J. Fluid Mech. (2010), *vol.* 654, *pp.* 207–231. © Cambridge University Press 2010
doi:10.1017/S0022112010000510

Modulational instability of Rossby and drift waves and generation of zonal jets

COLM P. CONNAUGHTON[1,2]†,
BALASUBRAMANYA T. NADIGA[3],
SERGEY V. NAZARENKO[2] AND BRENDA E. QUINN[2]

[1]Centre for Complexity Science, University of Warwick, Gibbet Hill Road, Coventry CV4 7AL, UK
[2]Mathematics Institute, University of Warwick, Gibbet Hill Road, Coventry CV4 7AL, UK
[3]Los Alamos National Laboratory, Los Alamos, NM 87545, USA

(Received 29 May 2009; revised 22 January 2010; accepted 26 January 2010;
first published online 5 May 2010)

We study the modulational instability of geophysical Rossby and plasma drift waves within the Charney–Hasegawa–Mima (CHM) model both theoretically, using truncated (four-mode and three-mode) models, and numerically, using direct simulations of CHM equation in the Fourier space. We review the linear theory of Gill (*Geophys. Fluid Dyn.*, vol. 6, 1974, p. 29) and extend it to show that for strong primary waves the most unstable modes are perpendicular to the primary wave, which correspond to generation of a zonal flow if the primary wave is purely meridional. For weak waves, the maximum growth occurs for off-zonal inclined modulations that are close to being in three-wave resonance with the primary wave. Our numerical simulations confirm the theoretical predictions of the linear theory as well as the nonlinear jet pinching predicted by Manin & Nazarenko (*Phys. Fluids*, vol. 6, 1994, p. 1158). We find that, for strong primary waves, these narrow zonal jets further roll up into Kármán-like vortex streets, and at this moment the truncated models fail. For weak primary waves, the growth of the unstable mode reverses and the system oscillates between a dominant jet and a dominate primary wave, so that the truncated description holds for longer. The two-dimensional vortex streets appear to be more stable than purely one-dimensional zonal jets, and their zonal-averaged speed can reach amplitudes much stronger than is allowed by the Rayleigh–Kuo instability criterion for the one-dimensional case. In the long term, the system transitions to turbulence helped by the vortex-pairing instability (for strong waves) and the resonant wave–wave interactions (for weak waves).

1. Introduction and motivation

Zonal flows are prominent features in the atmospheres of giant planets such as Jupiter and Saturn (see Simon 1999; Sanchez-Lavega, Rojas & Sada 2000; Galperin *et al.* 2004), the Earth's atmosphere (see Lewis 1988) and its oceans (see Galperin *et al.* 2004; Maximenko *et al.* 2008). Geophysical jets can regulate small-scale turbulence and transport processes via, for example, the 'Barotropic Governor' mechanism (James 1987). Zonal flows are also important in plasma turbulence of fusion devices (Diamond *et al.* 2005). There, they can also regulate the turbulence and suppress transport via

† Email address for correspondence: connaughtonc@gmail.com

a drift-wave/zonal-flow feedback loop (Balk, Nazarenko & Zakharov 1990*a*,*b*). The latter process is presently considered a candidate mechanism for the low-to-high (LH) confinement transitions in tokamaks discovered by Wagner *et al.* (1982) – an effect that is crucial for the success of future fusion devices. In this paper, we will mostly be concerned with geophysical applications, but we will briefly discuss implications of our results in the plasma physics context, and we will outline the modifications needed to be made to our study in order to make it more realistic for the nuclear fusion physics.

Two main zonal flow generation scenarios have been discussed in the literature. According to the first scenario, zonal flows are generated via an inverse energy cascade, which could be local or non-local (Balk *et al.* 1990*a*,*b*) (see also Horton & Ichikawa 1996 and references therein). The mechanism for such an inverse cascade is similar to that of two-dimensional Navier–Stokes turbulence (Kraichnan 1967), but the presence of the beta-effect makes this cascade anisotropic. This leads to a preferential transfer of energy into zonal flows at large scales rather than into round vortices as would be the case in Navier–Stokes turbulence. The beta-effect leads to three-wave resonant interactions that preserve an additional (to the energy and the potential enstrophy) quadratic invariant – zonostrophy (Balk, Nazarenko & Zakharov 1991; Balk 1991, 1997). An application of the standard Fjørtoft argument to the three invariants, the energy, potential enstrophy and zonostrophy leads to the conclusion that the energy can only be transferred to large zonal scales (Balk *et al.* 1991). This statistical argument is explained in detail by Nazarenko & Quinn (2009). The second mechanism of zonal flow generation, and the principal topic of this paper, is via modulational instability (MI) of a primary meridional Rossby/drift wave (Lorentz 1972; Gill 1974; Mima & Lee 1980; Manin & Nazarenko 1994; Smolyakov, Diamond & Shevchenko 2000; Onishchenko *et al.* 2004). In practice, such primary waves are excited as neutrally stable responses to weak external disturbances when the ambient vertically sheared state is subcritical with respect to baroclinic instability. On the other hand, such meridional structures can also appear as the most unstable modes of a supercritical (vertically sheared) background state (cf. the baroclinic instability).

