NONLINEAR PROPAGATION OF PLANET-GENERATED TIDAL WAVES

R. R. Rafikov
Princeton University Observatory, Peyton Hall, Ivy Lane, Princeton, NJ 08544; rrr@astro.princeton.edu

Received 2001 October 21; accepted 2002 January 1

ABSTRACT

The propagation and evolution of planet-generated density waves in protoplanetary disks is considered. The evolution of waves, leading to shock formation and wake dissipation, is followed in the weakly nonlinear regime. The 2001 local approach of Goodman and Rafikov is extended to include the effects of surface density and temperature variations in the disk as well as the disk cylindrical geometry and nonuniform shear. Wave damping due to shocks is demonstrated to be a nonlocal process spanning a significant fraction of the disk. Torques induced by the planet could be significant drivers of disk evolution on timescales of $10^6$–$10^7$ yr, even in the absence of strong background viscosity. A global prescription for angular momentum deposition is developed that could be incorporated into the study of gap formation in a gaseous disk around the planet.

Subject headings: planetary systems — planets and satellites: general — solar system: formation

1. INTRODUCTION

Tidal disk-companion interactions are important in a variety of astrophysical contexts ranging from the formation and evolution of protoplanetary systems to the origin of galactic spiral structure. Gravitational interaction between the disk and companion generates density waves in the disk (gaseous, stellar, or particulate) that carry angular momentum. This angular momentum is eventually deposited elsewhere in the disk, leading to the evolution of the disk as a whole.

In the case of protoplanetary systems, disk-planet interactions not only cause the migration of the planet (Goldreich & Tremaine 1980, hereafter GT80; Ward 1997) but can also drive noticeable evolution of the disk itself (Larson 1989; Goodman & Rafikov 2001, hereafter GR01). However, to change the state of the disk and orbit of the planet in these systems, it is necessary to somehow transfer angular momentum from the density waves launched by the perturber to the disk material, and this can only be accomplished by virtue of some damping process (Goldreich & Nicholson 1989).

Various mechanisms have been envisaged as possible sources for such damping. The most popular linear ones are viscosity in the disk (Takeuchi, Miyama, & Lin 1996) and radiative damping of the tidal perturbations (Cassen & Woolum 1996). Viscosity can dissipate tidal perturbations on scales smaller than the typical disk sizes only if it is large enough ($\alpha > 10^{-4}$; Drimmel 1996; Takeuchi et al. 1996; GR01), but it is difficult to identify a strong source of viscosity in protoplanetary disks. The most probable viscous mechanism in hot accretion disks—magnetohydrodynamic (MHD) turbulence driven by the magnetorotational instability (Velikhov 1959; Balbus & Hawley 1998)—probably does not operate in protoplanetary disk environments: the gas is too cold and weakly ionized throughout most of the disk (Jin 1996; Hawley & Stone 1998). Convection was put forward as another possible source of viscosity (Lin & Papaloizou 1980), but analytical arguments and numerical simulations cast serious doubt on the ability of this mechanism to produce outward angular momentum transport (Ryu & Goodman 1992; Balbus, Hawley, & Stone 1996).

The efficiency of radiative damping in protoplanetary disks is strongly reduced by dust opacity (Henning & Stognienko 1996), which leads to very high optical depths ($\tau \gtrsim 10^3$) and implies very small radiative losses from the disk surface. Waves could also be damped by radiative transfer in the plane of the disk, but this turns out not to be very important either (GR01). This implies that linear damping mechanisms cannot lead to the immediate transfer of the density-wave angular momentum to the disk material, contrary to the popular assumption (e.g., Ward & Hourigan 1989).

Nonlinear dissipation, namely, shock formation and its consequent damping, seems to represent a more efficient and almost inevitable process for transferring the angular momentum of the wave to the disk fluid. There are two reasons for this. First, the differential rotation of the background fluid causes the wavelength of the tidal perturbation to decrease as it travels away from the planet (Goldreich & Tremaine 1979), thus “compressing” the wave profile and facilitating its breaking. Second, the amplitude of the planet-induced wake is growing, at least in the planetary vicinity, as a consequence of the conservation of the angular momentum flux, and this speeds up the nonlinear evolution of the density wake. These processes working together can make the wave shock very rapidly if the initial wave amplitude is significant (i.e., if the perturber is massive enough). Also, for the shock to form, the nonlinearity is not required to be strong—even for almost linear waves, nonlinear effects accumulate and inevitably lead to the shock formation (if the linear wave dissipation described above is weak).

GR01 have performed a detailed study of the nonlinear evolution of planet-induced density waves in two-dimensional disks. They have demonstrated that for a small enough planetary mass, it is possible to separate the wake evolution into two distinct stages: linear generation, which is completed about $(1-2)h$ from the planet (here $h = c/\Omega$ is a typical scale length, which equals disk thickness in three dimensions, $c$ is the sound speed, and $\Omega$ is the disk angular frequency), and then nonlinear evolution of the wake causing it to shock. In the first stage, one can neglect all nonlinear effects (if the planet has sufficiently low mass) but include the tidal force leading to the wake generation. In the second stage, far from the planet, the situation is
reversed: planetary torques are not important anymore, but nonlinear effects accumulate and must be treated explicitly in the course of free wake propagation.

The distance that the wave has to travel before it shocks is typically several disk scale lengths $h$ from the perturber. It is proportional to $M_p^{-2/5}$, where $M_p$ is the mass of the planet (GR01). Thus, the wave damping is not immediate for low-mass planets, even if one considers nonlinear mechanisms. After the shock is formed, it dissipates, transferring its angular momentum to the mean flow and leading to disk evolution. In favorable conditions, the damping of the density waves can make the disk evolve on timescales comparable to those derived from observations ($10^3$–$10^7$ yr; see Hartmann et al. 1998).

The considerations of GR01 were in a certain sense local because of the assumed shearing-sheet approximation. It was shown that this approach is good for the description of the shock formation but probably not so accurate for studying the shock dissipation: in the shearing-sheet geometry, shock damping proceeds slowly, and at some point, the background fluid velocity can no longer be represented by just a uniform shear. The analysis of GR01 also assumed a disk with constant background surface density and sound speed and did not take into account the effects of the disk’s polar geometry on the evolution of the density-wave amplitude. These approximations naturally lead to a picture in which the wake itself and its damping pattern are symmetric on both sides of the planet.

The main purpose of this paper is to extend the analysis of GR01 by including the effects of radial surface density and sound-speed variations in the disk on the behavior of the weak tidal disturbances generated by a low-mass planet incapable of opening a gap in the disk. As a result, a quantitative description of the density wave damping is provided, which can be used in other studies (such as the gap formation in a gaseous disk; see Rafikov 2001). We consider a Keplerian rotation law (not a linearly sheared background flow) and a polar geometry to include the self-consistently important ingredients needed to provide a global picture of the nonlinear evolution of the density waves in non–self-gravitating disks. We restrict our attention to purely two-dimensional disks, thus completely disregarding vertical motions and related phenomena, such as wave-action channeling in the vertical direction (Lin, Papaloizou, & Savonije 1990; Lubow & Ogilvie 1998). We believe that this is a good approximation in passive, externally irradiated protoplanetary disks with high optical depths that should have almost isothermal vertical structure (Chiang & Goldreich 1997).

The paper is structured as follows. In § 2 we study the full system of fluid equations in cylindrical geometry. We provide in § 2.1 simple scaling laws for the behavior of the wake amplitude and shock-formation distance in the quasilinear approximation from rather general qualitative considerations. We then confirm these estimates by an accurate quantitative analysis in § 2.2. Wake properties are studied in § 3, in particular for the case of disks with power-law surface density and sound-speed radial dependencies (§ 3.2). Applications of our results are discussed in § 4.

2. DENSITY-WAVE STRUCTURE

We consider a system consisting of a gaseous non–self-gravitating disk rotating in the unperturbed potential $U_\star(r)$ ($r$ is the distance from the central object) and a planet of mass $M_p$ located at a distance $r_p$ from the center. The disk is assumed to be geometrically thin, and its unperturbed surface density $\Sigma_0$ and sound speed $c_0$ are both taken to be functions of $r$. Disk scale height in three dimensions is $h = c_0/\Omega$. We always denote by the subscript “$p$” various quantities evaluated at the position of the planet.

