Formation and Orbital Evolution of Planets

Wilhelm Kley

Institut für Astronomie & Astrophysik
Universität Tübingen, Morgenstelle 10, 72076 Tübingen, Germany
email: `wilhelm.kley@uni-tuebingen.de`

Abstract. The formation of planetary systems is a natural byproduct of the star formation process. Planets can form inside the protoplanetary disk by two alternative processes. Either through a sequence of sticking collisions, the so-called sequential accretion scenario, or via gravitational instability from an over-dense clump inside the protoplanetary disk. The first process is believed to have occurred in the solar system. The most important steps in this process will be outlined. The observed orbital properties of exoplanetary systems are distinctly different from our own Solar System. In particular, their small distance from the star, their high eccentricity and large mass point to the existence of a phase with strong mutual excitations. These are believed to be a result of early evolution of planets due to planet-disk interaction. The importance of this process in shaping the dynamical structure of planetary systems will be presented.

Keywords. planetary systems: formation, accretion disks, hydrodynamics

1. Introduction

Already in ancient times, it was recognized that planets are special objects in the heavens that are distinguished from the stars through their motion across the sky. Only much later, it was appreciated that planets are in orbit around our Sun, and the dynamical structure of the whole planetary system has been analyzed subsequently in much more detail. The most important facts are, that the Solar System planets are orbiting the Sun in nearly circular orbits, in the same direction and basically all in one plane, the ecliptic. Based on this information, Kant and Laplace developed in the 1750s a formation theory for the Solar System including the Sun and the planets. This model implied the collapse/contraction of a previously larger cloud, its subsequent flattening, and the condensation of the planets in the forming disk. Interestingly, these early ideas still form the basics of today's models of planet formation.

In these disks, planets can form along two different pathways. In the sequential accretion scenario, the growth proceeds bottom-up from small to large particles; this will be explained in more detail below. An alternative scenario of planet formation is the direct fragmentation of an unstable disk. This occurs if a local density perturbation is strong enough such that it begins to contract under its own self-gravity (Toomre 1964). Important for the occurrence of this gravitational instability is the mass of the disk and its cooling ability. For typical conditions around young stars, it is expected that only distant planets (e.g. around HR 8799) can form via this way.

2. The formation of the solar system

The protosolar nebula consisted of 99% gas and about 1% dusty material. In today's terms, this would be described as an accretion disk which transfers material inward to be accumulated by the protostar and transports angular momentum outward. The angular momentum transport in these disks is believed to be driven by magneto-rotational

turbulence (Balbus & Hawley 1993). The details and efficiency is calculated typically via 3D magneto-hydrodynamical simulations employing finite volume methods. For an up-to-date simulation involving full radiative transport, see Flaig *et al.* (2010) and Fig. 1.

