \left( \frac{\tau_s}{0.1} \right)^{-1} \left( \frac{M_\star}{M_\odot} \right) M_\odot. \]

(21)

When a planet is on an eccentric and inclined orbit, \(\Delta v\) is additionally contributed by the epicyclic motion \((e v_K\) and \(i \nu_K\) of the planet relative to its Keplerian velocity (Johansen et al. 2015; Liu & Ormel 2018; Ormel & Liu 2018).

The pebble mass accretion rate can be expressed as (Lambrechts & Johansen 2014; Morbidelli et al. 2015)

\[ \dot{M}_{PA} = \frac{2 \eta b_{\text{set}} \Delta v \Sigma_{\text{peb}}}{\sqrt{2 \pi}} \approx \frac{2 \sqrt{GM_{\text{stop}} \Delta v \Sigma_{\text{peb}}}}{\sqrt{2 \pi} H_{\text{peb}}} \]

(22)

where \(\Sigma_{\text{peb}} = \sqrt{2 \pi} H_{\text{peb}} / \rho_{\text{peb}}\). The scale height of the pebble disk is given by Youdin & Lithwick (2007),

\[ H_{\text{peb}} = \sqrt{\frac{\delta_d}{\rho_g}} H_g, \]

(23)

where \(\delta_d\) represents the coefficient of the gas diffusivity, which approximates to the efficiency factor of the turbulent viscosity \(\alpha_t\) when the disk turbulence is driven by the MRI (Johansen & Klahr 2005; Zhu et al. 2015)\(^9\). In order to distinguish the dominant accretion regime, one can numerically compare \(\dot{M}_{PA,2D/3D}\) to see which one is higher for given disk and planet parameters. From a physical perspective, whether the accretion is in 2D/3D is determined by the ratio between the pebble accretion radius \(b_{\text{set}}\) and the pebble scale height \(H_{\text{peb}}\) (Morbidelli et al. 2015). When the pebble accretion radius is larger than the pebble scale height, it is in the 2D accretion regime; otherwise, it is in the 3D accretion regime.

The disk pebble flux that bypasses the orbit of the planet is \(\dot{M}_{\text{peb}} = 2\pi r \Sigma_{\text{peb}} v_r\). The pebble accretion efficiency is defined as the probability of pebbles accreted by the planet, \(\varepsilon_{PA} = \dot{M}_{PA} / \dot{M}_{\text{peb}}\) (Guillot et al. 2014; Lambrechts & Johansen 2014). When the radial velocity of the gas is neglected in \(v_r\) (Eq. (6)), the efficiency can be written as

\[ \varepsilon_{PA,2D/3D} = \frac{2 \pi r \Sigma_{\text{peb}}}{\varepsilon_{PA,2D/3D}} \]

(24)

---

\(^8\) The above expression is consistent with eq. (25) of Visser & Ormel (2016).

\(^9\) It is worth noting that the non-ideal MHD effects, such as ambipolar diffusion and Hall Effect, also play important roles in distributing angular momentum of disk gas. The above processes crucially depend on the disk chemistry and the geometry and strength of the magnetic field (Bai & Stone 2013; Bai 2015, 2016; Gressel et al. 2015).
The pebble accretion timescale is therefore given by

\[
\tau_{PA} = \begin{cases} 
4 \times 10^{-3} \left( \frac{M_p}{0.02 M_\oplus} \right)^{1/2} \left( \frac{\tau_s}{0.1} \right)^{1/2} \left( \frac{\eta}{1.8 \times 10^{-3}} \right)^{1/2} [2D \text{ headwind}], \\
4 \times 10^{-3} \left( \frac{M_p}{0.02 M_\oplus} \right)^{2/3} \left( \frac{\tau_s}{0.1} \right)^{-1/3} \left( \frac{\eta}{1.8 \times 10^{-3}} \right)^{-1} [2D \text{ shear}], \\
4 \times 10^{-3} \left( \frac{M_p}{0.02 M_\oplus} \right) \frac{h_{\text{peb}}}{r} \left( \frac{\eta}{1.8 \times 10^{-3}} \right)^{-1} \left( \frac{\tau_s}{0.1} \right)^{-1} [3D], \\
\end{cases}
\]

where \( h_{\text{peb}} = H_{\text{peb}}/r \) is the pebble disk aspect ratio. The full expression of \( \varepsilon_{PA} \) including the eccentricity and inclination dependences can be found in Liu & Ormel (2018) and Ormel & Liu (2018). The pebble accretion efficiency is a crucial quantity in planet formation, since it corresponds to how efficiently the disk pebble mass can be converted into planet mass.

The pebble accretion timescale is therefore given by

\[
t_{PA} = \frac{M_p}{\varepsilon_{PA} \alpha_{PA}} = \begin{cases} 
5 \times 10^4 \left( \frac{M_p}{0.02 M_\oplus} \right)^{1/2} \left( \frac{\tau_s}{0.1} \right)^{1/2} \left( \frac{h_{\text{peb}}}{r} \right)^{1/2} \left( \frac{\eta}{1.8 \times 10^{-3}} \right)^{1/2} \left( \frac{10^{-4} M_\oplus \text{yr}}{M_p} \right)^{1/2} [2D \text{ headwind}], \\
5 \times 10^4 \left( \frac{M_p}{0.02 M_\oplus} \right)^{1/3} \left( \frac{\tau_s}{0.1} \right)^{1/3} \left( \frac{h_{\text{peb}}}{r} \right)^{1/3} \left( \frac{\eta}{1.8 \times 10^{-3}} \right)^{1/3} \left( \frac{10^{-4} M_\oplus \text{yr}}{M_p} \right)^{1/3} [2D \text{ shear}], \\
5 \times 10^4 \left( \frac{M_p}{0.02 M_\oplus} \right) \left( \frac{\eta}{1.8 \times 10^{-3}} \right) \left( \frac{10^{-4} M_\oplus \text{yr}}{M_p} \right)^{1/3} \left( \frac{h_{\text{peb}}}{r} \right)^{1/3} \left( \frac{10^{-4} M_\oplus \text{yr}}{r} \right) [3D], \\
\end{cases}
\]

