Spiral arms in scattered light images of protoplanetary discs: are they the signposts of planets?
Juhász, A.; Benisty, M.; Pohl, A.; Dullemond, C.P.; Dominik, C.; Paardekooper, S.J.

Published in:
Monthly Notices of the Royal Astronomical Society

DOI:
10.1093/mnras/stv1045

Citation for published version (APA):
Spiral arms in scattered light images of protoplanetary discs: are they the signposts of planets?

A. Juhász,†, M. Benisty, A. Pohl, C. P. Dullemond, C. Dominik and S.-J. Paardekooper

1 Institute of Astronomy, Madingley Road, Cambridge CB3 0HA, UK
2 Leiden Observatory, Leiden University, PO Box 9513, NL-2300 RA Leiden, the Netherlands
3 University Grenoble Alpes, IPAG, F-38000 Grenoble, France, CNRS, IPAG, F-38000 Grenoble, France
4 Institute for Theoretical Astrophysics, Heidelberg University, Albert-Ueberle-Strasse 2, D-69120 Heidelberg, Germany
5 Max-Planck-Institute for Astronomy, Königstuhl 17, D-69117 Heidelberg, Germany
6 Anton Pannekoek Institute for Astronomy, University of Amsterdam, Postbus 94249, NL-1090 GE Amsterdam, the Netherlands
7 Astronomy Unit, School of Physics and Astronomy, Queen Mary, University of London, Mile End Road, London E1 4NS, UK
8 DAMTP, University of Cambridge, Wilberforce Road, Cambridge CB3 0WA, UK

Accepted 2015 May 6. Received 2015 May 5; in original form 2014 October 21

ABSTRACT

One of the striking discoveries of protoplanetary disc research in recent years are the spiral arms seen in several transitional discs in polarized scattered light. An interesting interpretation of the observed spiral features is that they are density waves launched by one or more embedded (proto)planets in the disc. In this paper, we investigate whether planets can be held responsible for the excitation mechanism of the observed spirals. We use locally isothermal hydrodynamic simulations as well as analytic formulae to model the spiral waves launched by planets. Then H-band scattered light images are calculated using a 3D continuum radiative transfer code to study the effect of surface density and pressure scaleheight perturbation on the detectability of the spirals. We find that a relative change of \( \sim 3.5 \) in the surface density (\( \delta \Sigma / \Sigma \)) is required for the spirals to be detected with current telescopes in the near-infrared for sources at the distance of typical star-forming regions (140 pc). This value is a factor of 8 higher than what is seen in hydrodynamic simulations. We also find that a relative change of only 0.2 in pressure scaleheight is sufficient to create detectable signatures under the same conditions. Therefore, we suggest that the spiral arms observed to date in protoplanetary discs are the results of changes in the vertical structure of the disc (e.g. pressure scaleheight perturbation) instead of surface density perturbations.

Key words: scattering – planet–disc interactions – protoplanetary discs – circumstellar matter – stars: formation – infrared: stars.

1 INTRODUCTION

Today, more than a thousand extrasolar planets have been detected around main-sequence stars and these planetary systems show a great diversity in their architectures (Batalha et al. 2013). However, none of the claimed detections of planetary candidates in their native circumstellar disc (e.g. Huelamo et al. 2011; Kraus & Ireland 2012; Quanz et al. 2013) have been confirmed so far. The direct observation of a (proto)planet in a disc is difficult due to the large brightness contrast with the disc/star system that outshines it. None the less, a massive planet could be revealed by indirect signatures of its interaction with the disc. Planets more massive than Jupiter gravitationally interact with the surrounding gas, and carve a gap or a cavity in the disc (Crida, Morbidelli & Masset 2006). As a result of planet–disc interactions, a number of non-axisymmetric features can appear, such as spiral arms, warps (Facchini, Ricci & Lodato 2014) or an overall spatial discrepancy between small and large dust grains (Pinilla et al. 2012; de Juan Ovelar et al. 2013).

Hydrodynamic processes that do not involve planets can also lead to vortices or spirals (see Turner et al. 2014 for a review). Asymmetric features can result from gravitational instability (Lodato & Rice 2004; Boley & Durisen 2006; Rice et al. 2006; Baruteau, Meru & Paardekooper 2011) or tidal interaction with an external companion (Papaloizou, Nelson & Masset 2001). Recent numerical simulations also show that non-ideal magnetohydrodynamics can spontaneously lead to such non-axisymmetric structures (Kunz & Lesur 2013).

Asymmetric features have been recently observed in transition discs, which are thought to be at an advanced evolutionary stage.