The standard scenario for planet formation in the solar system in such disks, follows the *sequential accretion scenario*, i.e. the small dust particles in the gaseous disk collide with each other driven by Brownian and turbulent motions in the gas disk. For small relative velocities, the particles stick to each other through Van der Waals forces and form larger, fluffy and fractal aggregates. For a review on this early phase of planet formation see Blum & Wurm (2008). In the subsequent growth phase, the particles decouple from the gas because they move with a different velocity, since they do not feel the gas pressure. This leads to larger relative collision velocities such that mutual collisions might result in fragmentation rather than growth of objects. Additionally, these particles show a relatively fast inward drift towards the star. Together, these two obstacles to planetesimal formation (fragmentation and drift) are often referred to as the meter size barrier (Weidenschilling 1977, Nakagawa *et al.* 1986). Possible solutions to circumvent this problem are accumulation of particles in long-lived vortices (Klahr & Bodenheimer 2006), or turbulent eddies (Johansen *et al.* 2007), or different material properties. Concerning the last option, it was noticed some time ago that, for typical relative velocities, the collision of the rocks (basalt spheres) will lead to full fragmentation and destruction of the two bodies (Benz & Asphaug 1999). But, since it is known that pristine objects in the solar system (asteroids, comets) can have average densities far below the corresponding bulk density of the underlying material, it has been suggested that the initial growth might be facilitated through a high porosity of the material (Dominik *et al.* 2007). Indeed, recent simulations of aggregate collisions show that under suitable conditions this might lead to sticking and growth. A very suitable numerical method is a modified version of *SPH* which includes solid state physics (Libersky *et al.* 1993, Schäfer *et al.* 2007). Using this technique, Geretshauser *et al.* (2010) have first calibrated the code with experimental laboratory data and have then analyzed the outcomes of mutual collisions; see Geretshauser *et al.* (2011).
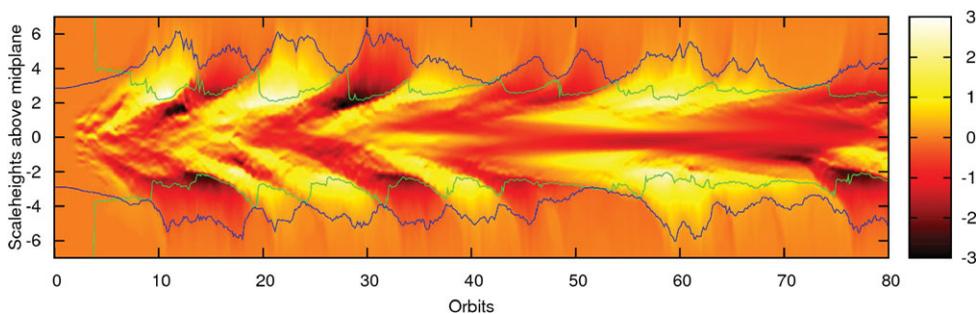
Figure 1. Space-time evolution of the horizontally averaged azimuthal magnetic field (B_y) for local fully radiative 3D magneto-hydrodynamical simulations of accretion disk turbulence. The green line denotes the location of the photosphere $(\tau = 1$ surface) and the blue line the location of the magnetosphere $(p_{\mathrm{gas}} = p_{\mathrm{mag}})$ of the disk. Figure adapted from Flaig *et al.* (2010).

After having overcome these initial obstacles, planetary cores may form via gravity-assisted growth where the geometrical cross section for the colliding partner can be greatly enhanced through gravitational focussing. Following the so called core-instability model, the gas giants can then form through rapid, runaway accumulation of gas onto the planetary cores (Perri & Cameron 1974), forming, for example, Jupiter and Saturn in the Solar System which are believed to have solid cores in their centers. During the

whole growth phase of planets, they remain embedded in the gaseous disk from which they form. This leads to gravitational interactions between the planet and the disk, which will be described in more detail further below. After reaching a critical mass of about $1/2\,M_{\mathrm{Jup}}$, an annular gap forms at the location of the planet, due to angular momentum transfer from the planet to the gas, which limits its further mass growth. This effect has been known for some time in the context of disk-satellite interaction (Goldreich & Tremaine 1980). Altogether, the sequential accretion scenario describes the formation and present physical and dynamical structure of the Solar System (inner rocky, outer gaseous planets, flatness, angular momentum, maximum mass) very well.

3. Extrasolar planetary systems

The orbital elements of the observed extrasolar planets are distinctly different from the Solar System. While major Solar System planets have a nearly coplanar configuration and orbits with small eccentricity, the exoplanet population displays large eccentricities, and many planets orbit their host star on very tight orbits. Recently, it has been discovered that inclined and retrograde orbits are quite frequent as well, at least for close-in planets. Historically, it was exactly the relatively 'calm' dynamical structure of the solar system that led to the hypothesis that planets form in protoplanetary disks. The discovery of the hot planets, which could not have been formed in-situ described by the above scenario due to the hot temperatures and limited mass reservoir, gave rise to the exploration of dynamical processes that are able to change the location of planets in the disk, accompanying the regular formation process.

