ON PLANET FORMATION AND MIGRATION[1]

William R. Ward
Jet Propulsion Laboratory
California Institute of Technology
Pasadena, CA 91109

Abstract. Some consequences of protoplanet migration for the planetary formation process are discussed, Migration of planet-sized objects can be caused by disk tidal torques. Two types of a migration are identified: a fast orbit decay relative to the disk (type I) at a rate proportional to the planet's mass, and a slow co-evolution with the disk (type 11) at a rate set by the disk's viscosity. Smaller (larger) protoplanets execute the former (latter), and a transition I - II occurs at a critical mass (Shiva limit). The fast mode is a precarious stage in the planet's development, in which orbital decay threatens to drive the embryo into the star, Comparison of growth and orbital decay times defines an upper mass limit for survival. Although it is possible that both the terrestrial and outer most planets of our solar system remained below this limit throughout the disk's lifetime, this possibility is not open to Jupiter and Saturn, Their existence is somewhat of a conundrum, and several possible mechanisms that could have cent ributed to their survival are considered.

October, 1997

[1]Based, in part, on an invited talk delivered at the First International origins Conference, held at Estes Park, Colorado. May 19-23, 1997.

1. introduction

The consequences of protoplanet migration due to disk tides is one on the most important, and yet most understudied, aspects of planetary formation. A protoplanet can gravitationally interact with its precursor disk at resonance sites. At Lindblad resonances, where the Doppler shifted forcing frequency, $m|\Omega - \Omega_{ps}|$, matches the natural oscillation frequency of the disk, $\kappa\sqrt{1 + mc/r\kappa}$, the disturbance in the surface density takes the form of a spiral wave that, in a pressure dominated disk, propagates away from the resonance zone as illustrated in Fig, 1 (e.g., Goldreich and Tremaine, 1979; Ward, 1986). The order m denotes the number of spiral arms in the wave, $\Omega(r)$ and $\kappa(r)$ are the disk's angular and epicycle frequencies, c is the gas sound speed, Ω_{ps} is the pattern speed of the potential term involved in the resonance, which, for a circular orbit, is simply the secondary's mean motion, Q,, The higher the order, the closer the resonance lies to the planet. Closer resonances are stronger up to m $\gtrsim r/h$, with h - c/Ω being the scale height of the gas disk, For higher order, the resonances "pile" up at a distance ~$2h/3$ and weaken precipitously (Artymowicz, 1993), Fig.2 illustrates resonance placement as a function of order. For each m order term, there is a Lindblad site inside and outside the perturber's orbits. The protoplanet's gravitational attraction for the non-axisymmetric mass distributions produced by the wave motions results in a reaction torque, For large protoplanets, this torque can be quite strong and affect substantial modifications of both the disk and the protoplanet's orbit, .

For years, I have promulgated this mechanism as an important feature in the formation and dynamical evolution of a planetary system (e.g., Ward, 1982, 1984, 1986a,b, 1988, 1989, 1993a,b and references therein), but the idea that planetary sized objects could migrate large distances seemed to be considered almost heretical, The discovery of several close stellar companions, in environments where the usual paradigm for planet formation would not predict planets to form, may constitute a proverbial "smoking gun" that large scale migrations of planets have occurred. This has lent new urgency to understanding the role of disk tides in the formation and ultimate survival of planetary systems (e.g., Lin, *et al.*, 1996; Ward, 1997a,b; Trilling *et al.* 1997).

II. Protoplanetary Migration

An aspect of this phenomenon I have recently stressed (Ward, 1997a,b) is that planet/disk interactions admit *two* types of migration, distinguished by their functional dependence on planet and disk characteristics, and by their times scales. Both types have appeared in the literature, but are not always distinquished from one another, This is done below.