† E-mail: juhasz@ast.cam.ac.uk

© 2015 The Authors
Published by Oxford University Press on behalf of the Royal Astronomical Society
They show a dip in their infrared (IR) spectral energy distribution (SED), indicative of a dust-depleted cavity (see Espaillat et al. 2014 for a review), that can sometimes be imaged (Andrews et al. 2011). High angular resolution images of their outer regions have indeed shown a variety of asymmetries in the submillimetre continuum tracing cold material (e.g. Casassus et al. 2013; van der Marel et al. 2013), and in the near-infrared (NIR) through scattered light that traces small hot dust in the disc surface layers (Fukagawa et al. 2006; Grady et al. 2013). For example, in the case of SAO206462, two spiral arms can be traced up from 28 to 140 au in NIR scattered light with a polarized intensity about 30 per cent larger than the background disc (Muto et al. 2012; Garufi et al. 2013). In HD 142527, six spiral arms can be traced in scattered light which start from the edge of its 142 au-radius gap (Avenhaus et al. 2014b). Submillimetre Atacama Large Millimetre/sub-millimetre Array (ALMA) observations indicate the presence of three CO arms extending up to 670 au, but only one is a radio counterpart of the IR arms (Christiaens et al. 2014). The only previous detection of spiral features in such wavelength range was in AB Aur (Piétu, Guilloteau & Dutrey 2005) but the spirals were found to be in counter rotation with the disc, indicating a likely origin in the late envelope infall. What is causing the gap or cavity in transition discs is still debated, but at least in some cases, it can be due to the dynamical interaction of one or multiple planets with the disc (Dodson-Robinson & Salyk 2011; Zhu et al. 2011). On the one hand, photoevaporation models fail to reproduce the properties of transition discs with high accretion (Owen & Clarke 2012) or NIR excess (such as HD 100546; Tatulli et al. 2011), while on the other hand, grain growth models could reproduce the IR dip seen in the SED, but not the millimetre images (Birnstiel, Andrews & Ercolano 2012). Thus, the presence of a planet has been often invoked to explain the observed asymmetries (Muto et al. 2012; Garufi et al. 2013; van der Marel et al. 2013).

While the presence of a planetary mass companion is a very intriguing explanation for the observed asymmetries, especially for spiral arms, no quantitative comparison between observation and theory has been done so far. Observational studies compared the morphology of spiral arms to those seen in hydrodynamic simulations, e.g. by fitting the spiral wake (Muto et al. 2012). However, it has not yet been investigated whether the observed amplitude of the spirals agrees with that in theoretical calculations. In this paper, we study the detectability of planet-induced spiral density waves in scattered light observations. Using a combination of 2D hydrodynamic and 3D radiative transfer codes, we test whether the spirals observed so far can be explained by planet-induced spiral density waves and we also make predictions for future observations with next generation instruments.

2 MODEL SETUPS

2.1 Hydrodynamic simulations

We use the 2D hydrodynamic code FARGO (Masset 2000) to study the shape and amplitude of the spiral density waves launched by an embedded planet. The input parameters of the simulations are summarized in Table 1. For all simulations with planet-to-star mass ratio of $M_{\text{pl}} \leq 10^{-7}M_{\ast}$, we use the same spatial grid. However, for the most massive case ($M_{\text{pl}} = 10^{-2}M_{\ast}$) we extend the outer boundary of the computational grid in order to minimize the effect of the outer boundary conditions on the structure of the disc around the gap (Kley & Dirksen 2006). The surface density distribution is assumed to be $\Sigma = \Sigma_0 (R/R_{\text{in}})^{-0.5}$ where $\Sigma_0 = 0.003$. Assuming $R_{\text{pl}} = 25$ au, the gas mass in our setup between 5.0 and 100 au is $0.01 M_{\ast}$.

A single planet is placed in the disc on a fixed circular orbit with a radius of unity in dimensionless units. In the reference model, the mass of the planet is $10^{-3}M_{\ast}$. The accretion on to the planet is switched off and the simulations are run for 1000 planetary orbits. We vary the mass of the planet and the disc aspect ratio to study their effect on the disc structure and on the density perturbation induced by the planet.

The simulations are locally isothermal not taking into account heating and cooling of the disc, which is a reasonable assumption if the disc is optically thin to its own radiation resulting in efficient cooling. Apart from surface density perturbations, temperature perturbations are also expected along the spirals due to the presence of shocks along the spiral. We study the effect of pressure scaleheight perturbations with analytic models in Section 4.2.2, but the full treatment of heating and cooling in the hydrodynamic simulations is the topic of a forthcoming paper.

2.2 Radiative transfer

We use the 3D radiative transfer code radmc-3d$^1$ to calculate NIR scattered light images. In the radiative transfer calculations, we use a spherical mesh $(r, \theta, \phi)$, where the radial and azimuthal grid are taken to be the same as in the hydrodynamical simulations. In the poloidal direction, we use $N_{\phi} = 250$ grid cells in total. To ensure that we resolve the upper layers of the disc, which is responsible for the observed scattered light, we place $N_{\theta} = \{10, 100, 30, 100, 10\}$ points to the $[0, \pi/2 - \theta_1], [\pi/2 - \theta_1, \pi/2 - \theta_2], [\pi/2 - \theta_2, \pi/2 + \theta_2], [\pi/2 + \theta_2, \pi/2 + \theta_1], [\pi/2 + \theta_1, \pi]$ intervals, respectively, where $\theta_1 = 9H_{\phi}/R$ and $\theta_2 = 3H_{\phi}/R$, with $H_{\phi}$ being the pressure scaleheight in the disc.

The density structure of the disc in the radiative transfer calculation is given as

$$\rho(R, z, \phi) = \frac{\Sigma(R, \phi)}{\sqrt{2\pi}H_{\phi}(R)} \exp \left( -\frac{z^2}{2H_{\phi}(R)^2} \right),$$

where $\Sigma(r, \phi)$ is the surface mass density, $H_{\phi}(R)$ is the pressure scaleheight, $R = \rho \sin \theta$ and $z = \rho \cos \theta$. The radial variation of the aspect ratio of the disc is described by a power law,

$$\frac{H_{\phi}(R)}{R} = \frac{H_{\phi}(R_{\text{pl}})}{R_{\text{pl}}} R^{2/3},$$

$^1$ http://www.ita.uni-heidelberg.de/dullemond/software/radmc-3d/
where $R_p$ is the radius of the planetary orbit and $\zeta$ is the flaring index. The aspect ratio and the flaring index was taken to be the same as in the hydrodynamic simulations.

For the radiation source, we assume a central star with parameters representative for that of a Herbig Ae star, $T_{\text{eff}} = 9500$ K, $R_* = 2.5 R_\odot$, $M_* = 2.0 M_\odot$. For simplicity, the stellar radiation field is described by blackbody emission.