We can see that \( t_{PA} \) is proportional to \( M_p^{1/2} \) or \( M_p^{1/3} \) in the 2D headwind or shear accretion regime, while \( \tau_{PA} \) is independent of the planet mass in the 3D regime.

### 5.3 Features

The first important feature is that pebble accretion is not a runaway process. From Equation (22) we can see that \( (dM/dt)/M \propto M^0 \) in the 3D regime, and \( (dM/dt)/M \propto M^{-1/2} \) or \( M^{-1/3} \) in the 2D headwind or shear regime. This means that the mass ratios among the growing bodies would approach the order of unity as the accretion proceeds. This is the feature of the orderly growth. Nevertheless, even not in a runaway mode, pebble accretion can still be very fast due to a large accretion cross section and a high flux of continuous feeding pebbles from the outer part of the disk (see discussions in Sect. 6).

The efficiency of pebble accretion is determined by the Stokes number of the pebbles. When pebble accretion is in the 3D regime, the efficiency increases with the Stokes number. This is because higher Stokes number pebbles sediment into a thinner vertical layer, increasing the number density of pebbles for being accreted. When the accretion is in the 2D regime, the efficiency decreases with \( \tau_s \). In this case, higher Stokes number pebbles drift faster and have a lower probability to be accreted by the planet (see fig. 4 of Ormel & Liu 2018). The above dependences hold when pebbles are marginally coupled to the disk gas (\( 10^{-3} \lesssim \tau_s \lesssim 1 \)). When \( \tau_s \) is much larger than the order of unity, gas drag is negligible and pebbles are more aerodynamic like planetesimals. In that circumstance, the actual accretion rate drops substantially (Ormel & Klahr 2010). Conversely, when \( \tau_s \lesssim 10^{-3} \), the pebbles are tightly coupled to the gas flow. Accretion rate is also very low in this geometric regime (Guillot et al. 2014). Therefore, the preferred Stokes number for pebble accretion ranges from \( 10^{-3} \) to 1, corresponding to 0.3 mm to 30 cm size particles at 5 AU in the MMSN model.

Pebble accretion is suppressed when the disk is highly turbulent (Morbidelli et al. 2015; Ormel & Liu 2018; Rosenthal et al. 2018). There are two reasons. First, pebbles are stirred up vertically due to the turbulent diffusion. In a strongly turbulent case, a vertically extended distribution of pebbles leads to less amount of them being accreted, and the corresponding efficiency is lower. This is similar to the effect of the smaller pebbles in the 3D regime discussed above. Second, the pebble’s random velocity also correlates with disk turbulence as \( \sqrt{\alpha_s \alpha_c} \) (Ormel & Cuzzi 2007). The impact velocity between the pebble and planet is additionally contributed by this turbulent-induced motion. The settling condition fails when the turbulent velocity is very high (Sect. 5.1), and therefore, the accretion is also significantly suppressed. Ormel & Liu (2018) obtained a pebble accretion efficiency formula by accounting for the stochastic turbulent velocity into the equation of motion for the pebble. Their results are in agreement with the pebble accretion measured from more realistic MHD simulations of Xu et al. (2017) as well as the VSI hydrodynamic simulations of Picogna et al. (2018).

We note that the trajectories of pebbles and accretion features could deviate from the above described paradigm that only steady-state shear gas flow is considered (Sect. 5.1). There are two additional effects. First, when pebbles are sedimented and accreted onto the planet, the potential energy of pebbles is transferred into frictional heat, which raises the temperature of surrounding gas. The deep gas layer close to the planet is more dynamically convective caused by this accretion-driven heating, which may affect the pebble accretion. Taking into account both adiabatic and convective models of pebble accretion in
hydrodynamic simulations, Popovas et al. (2018) found that even though an active mass mixing among different layers is indeed observed due to the vigorous gas motion, the net pebble accretion is not strongly affected, except for the smallest particles that are tightly coupled to the gas.

Second, when the pebbles fall into the region close to the planet with a temperature exceeding their evaporation temperature, these pebbles get vaporized, resulting in an enrichment of the planetary envelope rather than direct accretion onto the core (Alibert 2017; Brouwers et al. 2018; Valletta & Helled 2019). This enrichment increases the gas mean molecular weight of the envelope, resulting in a thinner, opaque envelope (Venturini et al. 2016). Meanwhile, hydrodynamic simulations indicated that the envelopes of gas for low-mass protoplanets are not in a steady-state but rather get replenished by the surrounding disk gas (Ormel et al. 2015; Fung & Dong 2015; Cimerman et al. 2017). It is unclear how largely the envelope enrichment process would be affected by the above gas recycling. The ablation of accreting pebbles with realistic radiative transfer plus gas replenishment models, and how these affect the core mass growth and gas accretion are an active research topics for future investigations.