The density disturbances in the disk induced by the growing protoplanet back-react onto the planet and exert gravitational torques which can change the orbital elements of the planet. It has been suggested that it is possible to bring a planet that has formed at large distances from the star to its close proximity (Ward 1986, Lin *et al.* 1996). This migration process has provided a natural explanation for the population of hot planets, and their existence has been considered as evidence for the migration process.

Another indication for a planetary migration process comes from the high fraction (nearly 20%) of configurations in a low order mean-motion resonance, within the whole sample of multi-planet systems. As the direct formation of such systems seems unlikely, only a dissipative process that changes the energy (semi-major axis) is able to bring planets from their initial non-resonant configuration into resonance. Since resonant capture excites the planetary eccentricities typically to large values in contradiction to the archetypical system GJ 876, it has been inferred that planet-disk interaction should lead to eccentricity damping (Lee & Peale 2002, Crida *et al.* 2008).

4. Planetary migration

An embedded object disturbs the ambient disk dynamically in two important ways: First, it divides it into an inner and outer disk separated by a co-orbital (horseshoe) region. Secondly, the propagating sound waves that are sheared out by the Keplerian differential rotation generate density waves in the form of spiral arms in the disk. The induced density structures in the co-orbital region and the spiral arms back-react on the planet and cause a change in its orbital elements. The main important parameter to be studied is the semi-major axis, i.e. the migration of the embedded planet.

Spiral arms: To put it simply, the spiral arms can be considered as density enhancements in the disk that 'pull' gravitationally on the planet. This gives rise to so-called *Lindblad torques* that change the planet's angular momentum. For circular orbits, the disk torque exerted on the planet is directly a measure of the speed and direction of

migration. The inner spiral forms a leading wave that causes a positive torque, while the outer wave generates a negative contribution. The combined effects of both spirals determine then the sign and magnitude of the total torque. A positive total torque will add angular momentum to the planet and cause outward migration. On the other hand, a negative torque will induce inward migration. It turns out that under typical physical disk conditions the contributions of the inner and outer spiral arm are comparable in magnitude. However, the effect of the outer spiral quite generally wins over the inner one causing the planet to migrate inward.

Corotation region: As viewed in the corotating frame, material within the corotation region performs so-called horseshoe orbits. Here, the gas particles come close to the planet, at the two ends of the horseshoe orbit and perform U-turns. They are periodically shifted from an orbit with a semi-major axis slightly larger than the planetary one to an orbit with slightly smaller value, and vice versa. Hence, at each close approach with the planet, there is an exchange of angular momentum between (co-orbital) disk material and the planet. The total corotation torque is then obtained by adding the contributions from both ends of the horseshoe. To obtain a net, non-zero torque, an asymmetry between the two U-turns has to exist. This requires non-vanishing radial gradients of vortensity and entropy across the corotation region (Baruteau & Masset 2008). For an ideal gas without friction or heat diffusion, mixing effects within the horseshoe tend to flatten out these gradients yielding a vanishing corotation torque, or so-called torque saturation.
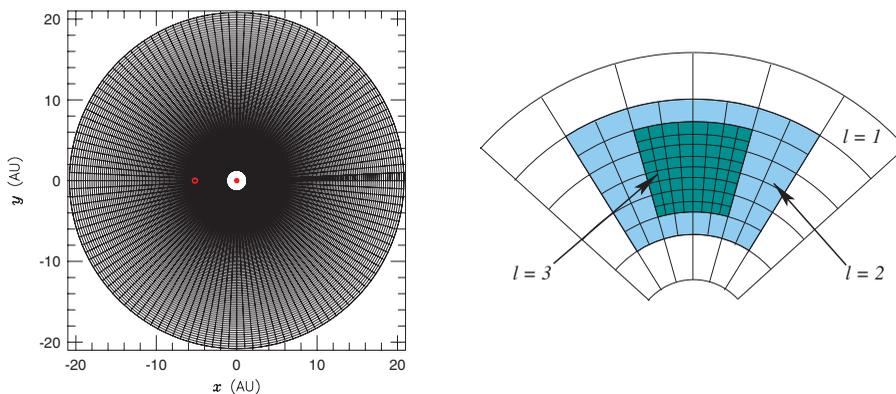
Figure 2. The numerical grid structure employed in numerical planet-disk simulations. **Left**: The global grid structure for a 128×128 grid. Indicated is the Roche-lobe of a Jupiter-type planet, located at $(-5.2, 0.0)$. **Right**: The structure of a nested grid, shown here with two sub-grid levels. The planet is located in the middle of the finest grid (G. D'Angelo).