The equations of motion and continuity for the two-dimensional disk read as always:

$$\frac{\partial \mathbf{v}}{\partial t} + \mathbf{v} \cdot \nabla \mathbf{v} = -\frac{1}{\Sigma} \nabla P - \nabla U ,$$

$$\frac{\partial \Sigma}{\partial t} + \nabla \cdot (\rho \mathbf{v}) = 0 .$$

Here $\mathbf{v}$ and $\Sigma$ are the fluid velocity and surface density, respectively, $P$ is the pressure, and the potential

$$U = U_\star - G \frac{M_p}{|r - r_p|} + U_i$$

consists of three contributions: the potentials of the central star and perturbing planet (we assume a Keplerian potential $U_\star = -GM_\star/r$, where $M_\star$ is the mass of the central star) and the indirect potential $U_i$, which can always be neglected here (because disk mass is much smaller than $M_\star$).

We assume that the pressure $P$ is related to the instantaneous value of the surface density $\Sigma$ by the locally polytropic law with a specific index $\gamma$:

$$P = P_0(r) \left( \frac{\Sigma}{\Sigma_0(r)} \right)^\gamma .$$

Then, the perturbed sound velocity is given by the usual expression,

$$c^2 = \frac{\partial P}{\partial \Sigma} = c_0^2(r) \left( \frac{\Sigma}{\Sigma_0(r)} \right)^{\gamma - 1} ,$$

meaning that $P_0(r) = \Sigma_0(r)c_0^2(r)/\gamma$. Clearly, $P_0(r)$ and $c_0(r)$ are unperturbed values of pressure and sound speed. The equation of state given by (4) does not force $c_0(r)$ to be related to $\Sigma_0(r)$, and this gives us additional flexibility in applications. The entropy of the disk fluid now varies with $r$, contrary to the equation of state with a radially independent polytropic constant.

2.1. General Considerations

Using basic physical principles such as conservation of angular momentum flux, it is possible to analyze the behavior of the wake in the quasilinear regime and determine when the wave shocks on a qualitative level. The simple scaling laws obtained in this way are confirmed below by a more rigorous analysis of the complete system of fluid equations.

Let us first concentrate on linear wake propagation. One can easily show that the solution of equations (1) and (2) in linearized form with the equation of state (4) yields, using the WKB approximation,

$$m^2 \left( \Omega - \Omega_p \right)^2 = \kappa^2 + k^2 c^2 ,$$

$$\frac{d}{dr} \left[ \frac{r c_0^3}{(\Omega - \Omega_p)\Sigma_0} (\Sigma - \Sigma_0) \right] = 0 ,$$

for non–self-gravitating disks. Here $m$ is the azimuthal wavenumber, $k$ is the radial wavenumber of the perturba-
tion, \( \Omega_n \) is the perturbation pattern angular frequency, 
\((\Sigma - \Sigma_0)_{m}\) is the \(m\)th harmonic of the surface-density 
perturbation, and
\[
\Omega^2 = \frac{1}{\Sigma_0} \frac{dU}{dr} + \frac{1}{\Sigma_0} \frac{dP_0}{dr}, \quad \kappa^2(r) = 4B(r)\Omega(r) , \tag{8}
\]
where
\[
B(r) = \Omega(r) + \frac{r d\Omega}{2 dr} . \tag{9}
\]
Equation (6) is the usual dispersion relation for small perturbations in 
a thin differentially rotating disk, and it shows that density waves behave 
basically like sound waves after propagating several scale lengths \(h\) from 
the perturber. Both equations (6) and (7) coincide with their analogs for 
the disks with constant entropy (Goldreich & Tremaine 1979).
One can also demonstrate that the total angular momentum flux 
across a cylinder of radius \(r\) carried by the \(m\)th harmonic of 
surface-density perturbation is given by
\[
F_{jm} = \pi \frac{r c^3}{(\Omega - \Omega_p)(\Sigma - \Sigma_0)_m^2} \tag{10}
\]
(where the same as the analogous expression of Goldreich 
& Tremaine 1979), which, when combined with equation (7), implies 
that the total angular momentum flux \(F_j\) is conserved.
If the nonlinear effects are fully neglected and \(F_j\) is strictly 
conserved as demonstrated above, then it follows from 
equation (10) that the magnitude of the surface-density 
perturbations scales with other flow variables as
\[
(\Sigma - \Sigma_0)^2 \propto \frac{\Sigma_0}{r c^3} F_j . \tag{11}
\]
This equation demonstrates how the amplitude of the wave 
varies in the linear regime, and we return to it in §2.2.

In reality, even if the perturbation is small, different 
points of the wake profile have different propagation velocities 
since the waveform constantly distorts. This nonlinear 
evolution leads to shock formation. In a shock, the 
angular momentum of the density wave gets transferred to 
the mean flow so that the scaling provided by equation (11) 
breaks down after the shock is formed. We can estimate 
when this happens.

Let us consider a part of the profile that initially has a 
phase \(\varphi = \varphi_0\). It evolves according to
\[
\left. \frac{\partial \varphi}{\partial r} \right|_{\varphi_0} = 2\pi \frac{\delta c(r)}{\lambda(r)c(r)} , \tag{12}
\]
where \(\delta c(r)\) is the perturbation of the propagation velocity 
\(c\), which is the sound velocity in our case (it is different for 
different \(\varphi_0\), and this is responsible for the profile distortion), 
and \(\lambda(r)\) is the current wavelength. The sound speed 
perturbation \(\delta c\) should be considered as the product of its 
initial value \(\delta c_0\), which is a function of \(\varphi_0\) only, and of 
\(\delta c/\delta c_0\), which depends only on \(r\) in the linear approximation.
Integrating equation (12) with respect to \(r\), one gets that
\[
\varphi = \varphi_0 + 2\pi \int_{r_0}^{r} \frac{\delta c(r)}{\lambda(r)c(r)} dr \tag{13}
\]
\((r_0)\) corresponds to the point where the wave is launched.

The wave shocks when its profile acquires an infinite slope 
after travelling some distance \(r_{sh}\): \(d\varphi/d\varphi_0 = 0\) at \(r = r_{sh}\) for 
some specific \(\varphi_0\) (for which the most negative value of the 
derivative \(d\varphi/d\varphi_0\) is reached). This gives us (using eq. [13]) the condition for \(r_{sh}\) in the following form:
\[
\int_{r_0}^{r_{sh}} \frac{k(r)\delta c(r)}{c(r)} dr = \text{const} . \tag{14}
\]
From the dispersion relation in equation (6), we get that 
\(k \propto (\Omega - \Omega_p)/c\), since \(\delta c/c \propto \Sigma/\Sigma_0\), we can use the scaling 
provided by equation (11) to finally obtain the shocking 
condition in the form
\[
\int_{r_0}^{r_{sh}} \left[ \frac{(\Omega - \Omega_p)^2}{r c^3} \right]^{1/2} dr = CF_j^{1/2} \propto M_p^{-1} , \tag{15}
\]
where \(C\) is some constant.

The scaling laws in equations (11) and (15) immediately 
provide important information about the wake propagation 
in the linear regime and the conditions needed for shock 
formation.

2.2. Basic Equations

In this section we study the behavior of the wake by solving 
the fluid equations for small perturbations in a weakly 
nonlinear regime. We work in the polar coordinate system 
rotating with the angular velocity of the perturber 
\(\Omega_p = [(1/r_p)(\partial U/\partial r)]_{r_0}^{1/2}\). In this coordinate frame, the 
flow is stationary (time independent). As always, we take 
the \(r\)-axis to be directed out from the central body and the 
\(\phi\)-axis to be directed in a prograde sense.

In this coordinate system, the equations of motion in the \(r\) 
and \(\phi\) directions are (Landau & Lifshitz 1959)
\[
v_r \partial_r v_r + \frac{v_\phi}{r} \partial_\phi v_r - \frac{v_\phi^2}{r} = - \frac{1}{\Sigma} \partial_r P + 2\Omega_p v_\phi + \Omega^2 r - \partial_\phi U , \tag{16}
\]
\[
v_r \partial_r v_\phi + \frac{v_\phi}{r} \partial_\phi v_\phi + \frac{v_\phi v_r}{r} = - \frac{1}{\Sigma} \partial_r P - 2\Omega_p v_r - \frac{1}{r} \partial_\phi U , \tag{17}
\]
and the continuity equation is
\[
\frac{1}{r} \partial_r (\Sigma r v_r) + \frac{1}{r} \partial_\phi (\Sigma v_\phi) = 0 . \tag{18}
\]
Here \(v_r\) and \(v_\phi\) are the fluid velocities in the \(r\) and \(\phi\) directions, 
respectively. Equations (16) and (17) include Coriolis 
and centrifugal forces.