The dust in the disc consists of 0.1 μm sized silicate grains, whose scattering matrix is calculated from the optical constants of astronomical silicates (Weingartner & Draine 2001) using Mie-theory. In reality dust grains are likely to have a wide size distribution, which we neglected here. Dust grains which are small compared to the wavelength scatter photons uniformly, while large grains have strongly forward peaking scattering phase function. This means that for discs viewed at low-inclination angles large grains scatter stellar photons mostly into the disc, and only a small fraction towards the observer, making the disc faint in scattered light (see Mulders et al. 2013). Therefore, as long as dust properties are uniform in the disc or at least there is no steep gradient in the dust size distribution, the main effect of a wider grain size distribution is to scale the total scattered flux of the disc. To maximize the computational efficiency, i.e. to have the largest number of photons scattered towards the observer also for low disc inclinations, we choose to use a single grain size of 0.1 μm.

We also assumed that the dust-to-gas ratio is uniform in the disc, thus we can model the dust distribution by simply scaling the gas distribution we calculate with FARGO. While the dust-to-gas ratio is known to get enhanced in the spirals excited by gravitational instability (see e.g. Rice et al. 2004), planet-induced spirals are unlikely to trap dust particles. Spiral waves launched by a planet are corotating with the planet, i.e. they will move with respect to the background gas flow. Dust particles can therefore only be trapped in the spirals if the dust accumulation time-scale is shorter than the time-scale for the spirals to pass by. As it was shown by Birnstiel, Dullemond & Pinilla (2013), the dust accumulation time-scale is hundred times longer than the local orbital period even in the most optimistic case, while the time-scale of the spiral pass by is only few orbital time-scale of the planet even very close to the planet. Thus particle trapping is only possible for structures which are either close to co-rotation with the local gas-flow, like e.g. spiral arms in gravitationally unstable discs, or for axisymmetric structures, like the pressure bump at the outer edge of the gap. Moreover, dust trapping works for only large grains which are at least partially decoupled from the gas (see e.g. Lyra et al. 2009; Pinilla et al. 2012; Birnstiel et al. 2013), while the micron- and submicron-sized dust particles, responsible for NIR scattered light, are rather well coupled to the gas. Thus, we think that the assumption of uniform dust-to-gas ratio in the disc does not affect the results of this paper.

The radiative transfer calculations contain two steps. First, the temperature structure of the dust is determined in a thermal Monte Carlo simulations then images are calculated including the full treatment of polarization. For the thermal Monte Carlo simulations, we use $1.6 \times 10^6$ photon packages while the $H$-band images are calculated using $6 \times 10^7$ photon packages. The square images are calculated at 1.65 μm, at an inclination angle of $10^\circ$ with 700 pixels along the edge, covering a linear scale of 250 au. This corresponds to a spatial resolution of 0.357 au pixel$^{-1}$. To convert linear scales to angular coordinates, we assume the source to be at 140 pc. RADMC-3D calculates images in the full Stokes vector ($I, Q, U, V$) from which the polarized intensity can be calculated as $PI = \sqrt{Q^2 + U^2}$.

These images represent the true sky brightness distribution, i.e. virtually at infinite spatial resolution. In reality, the observed images are the convolution of the true sky brightness distribution with the point spread function (PSF) of the telescope. To simulate the observations, the images need to be convolved with the telescope PSF. We assume the simplest PSF model, a Gaussian, which is a good representation of the Airy-disc, the core of a diffraction limited PSF. First, we investigate whether or not the already observed spirals in transitional discs can be explained with planet-induced density waves. For this purpose, we convolve the images with a PSF whose full width at half-maximum (FWHM) is taken to be 0.06 arcsec, representative for the observations with HI-CIAO/SUBARU presented in Muto et al. (2012). Then we make predictions for the observability of spirals in future observations with next generation instruments [e.g. SPHERE/Very Large Telescope (VLT) and telescopes [European Extremely Large Telescope (E-ELT)] by changing the size of the PSF the images are convolved with.

3 RESULTS FROM HYDRODYNAMIC SIMULATIONS

3.1 Morphology of the spirals

Low-mass planets which do not open gaps are known to drive a single one-armed spiral forming as a constructive resonance between waves launched at Lindblad-resonances both inside and outside of the planetary orbit (Ogilvie & Lubow 2002). However, for massive planets which open gaps, the wave propagation is highly non-linear the morphology of the spirals might be different. We studied the effect of the planet mass and disc aspect ratio on the morphology of the spirals.

To study the effect of the planet mass on the strength of the spiral perturbation, we run simulations with various planet masses ($M_p/M_\ast = \{6.25 \times 10^{-3}, 10^{-3}, 10^{-2}\}$). In Fig. 1, we show the surface density distribution of the disc for three different planet masses after 1000 planetary orbits. The spiral density waves are clearly visible in all cases. However, the structure of the spirals in the outer disc changes with the planetary mass. For a low-mass planet ($M_p/M_\ast = 6.25 \times 10^{-3}$), there is a well-defined, single spiral arm in the outer disc (see Fig. 1a). For a more massive planet ($M_p/M_\ast = 10^{-2}$), that opens a gap, there are two spiral arms. The two arms of the spiral are, however, not completely symmetric. The one that goes through the planet position has a smaller radial width and higher amplitude than the secondary spiral that is shifted in azimuth compared to the planet position by about 180° (see Fig. 1b). For the most massive planet in our simulation ($M_p/M_\ast = 10^{-2}$), the gap becomes eccentric and several spiral arms become visible (see Fig. 1c). The strength and sharpness of the secondary arm change smoothly with the planet mass.