Type I: If the protoplanet is not massive enough to open a gap in the disk, orbital drift is caused by systematically larger torques from outer Lindbald resonances (e.g., Goldreich and Tremaine, 1980; Ward, 1986, 1997a; Korycanski and Pollack, 1993). Since the torques on the planet from outer Lindblad resonances are *negative*, the object suffers orbital decay on a characteristic time scale

$$\tau_I \approx \frac{\varsigma_1}{\Omega}\left(\frac{M_\star}{M}\right)\left(\frac{M_\star}{\sigma_g r^2}\right)\left(\frac{c}{r\Omega}\right)^3 \tag{1}$$

where σ_g is the gas surface density, and $M, M_\star$ are the masses of the protoplanet and primary, respectively, The fractional torque asymmetry can be quite large, i.e. several tens of percent, resulting in lead coefficient, ς_1, of order unity (Ward, 1997a).

Type II: If the protoplanet opens a gap in the disk, it becomes linked into the overall angular momentum transport, It then co-evolves with the disk, with a drift time scale set by the disk's viscosity (e.g., Lin and Papaloizou, 1993),

$$\tau_{II} \approx \varsigma_2 \alpha \left(\frac{c}{r\Omega}\right)^2 r\Omega \tag{2}$$

assuming a Sakura-Sunyaev type viscosity law, v - αc 2/Q and a lead coefficient, ς_2, of order unity. Note that this rate is independent of the protoplanet or disk masses, as long as the conditions for maintaining a gap are met,

These two migration rates are comparable for a mass,

$$\frac{M_o}{M_\star} \sim \alpha(\frac{M_\star}{\sigma r^2})(\frac{c}{r\Omega})^5 \qquad (3)$$

which for α - 10 4 -103 in a minimum mass model of the solar nebula, yields a few tenths of an earth mass. We expect type I motion to be exhibited by small protoplanets, type 11 exhibited by large protoplanets. A key question is; at what mass does a transition take place? Is it near M_o where the rates are comparable, or does the transition threshold "over shoot" this value, so that there is a range of masses with orbital lifetimes *less* than the viscous evolution time scale of the disk?

Figure 3 shows a model calculation for protoplanet behavior at 5 AU in a minimum mass solar nebula (Ward, 1997a). The transition from I → II occurs at a mass that, for the range of conditions shown, can be approximated by

$$\frac{M_*}{M_\star} \approx \alpha^{2/3}(\frac{M_\star}{\sigma r^2})^{1/3}(\frac{c}{r\Omega})^3 \qquad (4)$$

and lies in the regime where type I decay is up to two orders of magnitude faster than type II. Thus, disk tides render the mass range $(M_o < M < M_*)$ an especially precarious stage in the growth of a planet. I have designated the upper limit M_* as the *Shivs* limit, after the Hindu god of destruction (Ward, 1997b). Figure 3 displays the characteristic decay times as a function of mass measured in earth masses, The curves are labeled by the strength of the turbulent viscosity. For the range of α shown, planetary embryos between --0,1 and 10 earth masses are in danger of decaying out of the disk,

III. Survival

How does a planetary system survive this process? The characteristic growth time of an embryo that has runaway in size from neighboring planetesimals is

$$\tau_{growth} \approx \Omega^{-1}\left(\frac{\rho_p R}{\sigma_d}\right) F_g^{-1} \tag{5}$$

where R is the embryo's radius, ρ_p is its density, σ_d is the surface density of accretable material, and F_g is the so-called enhancement factor due to gravitational focussing [e.g., see Lissauer and Stewart (1993) and Ward (1996) for recent readable reviews of solid body accretion]. Equating this to the type I timescale gives us the protoplanet size that will decay out of the disk before significantly more growth can occur (Ward, 1997b),

$$\frac{M_{crit}}{M_\star} \approx \left(\frac{f_d F_g}{|c_1|}\right)^{3/4}\left(\frac{M_\star}{\rho_p r^3}\right)^{1/2}\left(\frac{c}{r\Omega}\right)^{9/4} \tag{6}$$