GR01 have demonstrated that the system of equations 
(16)–(18) simplifies significantly if the mass of the perturber 
\(M_p\) is smaller than a characteristic mass \(M_1\) given by
\[
M_1 \equiv \frac{\Sigma_0}{2A} G , \tag{19}
\]
where \(A(r) = (r/2)d\Omega/dr\) is the Oort’s \(A\) constant that is 
related to \(B(r)\) by \(B = \Omega + A\) (eq. [9]). Then, the disk can be 
split into two distinct regions: an excitation region, within 
several scale lengths \(h_p = h(r_p) = c_0(r_p)/\Omega(r_p)\) from the 
planet, where one can neglect nonlinear effects, and a propagation 
region beyond several \(h_p\) from the planet, where the 
planetary potential is negligible and one can study wake 
evolution caused by nonlinear effects. It can also be shown
that planets with $M_p \ll M_1$ produce waves that are only weakly nonlinear—the surface-density perturbations associated with them are only a factor $\sim M_p/M_1$ of the unperturbed $\Sigma$. On the contrary, planets with masses larger than $M_1$ generate strongly nonlinear waves in the disk (Lin & Papaloizou 1993).

In the excitation region, numerous Lindblad resonances tidally excite density waves corresponding to different azimuthal harmonics and provide individual Fourier contributions to the angular momentum flux (GT80). Most of the angular momentum comes from the resonances with azimuthal wavenumbers $m \sim r_p/h_p \gg 1$ that lie close to the planet—at distances $\sim h_p$ from it. Harmonics with smaller $m$ are weaker because the tidal excitation they experience is reduced by their larger distance from the planet. Fourier components with $m$ higher than $r_p/h_p$ are strongly damped because of the so-called torque cutoff (GT80). Its nature can be qualitatively understood as follows: near the planet, less than $(2/3)h_p$ from it (in a Keplerian disk), the background fluid flow is subsonic, which does not allow a stationary perturber to excite sound waves (Landau & Lifshitz 1959). This cutoff strongly reduces the amount of the angular momentum flux carried by the corresponding harmonics.

GR01 studied the linear wave formation process in the shearing-sheet approximation, taking into account not only the contributions of individual resonances to the torque but also the phases of the Fourier harmonics of the surface-density perturbation that allowed them to obtain the shape of the wake in the linear theory. Here we simply use their results in our more general case because wake generation is essentially a local process, spanning only a few $h_p$ from the planet, where the shearing-sheet approximation gives a good representation of the disk velocity profile and $\Sigma_0$ and $c_0$ can be assumed constant. The solution of this problem in our more global setting with varying $\Sigma_0$ and $c_0$ will only be different from the previously obtained local solution by factors of the order of $h_p/r_p \ll 1$, which is of no interest to us here.

Thus, we can proceed immediately to studying the wake-propagation region. To do this, we use a simple extension of conventional perturbation theory capable of including a weak nonlinearity of the wave. We assume that $v_r$ and $v_\phi$ are given by

$$v_r = u, \quad v_\phi = v_0(r) + v, \quad \text{with} \quad v_0(r) = r[\Omega(r) - \Omega_p],$$

where $u$ and $v$ are the velocity perturbations, and we take $u, v \ll v_0$, since the shock is assumed to be weak and we are always several scale lengths away from the planet. We often write for brevity $\Delta \Omega = \Omega(r) - \Omega_p$.

Substituting equation (20) into equations (16)–(18), one obtains the following perturbation equations:

$$\Delta \Omega \partial_r u + u \partial_r u - 2v_v = -\left(\frac{1}{\Sigma} \partial_r P - \frac{1}{\Sigma} \partial_r P_0\right) + \frac{v^2}{r},$$

$$\Delta \Omega \partial_\phi v + u \partial_\phi v + 2B u = -\frac{r}{\Sigma} \partial_\phi P - \frac{vu}{r},$$

$$\Delta \Omega \partial_\phi \Sigma + u \partial_\phi \Sigma + \Sigma \partial_r u = -\frac{\Sigma u}{r} - \frac{v}{r} \partial_r \Sigma - \frac{\Sigma}{r} \partial_\phi v.$$

We have neglected $v$ everywhere in comparison with $\Delta \Omega r$ and made use of the fact that $\partial_r P_0 = 0$. In equations (21)–(23), terms quadratic in $u$ and $v$ are subdominant; however, we keep $u \partial_r u$ terms because they are the strongest drivers of nonlinear evolution. Also, we assume that $v \ll u$ and $\partial_\phi \ll \partial_r$, as a consequence of the tight-winding approximation. These assumptions are checked in the Appendix.

By introducing a new radial coordinate given by

$$\xi = \int_r^r \left[\Omega(r) - \Omega_p\right] dr,$$

this system is transformed into

$$\partial_\xi u + u \partial_\xi P + \frac{1}{\Sigma_0} \partial_\xi \Sigma_0 = \frac{20}{\Delta \Omega} \frac{v^2}{\Delta \Omega r},$$

$$\partial_\xi v + u \partial_\xi v + \frac{c^2}{\Delta \Omega r} \partial_\xi \Sigma = -\frac{2B}{\Delta \Omega} \frac{u v}{\Delta \Omega r},$$

$$\partial_\xi \Sigma + u \partial_\xi \Sigma + \Sigma \partial_\xi u = -\frac{\Sigma u}{\Delta \Omega r} - \frac{v}{\Delta \Omega r} \partial_\xi \Sigma - \frac{\Sigma}{\Delta \Omega r} \partial_\xi v.$$

The left-hand sides of equations (25) and (27) are similar to the usual system of equations describing isentropic one-dimensional gas motion (Landau & Lifshitz 1959), which possesses two invariants conserved on characteristics: Riemann invariants $R_\pm$. In our case, the nonzero right-hand sides of equations (25) and (27) cause these invariants to not be conserved exactly, but one can still use them in a slightly modified way. We extract from equations (25) and (27) two equations of evolution of the Riemann invariants:

$$[\partial_\xi + (u \pm c) \partial_\xi] R_\pm = -\left[\frac{1}{\Sigma} \partial_\xi P - \frac{1}{\Sigma_0} \partial_\xi P_0 - c \partial_\xi \frac{2c}{\gamma - 1} \right]$$

$$\pm cu \frac{\partial_\xi \Sigma}{\Sigma} + u \partial_\xi \frac{2c}{\gamma - 1},$$

$$\Delta \Omega v + \frac{v^2}{\Delta \Omega r} + \frac{cu}{\Delta \Omega r} \partial_\xi \ln \Sigma = -\frac{\Sigma u}{\Delta \Omega r} - \frac{v}{\Delta \Omega r} \partial_\xi \Sigma - \frac{\Sigma}{\Delta \Omega r} \partial_\xi v,$$

where

$$R_\pm = u \pm \frac{2c}{\gamma - 1}.$$
\[ \eta \equiv \frac{r_p}{l_p} \left[ \phi + \text{sign}(r - r_p) \int_{r_p}^{r} \frac{\Omega(r') - \Omega_p}{c_0(r')} \, dr' \right], \quad (33) \]
\[ g(r) \equiv \frac{2^{1/4}}{r_p c_0^3} \frac{r \Sigma_0 c_0^3}{(\Omega - \Omega_p)} \left( \frac{r \Sigma_0 c_0^3}{|\Omega - \Omega_p|} \right)^{1/2}. \quad (34) \]

Here \( l_p = c_p/|2a(r_p)| = (2/3)h_p \) is a Mach 1 length (distance from the planet where the Keplerian shear makes fluid velocity equal to \( c_p \)).

In the immediate vicinity of the planet (but still several \( h_p \) from it), the definitions in equations (31)–(34) reduce to
\[ \chi \to \frac{\gamma + 1}{2^{1/4}} \left| \frac{x}{l_p} \right| \left( \frac{\Sigma - \Sigma_0}{\Sigma_0} \right), \quad t \to \frac{2^{3/4}}{5} \left| \frac{x}{l_p} \right|^{5/2}, \]
\[ \eta \to \frac{y}{l_p} + \frac{x^2}{2 l_p^2} \text{sign}(x), \quad (35) \]
where \( x = r - r_p \) and \( y = r_p \phi \). These equations coincide with analogous expressions in GR01.