Variations in the aspect ratio of the disc, i.e. the ratio of the pressure scaleheight and the radius, changes the temperature and the sound speed in the disc, which in turn affects the propagation of sound waves through the disc. The angle between the radial direction and the tangent of the spiral, the complement of the pitch angle, is inversely proportional to the sound speed and the pressure scaleheight (Rafikov 2002). Thus increasing aspect ratio increases the pitch angle, making the spiral pattern more open (see Fig. 1 top row and bottom row). The strength of the spirals, which are offset in azimuth with respect to the position of the planet, decreases with increasing aspect ratio. As can be seen in Fig. 1b, the secondary spiral is clearly visible for an aspect ratio of $H_\parallel/R = 0.05$ in case of a planet mass of $10^{-3}M_\ast$. If the aspect ratio is increased to $H_\parallel/R = 0.1$, the secondary spiral weakens to a barely recognizable level. By increasing the aspect ratio from 0.05 to 0.1 for a planet mass of...
The perturbation induced by low-mass planets, for which $M_{pl}/M^* \ll (H_p/R)^3$, can be calculated by decomposing the potential of the planet into Fourier harmonics azimuthally and solving the resulting ordinary differential equations for the linear perturbations numerically (e.g. Korycansky & Pollack 1993). The resulting ordinary differential equations for the linear perturbation in radial velocity. Since these quadratic terms have twice the azimuthal wavenumber of the corresponding linear terms, the second-order solution will only involve those with values of $m$ odd. A second effect that sets in as $M_{pl}/M^*$ approaches $(H_p/R)^3$ is that the width of the horseshoe region becomes comparable to the scaleheight of the disc (Paardekooper & Papaloizou 2009). This means that the highest order resonances, located approximately one scaleheight away from the planet, can no longer contribute to the wake. In the limit of $M_{pl}/M^* \gg (H_p/R)^3$, the $m = 2$ becomes dominant, resulting again in a two-armed spiral. We leave a detailed description of this mechanism to a forthcoming paper (Paardekooper et al., in preparation). The shifted spiral arm can be clearly identified for example in Fig. 1(b).

This second spiral can only be comparable in magnitude to the original spiral wave if $M_{pl}/M^* \sim (H_p/R)^3$. For smaller planets, only a one-armed spiral will be visible. Note that this criterion is similar to the thermal criterion for gap opening (e.g. Crida et al. 2006). We therefore expect only gap-opening planets to exhibit two spiral arms, which is confirmed by Fig. 1. For $M_{pl}/M^* \gg (H_p/R)^3$, additional effects kick in. For example, it is well known that for high enough radial location where constructive interference can occur; hence a one-armed spiral.

However, when $M_{pl}/M^*$ approaches $(H_p/R)^3$, linear theory breaks down (Korycansky & Papaloizou 1996). One can probe this regime by going beyond linear theory, and include second-order perturbations (Artymowicz & Lubow 1992). These perturbations do not couple to the planet potential directly, since the potential is completely specified at linear order, but instead to quadratic terms in for example the linear perturbation in radial velocity. Since these quadratic terms have twice the azimuthal wavenumber of the corresponding linear terms, the second-order solution will only involve even numbers of $m$. This selection rule allows for interference at two values of $\varphi$ at any radial location: one at the original location of the one-armed spiral, and one shifted by $\pi$. A second effect that sets in as $M_{pl}/M^*$ approaches $(H_p/R)^3$ is that the width of the horseshoe region becomes comparable to the scaleheight of the disc (Paardekooper & Papaloizou 2009). This means that the highest order resonances, located approximately one scaleheight away from the planet, can no longer contribute to the wake. In the limit of $M_{pl}/M^* \gg (H_p/R)^3$, the $m = 2$ becomes dominant, resulting again in a two-armed spiral. We leave a detailed description of this mechanism to a forthcoming paper (Paardekooper et al., in preparation). The shifted spiral arm can be clearly identified for example in Fig. 1(b).

This second spiral can only be comparable in magnitude to the original spiral wave if $M_{pl}/M^* \sim (H_p/R)^3$. For smaller planets, only a one-armed spiral will be visible. Note that this criterion is similar to the thermal criterion for gap opening (e.g. Crida et al. 2006). We therefore expect only gap-opening planets to exhibit two spiral arms, which is confirmed by Fig. 1. For $M_{pl}/M^* \gg (H_p/R)^3$, additional effects kick in. For example, it is well known that for high enough radial location where constructive interference can occur; hence a one-armed spiral.
Spiral arms in protoplanetary discs

Figure 2. Simulated H-band polarized intensity images with the surface density given by hydrodynamic simulations with a planet-to-star mass ratio of 0.001 for different values of the disc aspect ratio ($H_p/R$) and flare index ($\zeta$). The images are normalized to the highest intensity outside of a 0.18 arcsec radius (25.2 au at 140 pc distance) from the central star. The top row shows images at original resolution as calculated by RADMC-3D, while the bottom row shows the images convolved with a circular Gaussian beam with an FWHM of 0.06 arcsec. While spiral arms are clearly visible in the original, quasi-infinite resolution (top row), no spiral arm is visible after convolving the images with a 0.06 arcsec PSF (bottom row) in any of the images.

planet masses the outer disc can become eccentric (Papaloizou et al. 2001; Kley & Dirksen 2006). The eccentricity of the disc comes with its own surface density perturbations that interfere with the two spirals (see Fig. 1c).