These mechanisms are unlikely to be exclusive in practice and both may coexist under some conditions. The extent to which one mechanism dominates over the other is determined by the parameter regime and configurational details. If the parameter regime were to be such that meridional Rossby waves are excited in an ambient state that is subcritical with respect to baroclinic instability or if weak baroclinic instability resulted in meridional Rossby wave-like structures, zonal flows would presumably result from the MI mechanism, whereas if the parameter regime were to be such that the baroclinic instability resulted in more isotropic eddies at the Rossby deformation radius, the cascade scenario would likely be more relevant. In our purely barotropic model, these effects are modelled by the initial condition. A narrow initial spectrum of the waves and large initial amplitude promotes the modulational instability mechanism leading to fast zonal flow generation bypassing the turbulent cascade stages. On the other hand, for broad initial spectra, the cascade scenario is likely to be more relevant. There is an analogy with the turbulence of surface gravity waves on water where the inverse cascade and modulational (Benjamin–Feir) instability (Benjamin & Feir 1967) can compete with each other in the generation of long waves (Onorato *et al.* 2001). A quantitative measure, called the Benjamin–Feir index, was suggested to estimate probability for triggering the modulational instability (Onorato *et al.* 2001; Janssen 2003). Developing a similar approach for the

Rossby/drift wave system would also be useful. However, we will leave this interesting subject for future studies, and in the present paper we are only concerned with the modulational instability of a monochromatic wave.

We will start by revisiting the linear theory of the modulational instability, which was first analysed by Lorentz (1972) and then treated in great detail in a beautiful paper by Gill (1974). Using numerical and semi-analytical calculations, we highlight the most important properties of Gill's theory. In particular, we will see how the character of instability changes with the strength of the primary wave: from being the classical hydrodynamic instability of the sinusoidal (Kolmogorov) shear flow for large amplitudes (Arnold & Meshalkin 1960) to becoming a (three-wave) decay instability of weakly nonlinear waves for small amplitudes (Sagdeev & Galeev 1969). We will also study the effect of the finite Rossby/Larmor radius on the instability.

We will then study the nonlinear stage of the modulational instability with direct numerical simulations (DNS), comparing them with the solutions of the four-mode truncated (4MT) and the three-mode truncated (3MT) systems.We find that at the nonlinear stage, for strong primary waves the growth of the zonal mode deviates from the truncated dynamics and the zonal flow tends to focus into narrow jets, as was theoretically predicted by Manin & Nazarenko (1994). These zonal jets subsequently become unstable and acquire the interesting two-dimensional structure of a double (Kármán-like) vortex street. The vortex street appears to be more stable than a plane parallel shear flow with the same zonal profile (McWilliams 2006) and persists for a relatively long time until (possibly due to dissipation) a vortex pairing instability sets in and triggers a transition to turbulence (McWilliams 2006). As the nonlinearity of the primary wave is decreased we find that there is a level of nonlinearity below which this sequence of events changes. For sufficiently weak primary waves, the jet strength reaches a maximum which is still stable. After this maximum is reached, the jet amplitude starts decreasing again, continues to follow the truncated dynamics and avoids the roll-up into vortices. This reversal of the jet growth, particularly the maximum jet strength, is well predicted by nonlinear oscillatory solutions of the 4MT, and often by the 3MT, equations. The latter are relevant for non-degenerate (in a sense that we shall explain) resonant wave triads. Once the full system deviates from the solutions of the truncated system, as it inevitably does, it sometimes continues to exhibit oscillatory behaviour for a while in the weak nonlinearity cases. These subsequent oscillations have different periods, however, and are often rather irregular.

Along the way, we will examine the relative performance of the 3MT versus 4MT models, thereby clarifying possible confusions on whether the principal mechanism of the modulational instability is three-wave or four-wave.

2. The model

Geophysical and plasma zonal flows are often mentioned together because of the same simplified nonlinear partial differential equation (PDE) that was suggested for their description, namely, the Charney–Hasegawa–Mima (CHM) equation (Charney 1949; Hasegawa & Mima 1978):

$$\partial_t (\Delta\psi - F\psi) + \beta\partial_x\psi + J[\psi, \Delta\psi] = \nu_n(-\Delta)^n\psi, \tag{2.1}$$

where ψ is the streamfunction, $F = 1/\rho^2$ with ρ being the deformation radius in the geophysical fluid dynamics (GFD) context and the ion Larmor radius at the electron temperature in the plasma context, β is a constant proportional to the latitudinal gradient of the vertical rotation frequency or of the plasma density in the GFD and

plasma contexts, respectively. We introduced notation for the Jacobean operator,

$$J[f, g] = (\partial_x f)(\partial_y g) - (\partial_y f)(\partial_x g). \tag{2.2}$$

In the GFD context, the x-axis is in the west–east and the y-axis is along the south–north directions, respectively. In plasmas, the y-axis is along the plasma density gradient and the x-axis is, of course, transverse to this direction. In both plasma and GFD applications, the CHM model represents only the simplest of the useful nonlinear models. In each case, there exists a hierarchy of models with increasing degree of realism, achieved at a cost of increasing complexity. In plasmas, the next level model is the modified Hasegawa–Mima model that improves the description of the electron response (Dorland *et al.* 1990). In GFD the next level of description is the two-layer model which includes baroclinic effects (McWilliams 2006). We will discuss the prospects of extending our study to these systems in the last section. Until then, however, we shall focus only on the CHM model and the interplay between the two fundamental effects it contains: hydrodynamic nonlinearity and the drift (β) effect.