The tidal perturbation launched by the planet follows a nearly parabolic path in the immediate vicinity of the planet where the shear can be assumed constant. Farther from the perturber, where the shear is no longer uniform, the density wave has a spiral shape. Its pattern is described by the equation
\[ \phi = \phi_0 - \text{sign}(r - r_p) \int_{r_p}^{r} \frac{\Omega(r') - \Omega_p}{c_0(r')} \, dr'. \quad (36) \]

Depending on the conditions in the disk, this spiral can wind up several times around the center before the wave damps. One can see from equation (36) that the perturbation is indeed tightly wound if \( |r - r_p| \gg h_p \) (this could, however, be violated farther away from the planet for some profiles of \( \Sigma_0 \) and \( c_0 \); see § 3.2).

Now, if we disregard the nonlinear evolution entirely, it follows from equation (30) that \( \chi = f(\eta) \) \( [f(x)] \) is an arbitrary function of the argument \( x \), or
\[
\begin{align*}
(S_\Sigma - S_0)_{\text{lin}} & \propto \frac{\Sigma_0}{g(r)} f(\eta) \\
& \propto \frac{\Sigma_0 (\Omega - \Omega_p)}{r c_0^3} \left( \frac{\Sigma_0 (\Omega - \Omega_p)}{|\Omega - \Omega_p|} \right)^{1/2} \\
& \times f \left( \phi + \text{sign}(r - r_p) \int_{r_p}^{r} \frac{\Omega(r') - \Omega_p}{c_0(r')} \, dr' \right). \quad (37)
\end{align*}
\]

This means that in the linear regime, the waveform propagates along the wake (whose location is given by the condition \( \eta = \text{const} \)), and its amplitude scales with the distance from the planet in complete agreement with equation (11).

3. NONLINEAR EVOLUTION OF THE WAKE

Burger’s equation (30) is probably the simplest partial differential equation able to exhibit a shock-formation phenomenon (Whitham 1974). For the reasons outlined in § 2.1, the profile of the density wave produced by the linear generation mechanism is constantly distorted in the course of its propagation away from the planet, so that finally the waveform breaks to become double valued. This implies that a shock must be formed at this point. The distance that the density wave travels before it shocks depends on the initial shape and the amplitude of the wake.

To study shock formation and propagation quantitatively, we need to solve equation (30) with the initial condition given by the solution of the linear wake generation problem. As we have mentioned in § 2.2, since all our variables reduce to the corresponding variables of GR01 in the limit \( r \to r_p \), the whole linear-generation problem reduces to the one solved before: shock excitation in the shearing-sheet approximation. Thus, we might use the solution for the wave shape \( f_0(t) \) calculated in GR01 as the initial condition in our more general case:
\[ \chi(M_p, t = t_0, \eta) = \frac{M_p}{M_1} f_0(\eta) \quad (38) \]

\( (M_1 \) is defined in eq. [19]). Here the initial profile \( \chi(M_p, t = t_0, \eta) \) is taken not at \( t = 0 \) but at some \( t_0 > 0 \) because in the linear regime, the wake has to propagate some distance from the planet before it fully forms. Thus, \( t_0 \) is a matching boundary where one switches from the linear to the weakly nonlinear regime, and following GR01, we take \( t_0 \approx 1.89 \), corresponding to the distance from the planet \( \chi_{0} = 2h_p = (4/3)h_p \). One can usually neglect \( t_0 \) once the wave has travelled several \( h_p \) away from the planet.

The function \( f_0(\eta) \) represents the shape of the initial profile at the moment \( t = t_0 \) for \( M_p = M_1 \) (see eq. [19]). It was calculated by GR01 in the shearing-sheet approximation and is reproduced for convenience in Figure 1 (top). The factor \( M_p/M_1 \) in equation (38) rescales the amplitude of the wake for an arbitrary mass of the planet.

In terms of the variables \( \chi, \eta, \) and \( t \), our nonlinear problem in equation (30) is identical to the one studied in the
shearing-sheet approximation (including the initial condition in eq. [38]). This means that the result of GR01 for \( \chi(M_1, t, \eta) \) in terms of variables \( \chi, \eta, \) and \( t \) is also a solution of our more general problem in terms of these variables. One needs only to express them in terms of \( r \) and \( \phi \) using equations (31)–(34) and rescale to an arbitrary mass \( M_p \) in the following way: suppose that \( \chi_1(t - t_0, \eta) = \chi(M_1', t - t_0, \eta) \) is a solution of equation (30) with the boundary condition (38) and \( M_p = M_1' \) calculated in GR01. Then one can easily see that the solution for an arbitrary \( M_p \) can be written as

\[
\chi(M_p, t - t_0, \eta) = \frac{M_p}{M_1'} \chi\left(\frac{M_p}{M_1'}(t - t_0), \eta\right).
\]  

(39)

This reduction to the previously studied case is a very convenient feature of the analysis because it allows one to use all results previously obtained by GR01 in our more general setting. For instance, it was found before that the shock has to form at a certain value of \( t \) as given by the function \( f_0(\eta) \) see Fig. 1 (top). The corresponding radial distance from the planet \( r_{sh} - r_p \) at which this happens can be found using equations (32) and (34): \( r_{sh} - r_p \approx 0.93 \left(\frac{M_p}{M_1}\right)^{2/3} \left(\frac{\rho_0}{\rho_{cr}}\right)^2 \) (we have neglected \( t_0 \) here and used the local approximation). Based on what we have previously said, the shock has to form at the same value of \( t_{sh} \) in our global case, and for a fixed \( M_p \) this leads to the condition in equation (15) found before on the basis of qualitative considerations.

After the wave shocks, it starts transferring its angular momentum to the disk mean flow, which leads to the damping of the wave amplitude. The initial profile shown in Figure 1 (top) has both positive and negative parts; thus, it is destined to evolve asymptotically into an “N-wave” profile (Landau & Lifshitz 1959; Whitham 1974). In this regime, the amplitude of \( \chi \) decays as \( t^{-1/2} \), whereas it was constant near the planet before forming a shock. The width of the spiral perturbation, which was \( \approx h_p \) initially, grows as \( t^{1/2} \) in the N-wave stage.

### 3.1. Angular-Momentum Flux

Summing up all the angular momentum flux contributions given by equation (10) and applying Parseval’s theorem, one can show that the total angular momentum flux of the density wave \( F_j \) is given by

\[
F_j(r) = \frac{c_s^2(r)r}{\Omega(r) - \Omega_{\infty}} \int_{-\infty}^{\infty} (\Sigma - \Sigma_0)^2 d\phi.
\]  

(41)

Using the definitions in equations (31) and (34), we obtain that in terms of the variables \( t \) and \( \eta \),

\[
F_j(t) = \frac{2^{3/2} c_s^2 p \Sigma_p}{(\gamma + 1)^3} \Phi(M_p, t),
\]

where

\[
\Phi(M_p, t) = \int \chi^2(M_p, t, \eta) d\eta = \left(\frac{M_p}{M_1'}\right)^2 \Phi(M_1', \frac{M_p}{M_1'} t),
\]  

(42)

exactly like in GR01 (we neglected \( t_0 \) here). From equation (30), one can easily see that without shocks, \( \Phi(M_p, t) = \text{const} \) and angular momentum flux is conserved.

After the shock forms, \( \Phi \) starts to decrease. Asymptotically, in the N-wave regime it falls as \( \Phi(M_p, t) \propto t^{-1/2} \). The behavior of \( \Phi \) as a function of \( t \) for \( M_p = M_1 \) was calculated in GR01 and is shown in Figure 1 (bottom). What happens with the angular momentum flux of the wave in the disk depends on how \( t \) is related to the distance traveled by the wave. If one assumes a uniform shear, then according to equation (35), \( t \propto |r - r_p|^{3/2} \), and at infinity the wave damps completely. Real disks are different—they are always finite, dissipation can be incomplete, and the damping pattern is asymmetric between the inner and outer disks. We demonstrate below that for some background surface density and sound-speed profiles, \( t(r = \infty) \) or \( t(r = 0) \) are finite. This means that a density wave that has managed to propagate to the outermost parts of the disk, or to its center, still carries some undamped angular momentum flux. It could even happen that \( t(r = \infty) < t_{sh} \) or \( t(r = 0) < t_{sh} \), meaning that the wave does not shock before it reaches the outer or inner edge of the disk. The wave could then be reflected, and part of its angular momentum could be carried back to the planet. This might have important consequences for planetary migration, and we dwell on this more in § 4.