5.4 Applications

Solar System
The pebble accretion scenario has been used to explain the formation of the Solar System. Based on the fact that icy pebbles drift across the water-ice line and sublimate into small silicate pebbles, Morbidelli et al. (2015) inferred that the growth of protoplanets is in a slow 3D accretion interior to the ice line and a fast 2D accretion exterior to the ice line. This results in low-mass progenitors of terrestrial planets in the inner disk regions and massive cores of giant planets in the outer disk regions, in agreement with the architecture of the Solar System. Levison et al. (2015) found that before efficient pebble accretion commences, an early phase of velocity stirring of protoplanets is required. Protoplanets then evolve into a bimodal mass distribution. Due to dynamical friction, the massive protoplanets have low velocity dispersions. After this phase, only these massive protoplanets can undergo rapid pebble accretion and grow to giant planet cores. The above self-excitation process is essential to explain why only a few giant planet cores formed in the Solar System. If there were no such pre-stirring phase, hundreds of Earth-mass objects would form instead (Kretke & Levison 2014). Johansen et al. (2015) proposed the formation of asteroids and Kuiper belt objects by accreting chondrules, which are millimeter-sized spherules commonly found in primitive meteorites. Furthermore, formation of the Galilean satellites has also been recently proposed to be aided by pebble accretion (Shibaike et al. 2019; Ronnet & Johansen 2020).

Exoplanetary Systems
The pebble accretion scenario is also widely adopted to explain the formation of super-Earths and gas giants in exoplanetary systems. Lambrechts & Johansen (2014) constructed a dust growth and pebble drift model to investigate the formation of giant planet cores by pebble accretion. Bitsch et al. (2015) studied the influence of disk radial distance and evolutional phase on the growth and migration of a single protoplanet, and Johansen et al. (2019) focused on conditions under which growth can overcome migration. Numerous recent studies have incorporated the pebble accretion model into an N-body code to investigate the impact of pebble accretion on final planetary architectures (Matsumura et al. 2017; Lambrechts et al. 2019; Bitsch et al. 2019; Izidoro et al. 2019; Liu et al. 2019b; Schoonenberg et al. 2019; Coleman et al. 2019; Wimarsson et al. 2020; Ogihara et al. 2020). For instance, Lambrechts et al. (2019) demonstrated that the final type of planetary system (terrestrial planets or super-Earths) is crucially determined by the pebble flux, or equivalently, the total mass of the pebble reservoir in a protoplanetary disk.

Systems around Low-mass Stars
Studies of pebble accretion are not only limited to the systems around solar-mass stars, but also are generalized to stellar hosts of different masses. Ormel (2017) proposed a formation scenario for TRAPPIST-1 (Gillon et al. 2016, 2017) and other compact systems around very low-mass M-dwarf stars. In their scenario, protoplanets form by the streaming instability at the water-ice line. These protoplanets subsequently undergo inward migration and accrete pebbles to reach their final masses. Ormel (2017) suggested that all TRAPPIST-1 planets of roughly Earth-mass could be an indication of the planet mass regulated by pebble isolation. The follow-up numerical simulations by Schoonenberg et al. (2019) verified that these forming planets have ≈10% water mass fractions, consistent with bulk density measurements and interior modeling of the TRAPPIST-1 planets (Grimm et al. 2018; Unterborn et al. 2017; Dorn et al. 2018).

Liu et al. (2019a) investigated pebble-driven planet formation around stars of masses from 0.08 $M_\odot$ to 1 $M_\odot$. Figure 7 illustrates the planet populations from observations and the population synthesis model of Liu et al. (2019a). It should be noted that the observed sample is adopted from different surveys and uncorrected for any selectional bias. The figure nevertheless illustrates important features. First, a paucity of giant planets but not super-Earths is found around stars with masses below $\lesssim0.1–0.2$ $M_\odot$ (Figure 7a). Since larger planets are easier to detect than smaller planets for the same mass stellar
The validity of this positive correlation also needs to be unbiased and complementarily tested by current and future RV programs for a wider mass range of stellar hosts, such as High Accuracy Radial Velocity Planet Searcher, (HARPS, Mayor et al. 2003, 2011; Udry et al. 2019), Echelle Spectrograph for Rocky Exoplanet and Stable Spectroscopic Observations (ESPRESCO, Pepe et al. 2010), Planet Finder Spectrograph, (PFS, Crane et al. 2010; Feng et al. 2019), and Calar Alto high-Resolution search for M dwarfs with Exoearths with Near-infrared and optical Echelle Spectrographs (CARMENES, Quirrenbach et al. 2016).

Importantly, Liu et al. (2019a) proposed that the characteristic mass (core mass) of super-Earths is set by the pebble isolation mass, which increases linearly with the mass of the stellar host. The pebble isolation mass can be written as (Liu et al. 2019a)

$$M_{iso} = 25 \left( \frac{M_\star}{1 \ M_\odot} \right) \left( \frac{h_g}{0.05} \right)^3 M_\oplus = 25 \left( \frac{M_\star}{1 \ M_\odot} \right)^{4/3} \left( \frac{\dot{M}_{\text{vis}}}{6 \times 10^{-8} M_\odot \text{ yr}^{-1}} \right)^{2/3} .$$

(26)