4.1. *Migration in isothermal disks*

Small mass planets do not alter the global disk structure significantly, in particular they do not open a gap within the disk. Hence, the combined effect of Lindblad and corotation torques can be calculated for small planetary masses using a *linear analysis*. The outcome of such linear, no-gap studies has been termed *Type-I migration*. Due to the complexity of considering heat generation and transport in disks, these linear studies have relied nearly exclusively on simplified, locally isothermal disk models. Here, the temperature is assumed to be independent of height and is given by a pre-described function of radius, $T = T(r)$. Typically, it is assumed that the relative scale height H of the disk is a constant, $H/r = const.$, yielding $T \propto r^{-1}$. The total torques Γ_{tot} is given as the sum of Lindblad and corotation torques $\Gamma_{\mathrm{tot}} = \Gamma_{\mathrm{L}} + \Gamma_{\mathrm{CR}}$. The speed of the induced linear, type-I

migration scales inversely with the disk temperature (i.e. disk thickness) as $\propto (H/r)^{-2}$, linear with the planet mass $\propto m_{\mathrm{p}}$, and with the disk mass $\propto m_{\mathrm{d}}$. Linear models have been calculated for flat 2D disks as well as full 3D configurations. The problem of 2D simulations lies in taking into account approximately the neglected vertical stratification of the disk, which is typically done through a smoothing of the gravitational potential near the planet. Additional problems arise when considering radial gradients.

The linear studies have been supplemented recently by many numerical studies in 2 and 3 dimensions. Here, the disk is typically modelled using as a viscous gas, and the Navier-Stokes equations are solved using grid-based methods. To verify the correctness of the results, a detailed *code comparison* project was conducted a few years ago; see De Val Borro *et al.* (2006). A very important ingredient to speed up such simulations is the `FARGO`-method introduced by Masset (2000). To increase resolution around the planet, a nested-grid structure has been employed successfully in 2D and 3D; see D'Angelo *et al.* (2002, 2003) and Fig. 2. New full 3D, nested grid locally isothermal hydrodynamic simulations of planet-disk interaction give very good agreement with previous 3D linear results (Tanaka & Ward 2004) and yield the following form for the total torque for small mass planets below about 10 M_{Earth} (D'Angelo & Lubow 2010)

$$\Gamma_{\mathrm{tot}} = -(1.36 + 0.62\alpha_{\Sigma} + 0.43\alpha_T) \left(\frac{m_{\mathrm{p}}}{M_*}\right)^2 \left(\frac{H}{r_{\mathrm{p}}}\right)^{-2} \Sigma_{\mathrm{p}}\, r_p^4\, \Omega_{\mathrm{p}}^2. \tag{4.1}$$

In eq. (4.1), the index p refers to the planet, α_{Σ} and α_T refer to the radial variation of density and temperature, such that $\Sigma(r) \propto r^{-\alpha_{\Sigma}}$ and $T(r) \propto r^{-\alpha_T}$. One should keep in mind that these results depend on the magnitude of the viscosity.