3.3 Images from hydrodynamic simulations

We use the surface density distribution ($\Sigma(R, \phi)$) calculated by FARGO to describe the disc structure in equations (1) and (2). The calculated scattered light images in H-band polarized intensity for a planet-to-star mass ratio or 0.001 are presented in Figs 2(a), (b), (c) and (d). Since we are interested in the observability of the spirals in the outer disc, we normalize the images to the highest surface brightness outside of the planetary orbit (25 au or ~0.179 arcsec at 140 pc distance) and we set the dynamic range of the images to 20 similar to the images of HD 135344B presented by Muto et al. (2012).

The spiral arms are clearly visible in the images calculated by RADMC-3D (see Figs 2a, b, c and d). The contrast between the spiral and the background disc is higher for lower aspect ratio (see Figs 2a and b), than for a larger one (see Figs 2c and d), as expected on the basis of the surface density perturbation. Changing the flaring index affects the brightness of the disc in the outer regions, but does not affect the visibility of the spirals. The increase of the pressure scaleheight increases the pitch angle, i.e. the tightness, of the spirals making them more open. However, the increase of the pressure scaleheight also decreases the amplitude of the spirals making them even harder to observe.

As can be seen in Figs 2(e), (f), (g) and (h) after convolving the images with a 0.06 arcsec PSF, representative for the reported observations of transitional discs with 8–10 m-class telescope in the H band, the spiral features are not visible anymore. This is true even for the case of a very massive planetary companion with a planet-to-star mass ratio of 0.01. Therefore, it seems to be very unlikely that the spiral arms observed to date in transitional discs in scattered light are caused by pure surface density perturbations.

4 RESULTS FROM ANALYTIC MODELS

4.1 Analytic model for spiral perturbation

Since the spirals are not visible in the PSF-convolved images, we calculated based on the hydrodynamic simulations, we use the analytic models to test the contrast/amplitude requirement for the observability of the spirals. Our model is based on the WKB solution for propagation of planet-induced density waves in power-law discs given by Rafikov (2002). We assume that any perturbed variable in the disc can be expressed in the form

$$X(R, \phi) = X(R) \left(1 + \frac{\delta X(R, \phi)}{X(R)}\right).$$

The first term on the right-hand side describes the symmetric unperturbed part, that is assumed to have the form of a power-law $X(r) = X_0(R/R_0)^n$ with $R_0$ being the orbital radius of the planet. $\delta X(R, \phi)$ contains the non-axisymmetric perturbation caused by the planet. The spiral perturbation function $\delta X(R, \phi)$ is described by a Gaussian in the radial direction around the centre of the spiral wake whose amplitude decreases with the radial distance from the star

$$\delta X(R, \phi) = A \left(\frac{R}{R_0}\right)^{\nu \mu} \exp\left(-\frac{(R - R_0(\phi))^2}{\sigma_{\text{sp}}^2}\right).$$

The constant $A$ sets the amplitude of the perturbation at the position of the planet, $\sigma_{\text{sp}}$ is the radial width of the spiral and $R_0(\phi)$ describes
the spiral wake. We calculate $R_0(\phi)$ from the wake equation of Rafikov (2002) used also by Muto et al. (2012),

$$\phi(R) = \phi_{pl} - \frac{\text{sgn}(R - R_{pl})}{h_{pl}} \left( \frac{R}{R_{pl}} \right)^{1+\beta} \left\{ \frac{1}{1+\beta} - \frac{1}{1-\alpha + \beta} \left( \frac{R}{R_{pl}} \right)^{-\alpha} \right\}$$

$$+ \frac{\text{sgn}(R - R_{pl})}{h_{pl}} \left( \frac{1}{1+\beta} - \frac{1}{1-\alpha + \beta} \right),$$

(5)

where $R_{pl}$ is the radius of the planetary orbit, $\alpha$ is the power exponent of the rotation angular frequency ($\Omega(r) \propto r^{-\alpha}$), $\beta$ is the power exponent of the radial distribution of the sound speed ($c_s \propto r^{-\beta}$), $h_{pl} = h_0(R_{pl})/R_{pl}$ is the aspect ratio of the disc at $R_{pl}$ and $\phi_{pl}$ is the azimuthal coordinate of the planet.

Our goal is to derive a lower limit for the relative perturbation amplitude ($\delta X(r, \phi)/X(r)$) required to detect the spirals in scattered light observations and compare that to the numerical simulations. To keep our analytic model as realistic as possible, we fit the 2D surface density distribution of the disc in our reference hydrodynamic simulation with equations (3)–(5). This way we can ‘calibrate’ the shape of the perturbation function.

Equations (4) and (5) describe the 2D disc structure with 10 parameters. We fix $\alpha$, $\beta$, $R_{pl}$, $\phi_{pl}$ to the values we used in the hydrodynamic simulations and we fit $h_0$, $\Sigma_0$, $p$, $A$, $q$ and $\sigma_{q0}$. We use an eyeball fit to the surface density distribution and do not use sophisticated optimization algorithms. The fitted values to the reference hydrodynamic model are summarized in Table 2. While value of $h_0$ was also a known (input) parameter in the hydrodynamic simulation, we need to increase it by 33 per cent in order to match the spiral wake in the hydrodynamic simulation far away from the planet in the outer disc. Finally, we reduce the surface density by a factor of $10^{-4}$ between 15 and 30 au to mimic the presence of a gap.

In Figs 3(a)–(c), we compare the fitted analytic surface density distribution to the hydrodynamical simulations. The analytic model reproduces the structure of the disc, the shape and amplitude of the spirals reasonably well in the outer disc. In the inner parts, the analytic model overestimates the surface density by a factor of 2 (see Fig. 3c). The higher surface density in the analytic models may increase the shadowing of the outer disc (Juhász et al. 2007). Since we are interested only in the contrast of the spiral arms with respect to the background disc, the increased shadowing by the inner disc would not affect our results as long as the shadowing is axisymmetric. We have verified that the spiral perturbation in the inner disc, as described in the analytic model, is too weak and too tightly wound to cause any non-axisymmetric shadowing.