In the above equation, we neglect $\alpha$ and $\eta$ dependencies on $M_{iso}$. $M_{iso}$ is the fiducial disk accretion rate around the solar-mass star, and $h_g$ is derived from the viscously heated disk model by assuming that the disk accretion rate scales with the stellar mass squared. Figure 7(b) shows the resultant planets generated from the population synthesis model, and the solid line refers to the pebble isolation mass (Eq. (26)). We clearly see that the super-Earths reaching $M_{iso}$ increase with their stellar masses, approximately Earth-mass terrestrial planets around stars of $0.1 \ M_\odot$ and $10-20 \ M_\odot$ around solar-mass stars. We also note that $M_{iso}$ decreases when the disk evolves and $h_g$ as well as $M_\star$ decline. For instance, planets around solar-mass stars reach $M_{iso}$ of $7 \ M_\oplus$ when $\dot{M}_{\text{vis}} = 10^{-8} M_\odot \text{ yr}^{-1}$. Since $M_{iso}$ is so low around late M-dwarfs, massive gas giants are unlikely to form in such systems through the pebble accretion planet formation channel. Liu et al. (2020) further applied this approach for even lower-mass central hosts and found the above linear mass scaling holds for planets around brown dwarfs.

6 COMPARISON BETWEEN PEBBLE ACCRETION AND PLANETESIMAL ACCRETION

We compare the efficiency of pebble accretion and planetesimal accretion in Section 6.1. The stellar mass and
We highlight the importance of incorporating these two accretion mechanisms for planet formation in Section 6.3.

6.1 Why Pebble Accretion is More Efficient than Planetesimal Accretion

Pebble accretion has gained attention since it is a more efficient growth process compared to planetesimal accretion. This comparison can be demonstrated from the following two aspects: the accretion cross section and total mass of feeding materials.

For planetesimal accretion, the gravitational focusing factor reaches its maximum value when \( \delta v = v_H \) (see Sect. 4.1). Therefore, the ratio between accretion radius and physical radius is expressed as

\[
b_{PA}/R = \sqrt{1 + \left( \frac{v_{esc}}{R_H \Omega_K} \right)^2} \approx 2 \left( \frac{M_p}{M_*} \right)^{1/6} \left( \frac{a}{R} \right)^{1/2} \left( \frac{\rho_*}{5 \text{ g cm}^{-3}} \right)^{1/2} \left( \frac{M_*}{M_\odot} \right)^{-1/6}.
\]

(27)

In the shear-dominated pebble accretion regime \( (\Delta v \approx b_{peb} \Omega_K) \), from Equation (17) the accretion radius reads

\[
b_{PA} \approx \sqrt{\frac{\Sigma_{pl} \dot{M}_{\text{stop}}}{\Delta v}} \approx \tau_s^{1/3} R_H.
\]

(28)

The pebble accretion radius can maximally approach the planet Hill radius when \( \tau_s \) reaches the order of unity. The enhanced factor is given by

\[
b_{PA}/R \approx \tau_s^{1/3} \left( \frac{R_H}{R} \right) \approx 230 \tau_s^{1/3} \left( \frac{a}{1 \text{ AU}} \right) \left( \frac{\rho_*}{5 \text{ g cm}^{-3}} \right)^{1/3} \left( \frac{M_*}{M_\odot} \right)^{-1/3}.
\]

(29)

From Equations (27) and (29), we can see that generally the pebble accretion has a larger accretion cross section compared to the planetesimal accretion. This feature can be seen in Figure 8, which illustrates the trajectories of planetesimals (left) and pebbles (right) during their interactions with an Earth-mass planet. Pebbles lose angular momentum more efficiently and sediment towards the planet by a combination of gas drag and gravitational force. The above differences are more pronounced when the Stokes number of the pebbles is closer to the order of unity, and the planet is further out and/or around less massive stellar hosts.

In addition, since planetesimals (~100 km) are weakly affected by gas, the orbits of the planetesimals are relatively fixed. The planetary bodies only accrete their local planetesimals. The maximum mass that the planet can reach (planetesimal isolation mass) is written as

\[
M_{pl,\text{iso}} = 2\pi a \Delta \alpha \Sigma_{pl}
\]

where the width of the feeding zone \( \Delta \alpha \sim 10 R_H \). For the MMSN, the planet grows maximally to Mars mass in the inner terrestrial planet region and a few Earth masses beyond the water ice line (e.g., Ida & Lin 2004a).

On the contrary, pebble accretion is not limited to the local pebble density. The majority of the dust reside in the outer part of the disk. Pebbles, which grow from dust grains, drift inwardly from the outer region of the disk. Due to the mobility of pebbles, there is no concept of feeding zone for pebble accretion. The feeding materials correspond to all pebbles that are able to bypass the orbit of the planet. The planets stop their core growth only when they reach the pebble isolation mass, which has a similar scaling as the gap opening mass (Eq. (20)). Therefore, in contrast to planetesimal accretion, the final core mass of the planet by pebble accretion mainly depends on the mass of the central star and the aspect ratio of the disk, but not the local density of solids.

6.2 Stellar Mass and Radial Distance Dependence

Here we qualitatively discuss how these two accretion scenarios depend on the mass of the stellar host and the orbital distance.