For larger planet masses, a gap is opened in the disk because the planet transfers angular momentum to the disk, positive exterior and negative interior to the planet. The depth of the gap that the planet carves out depends for given disk physics (temperature and viscosity) only on the mass of the planet. Because the density in the co-orbital region is reduced, the corotation torques are strongly affected and are no longer of any importance for larger planet masses. For very large masses, even the Lindblad torques are reduced yielding a slowing down of the planet. This non-linear regime has been coined the *Type-II* regime of planetary migration; here the drift of the planet is dominated by the disk's viscous evolution.

4.2. *Migration in radiative disks*

The migration rates obtained through the analyses mentioned above have resulted in approximate formulae for the migration speed $\dot{a}$ of a planet (e.g. from eq. 4.1) that are frequently used in *population synthesis models*. These are growth models of planets that include disk evolution and planet migration in parameterized form, which can be used to calculate the final location of many planets in the mass/semi-major axis diagram and compare the results statistically with the observed distribution. The results indicate in particular, that the migration for small mass planets in the type-I regime is by far too fast to account for the observed distribution; the majority of planets would have been lost to the star, see Ida & Lin (2008) and Mordasini *et al.* (2009). Only a significant reduction in the type-I migration speed gives satisfactory results.

Here, we shall concentrate on a *physical* improvement to the models, that represent a possible solution to the type-I migration problem: the inclusion radiative effects. Extending the previous isothermal models requires the incorporation of a heating and cooling mechanism. The importance of radiative diffusion has first been pointed out by Paardekooper & Mellema (2006), and recent papers quote approximate formulas for viscous and diffusive disks (Masset & Casoli 2010, Paardekooper *et al.* 2011). To

demonstrate the effect for two-dimensional flat disks, we show results for a planet-disk simulation where we include viscous heating, local radiative cooling, as well as diffusive radiative transport in the disk's plane. The energy equation then reads

$$\frac{\partial \Sigma c_v T}{\partial t} + \nabla \cdot (\Sigma c_v T \mathbf{u}) = -p \nabla \cdot \mathbf{u} + D - Q - 2H \nabla \cdot \vec{F} \qquad (4.2)$$

where Σ is the surface density, T the midplane temperature, p the pressure, D the viscous dissipation, Q the radiative cooling, and $\vec{F}$ the radiative flux in the midplane. Models where the various contributions on the right hand side of eq. (4.2) were selectively switched off and on, have been constructed by Kley & Crida (2008).
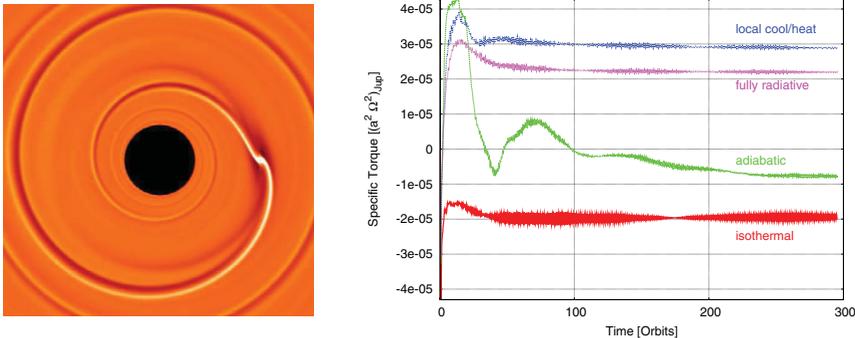
Figure 3. The evolution of a 20 M_{Earth} planet embedded in a disk with 0.01 solar masses. **Left**: Spiral arm structure (Courtesy, F. Masset). **Right**: Time evolution of the torque acting on the planet. From bottom to top the curves indicate simulations where *i)* no energy equation (isothermal), *ii)* only the first term on the rhs. of eq. (4.2) (adiabatic), *iii)* all terms on the rhs. (fully radiative), and *iv)* all but the last term on the rhs. (heating/cooling), have been used (after Kley & Crida, 2008).