where $f_d \equiv \sigma_d/\sigma_g$ is the ratio of solids to gas. The enhancement factor from stirring (Lissauer, 1985; Ida and Makino, 1993) is roughly,

$$F_g \approx V_{esc}/R_H\Omega \equiv F_{stir} \approx 10^3(r/AU) \tag{7}$$

where V_{esc} denotes the embryo's escape velocity, and R_H is its Hill radius. With this, eqn (6) can be recast as

$$\frac{M_{crit}}{M_\oplus} \approx \left(\frac{r}{5AU}\right)^{3/8}\left(\frac{T}{150K}\right)^{9/8}\left[\varsigma_1\left(\frac{f_d}{.01}\right)\left(\frac{F_g}{F_{stir}}\right)\right]^{3/4} \tag{8}$$

Table I shows several masses and time scales for three zones in a minimum mass model of the solar nebula, including the so-called runaway accretion limit (e.g., Wetherill, 1990; Lissauer and Stewart, 1993)

$$\frac{M_{run}}{M_\star} \equiv 3^{1/4}\left(\frac{8\pi\sigma r^2}{M_\star}\right)^{3/2} \tag{9}$$

and characteristic growth time, $\tau_{run} \equiv \tau_{growth}(M_{run})$ The critical mass is nearly the same in each zone. Notice, however, that M_{run} is smaller than M_{crit} in the terrestrial zone When this happens, the table gives the decay time for the runaway mass instead, which for the terrestrial zone, is comparable to the disk life, This suggests that the final assembly of these planets *could* have post-dated the nebula, For the outer most planets, it is possible that accretion was slow enough that the critical mass may have not been achieved during the disk's lifetime. However, neither of these alternatives are available for the giant planets, which seem to be a *conundrum*. They must form in the presence of the gas, and yet, *the predicted decay time of their cores is much less than the generally assumed lifetime of the disk!*

IV. Discussion

It seems likely that the key to survival is the *onset* of the gas accretion phase. The Bondi rate for gas accretion predicts a growth time $M/\dot{M}$ that is less than τ_I by a factor $(c/Al)^z << 1$ (e.g., Ward, 1989), The rapid growth may enable the protoplanet to initiate gap formation by directly cannibalizing the disk, and/or exceed the Shivs mass for transition to the slower type II migration. For a $15M_{\oplus}$ core at 5 AU, a combination of parameters is needed such that: $\varsigma_1(f_d/.01)(F_g/F_{stir}) \approx 37$. This seems difficult to achieve, but I shall examine some possibilities in turn:

1. The torque coefficient, ς_1 - 0(1), maybe underestimated in current models that treat the disk as 2D. Thus, it is very important that continued improvement of the theory of disk-planet interaction and protoplanet migration be made,

2. The disk viscosity could be much lower than usually assumed (i.e., α « 10'4), so that the transition from type 1 to type 11 migration occurred at $<10M_{\oplus}$. Vertical convection in the disk is usually cited as the probable cause of an effective turbulent viscosity. However, doubt has been cast on this assumption by recent work by Stone, *et al. (1997)* that finds that vertical convection in the disk does not couple well to the horizontal motion. Thus, energy cannot be drawn from orbital motion to sustain the turbulence. In a nearly inviscid disk, gap formation must be prevented by the motion of the protoplanet, and the Shiva mass reverts to the inertial mass M_I - $\pi\sigma h^2(h/r)$ - $few \times M_{\oplus}$, first described by Hourigan and Ward (1984; see also Ward and Hourigan, 1989). This is at odds with Lin and Paploizou's proposal that disk truncation is what

determines the final mass of the giant planets, A possible resolution could be found in recent work by Artymowicz and Lubow (1996) that suggests it may be possible for an object occupying a gap to continue to accrete gas from streams emanating from the gap edges.