First, we focus on the stellar mass dependence. Based on Equation (15), the growth timescale for planetesimal accretion is given by \( t_{PA} \propto \Sigma_{pl}^{-1} M_*^{-1/2} \). We have \( t_{PA} \propto M_{peb}^{-1} M_*^{1/2} \) in the 2D headwind regime from Equation (25). Since observations indicate lower-mass disks around less massive massive stars (Natta et al. 2006; Andrews et al. 2013; Pascucci et al. 2016), the planetesimal surface density and pebble flux are expected to decrease with the mass of the stellar host. As a result, both pebble accretion and planetesimal accretion tend to be slower as \( M_* \) decreases. Furthermore, assuming that the planetesimal surface density and pebble flux have the same stellar mass dependence, we find that the ratio of these two timescales \( (t_{PA}/t_{PlA}) \) increases with \( M_* \). This in turn indicates that the decreasing rate of planetesimal accretion is faster than that of pebble accretion as \( M_* \) decreases. In other words, pebble accretion is even more pronounced compared to planetesimal accretion for the planet growth around lower mass stars than higher mass stars. The above conclusion holds as well when adopting \( t_{PA} \) for the 2D shear regime (Eq. (25)) or \( t_{PlA} \) corresponding to the oligarchic growth (Eq. (14)).

The radial distance dependence is explained as follows. Based on Equation (15), the planetesimal accretion timescale is written as \( t_{PlA} \propto \Sigma_{pl}^{-3/2} \). Since the disk surface density additionally decreases with \( r \), \( t_{PlA} \) is expected to be a strongly increasing function of \( r \). Planetesimals spend a much longer time growing...
their masses at further out disk locations. Moreover, as discussed in Section 4.4, planet-planet encounters result in scatterings/ejections instead of collisions at large orbital distances. Therefore, the growth by planetesimal accretion is strongly suppressed or even quenched at large orbital distances. On the other hand, for pebble accretion, $t_{\text{PA}} \propto \eta$ in the 2D shear regime. Although the growth by pebble accretion also turns out to be slower at further out disk locations, the radial distance dependence is weaker than planetesimal accretion. Therefore, pebble accretion is more appealing for the formation of distant massive planets.

To summarize, when comparing these two accretion scenarios, we find that pebble accretion becomes more attractive than planetesimal accretion when the stellar host is less massive and/or the accretion occurs at a larger orbital distance.

### 6.3 A Hybrid Accretion of Pebbles and Planetesimals

Despite the above distinctive differences, we want to raise the point that pebble accretion and planetesimal accretion are not two isolated, mutually exclusive growth channels. They are nevertheless likely to be connected and complementary for planet growth.

On the one hand, the above two mechanisms can operate concurrently and jointly contribute to the mass increase. On the other hand, they also compete at certain levels since both pebbles and planetesimals are basic components of solids in protoplanetary disks (Schoonenberg et al. 2019). For instance, the streaming instability converts pebbles into planetesimals once the triggering condition is satisfied. After that, these forming planetesimals accrete surrounding planetesimals as well as pebbles that continuously drift from the outer region of the disk. The following mass growth is a combination of accreting pebbles and planetesimals (Liu et al. 2019b; Schoonenberg et al. 2019). We can consider two types of extreme situations. In one circumstance, the streaming instability is extensively triggered and the majority of the solids are in the form of planetesimals. The following growth is naturally led by planetesimal accretion. In the other circumstance, when the streaming instability is modestly triggered, the dominant solid masses are still in pebbles. Therefore, pebble accretion is the central mechanism for the growth of the cores. As can be expected, a more general pattern is that pebble and planetesimal accretions co-operate together to feed the planet growth.

In addition, the above two mechanisms may occur at different evolutionary stages (Alibert et al. 2018; Venturini & Helled 2020) and/or in different disk regions (Ormel et al. 2017). For example, when a planet reaches the pebble isolation mass, the inwardly drifting pebble flux is truncated. Pebble accretion and planetesimal accretion

---

**Fig. 8** Illustration of planetesimal-planet (left) and pebble-planet (right) encounters, viewed in a co-moving frame with a central planet. The filled circle signifies an Earth-mass planet and the dashed circle indicates the planet Hill sphere. Two blue lines represent the trajectories of planetesimals or pebbles from the two different impact distances. In the above illustration, planetesimals are deflected (left) while pebbles get accreted (right). We note that the central filled circle is not scaled with the physical size of the planet. We set a zero-drag force for the planetesimal-planet interaction and pebbles of $\tau_s = 1$ for the pebble-planet interaction, corresponding to $\approx 10$ cm size at 5 AU in the MMSN model. The black dots represent the snapshots of planetesimals or pebbles, whereas the green and orange arrows correspond to the planet gravitational force and gas drag respectively. The direction and length of the arrow represent the direction and magnitude of the force respectively. Aided by gas drag, pebbles can be more efficiently accreted by the planet compared to planetesimals.
occur for planets whose orbits are beyond the gap-opening planet. Interior to this massive planet, collisions between planetesimals and protoplanets are the only pathway for core mass growth. This might provide an explanation for the mass dichotomy between the inner low-mass terrestrial planets and outer massive giant planets in the Solar System (Ormel et al. 2017). All in all, the detailed exploration based on the concept of hybrid planetesimal and pebble accretion is an active research area and requires future studies.

7 SUMMARY AND FUTURE OUTLOOK

In this review, we have recapitulated the current states of exoplanet demographics and disk observations (Sect. 1). The planet formation theories have been overviewed chronologically, including dust coagulation and radial drift (Sect. 2), planetesimal formation (Sect. 3), and subsequent planetesimal growth by planetesimal accretion (Sect. 4) and pebble accretion (Sect. 5). Importantly, we have discussed how different planet formation models fit into observations in each growth stage.