The effect of this procedure on the resulting torque is shown in the right panel of Fig. 3. The basis for all the models is the same equilibrium disk model constructed using all terms on the rhs. of eq. (4.2) and no planet. Embedding a planet of $20 M_{\mathrm{Earth}}$ yields in the long run *positive torques* only for the radiative disks, where the maximum effect is given when only viscous heating and local radiative cooling are considered. The left hand panel of Fig. 3 shows the global spiral arms structure of the embedded planet. The strength of this positive corotation effect also scales with the square of the planet mass up to about 20 to 25 M_{Earth}. Beyond this mass, gap opening begins and only the Lindblad torques remain, and the planets begin to migrate inwards again. In full 3D radiative simulations, the results are qualitatively the same (Kley *et al.* 2009), but interestingly the full 3D results show an even stronger effect. New population synthesis models based on the modified migration rates seem to show better agreement with the observational data set (Mordasini *et al.* 2011).

5. Eccentricity and inclination

In addition to a change in semi-major axis, planet-disk interaction will modify the planetary eccentricity (e) and inclination (i) as well. For small e and i, linear analysis has indicated exponential damping (Ward 1988, Ward & Hahn 1994). Extending previous isothermal numerical studies by Cresswell *et al.* (2007), fully 3D radiative disk simulations have been performed recently to study the evolution e and i. The results indicate that both are damped for all planetary masses (Bitsch & Kley 2010, 2011). For

small values of e and i that are below about $2H/r$, the damping occurs exponentially on timescales comparable to the linear estimates by Tanaka & Ward 2004. For larger values, the damping is slowed down and follows approximately $\dot{e} \propto e^{-2}$, and for the damping of i, an identical relation holds. Interestingly, the presence of outward migration is coupled to the magnitude of e and i. Outward migration only occurs for eccentricities smaller than about 0.02, and inclinations below about 4°. This reduction is due to the fact that for non-circular orbits the flow structure in the corotation region becomes strongly time dependent and no stationary corotation torque can develop.

6. Summary

The formation of planets takes place in protoplanetary disks via sequential accretion or the gravitational instability. In the first process, the initial growth is not yet fully understood. Close in planets have migrated inwards from their birthplace either through disk migration or tidal shrinking. The eccentric and inclined population of planets has been created by dynamical scattering processes.

References

Balbus, S. A. & Hawley, J. F. 1993, *ApJ*, 367, 214
Baruteau, C. & Masset, F. 2008, *ApJ*, 672, 1054
Benz, W. & Asphaug, E. 1999, *Icarus*, 142, 5
Bitsch, B. & Kley, W. 2010, *A&A*, 523, A30 ; 2011, *A&A*, 530, A41
Blum, J. & Wurm, G. 2008, *ARA&A*, 46, 21
Cresswell, P., Dirksen, G., Kley, W., & Nelson, R. P. 2007, *A&A*, 473, 329
Crida, A., Sándor, Z., & Kley, W. 2008, *A&A*, 483, 325
D'Angelo, G. & Lubow, S. H. 2010, *ApJ*, 724, 730
D'Angelo, G., Henning, Th., & Kley, W. 2002, *A&A*, 385, 647 ; 2003, *ApJ*, 599, 548
De Val Borro *et al.*, 2006, *MNRAS*, 370, 529
Dominik, C., Blum, J., Cuzzi, J. N., & Wurm, G. 2007, *Protostars and Planets V*, Eds. B. Reipurth, D. Jewitt, and K. Keil, p783-800
Flaig, M., Kley, W., & Kissmann, R. 2010, *MNRAS*, 409, 1297
Geretshauser, R. J., Meru, F., Speith, R., & Kley, W. 2011, *A&A*, 531, 166
Geretshauser, R. J., Speith, R. Güttler, C., Krause, M., & Blum, J. 2010, *A&A*, 513, 58
Goldreich, P. & Tremaine, S. 1980, *ApJ*, 241, 425
Ida, S. & Lin, D. N. C. 2008, *ApJ*, 673, 487
Johansen, A., Oishi, J. S., Mac Low, M.-M., Klahr, H., Henning, T., & Youdin, A. 2007, *Nature*, 448, 1022
Klahr, H. & Bodenheimer, P. 2006, *ApJ*, 639, 432
Kley, W., Bitsch, B., & Klahr, H. 2009, *A&A*, 506, 971
Kley, W. & Crida, A. 2008, *A&A*, 487, L9
Lee, M. H. & Peale, S. J. 2002, *ApJ*, 567, 596
Libersky, L. D., Petschek, A. G., Carney, T. C., Hipp, J. R., & Allahdadi, F. A. 1993, *Journal Computational Physics*, 109, 67
Lin, D. N. C., Bodenheimer, P., & Richardson, D. C. 2006, *Nature*, 380, 606
Masset, F. S. 2000, *A&AS*, 141, 165
Masset, F. S. & Casoli, J. 2010, *ApJ*, 723, 1393
Mordasini, C., Alibert, Y., Benz, W., & Naef, D. 2009, *A&A*, 501, 1161
Mordasini, C., Dittkrist, K., Alibert, Y., Klahr, H., Benz, W., & Henning, T. 2011, *ArXiv*, 1101.3238