3. Solid material being accreted may be stirred in a non-isotropic manner, since perturbations from the protoplanet are primarily horizontal. The enhancement factor would be increased to F_g - $(e/I)F_{stir}$ (e.g., Ward 1989, 1996; Lissauer and Stewart, 1993). Also, there is a potential feedback loop in density wave assisted accretion because larger objects migrate faster and can overtake other material (Ward, 1989). Ward and Hahn (1995) and Araki and Ward, (1996) showed that disk tidal torques can prevent the isolation of a runaway embryo in the giant planet zone, which through its mobility can breach the runaway limit and penetrate new, undepleted regions of the disk. Disk tides will decay eccentricities and inclinations (Ward, 1993b), but encounters can occur due to the differential migration rates among a collection of various sized masses. Ward (1989, 1993b) argued that, if the disk is thinner than the effective enhanced radius, S - $R\sqrt{F_g}$, the problem is essentially 2D, and the characteristic growth time is

$$\tau_{swp} \sim \tau_I\left(\frac{M}{2\pi\sigma_d r^2}\right) \quad (1o)$$

which is mass independent and very short, i.e., - 10^5 years. The key question is whether the disk can remain flat enough during the growth of a dominant protoplanet for something like this rate to apply,

4. Current models may overestimate critical core size for gas accretion The critical size might be lowered if the extended atmosphere developing about the protoplanet has a low grain opacity and/or the mean molecular weight of the atmosphere is raised by a high water content (Stevenson, 1984). Recent modeling by Pollack, et *al.* (1997), however, still predicts that rapid gas accretion for Jupiter does not occur until the planet has achieved several earth masses in size

5. The solid-to-gas ratio, f_d, may be enhanced $O(1O)$ at the snowline due to diffusive redistribution of water vapor (Stevenson and Lunine, 1988), Although this would only work for a planet near that boundary (i. e., Jupiter), the formation of one giant planet greatly improves the

chance of forming a second (e. g., Saturn) because the first object can inhibit the tidal decay of embryos further out in the disk by capturing them at various Lindblad resonances (Beauge, *et* al. 1994; Hahn and Ward, 1996),

Finally, we must caution that even if a protoplanet succeeds in converting to type II behavior, its survival is not assured, The viscous evolution of the disk can drive these objects into the primary as well, albeit on a longer time scale. Lin and co-workers have suggested that loss of such planets may be a common occurrence in early formation of a planetary system, perhaps including our own, The detected close companions are large enough that gap formation seems probable, and that any migration they have suffered has been that oft ype 11. Consequently, it also seems likely that many objects have met their demise this way, Trilling *et al. (1997)* have described the terminal period of such a giant planet's life that includes a stage of Roche lobe overflow, during which the planet is stripped of its gas prior to plunging into the primary.

On the other hand, ifit turns out that the α is very low (see point #2 above) gas giants may *not* migrate that much over the life of the disk, Is this necessarily cent rary to the existence of close companions? Perhaps not, A low density zone maintained by the star via magnetic coupling to the disk (Lin, *et* al., 1996) would stabilize solid embryos undergoing type I decay as well as gas giants undergoing type 11 decay. Solid embryos trapped in this cavity could subsequently accrete into a single planet enriched in CHON compared to a typical gas giant (Ward, 1997b). Consequently, we cannot rule out "in situ" accretion of close companions, although in the case of tau Bootis b (- 3,87 Jupiter masses), the amount of solid material seem would require a very large disk.

V. Conclusions

It is conceivable that both the terrestrial planets and outer most planets could have remained small enough during the nebula lifetime to avoid severe orbital decay from disk tidal torques, but Jupiter and Saturn may owe their survival to the onset of gas accretion. The conditions for this seem rather marginal, however, suggesting that not all forming planetary systems may produce such objects. Protoplanets and solid cores may sometimes be lost to the primary, especially due to rapid type I orbital decay relative to the disk

The good news is: once a giant planet (e.g., Jupiter) forms and converts to type II behavior, the chance of forming a second gas giant (e.g , Saturn) is greatly improved, since the first object can capture other decaying embryos at outer Lindblad resonances (Beauge *et al.*, 1994; Hahn and Ward, 1996), The bad news is: event ype II motion (co-evolution with the disk) can be destructive if the nebula is not dissipated in a timely manner, Close stellar companions may be evidence of significant orbital migration

A detailed model of concurrent accretion of solids and gas by proto-Jupiter (e.g., Pollack *et al.*, 1996) together with radial migration from disk tides (e. g., Ward, 1997) is a high priority problem to understand the survival of planetary systems embedded in their precursor disks.