Lastly, we propose some open questions in this field, which are existing topics with disputed interpretations and unsolved puzzles that require future studies. These questions are summarized as follows:

– What is the characteristic size of solid particles in protoplanetary disks? Are these particles mm-cm size as suggested by spectral index observations, or 100 μm as inferred from polarization measurements? On the one hand, the answer itself is valuable, since it needs to test the validity of different model interpretations and the built-in assumptions. On the other hand, the answer is also related to the subsequent planet formation processes such as the streaming instability and pebble accretion, which essentially depend on the size of solid particles.

– Current streaming instability and pebble accretion studies are limited to the systems around single stars. Since the gas streamline in protoplanetary disks around binaries significantly deviates from static, axisymmetric cases around single stellar hosts, it still remains to be seen how these gas-assisted formation mechanisms operate in this highly-perturbed binary environment.

– How can we distinguish planetesimal accretion and pebble accretion mechanisms from an observational perspective? Namely, what would be the key observational signatures that result from different formation channels (see e.g., Brügger et al. 2020)? Can we really claim which types of systems can be uniquely produced by pebble accretion, or vice versa?

– Recent impact models found that the accreting pebbles as well as small planetesimals may get vaporized on their way to the planet interior due to the increased thermal ablation and friction (see Sect. 5.3). The materials enrich the planetary atmospheres rather than directly impact the solid cores. The resultant core mass is substantially lower than the traditional critical mass that a planet requires for initiating rapid gas accretion (Brouwers & Ormel 2020). How this mass deposition process affects the compositional and thermal structures of envelopes, and further influences the gas accretion are not well understood. In addition, small bodies in protoplanetary disks feel head wind from the disk gas and suffer the surface shear stress (Paraskov et al. 2006; Schaffer et al. 2020). The influence of wind erosion during the above two accretion processes is also poorly investigated.

– How early can planets form? There seems to be evidence of large grains existing in the early Class I disk phase (Harsono et al. 2018). Based on the isotope measurements of meteorites, the core of Jupiter with \( \sim 20 \, M_{\oplus} \) is suggested to form early within 0.5–1 Myr (Krijger et al. 2017). Furthermore, if the rings and gaps in the HL Tau disk are induced by planets, it then indicates that planets of sub-Jupiter masses might already form at large orbital separations within 1 Myr in such a young system. Putting all clues together, can we speculate that planet formation is more rapid than what we thought before?

– We have already learned much about of the planet-related properties from exoplanet demographics: occurrence rate, mass, radius and corresponding dependencies on the mass and metallicity of the stellar host. The next step is to understand the planetary-system-related properties, i.e., the architectures. Can different populations of planets co-exist with each other under certain conditions, or the formation of one type of planets inhibit the growth of the others? What would be the architecture of these systems? Some of these questions have already been pointed out by Zhu & Wu (2018) and Masuda et al. (2020). In order to further address these questions, both follow-up observations and dedicated numerical modelings are needed.

– As raised by Murchikova & Tremaine (2020), is planet formation really an independent and generalized process such that planets do not know each other and the environment where they reside?
**Table 1 List of Notations**

| Symbol | Description |
|--------|-------------|
| $v_g$ | gas velocity with azimuthal and radial components |
| $v$ | particle’s velocity with azimuthal and radial components |
| $v_{peb}$ | pebble velocity |
| $v_{K}$ | Keplerian velocity |
| $v_{th}$ | gas mean thermal velocity |
| $v_{set}$ | settling velocity of pebble onto planet |
| $v_{esc}$ | planet escape velocity |
| $\Delta v$ | relative velocity between gas and particle |
| $\delta v$ | relative velocity between large and small planetesimals |
| $\Omega_K$ | Keplerian angular frequency |
| $\eta$ | gas headwind prefactor |
| $P_g$ | gas pressure |
| $T_g$ | gas temperature |
| $T_d$ | dust temperature |
| $H_g$ | gas disk scale height |
| $H_{peb}$ | pebble disk scale height |
| $H$ | vertical height of small planetesimals |
| $h_g$ | gas disk aspect ratio |
| $h_{peb}$ | pebble disk aspect ratio |
| $\Sigma_g$ | gas surface density |
| $\Sigma_d$ | dust surface density |
| $\Sigma_{peb}$ | pebble surface density |
| $\rho_g$ | gas volume density |
| $\rho_{peb}$ | pebble volume density |
| $Z$ | disk metallicity |
| $Q_T$ | Toomre parameter |
| $\gamma$ | self-gravity parameter |
| $G$ | gravitational constant |
| $F_{drag}$ | hydrodynamic gas drag force |
| $a_{drag}$ | hydrodynamic gas drag acceleration |
| $C_D$ | coefficient of gas drag force |
| $R_o$ | particle’s Reynolds number |
| $\nu_{mol}$ | kinematic molecular viscosity |
| $\nu_t$ | turbulent viscosity |
| $\sigma_{mol}$ | collisional cross section of gas molecule |
| $\lambda_{mfp}$ | gas mean free path |
| $m_{mol}$ | mass of gas molecule |
| $m_H$ | mass of hydrogen atom |
| $\mu$ | gas mean molecular weight |
| $\rho_{**}$ | internal density of solid body |
| $r$ | disk radial distance |
| $R$ | radius of solid body |
| $R_H$ | Hill radius |
| $a$ | semimajor axis |
| $t_{stop}$ | stopping time |
| $\tau_a$ | Stokes number or dimensionless stopping time |
| $\tau_\nu$ | optical depth (at the corresponding frequency $\nu$) |
| $I_{\nu}$ | disk emission intensity |
| $F_{\nu}$ | observed flux |
| $\kappa_\nu$ | disk opacity |
| $B_\nu$ | Plank function |
| $\omega_\nu$ | single-scattering albedo |
| $c$ | light speed |
| $\delta_\lambda$ | coefficient of gas diffusivity |
| $\alpha_{t}$ | efficiency factor of turbulent viscosity |
| $\alpha$ | spectral energy index |
| $\beta$ | dust opacity index |
| $\beta_{ISM}$ | opacity index of ISM dust |
| $\beta_{disk}$ | opacity index of protoplanetary dust |
| $k_B$ | Boltzmann constant |
| $k_Z$ | gradient of disk surface density |
| $k_T$ | gradient of disk temperature |
| $M_*$ | stellar mass |
| $M_{pl}$ | characteristic mass of planetesimal |
| $M_P$ | planet mass |
| $M_{LM}, m$ | mass of large and small planetesimals |
| $R_{M}, R_m$ | radius of large and small planetesimals |
| $\sigma$ | collisional cross section of large planetesimal |
Table 1  Continued.