Nakagawa, Y., Sekiya, M., & Hayashi, C. 1986, *Icarus*, 67, 375
Paardekooper, S.-J. & Mellema, G. 2006, *A&A*, 459, L17
Paardekooper, S.-J., Baruteau, C., & Kley, W. 2011, *MNRAS*, 410, 293
Perri, F. & Cameron, A. G. W.. 1974, *Icarus*, 22, 416
Schäfer, C., Speith, R., & Kley, W. 2002, *A&A*, 470, 733
Tanaka, H. & Ward, W. R. 2004, *ApJ*, 602, 388
Toomre, A. 1964, *ApJ*, 139, 1217
Ward, W. R. 1986, *Icarus*, 67, 164
Ward, W. R. 1988, *Icarus*, 73, 330
Ward, W. R. & Hahn, J. M. 1994, *Icarus*, 110, 95
Weidenschilling, S. J. 1977, *MNRAS*, 180, 57

Discussion

J. GRYGAR: Are there estimates how much time elapses to migrating Jupiters from their birthplaces to close orbits of hot Jupiters?

W. KLEY: Yes, the migration formulae and the numerical simulations yield timescales of typically 10^4 orbital periods.

M. MONTGOMERY: Have you considered adding asteroids in the gap which will eventually meet at L4 and L5 and thus may help in taking away energy for planet migration?

W. KLEY: Asteroids have not been taken into account in migration simulations, as their mass is too small to prevent or slow down migration significantly. Note, that the whole coorbital horseshoe region with all its mass is dragged along with the inward motion of the planet.

J. SOUTHWORTH: I think it is premature to suggest that many planetary systems are coplanar. Kepler-II tells us that they can be coplanar, but planetary systems which are not coplanar will not be picked up by Kepler as the planets in questions will not transit. On the plus side, as Kepler data continue to accrue we will be able to detect such planets by their dynamical effects on those planets in the system which do transit.

W. KLEY: Yes, it is difficult to decide presently what fraction of planetary systems will be nearly coplanar. But the existing data indicate that at least some systems will have undergone migration.

S. HINKLEY: Is it a coincidence that HR8799, Fomalhaut and Pictoris (the systems with directly imaged planets and presumably formed through gravitational instability) are all found orbiting A-stars which presumably possess more massive disks?

W. KLEY: Yes, the mass of a disk is important as the instability scales with the disk mass. But the cooling and heating (irradiation) of the disk have to be taken into account.

K. BJORKMAN: The requirement for forming a planet via gravitational instability is that the surface density be large enough, so if it is possible to have an unusual surface density distribution (say by infall into the outer disk) then a large total disk mass may not be required.

W. KLEY: That's correct, but typically it is assumed that the disk structure (and infall) vary relatively smoothly with radius, and then the total mass is again a suitable criterion.