Acknowledgments

The research described in this paper was carried out at the Jet Propulsion Laboratory, California Institute of Technology under contract with the Nationa! Aeronautics and Space Administration.

References

Araki, S., and W. R. Ward, 1996. *Lunar & Planet. Sci.* **XXVII**.

Beauge, C. S., S. J. Aarseth, and S. Ferraz-Mello, 1994. *Mon. Not. R Astron. Soc. 270,21.*

Goldreich, P., and S. Tremaine, 1980, *Astrophys. J. 241, 425-441*. . .

Hahn, J. M., and W. R. Ward, 1996, *Lunar & Planet. Sci.* **XXVII.**

Ida, S., and D. N. C. Lin, 1996, *Astrophys. J. 112*, 1239.

Korycansky, D. G., and J. B. Pollack 1993. *Icarus* **102, 150-165.**

Lissauer, J. J., and G. R. Stewart, 1993. In *Protostars and Planets III* (eds. E. H. Levy and J. I. Lunine: University of Arizona Press) pp. 1061-1088.

Shu, F. H., 1984. In *Planetary Rings*. (R Greenberg and A. Brahic, Eds.) pp. 513-561, University of Arizona Press, Tucson.

Stern, S. A, and J. E. Colwell, **1997.** *Astrophys. J. Lett.*

Ward, W R, 1982. *lunar& Planet Sci.* **XIII**, 831

Ward, W. R., 1984, In *Planetary Rings* (Eds. Brahic and Greenberg, Univ. Arizona Press), pp.

660-684

Ward, W. R. 1986a. *Icarus* **67**, 164-180.

Ward, W. R., 1986b. In *Proc. Solid Bodies Outer Solar System Conf.* (Vulcano, Italy, Sept. 1985), pp. 1-4.

Ward, W. R. 1988. *Icarus 73,330-348.*

Ward, W. R., 1989. *Astrophys. J. Lett.* 34S, 1.99.

Ward, W. R. 1993a. In *Proc. 7th Florida Workshop on Non-linear Dynamics*, pp. 314-323