### 3.2. Power-Law Disks

We now consider the case of power-law disks, i.e., \( \Sigma_0 \) and \( c_0 \) are assumed to be some powers (usually negative) of \( r \), to illustrate the results obtained before and test some of our assumptions. We consider a Keplerian disk with \( U_0 \propto r^{-1} \) and take \( \Sigma_0 = \Sigma_p(r/r_p)^{-\delta} \), \( c_0 = c_p(r/r_p)^{-\nu} \), \( \delta > 0 \), and \( \nu > 0 \).

Using the definitions in equations (31), (32), and (34), we find that

\[
t = \frac{3}{2^{5/4}} \left(\frac{r_p}{h_p}\right)^{5/2} \int_{r_p}^{r_0} \left|\chi^2/2 - 1\right|^{5/2} d\phi \]  

\[
= \left[\left(\frac{r/r_p}\right)^{5/2} - \left(\frac{r/r_p}\right)^{5/2+1}\right],
\]  

(43)

and the wake equation

\[
\phi = \phi_0 - \text{sign}(r - r_p) \frac{r_p}{h_p} \left[\left(\frac{r/r_p}\right)^{5/2} - \left(\frac{r/r_p}\right)^{5/2+1}\right].
\]  

(44)

Using equation (43), let us first consider the inner \( (r < r_p) \) part of the disk. One can easily see that \( t \rightarrow \infty \) as \( r \rightarrow 0 \) if \( \rho = 5\nu + \delta < \rho_{cr} = 7/2 \). This means that the density waves launched by the planet at \( r_p \) in the inner disk span the whole range of \( t \) and thus damp completely upon reaching the center. Of course, real disks always have an inner cavity, presumably formed by the protostellar magnetospheric activity, where the gas is absent, so we only have to consider wave propagation up to this inner edge. Depending on the location of the planet, this inner disk boundary may be very close to the center, so we consider wave propagation down to \( r = 0 \).

For \( \rho > \rho_{cr} \), however, \( r = 0 \) corresponds to a finite \( t \), and if \( t(0) < t_{sh} \), the wave does not shock at all as \( r \) decreases to
0. Indeed, for \( \rho = 5\nu + \delta > \rho_{\text{cr}} = 7/2 \), one obtains that

\[
t(0) = \frac{1}{2^{1/4}} \frac{r_p}{h_p} \frac{5/2}{B \left( \frac{5\nu + \delta}{3} - \frac{7}{6} \right)},
\]

where \( B(\alpha, \beta) \) is a beta function. Using equation (40) and neglecting \( t_0 \) in it, we get that if

\[
M_p < M_x \equiv M_1 0.94 \left( \frac{h_p}{r_p} \right)^{5/2} \left[ B \left( \frac{5\nu + \delta}{3} - \frac{7}{6} \right) \right]^{-1},
\]

then the wave does not shock on the first passage to the disk center. For this to happen, \( M_p \) should be sufficiently small: assuming that \( B(\alpha, \beta) \sim 1 \) and \( r_p/h_p \gg 1 \), one needs \( M_p \lesssim M_1 (h_p/r_p)^{5/2} \) for this to occur. If \( \rho > \rho_{\text{cr}} \), but the condition in equation (46) is not fulfilled (if the planet is massive enough), the wave shocks as it travels toward the disk center, but it does not damp completely upon reaching it. Part of the angular momentum flux could be transferred back to the planet in this case.

One can easily see that in the outer disk, this problem never arises: for \( \delta > 0 \) and \( \nu > 0 \), the wave always shocks as it propagates outward, if we forget about the outer boundary of the disk. In Table 1 we summarize final outcomes of wave propagation in the inner and outer parts of the disk for different values of \( \rho \). In Figure 2 we plot the behavior of \( t(r) \) for different values of \( \rho \) in the inner and outer parts of the disk using equation (43). In the minimum-mass solar nebula (MMSN), \( \delta = 3/2 \) and \( \nu = 1/4 \) (Hayashi 1981), meaning that \( \rho = 11/4 < \rho_{\text{cr}} \). Thus, in the MMSN, tidal waves always shock and are damped on both sides of the disk.

It is interesting to see how the shock damping is distributed in the disk. Using the asymptotic behavior of \( \Phi(r) \) for large \( r \), we find that

\[
F_J \propto \left( \frac{r}{r_p} \right)^{(7-2\rho)/8}
\]

as \( r \to 0 \). As \( \rho \) varies from 0 to \( \rho_{\text{cr}} \), the power-law index here changes from 7/8 to 0 (it is equal to 3/16 for the MMSN), which implies that for some \( \Sigma_0 \) and \( c_0 \) profiles, quite a lot of the residual angular momentum is transferred to the disk close to its center. This might significantly enhance the accretion rate there. In the outer disk,

\[
F_J \propto \left( \frac{r}{r_p} \right)^{-(\rho+1)/4}
\]

for \( r \to \infty \). For shallow profiles of \( \Sigma_0 \) and \( c_0 \), the nonlocality of the damping could be important in the outer disk too. The behavior of the dimensionless angular momentum flux \( \Phi(r) \) is shown in Figure 3 for different values of \( \rho \).

Table 1

| Wave Behavior in the Infinite Disk for Different Values of \( \rho > 0 \) |
|---------------------|---------------------|---------------------|
| Outcome of the Wave Propagation | Inner Disk | Outer Disk |
| Wave shocks, damps completely | \( 0 < \rho \leq \rho_{\text{cr}}, \text{any } M_p \) | Any \( \rho > 0 \), any \( M_p \) |
| Wave shocks, but does not damp completely | \( \rho > \rho_{\text{cr}}, M_p \geq M_x \) | \( \rho > \rho_{\text{cr}}, M_p < M_x \) |
| Wave does not shock and does not damp | \( \rho > \rho_{\text{cr}}, \text{any } M_p \) | \( \rho > \rho_{\text{cr}}, \text{any } M_p \) |

Fig. 2.—Dependence of the dimensionless variable \( t \) [multiplied by \( (h_p/r_p)^{5/2} \)] on the distance from the central body \( r \) for several values of the power-law index \( \rho = 5\nu + \delta \) (see the beginning of § 3.2). The left panel is for \( r < r_p \), and the right one is for \( r > r_p \). Curves are labeled by the corresponding values of \( \rho \). Note a universal behavior of \( t(r) \) near \( r = r_p \), where the shearing-sheet approximation is valid. In the inner part of the disk, \( t(r) \) is increasing to infinity as \( r \to 0 \) (which implies a complete dissipation of the shock) only for \( \rho < \rho_{\text{cr}} = 7/2 \). In the outer disk, the wave shocks and damps for all values of \( \rho \) as \( r \to \infty \). The dashed curve is for the MMSN.
Let us now test the validity of some of our assumptions that were used in the analysis presented in § 2.2. The tight-winding approximation is one of them. From equation (36), we find that
\[ \tan \theta = r \frac{\Omega(r) - \Omega_\circ}{c_0(r)} = \frac{r_p}{h_p} \left( \frac{L}{r_p^2} \right)^{\nu+1} \left[ \left( \frac{r_p^8}{r^8} \right)^{3/2} - 1 \right], \] (49)
where \( \theta \) is the angle between the radial direction and the tangent of the spiral and the last equality in equation (49) is for a Keplerian power-law disk. In the immediate vicinity of the planet, for \( |r - r_p| \sim h_p \), this angle is small and the spiral pattern is not tightly wound. However, this is the region of the wake generation, where the free nonlinear propagation approach is not applicable anyway. After travelling several \( h_p \) from the planet, the wake becomes tightly wrapped by the Keplerian shear if the disk thickness is small \( (r_p/h_p \gg 1) \).

In the outer disk, \( \theta \rightarrow \pi/2 \) asymptotically as \( r \rightarrow \infty \) (the spiral winds up). This happens because in the frame rotating with the planet, outer parts of the disk move with the angular speed \( -\Omega_p \), giving rise to a large linear velocity, while \( c_0 \) decreases with growing \( r \). Inside the planet’s location, \( \tan \theta \approx (r_p/h_p)(r/r_p)^{\nu-1/2} \) as \( r \to 0 \). Thus, for \( \nu > 3/2 \), the spiral pattern unwinds in the inner regions of the disk and the tight-winding approximation might become inapplicable. However, if \( r_p/h_p \) is large enough, the wave could still reach the inner disk edge before this effect becomes important (in fact, the spiral pattern unwinds only if \( h \sim r \), which is usually not the case in the inner part of the disk).