| Symbol | Description |
|--------|-------------|
| $n_{\text{sm}}, n_{\text{LM}}$ | number density of small and large planetesimals |
| $\Sigma_{\text{sm}}, \Sigma_{\text{LM}}$ | surface density of small and large planetesimals |
| $t_{\text{set}}$ | settling timescale of pebble onto planet |
| $t_{\text{acc}}$ | pebble-planet encounter timescale |
| $t_{\text{PA}}$ | pebble accretion timescale |
| $t_{\text{PA}}$ | planetesimal accretion timescale |
| $t_{\text{grow}}$ | mass growth timescale |
| $t_{\text{vis}}$ | viscous stirring timescale |
| $t_{\text{e,GM}}$ | gas eccentricity damping timescale |
| $t_{\text{og}}$ | oligarchic growth timescale |
| $t_{\text{g}}$ | runaway growth timescale |
| $C_{\text{RG}}$ | numerical prefactor of runaway growth |
| $\Phi$ | ratio between planetary escape velocity and Keplerian velocity |
| $\ln \Lambda$ | Coulomb factor |
| $b$ | impact parameter during pebble-planet encounter |
| $b_{\text{set}}$ | pebble accretion radius in the settling regime |
| $b_{\text{PA}}$ | pebble accretion radius |
| $b_{\text{PA}}$ | planetesimal accretion radius |
| $f_{\text{fit}}$ | fitting function between settling and ballistic regimes |
| $f_{\text{gf}}$ | gravitational focusing factor |
| $M_{\text{onset}}$ | planet mass for onset of efficient pebble accretion |
| $M_{\text{th}}$ | planet mass for termination of pebble accretion |
| $M_{\text{hw/sh}}$ | planet mass between headwind and shear-dominated pebble accretion regimes |
| $b_{\text{PA}}$ | gas disk accretion rate |
| $M_{\text{d}}$ | gas disk accretion rate around solar-mass star |
| $M_{\text{peb}}$ | disk pebble flux |
| $M_{\text{PA}}$ | pebble accretion rate |
| $\varepsilon_{\text{PA}}$ | pebble accretion efficiency |

Or is planet formation universally set by a few key physical processes related to disk and host properties (Liu et al. 2019a), and therefore, the imprinted planetary systems may contain some degrees of intra-system/inter-system similarity (Millholland et al. 2017; Wang 2017; Weiss et al. 2018; Wu 2019)?

With a rapidly increasing number of characterized exoplanets and protoplanetary disks, we have greatly improved the statistics of planet formation, in the context of both initial conditions and final products. The theoretical study of planet formation has been advanced enormously in the last decade. The validity of modern planet formation scenarios needs to match various observational constraints. Future work is also required to account for the consistency with findings from ongoing/upcoming space and ground-based facilities, such as TESS and ALMA.

Acknowledgements We thank editor Wing-Huen Ip and the anonymous referee for useful comments and suggestions. We also thank Anders Johansen, Michiel Lambrechts, Joanna Drążkowska, Chris Ormel, Rixin Li, Jiwei Xie, Shangfei Liu, Wei Zhu and Feng Long for proof-reading the manuscript and providing helpful suggestions. B.L. greatly appreciates the valuable discussions with Gijs Mulders, Daniel Harsono, Joanna Drążkowska and Chris Ormel as a PPVII chapter proposal team.

In addition, B.L. wishes to express the deepest gratefulness to Adam Showman (1968–2020) for his inspiration and guidance in B.L.’s early academic career. Adam is an amazing scientist and a tremendous mentor. His scientific contribution as well as great personality will keep influencing and being remembered by the community. Lastly, B.L. feels especially grateful to his girlfriend, Jing Yang, for her spiritual support during the writing period. In these tough COVID-19 pandemic times, many things that formerly seemed extraordinary have now become ordinary. The work would not be done successfully without her dedicated encouragements.

B.L. is supported by the European Research Council (ERC Consolidator Grant 724687-PLANETESYS), the Swedish Walter Gyllenberg Foundation, and start-up grant of Bairen program from Zhejiang University. J.J. is supported by the B-type Strategic Priority Program of the Chinese Academy of Sciences (Grant No. XDB41000000), the National Natural Science Foundation of China (Grant Nos. 12033010 and 11773081), CAS Interdisciplinary Innovation Team and Foundation of Minor Planets of the Purple Mountain Observatory.