Ward, W. R. 1993b. *Icarus* **106,274.**

Ward, W. R 1997a. *Icarus* **126,** *261.*

Ward, W. R., 1997b. *Astrophys. J. Lett.* In press

Ward, W. R., and J. M. Hahn, 1995. *Astrophys. J. Lett. 440, L25.*

TABLE I

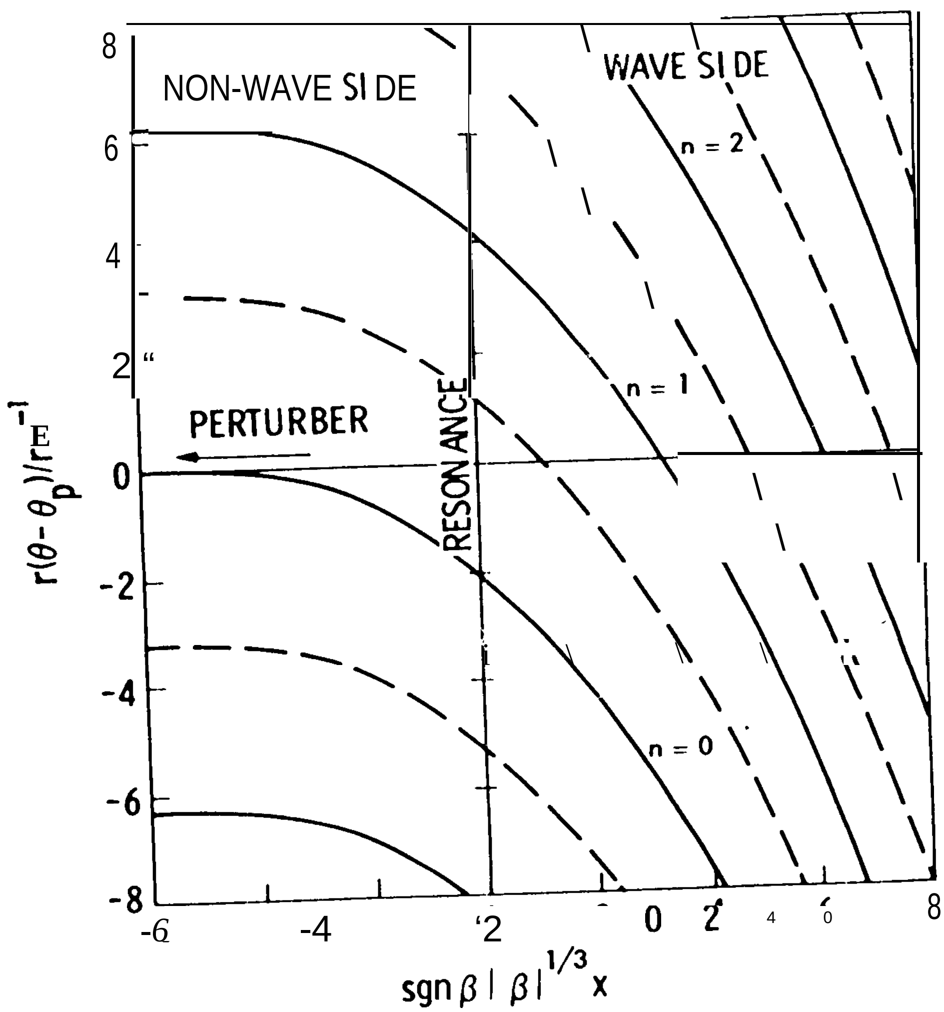
Distance	1 AU	5 AU	20 AU
$\sigma_d\,(g/cm^2)$	7	4	0,3
$\alpha_d \equiv \sigma_d/\sigma$	0.003	0,01	0.01
T (K)	700	150	75
M_{crit}	$3\,M_\oplus$	$M_\oplus$	$0.8\,M_\oplus$
M_{run}	$0{,}04\,M_\oplus$	$2\,M_\oplus$	$3\,M_\oplus$
τ_{run} (years)	$few \times 104$	$few \times 10^5$	$few \times 10'$
τ_I **(years)**	$few \times 10^6$	$few \times 10^5$	$few \times 10^6$