We use the above described formalism for the spiral perturbation to calculate a series of models with analytic surface density perturbation. In reality not only density but also thermal perturbation is expected along the spiral wake. More than a few scaleheights away from the planet shocks form along the spiral wake which will introduce also thermal perturbation along the spirals (see e.g. Goodman & Rafikov 2001; Rafikov 2002). Since the pressure scaleheight of the disc depends on the local temperature, the temperature perturbation in turn will cause changes in the vertical density structure. We study the effect of perturbation in the vertical density structure on the scattered light images by applying the above described analytic formalism for the pressure scaleheight and run a series of models with various perturbation amplitude.

### 4.2 Images with analytic perturbations

#### 4.2.1 Surface density perturbations

We take the analytic model for the spiral perturbation described in Section 4.1 and whose free parameters are fitted to match the shape of the spirals in the hydrodynamic simulations, and change the amplitude of the spiral perturbation by changing the value of $A$ in equation (4). The value of $A$ sets the relative change of the perturbed variable ($\delta X/X$) at the position of the planet. Massive
planets can open a gap where the perturbation may not be visible due to the low surface density. Therefore, we also calculate the relative change of the perturbed variable at 270° azimuth angle from the planet position, at the point where the spiral first crosses the negative y-axis outside of the planetary orbit. We denote this quantity as $(\delta X/X)_{270}$. In all our models, this position is already located in the outer disc and not in the gap, thus the spiral can already be observed. For the parameters given in Table 2, the value of $(\delta X/X)_{270}$ is 0.44.

In Figs 4 and 5, we show the calculated $H$-band polarized intensity images for the case of surface density perturbations, and the radial cross-section of the images, respectively. It can be seen that the relative change of the surface density along the spirals needs to be at least 3.5 at 270° azimuth angle to make the spirals visible in scattered light images if the FWHM of the PSF is 0.06 arcsec. This is a factor 8 times stronger perturbation than what we see in our hydrodynamical simulations for a planet mass of $10^{-3} M_\ast$.

4.2.2 Pressure scaleheight perturbations

We use the same perturbation function that we used in Section 4.2.1 for the surface density, but we apply it to the pressure scaleheight. Similar to the surface density perturbations, we change the strength of the two-dimensional perturbation function to study the detectability of the spirals. The resulting images for different values for the relative perturbation amplitude are shown in Figs 4(e)–(h). One needs a relative change in the pressure scaleheight of at least $\sim 0.2$ in order to detect the spirals in observations with a PSF whose FWHM is 0.06 arcsec.

The radial cross-section of the images is presented in Fig. 6. By comparing Figs 5 and 6, it can be seen that to achieve a certain contrast between the spiral features and the background disc we need a factor of about 16–20 smaller relative perturbation in case of pressure scaleheight variation, than for surface density perturbation.

It is expected that the same relative perturbation of the pressure scaleheight causes higher density perturbation in the upper layers (4$H_p$ and above) of the disc compared to surface density perturbation. Surface density variations translate to the same amount of relative perturbations in the volume density of the dust and gas independently of the vertical height above the disc mid-plane. In contrast, perturbation in the pressure scaleheight changes the volume...
Figure 6. Cross-section of the calculated images with different pressure scaleheight perturbations at original resolution, as calculated by 
radmc-3d (Panel a), and after convolved with a 0.06 arcsec circular Gaussian PSF (Panel b).

\[ \frac{\rho(R, z) + \delta\rho(R, z)}{\rho(R, z)} = \frac{H_p(R)}{H_p(R) + \delta H_p} \exp \left( \frac{z^2}{2H_p(R)} - \frac{z^2}{2(H_p(R) + \delta H_p)} \right) \]

According to equation (6), 20 per cent relative perturbation in the pressure scaleheight introduces the same volume density perturbation at the bottom of the disc atmosphere at 4\(H_p\) above the disc mid-plane as relative perturbation of a factor of \(\sim 8\) in the surface density. Since the variation in the surface brightness in scattered light images of protoplanetary discs reflects volume density variation in the upper layers of the disc, we expect pressure scaleheight perturbation to be more effective in causing observable, i.e. high contrast, perturbations compared to surface density variations.

5 EFFECT OF SPATIAL RESOLUTION

The dramatic decrease of the contrast between the spiral features and the background disc is caused by the convolution of the images with the telescope PSF. As mentioned in Section 2.2, the images calculated by 
radmc-3d have a resolution of 0.357 au pixel\(^{-1}\) corresponding to 2.55 \(\times 10^{-3}\) arcsec pixel\(^{-1}\), sufficient to resolve the spiral arms. These images indeed reflect the contrast between the spirals and the disc we see in the surface density. In the PSF convolved images, however, the width of the spiral arms is significantly less than that of the PSF. In this case, the contrast between the spirals and the background disc is lowered by the fraction of the PSF covered by the spiral.

For this very reason, the detectability criterion for the perturbation amplitude depends on the ratio between the width of the spirals and the size of the PSF. For a source at a given fixed distance from the observer, the contrast between the disc and the spiral increases with increasing resolution as long as the spirals are not spatially resolved. If, however, we fix the resolution, i.e. the size of the telescope, the contrast between the spiral and the disc increases with decreasing distance to the source. In Fig. 7, we investigated the detectability of spiral surface density perturbations with next generation instruments (SPHERE/VLT and telescopes (E-ELT). We take the images, we calculated in Section 3.3 based on hydrodynamical simulations of an embedded planet with a mass of \(10^{-3}M_\odot\), and convolve them with various sized PSFs expected in observations of these instruments. We assume that the source is located at a fixed 140 pc distance.