Geometrical effects may also be of some importance. As we mentioned before, in the asymptotic regime, the wake width increases as \( t^{1/2} \). At the same time, if the pattern of the wake winds up, the distance between the consecutive wave crests (at a fixed polar angle \( \phi \)) decreases. At some point, the “rear” shock front of the N-wave profile comes so close to the “forward” shock front of the profile, lagging by \( 2\pi \) in \( \phi \), that our approximation of an almost constant Riemann invariant \( R_+ \) becomes poor because the change of \( R_+ \) during the shock crossing gets comparable to the change of \( R_+ \) following the shock (see the discussion in § 2.2).

From Figure 2 of GR01, one can find that the width of the shock in the \( \eta \)-coordinate in the asymptotic region \( t \to \infty \) is \( \Delta \eta \approx 2.3t^{1/2} \). Obviously, the “front-touching” phenomenon occurs when \( (r_p/h_p)\Delta \eta \gtrsim 2\pi \) (see eqs. (A10) and (A15)), that is, when \( t(r_p/h_p)^{3/2} \gtrsim 17(h_p/r_p)^{1/2} \). This particular form of the condition is used because it allows direct comparison with Figure 2. One can see from this figure that for \( r_p/h_p = 20 \) and any \( \rho > 0 \), one can propagate as far as \( r/r_p \approx 0.3 \) in the inner disk and \( r/r_p \gtrsim 4 \) in the outer, and still not encounter this “front touching.” In colder disks with higher \( r_p/h_p \), this limitation becomes more stringent, but even in this case, there is a significant region of applicability of our analysis not too far from the planet where \( R_+ \) could be assumed approximately constant.

We also need to mention that our treatment of the nonlinear evolution essentially neglects linear dispersion of the wave profile. Near the planet, wave dispersion is strongest, but it is properly taken into account there in the linear calculations of GR01. Farther away from the planet, according to equation (6), dispersion rapidly becomes less important in comparison with nonlinear steepening (if \( M_p \) is not very small) and can be ignored.

4. DISCUSSION AND APPLICATIONS

It is quite possible that in the presence of vertical temperature gradients, the channeling of the wave action into the vertical direction can damp density waves in the disk atmosphere more efficiently than we have found here (Lin et al. 1990; Lubow & Ogilvie 1998). It seems reasonable, however,
that in passive disks heated by their central stars, thermal stratification in the $z$-direction must be small because of their high optical depth. This strongly diminishes the wave action channeling into the disk atmosphere (Ogilvie & Lubow 1999) and leads to an almost two-dimensional picture of the wave propagation in the disk supporting the validity of our consideration.

Throughout our analysis, we assumed the wave nonlinearity to be weak, meaning that $(\Sigma - \Sigma_0)/\Sigma_0$ is small. From equation (38), we see that if $M_p \gtrsim M_1$, the wave is nonlinear from the very beginning, and it shocks immediately (see eq. [40]). This means that the separation of the disk into two distinct regions where one can neglect either planetary torques or nonlinearity of the wave does not exist. Also, for $M_p \gtrsim M_1$, a gap in the disk can form around the planet (Lin & Papaloizou 1993). Thus, as we have mentioned before in § 2.2, our analysis is applicable only for small-mass planets: $M_p \lesssim M_1$.

We have seen in § 3.2 that wave damping can be a nonlocal process and that part of the wave action could reach the disk edge undamped. This incomplete damping is important for the question of the planetary migration. The migration speed and direction depend sensitively on a delicate balance between the amounts of the angular momentum that the planet deposits into the inner and outer parts of the disk. If the density waves dissipate completely and transfer all their angular momentum to the disk fluid, then the only difference in torques acting on both sides of the disk is due to the surface density and temperature gradients in the disk and to asymmetries in the locations of the inner and outer Lindblad resonances (Ward 1986). The amount of the resultant torque that leads to the orbital evolution of the planet is only $\sim h_p/r_p$, compared to the magnitude of one-sided torque (GT80). Interaction with the outer part of the disk is usually stronger than with the inner one, leading to an inward migration (Ward 1986).

Let us now assume that tidal perturbations are not damped completely upon reaching the disk edge and are reflected from it. The remaining waves will be dissipated on the way back to the planet, but some of them might survive and interact gravitationally with the planet, returning to it some of the initially launched angular momentum. If this effect is able to return to the planet about $h_p/r_p$ of the one-sided angular momentum, then the migration could be strongly modified (Tanaka & Ward 2000 studied a similar effect caused by asymmetries in wave damping). Consider, for instance, a planet sitting close to the outer edge of the disk, but still several $h_p$ from it (otherwise strong asymmetries in the torque would be produced when the wave was still forming). A tidal wave launched in the inner disk might be completely dissipated because of the large distance it has to travel in the disk. At the same time, a wave in the outer disk might not shock at all before being reflected from the disk edge, and could bring a significant amount of the angular momentum back to the planet. Thus, interaction with the outer part of the disk would now be weaker than with the inner one, and migration would change its direction—the planet would move outward. In the same way, one could show that planets near the inner edge of the disk would tend to move inward faster if the waves were incompletely damped in the inner part of the disk. One can roughly describe this process by saying that the planet tends to be pushed out from the disk toward its closest boundary. The same picture holds for wave reflection off the edges of gaps formed by giant planets. These conclusions depend on a lot of assumptions, such as the details of the reflection process and the gravitational interaction of the reflected wave with the planet, which certainly deserve further study. Whether this process is an interesting issue for the question of planet formation and survival in the course of migration depends on the relevant timescales. Nevertheless, incomplete damping of the density waves introduces an additional degree of freedom on which the migration process depends and that could be important in some systems.

Deposition of the wave angular momentum into the disk fluid leads to the evolution of the disk itself (Larson 1989), Spruit (1987) and Larson (1990) found that the action of shocked density waves is equivalent to an appreciable viscosity with the corresponding dimensionless $\alpha$-parameter (Shakura & Sunyaev 1973) reaching $\sim 10^{-4}$–$10^{-3}$. However, they did not specify the source of the tidal perturbation, a deficiency remedied in GR01. There it was demonstrated that if all the solids in the disk were deposited into a population of Earth-sized objects, then an effective viscosity $\alpha \sim 10^{-4}$–$10^{-3}$ would again be produced.

Our results allow one to study the global evolution of the disk affected by all the planets present in it. The theory of time-dependent accretion disks (Pringle 1981) states that the mass accretion rate at each point in the disk is uniquely determined by the divergence of the angular momentum flux carried by the waves:

$$M = \left[ \frac{\partial}{\partial r} \left( r^2 \Omega \right) \right]^{-1} \frac{\partial F_j}{\partial r}.$$ (50)

Since $F_j$ depends on the distance travelled by the wave in a complex way (see eq. [42]), one should expect the $M$ produced by a single planet to be a function of $r$. Using equations (42), (43), and (50), we can calculate the accretion rate in the disk at a distance $r$ from the center produced by a single planet located at $r = r_p$:

$$M(r) = \text{sign}(r-r_p) \frac{29/4}{(\gamma + 1)^{1/2}} \frac{(h_p/r_p)^{1/2} \Omega(r_p)\Sigma_0(r_p)r_p^2}{M_p/M_1}$$
$$\times \left( \frac{M_p}{M_1} \right)^{3/2} \left( \frac{r}{r_p} \right)^{3/2} - 1 \left( \frac{r}{r_p} \right)^{-\gamma/2-9/4}$$
$$\times \Phi'(\frac{M_1}{M_p})^{1/2} \left( \frac{r}{r_p} \right).$$ (51)

Here $\Phi'(x) \equiv d\Phi(x)/dx$. Note that because $\Phi' < 0$, the accretion rate is positive inside the orbit of the planet (inflow) and negative outside of it (outflow)—the perturber tries to repel the surrounding gas.

In Figure 4 we plot the accretion rate in the disk due to the planetary torques from eight planets of the solar system (excluding Pluto) using equation (51). It is assumed that (Hayashi 1981)

$$\Sigma_0(r) = (1700 \text{ g cm}^{-2})r_{AU}^{-3/2}, \quad c_0(r) = (1.2 \text{ km s}^{-1})r_{AU}^{-1/4}$$

($r_{AU}$ is the distance from the center measured in AU), and for the masses of the giant planets, we take only the masses of their rocky cores: $15 M_\oplus$ for Jupiter and Saturn and 10 $M_\oplus$ for Uranus and Neptune (otherwise, they are likely to open gaps).
and sound speed have arbitrary radial distributions. They should only vary on scales larger than the perturbation wavelength for our tight-winding approximation to be valid (since in differentially rotating disks the radial wavenumber grows rapidly with distance from the planet, this condition does not pose serious restrictions). Thus, one can not only calculate the instantaneous accretion rate at each point in the disk but also study the self-consistent temporal evolution of the disk driven by the planetary torques. This is very important, for example, for studying the gap formation around the planet in gaseous disks, and we are going to investigate this process in a future study (Rafikov 2001).