Figure Captions

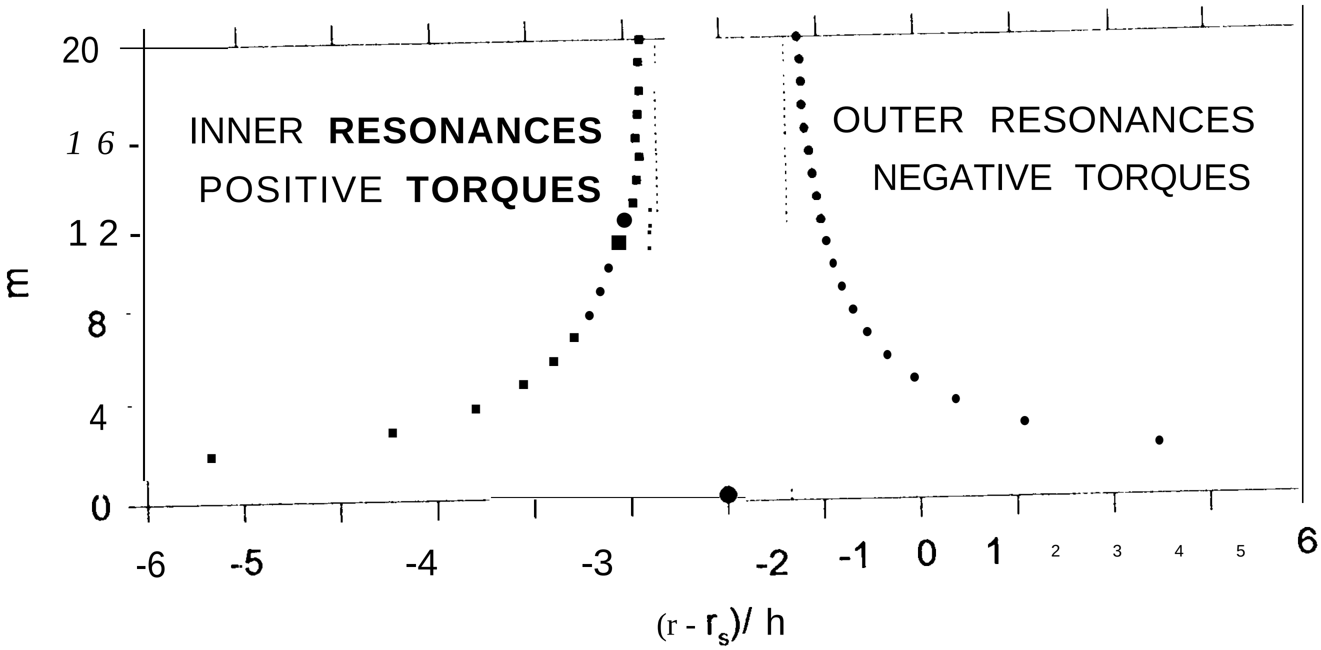
Fig. 1, Spiral wave format an m-order Lindblad resonance in a pressure dominated disk The frame of reference rotates with the perturber, which lies on the -x axis. Solid (dashed) curves are surface density maxima (minima), Waves propagate to **the** right, away from the protoplanet and **wind** up producing a spiral pattern with decreasing wavelength. The vertical axis is normalized to r/m; the horizontal axis is normalized to the collective scale, $|\beta|^{-1/3}$ where $\beta \equiv -3mr^2\Omega\Omega_p/c^2$. (Figure taken from Ward, 1986.)

Fig. 2. Lindblad resonance sites relative to the protoplanet's position with distance measured in disk scale heights. The vertical axis is the resonance order, m, which gives the number of arms in the spiral wave, Higher order resonances lie closer to the planet, but pressure effects exclude resonances from the so-called torque cut-off zone, $\pm 2h/3$, on each side oft he orbit.

Fig. 3. Characteristic decay times, $\tau \equiv r/\dot{r}$, as a function of protoplanet mass, measured in earth masses. Curves are labeled by the viscosity parameter, $\alpha \equiv \nu/c^2\Omega^{-1}$, and constructed for a minimum mass solar nebula at a distance of 5 AU. The time scale decreases inversely with mass (type I) until it reaches a threshold size M_* (Shiva limit), past which the time scale increases (type II) as a gap progressively opens. When a gap is fully established, the planet co-evolves with the disk with a time scale inversely proportional to α.



NON-WAVE SIDE
WAVE SIDE
n = 2
n = 1
n = 0
PERTURBER
RESONANCE
$r(\theta - \theta_p)/r_E^{-1}$
$\mathrm{sgn}\,\beta\,|\beta|^{1/3}x$
8
6
4
2
0
-2
-4
-6
-8
-6
-4
-2
0
2
4
6
8


INNER RESONANCES
POSITIVE TORQUES
OUTER RESONANCES
NEGATIVE TORQUES
20
16
12
8
4
0
m
-6
-5
-4
-3
-2
-1
0
1
2
3
4
5
6
(r - r_s)/ h