SPHERE on VLT will deliver diffraction limited images at optical and NIR wavelengths. In the \(H\)-band, the FWHM of the Airy-disc\(^2\) in SPHERE observations is \(\sim 0.04\) arcsec. This resolution is not yet sufficient to detect pure surface density perturbations along the spirals in discs at 140 pc distance (see Fig. 7 top row). One can increase the resolution by almost a factor of 2 if one observes the disc in the \(R\) band instead of the \(H\) band. SPHERE images in the \(R\) band will have high enough resolution (FWHM = 0.02 arcsec) to detect the spirals (see Fig. 7 middle row). Since the size of the PSF is still larger than the width of the spirals (\(\sim 0.013\) arcsec), the contrast between the spiral and the disc is still lower in the scattered light images than in the surface density. Only E-ELT class telescopes can deliver high enough resolution images (FWHM = 0.008 arcsec) to spatially resolve the spirals (see Fig. 7 bottom row). E-ELT class telescopes will be able to directly measure the true density contrast between the spirals and the background disc.

6 MULTIWAVELENGTH OBSERVATIONAL DIAGNOSTICS

In the previous sections, we showed that the amplitude of the surface density perturbations induced by planets along the spirals is too low to explain the contrast in scattered light images. Thus the observed spirals may represent perturbations in the upper layers of the disc e.g. by variations of the pressure scaleheight that changes the density preferentially high above the mid-plane. Here, we discuss observational diagnostics supplementary to scattered light images to constrain the origin of the observed spiral arms.

One possible test is to fit the spiral wake in the scattered light images and the SED simultaneously. The pitch angle of the spirals is determined by the sound-speed (i.e. temperature) distribution in the disc, which in turn sets the pressure scaleheight. Thus by fitting the spiral wake, we can constrain the pressure scaleheight, i.e. the vertical structure of the disc, as a function of radius. The scaleheight and its variation as a function of radius (i.e. the flaring index) determines also the luminosity of the disc and the slope of the SED from mid- to far-infrared wavelengths, respectively. If the observed spiral arms are caused by planets, there must be a solution for the pressure scaleheight and sound speed distribution that fits both the SED and the spiral wake in the scattered light images. The lack of such solution would imply that the spirals are not caused by planetary mass companions in locally isothermal gravitationally stable discs.

There are also methods to test observationally whether the spiral arms are perturbation in the surface density or only confined to the upper layers of the disc, by comparing submillimetre continuum images to those in scattered light. At submillimetre wavelengths,
Spiral arms in protoplanetary discs

Figure 7. Effect of the spatial resolution on the detectability of spiral arms with pure surface density perturbation. The images are calculated from hydrodynamic simulations for a planet-to-star mass ratio of 0.001. Panels in different columns show models with different values for the aspect ratio ($H_0/R$) and the flaring index ($\zeta$), which is indicated in the top-left corner of each panel. The three rows show the images convolved with Gaussian PSFs representative in size for SPHERE/VLT in the $H$-band (FWHM = 0.04 arcsec, top row), SPHERE/VLT in the $R$-band (FWHM = 0.02 arcsec, middle row) and E-ELT in the $H$-band (FWHM = 0.008 arcsec, bottom row). To detect pure surface density perturbation along the spirals at 140 pc distance, one needs an extreme high spatial resolution, where the FWHM of the PSF is 0.02 arcsec or smaller.

In contrast to surface density perturbation, scaleheight perturbation does not change the column density of the dust and gas, it merely re-distributes material along the vertical direction. However, scaleheight variations are linked to temperature variation and we may observe variation of the temperature along the spirals. Since $H_0 \propto \sqrt{T}$, the temperature variation is $(1 + \delta T/T) = (1 + \delta H_0/H_0)^2$. This temperature perturbation will be identical to the variation of the flux at submillimetre wavelengths due to the linear dependence of the flux on the temperature in the Rayleigh–Jeans approximation. As we showed in Section 4.2.2, a given contrast in scattered light images between the spirals and the disc requires a much lower scaleheight/temperature variation than surface density perturbation.

Thus, in case of pressure scaleheight perturbation, we expect to see a significantly lower contrast between the spirals and the background disc in the submillimetre continuum than in NIR polarized intensity. Note, that the images in submillimetre and NIR should have the same angular resolution for a straightforward comparison.

To demonstrate this test, we take the models with analytical surface density and pressure scaleheight perturbations from Sections 4.2.1 and 4.2.2, respectively, and calculated images at 880 $\mu$m, representative for ALMA Band 7 observations. We choose the amplitude of the perturbation such that both perturbation type (surface density and scaleheight) should cause a similar contrast signature in the NIR. For the model with pressure scaleheight variation, we changed (i.e. increased) the dust temperature calculated by RADMC-3D by $(1 + \delta H_0/H_0)^2$ along the spirals before raytracing to simulate the thermal perturbation required for the imposed scaleheight variations. We neglected the thermal perturbation in the calculation of the NIR images, as thermal emission is negligible in the NIR at several tens of astronomical units distance from the central star. Then we convolve both the NIR polarized intensity images and the submillimetre images with a 2D Gaussian kernel with an FWHM of 0.06 arcsec. We note that 0.06 arcsec angular resolution can already be achieved with ALMA in Cycle 3. The

we see the thermal emission of the dust. If the disc is optically thin, which it is in most cases outside of about 5–10 au, the submillimetre emission probes the total column density of the dust. If the spirals observed in scattered light are caused by surface density perturbation, one should observe the same contrast between the spirals and the background disc at both wavelengths.