![Figure 4](image-url)  
**Fig. 4.** Dependence of the planet-induced accretion rate $\dot{M}$ (in $M_\odot$ yr$^{-1}$) on the distance $r$ (in AU) in the MMSN. The torques produced by the eight major planets (only masses of the rocky cores are taken for giants) are included here, and the calculation is done using the usual MMSN parameters (see eq. [52]). Positive values of $\dot{M}$ (inflow toward the center) are displayed in the top panel, while negative ones (meaning the outflow) are in the bottom panel. Filled circles denote the locations of the planets. One can see that in some positions in the nebula, $\dot{M}$ could reach more than $10^{-8} M_\odot$ yr$^{-1}$, leading to a significant surface-density evolution there on timescales of $\sim 10^6$ yr.

One can see that in some locations in the nebula, $\dot{M} \sim (1-5) \times 10^{-8} M_\odot$ yr$^{-1}$, suggesting a significant surface-density evolution there. Indeed, the total disk mass inside Neptune’s orbit calculated using equation (52) is only $\approx 0.007 M_\odot$. Averaging $\dot{M}$ over the bulk of the disk, one obtains $\langle \dot{M} \rangle \approx (2-4) \times 10^{-9} M_\odot$ (depending on whether one averages $\dot{M}$ or $|\dot{M}|$), which implies the typical dispersal time of the MMSN by planetary torques $\approx (2-3) \times 10^6$ yr, in rough agreement with observations (Hartmann et al. 1998). For more massive disks, evolution could be more rapid, because the mass contained in the planets is increased (for a fixed disk metallicity): planets could be more massive, meaning a higher ratio $M_p/M_1$ (see eq. [51]), or simply be more numerous. This leads to stronger accretion so that the timescales of $\lesssim 10^6$ yr could be typical. Note, however, that at some point this increase in accretion rate is stopped by the tendency of massive planets to open gaps in the disk. This could drastically reduce their influence on the nebular evolution. Note also that since $\dot{M}$ is very inhomogeneous radially, planet-driven disk evolution must be highly time dependent.

A very useful feature of our analysis presented in § 2.2 is that it can be applied to disks in which the surface density and temperature profiles described here could easily be incorporated into other problems such as the gap formation around massive planets or planet-driven global evolution of the surface density in the disk.

5. CONCLUSIONS

The tidal interaction of a gas disk with a planet embedded in it is important for the question of the orbital evolution of the planet as well as for the fate of the disk itself. In this paper we presented a global description of the evolution of density waves in vertically isothermal disks where two-dimensional fluid equations provide the proper approximation. The disk surface density and sound speed are allowed to vary independently with radius, which distinguishes this analysis from GR01.

Our quantitative results are not very different from those obtained by GR01: for low-mass planets, surface-density perturbations are weakly nonlinear, and they inevitably shock (in the absence of strong linear dissipation) after propagating several disk scale lengths $h_p$ from the planet. This is enough to make density wave damping a nonlocal process. Subsequent nonlinear dissipation of the wave transfers its angular momentum to the disk and is intrinsically asymmetric, which could be important for planet migration. Disks evolve because of this angular momentum flux deposition, and in the absence of other viscous mechanisms, this could be the only driver of their evolution. We have demonstrated that for parameters similar to those of the solar system, the tidal perturbations alone could produce spatially nonuniform and time-dependent evolution with average accretion rates $\dot{M} \sim 10^{-9}-10^{-8} M_\odot$ yr$^{-1}$ (yielding a typical timescale of $\sim 10^6-10^7$ yr). In protoplanetary systems with more favorable conditions, the disk evolution may be even stronger.

The prescription for the angular momentum deposition by planets in disks with arbitrary surface density and temperature profiles described here could easily be incorporated into other problems such as the gap formation around massive planets or planet-driven global evolution of the surface density in the disk.

I am indebted to Jeremy Goodman and Scott Tremaine for illuminating discussions and a lot of thoughtful comments on the manuscript. Financial support provided by the NASA Origins Program under grant NAG 5-8385 and NASA grant NAG 5-10456 is gratefully acknowledged.
APPENDIX

REDUCTION OF THE EQUATION FOR \( R_+ \) TO BURGER’S EQUATION

The equation for \( R_+ = u - 2c/\gamma = 2(c_0 - 2c)/\gamma \) reads

\[
\left[ \partial_\phi + (u - c) \partial_t \right] R_- = - \left( \frac{1}{\Sigma} \partial_t P - \frac{1}{\Sigma_0} \partial_t P_0 - c \partial_t \frac{2c}{\gamma - 1} - cu \frac{\partial_\phi \Sigma}{\Sigma} - u \partial_\phi \frac{2c}{\gamma - 1} \right) \\
+ \frac{2\Omega}{\Delta \Omega} v + \frac{v^2}{\Delta \Omega^2} + cu \frac{\partial_\phi \Sigma}{\Sigma} + \frac{c}{\Delta \Omega^2} \partial_\phi \ln \Sigma + \frac{c}{\Delta \Omega^2} \partial_\phi v .
\]

We find it useful to introduce a new function

\[
\psi = \frac{\gamma + 1}{\gamma - 1} \leq 1 .
\]

Using this definition and equation (5), one can find that up to the second order in \( \psi \),

\[
\frac{\Sigma - \Sigma_0}{\Sigma_0} = \frac{2}{\gamma + 1} \psi + \frac{3 - \gamma}{(\gamma + 1)^2} \psi^2 .
\]

Also, from the conservation of the Riemann invariant \( R_+ \), we have

\[
u = \frac{2c_0 - c}{\gamma - 1} = -2 \frac{c_0}{\gamma + 1} \psi.
\]

Now, we want to relate \( \partial_t P \) to the derivatives of \( \Sigma - \Sigma_0 \). When doing this, we have to remember that the relation in equation (4) holds true only in the reference frame comoving with the fluid. Thus, the Eulerian increment of any quantity \( \Delta \Sigma \) has to be related to its Lagrangian counterpart \( \Delta L \) by \( \Delta \Sigma = \Delta L - dU \), where \( dL \) is a Lagrangian displacement. Obviously, \( \vec{v} \cdot d = - (\Sigma - \Sigma_0)/\Sigma_0 \), and since the disk is assumed to be thin and the spiral pattern is tightly wound, \( \partial_t d \approx \vec{v} \cdot d = -(\Sigma - \Sigma_0)/\Sigma_0 \). Expanding the Lagrangian increment of \( P \) up to the second order in \( \Sigma - \Sigma_0 \), we obtain

\[
\partial_t P = \partial_t P_0 + \partial_t \left[ c^2 (\Sigma - \Sigma_0) \right] + \frac{\gamma - 1}{\Sigma_0} \partial_\phi (\Sigma - \Sigma_0) + P \frac{\Sigma - \Sigma_0}{\Sigma_0} \partial_\phi \ln \Sigma + \frac{c}{\Delta \Omega^2} \partial_\phi v - \frac{2c_0^2}{\gamma + 1} \psi \partial_\phi \ln c_0 - \frac{2c_0^2}{\gamma + 1} \psi \partial_\phi \ln \Sigma_0 .
\]

For a polytropic equation of state with a fixed polytropic constant (entropy in the disk is independent of \( r \)), the last term in equation (A5) is absent, and the derivative of the pressure can be taken without worrying about the Lagrangian displacement.

Since the tidal perturbation is assumed to be weak and tightly wrapped, the most important nonlinear terms responsible for the wave evolution are proportional to \( \psi \partial_\phi \psi \), while terms such as \( \psi^2 \) can be disregarded. Substituting equations (A3), (A4), and (A5) into equation (A1), we get after lengthy but straightforward calculation that

\[
\partial_\phi \psi = c_0 (1 + \psi) \partial_\phi \psi = \frac{\gamma - 1}{4c_0} \left( 2 \frac{\Omega \Delta \psi}{\Delta \Omega^2} + \frac{c}{\Delta \Omega^2} \partial_\phi \ln \Sigma + \frac{c}{\Delta \Omega^2} \partial_\phi v - \frac{2c_0^2}{\gamma + 1} \psi \partial_\phi \ln c_0 - \frac{2c_0^2}{\gamma + 1} \psi \partial_\phi \ln \Sigma_0 \right) .
\]

We need to expand \( \Sigma - \Sigma_0 \) and \( \partial_t P \) up to the second order in the perturbation to make sure that all the terms proportional to \( \partial_\phi \psi \) and \( \partial_\phi \partial_\phi \psi \) cancel out in the right-hand side of equation (A1).