References

Abod, C. P., Simon, J. B., Li, R., et al. 2019, ApJ, 883, 192
Adachi, I., Hayashi, C., & Nakazawa, K. 1976, Progress of Theoretical Physics, 56, 1756
Agnor, C., & Asphaug, E. 2004, ApJL, 613, L157
Agnor, C. B., Canup, R. M., & Levison, H. F. 1999, Icarus, 142, 219
Wang, S., & Ji, J. 2017, AJ, 154, 236
Wang, S., Ji, J., & Zhou, J.-L. 2012, ApJ, 753, 170
Ward, W. R. 1997, Icarus, 126, 261
Weidenschilling, S. J. 1977, MNRAS, 180, 57
Weidenschilling, S. J. 1980, Icarus, 44, 172
Weiss, L. M., Marcy, G. W., Petigura, E. A., et al. 2018, AJ, 155, 48
Wetherill, G. W., & Stewart, G. R. 1989, Icarus, 77, 330
Whipple, F. L. 1972, in From Plasma to Planet, ed. A. Elvius, 211
Wimarsson, J., Liu, B., & Ogihara, M. 2020, MNRAS, 496, 3314
Winn, J. N. 2010, arXiv e-prints, arXiv:1001.2010
Winn, J. N., & Fabrycky, D. C. 2015, ARA&A, 53, 409
Winn, J. N., Sanchis-Ojeda, R., & Rappaport, S. 2018, New Astron. Rev., 83, 37
Wittenmyer, R. A., Wang, S., Horner, J., et al. 2020, MNRAS, 492, 377
Wright, J. T., Marcy, G. W., Howard, A. W., et al. 2012, ApJ, 753, 160
Wright, J. T., Upadhyay, S., Marcy, G. W., et al. 2009, ApJ, 693, 1084
Wu, Y. 2019, ApJ, 874, 91
Wu, Y., & Murray, N. 2003, ApJ, 589, 605
Xie, J.-W. 2014, ApJ, 786, 153
Xie, J.-W., Payne, M. J., Thébault, P., Zhou, J.-L., & Ge, J. 2010a, ApJ, 724, 1153
Xie, J.-W., Zhou, J.-L., & Ge, J. 2010b, ApJ, 708, 1566
Xie, J.-W., Dong, S., Zhu, Z., et al. 2016, Proceedings of the National Academy of Science, 113, 41
Xu, Z., Bai, X.-N., & Murray-Clay, R. A. 2017, ApJ, 847, 52
Yang, C.-C., Johansen, A., & Carrera, D. 2017, A&A, 606, A80
Yang, C.-C., Mac Low, M.-M., & Johansen, A. 2018, ApJ, 868, 27
Yang, H., Li, Z.-Y., Looney, L., & Stephens, I. 2016, MNRAS, 456, 2794
Yang, J.-Y., Xie, J.-W., & Zhou, J.-L. 2020, AJ, 159, 164
Yin, Q., Jacobsen, S. B., Yamashita, K., et al. 2002, Nature, 418, 949
Youdin, A., & Johansen, A. 2007, ApJ, 662, 613
Youdin, A. N., & Goodman, J. 2005, ApJ, 620, 459
Youdin, A. N., & Lithwick, Y. 2007, Icarus, 192, 588
Yu, C., Li, H., Li, S., Lubow, S. H., & Lin, D. N. C. 2010, ApJ, 712, 198
Zanazzi, J. J., & Lai, D. 2018, MNRAS, 478, 835
Zhang, H., & Zhou, J.-L. 2010, ApJ, 714, 532
Zhang, K., Blake, G. A., & Bergin, E. A. 2015, ApJL, 806, L7
Zhang, K., Bosman, A. D., & Bergin, E. A. 2020, ApJL, 891, L16
Zhang, S., Zhu, Z., Huang, J., et al. 2018, ApJL, 869, L47
Zhang, X. 2020, RAA (Research in Astronomy and Astrophysics), 20, 99
Zhang, X., Liu, B., Lin, D. N. C., & Li, H. 2014, ApJ, 797, 20
Zhao, Y. et al. 2020, Nature Astronomy, in press
Zheng, X., Kouwenhoven, M. B. N., & Wang, L. 2015, MNRAS, 453, 2759
Zheng, X., Lin, D. N. C., & Kouwenhoven, M. B. N. 2017, ApJ, 836, 207
Zhou, J.-L., & Lin, D. N. C. 2007, ApJ, 666, 447
Zhou, J.-L., Lin, D. N. C., & Sun, Y.-S. 2007, ApJ, 666, 423
Zhou, J.-L., Xie, J.-W., Liu, H.-G., Zhang, H., & Sun, Y.-S. 2012, Research in Astronomy and Astrophysics, 12, 1081
Zhu, W. 2019, ApJ, 873, 8
Zhu, W., Petrovich, C., Wu, Y., Dong, S., & Xie, J. 2018, ApJ, 860, 101
Zhu, W., Wang, J., & Huang, C. 2016, ApJ, 832, 196
Zhu, W., & Wu, Y. 2018, AJ, 156, 92
Zhu, Z., Dong, R., Stone, J. M., & Rafikov, R. R. 2015, ApJ, 813, 88
Zhu, Z., Nelson, R. P., Dong, R., Espaillat, C., & Hartmann, L. 2012, ApJ, 755, 6
Zhu, Z., Stone, J. M., & Bai, X.-N. 2015, ApJ, 801, 81
Zhu, Z., Zhang, S., Jiang, Y.-F., et al. 2019, ApJL, 877, L18
Zsom, A., Ormel, C. W., Güttler, C., Blum, J., & Dullemond, C. P. 2010, A&A, 513, A57