3 The disc might become optically thick in submillimetre continuum in dust traps in the cores of anticyclonic vortices.
resulting images are presented in Fig. 8. As can be seen, the contrast in the scattered light and submillimetre thermal emission is the same for the surface density perturbation. In case of pressure scaleheight variation, the contrast between the spirals and the disc in the submillimetre is so low that the spirals are barely visible.

7 SUMMARY AND CONCLUSIONS

We investigated whether or not the spiral arms seen in protoplanetary discs in polarized scattered light (Muto et al. 2012; Grady et al. 2013; Avenhaus et al. 2014b,a) can be interpreted as density waves launched by one or more planets under the assumption that the disc is locally isothermal. There are several important properties of the observed spirals that we can compare to the properties of planet-driven spiral density waves, (i) the number of spiral arms observed and their relative strength, (ii) the pitch angle of the spiral and (iii) the contrast between the spiral and the background disc.

We use 2D hydrodynamic simulations to investigate the morphology of the spirals driven by single planets. We find that two or even more armed spirals can be excited by a single planet. The number of spirals increases for higher planet masses and for lower aspect ratios. These spirals are not symmetric, though, as they have different amplitude and width. More than one spiral arm is only seen in simulations where the planet is able to open a well-defined gap. Thus to explain the presence of the symmetric two-armed spiral seen in HD 135344B (Muto et al. 2012) two companions are required. Whether or not such configuration (combination of planet masses and orbital parameters) is dynamically stable, how long it can be sustained, i.e. how likely it is to observe such system, and whether or not the amplitude of the spirals could be reproduced with the two companions is yet to be investigated.

All spirals seen in scattered light observations seems to be open, i.e. the pitch angle seems to be large, which bears important consequences for the vertical structure of the disc. On one hand, the more open the spiral the higher the aspect ratio of the disc needs to be and also the smaller the amplitude of the spiral will become. On the other hand, the aspect ratio of the disc, i.e. the pressure scaleheight and its radial variation, sets the amount of reprocessed stellar radiation, which in turn will determine the luminosity of the disc, the shape of the SED from near- to far-infrared wavelengths as well as the absolute surface brightness of the disc. Thus simultaneous fitting of the spiral wake and the SED with a single $H_{\nu}(R)$ curve can be a powerful tool to test whether or not the spirals are driven by an embedded planet.

Our simulations show that planets by themselves in gravitationally stable locally isothermal discs cannot create strong enough surface density perturbation along the spirals that could be observed in NIR scattered light with current 8 m class telescopes. Planets in such discs can create a relative surface density perturbation of the order of several tens of percent at most. This contrast is however lowered to below the detection limit by the convolution with the telescope PSF, which is significantly larger than the width of the spirals. We use analytic models to estimate that a surface density perturbation of a factor of 3.5 or higher above the background disc is required to create spiral density waves which are observable at a spatial resolution of 0.06 arcsec. Such high amplitude perturbation is unlikely to be caused by planetary mass companions.

We also test the effect of pressure scaleheight perturbation along the spirals to the scattered light images. The variation in the volume density, caused by pressure scaleheight perturbations, increases monotonically with height above two scaleheights. One needs a factor of $\sim 16$ of the relative change in the scaleheight to create the same contrast signature in the scattered light images as any given surface density perturbation. In order for the spirals to be detectable in images convolved with a PSF with an FWHM of 0.06 arcsec, a relative change in the pressure scaleheight of $\sim 0.2$ is required above the background disc. We therefore suggest that the spiral arms observed in protoplanetary discs are caused by variations in the vertical structure of the disc (e.g. pressure scaleheight perturbation) instead of surface density perturbation.

We also investigate the effect of spatial resolution on the detectability of surface density perturbations along the spirals. We predict that spiral surface density perturbations cannot be detected with current 8 m class telescopes in the $H$ band, even with state-of-the-art high-contrast imagers, such as SPHERE on VLT. While the detection of planet-induced spiral surface density perturbation...
is possible in the $R$ band with 8 m class telescopes, an E-ELT class telescope is required to fully resolve the spirals at NIR wavelengths and measure the true density contrast between the spiral and the background disc.

Finally, we suggest two types of tests to study the origin of the spirals. To study whether the spirals seen in scattered light represent surface density perturbation or scaleheight variation we can compare submillimetre and scattered light images. In case of surface density perturbation, the variation of the density is independent of the height above the disc mid-plane. Therefore, the observed contrast between the spiral and the disc should be the same in submillimetre as in scattered light if the spatial resolution is the same. For scaleheight variation, we should observe a significantly lower contrast in submillimetre as in scattered light due to the lower amplitude of the thermal perturbation required to produce the observed scattered light signatures. For planet-induced spirals, the pitch angle of the spirals is determined by the sound-speed distribution in the disc. The sound-speed depends on the temperature, which in turn will determine the scaleheight and the flaring index of the disc. The flaring index and pressure scaleheight can be constrained by fitting the SED of the disc, while the pitch angle can be derived by fitting the spiral wake in the scattered light images. Thus a simultaneous fit of the SED and the spiral wake can help to understand whether planets can be responsible for the observed spirals and whether spiral arms in protoplanetary discs are indeed a signposts of on-going planet formation.

ACKNOWLEDGEMENTS

We are grateful to Paola Pinilla, Cathie Clarke and Giovanni Rosotti of on-going planet formation. We thank the anonymous referee for the insightful comments that helped to improve our paper.

REFERENCES

Casassus S. et al., 2013, Nature, 493, 191