To an adequate approximation, we can express \( \partial_\phi v \) from equation (26) as

\[
\partial_\phi v \approx - \frac{2B}{\Delta \Omega(r)} u .
\]

Here we neglect the nonlinear terms \( u \partial_\phi v \) and \( v/\Delta \Omega^2 \); the term with \( \partial_t \Sigma \) only slightly changes the propagation velocity of the wake and thus is disregarded too.

In equation (26), the terms that are in phase with \( \nu \) or \( \psi \) are important for the amplitude of the perturbation evolution, while those out of phase only slightly affect the characteristic velocity but not the wave amplitude (as we noted before). The derivative with respect to \( \phi \) changes phase by \( \pi/2 \); thus, terms with \( \partial_\phi \) in equation (26) can be considered separately from the others. We can integrate them over \( \phi \) to obtain

\[
\psi \approx -2 \frac{c_0^2}{\Delta \Omega(r)} \frac{1}{\gamma + 1} \nu .
\]

This result, combined with equation (A4), confirms that \( \nu \ll \nu \) (see § 2.2).

Using the relations in equations (A4), (A7), and (A8), we finally get that

\[
\partial_\phi \psi - c_0 (1 + \psi) \partial_\phi \psi = \psi \left( \frac{c_0}{\Delta \Omega} \left[ \frac{1}{r} + \frac{1}{2} \frac{\partial \ln \Sigma_0}{\partial \ln r} + \frac{3}{2} \frac{\partial \ln c_0}{\partial \ln r} - \frac{A(r)}{\Delta \Omega} \right] + O(\psi^2) \right) .
\]

The local approximation of GR01 can be retrieved now by assuming \( \Sigma_0 = \text{const} \) and \( c_0 = \text{const} \) and expanding \( \Delta \Omega \) to first order in terms of \( r - r_p \).
All effects of the nonlinear wake evolution are embodied in the term $\psi\partial_\psi\psi$ in the left-hand side of equation (A9). We now reduce this equation to the conventional Burger’s equation following the approach outlined in GR01. First, we make a change of independent variables from $\phi$, $\xi$ to $\phi'$, $\eta_1$, given by the relations
\[ \int \frac{d\xi}{c(\xi)} = \eta_1 - \phi', \quad \phi' = \phi'. \] (A10)
Here $\phi'$ is the azimuthal distance along the wake while $\eta_1$ represents the displacement from the wake center in the $\phi$-direction. Since the wake is narrow, $\eta \ll \phi$ and we can consider $c_0$ to be a function of $\phi'$ only: $c_0(\xi) = c_0(\phi)$. This transforms equation (A9) into (we drop the prime from $\phi'$)
\[ \partial_\phi \psi - \psi \partial_\eta \psi = -\psi \frac{1}{g(\phi)} \partial_\psi g(\phi), \] (A11)
where the function $g(\phi)$ is defined by
\[ \frac{1}{g(\phi)} \partial_\phi g(\phi) = -\frac{c_0(\phi)}{A(\phi)} \left[ \frac{1}{5} \frac{1}{3} \frac{3}{2} \partial_\phi \ln c_0 + \frac{3}{2} \partial_\phi \ln r + \frac{2}{A(\phi)} \partial_\phi \ln r \right]. \] (A12)
Introducing a new function $\chi(\phi) = g(\phi)\psi$, we rewrite equation (A11) as
\[ \partial_\phi \chi = \chi \frac{1}{g(\phi)} \partial_\eta \chi. \] (A13)
Changing from $\phi$ to a new variable $t_1$, given by
\[ t_1 = \int_0^\phi \frac{d\phi'}{g(\phi')}, \] (A14)
we arrive at Burger’s equation
\[ \partial_\eta \chi - \chi \partial_\eta \chi = 0. \]
Both $t_1$ and $\eta_1$ are dimensionless, and we find it useful to rescale them to $t$ and $\eta$ in the following way:
\[ \eta = \frac{r_p}{l_p} \eta_1, \quad t = \frac{r_p}{l_p} t_1, \] (A15)
where $l_p = c_0/[2A(\rho)]$ is a Mach 1 length at the position of the planet (in the limit $r \to r_p$ this rescaling makes our $\eta$ identical with the $\eta$ used in GR01). Thus, we obtain equation (30).

Finally, we calculate the behavior of the function $g$. From the definitions in equations (24) and (A10), one finds that $d\phi = -(\Delta\Omega/c_0) \, dr$. Using this and the definition of Oort’s parameter $A$, one can easily integrate equation (A12) to find
\[ g = C_1 \left( \frac{\Sigma_0}{\Delta\Omega} \right)^{1/2}, \quad C_1 = \text{const}. \] (A16)
The arbitrary constant $C_1$ should be chosen in such a way that in the limit $r \to r_p$, our definitions of $\chi$ and $t$ reduce to the analogous expressions of GR01 (given by eq. [35]). One achieves this by taking
\[ C_1 = \frac{2^{1/4}}{r_p c_p \Sigma_0^{1/2}} \] (A17)
(see eqs. [34] and [35]).

REFERENCES

Balbus, S. A., & Hawley, J. F. 1998, Rev. Mod. Phys., 70, 1
Balbus, S. A., Hawley, J. F., & Stone, J. M. 1996, ApJ, 467, 76
Cassen, P., & Woolfum, D. S. 1996, ApJ, 472, 789
Chiang, E. I., & Goldreich, P. 1997, ApJ, 490, 368
Drummond, R. 1996, MNRAS, 282, 982
Goldreich, P., & Nicholson, P. D. 1989, ApJ, 342, 1075
Goldreich, P., & Tremaine, S. 1979, ApJ, 233, 857
———. 1980, ApJ, 241, 425 (GT80)
Goodman, J., & Rafikov, R. R. 2001, ApJ, 552, 793 (GR01)
Hartmann, L., Calvet, N., Guibbing, E., & D’Alesio, P. 1998, ApJ, 495, 385
Hawley, J. F., & Stone, J. M. 1998, ApJ, 501, 758
Hayashi, C. 1981, Prog. Theor. Phys. Suppl., 70, 35
Henning, T., & Stognienko, R. 1996, A&A, 311, 291
Jin, L. 1996, ApJ, 457, 798
Landel, L. D., & Lifschitz, E. M. 1959, Fluid Mechanics (Oxford: Pergamon)
Larson, R. B. 1981, In The Formation and Evolution of Planetary Systems, ed. H. A. Weaver & L. Danly (Cambridge: Cambridge Univ. Press), 31
Larson, R. B. 1990, MNRAS, 243, 588
Larson, R. B. 1993, in Protostars and Planets III, ed. E. H. Levy & J. I. Lunine (Tucson: Univ. Arizona Press), 749
Lin, D. N. C., Papaloizou, J. C. B., & Savonije, G. J. 1990, ApJ, 364, 326
Lubow, S. H., & Ogilvie, G. I. 1998, ApJ, 504, 983
Ogilvie, G. I., & Lubow, S. H. 1999, ApJ, 515, 767
Pringle, J. E. 1981, ARA&A, 19, 137
Rafikov, R. R. 2001, ApJ, 572, in press (astro-ph/0110540)
Ryu, D., & Goodman, J. 1992, ApJ, 388, 438
Shakura, N. I., & Sunyaev, R. A. 1973, A&A, 24, 337
Spruit, H. C. 1987, A&A, 184, 173
Takeuchi, T., Miyama, S. M., & Lin, D. N. C. 1996, ApJ, 460, 832
Tanaka, H., & Ward, W. R. 2000, BAAS, 32, 55.04
Velevich, E. P. 1959, Sov. Phys. JETP, 36, 995
Ward, W. R. 1986, Icarus, 67, 164
———. 1997, Icarus, 126, 261
Ward, W. R., & Hourigan, K. 1989, ApJ, 347, 490
Whitham, G. B. 1974, Linear and Nonlinear Waves (New York: Wiley)