Note that we include in (2.1), a hyper-viscous dissipation term having integer degree $n \geqslant 2$ with a small positive coefficient, ν_n. All numerical calculations in this article were done with a standard pseudo-spectral code using a third-order Runge–Kutta time integration algorithm with integrating factors for the linear terms. All presented results were obtained at 1024^2 resolution and hyper-viscosity was used to damp the high wavenumbers. We checked that the large-scale behaviour that we study here was relatively insensitive to the details of this small-scale dissipation by varying the resolution and hyperviscosity parameters. In this paper, we compute the initial value problem, hence no forcing term in (2.1).

Introducing the Fourier transform of the streamfunction, $\psi_{\boldsymbol{k}} = \int \psi(\boldsymbol{x}) \mathrm{e}^{-\mathrm{i}(\boldsymbol{k}\cdot\boldsymbol{x})}\, \mathrm{d}\boldsymbol{x}$, the CHM equation, (2.1), ignoring the hyper-viscosity term for now, is equivalent to the following:

$$\partial_t \psi_{\boldsymbol{k}} = +\mathrm{i}\,\omega_{\boldsymbol{k}}\, \psi_{\boldsymbol{k}} + \frac{1}{2} \sum_{\boldsymbol{k}_1, \boldsymbol{k}_2} T(\boldsymbol{k}, \boldsymbol{k}_1, \boldsymbol{k}_2)\, \psi_{\boldsymbol{k}_1}\, \psi_{\boldsymbol{k}_2}\, \delta(\boldsymbol{k} - \boldsymbol{k}_1 - \boldsymbol{k}_2), \tag{2.3}$$

where

$$\omega_{\boldsymbol{k}} = -\frac{\beta k_x}{k^2 + F}, \tag{2.4}$$

$$T(\boldsymbol{k}, \boldsymbol{k}_1, \boldsymbol{k}_2) = -\frac{(\boldsymbol{k}_1 \times \boldsymbol{k}_2)_z \left(k_1^2 - k_2^2\right)}{k^2 + F} \tag{2.5}$$

and $\boldsymbol{k} = (k_x, k_y)$ and $k = |\boldsymbol{k}|$. The system clearly supports linear waves, known as Rossby or drift waves, in the GFD and plasma contexts respectively. They have the anisotropic dispersion relation given by (2.4). The structure of the nonlinear interaction, (2.5), is such that the nonlinear term vanishes for a monochromatic wave. Hence, Rossby waves are actually exact solutions of the full CHM equation. Much of this paper will focus on the stability properties of these solutions.

Typically, the waves arise either because of a primary instability, e.g. the baroclinic instability in GFD (McWilliams 2006) or because of the ion-temperature-gradient (ITG) instability in fusion plasmas (Rudakov & Sagdeev 1961), or as a response to some external forcing. The instability is not included in the CHM equation, and it could be modelled by simulating CHM with an initial condition or introducing a linear forcing term on the right-hand side mimicking the linear instability (this would

not take into account the nonlinear mechanisms in the wave forcing). It is interesting that the GFD–plasma analogy extends to the instabilities too in which the most unstable mode is 'meridional' (i.e. along the x-axis) and concentrated at the scales of the order of ρ. Thus, in most of our considerations below we will consider the initial (primary) wave which is purely meridional.

3. Spectral truncations

We shall use spectral truncations of (2.3) in our study of the stability properties of Rossby waves. They provide approximations of an intermediate degree of complexity between monochromatic waves and the full PDE. At this stage, such truncations should be viewed as *ad hoc* since, in reality, all triads are coupled together in (2.3). Their usefulness will be determined by comparison with DNS solutions of the full system (2.3). We shall consider two natural truncations: the three-mode truncation (3MT) and the four-mode truncation (4MT).

3.1. *Three-mode truncation*

The simplest such truncation is to restrict the right-hand side of (2.3) to a single triad containing only three modes that we denote by $\boldsymbol{p}$, $\boldsymbol{q}$ and $\boldsymbol{p}_- = \boldsymbol{p} - \boldsymbol{q}$. In our convention, $\boldsymbol{p}$ corresponds to the wavenumber of the primary wave with amplitude $\Psi_{\boldsymbol{p}}$. We construct the truncated equations by taking each wave vector in the triad in turn and assigning it to be $\boldsymbol{k}$ in (2.3), enumerating all ways of assigning the others and their negatives to $\boldsymbol{k}_1$ and $\boldsymbol{k}_2$ on the right-hand side and neglect any terms that involve $\psi_{\boldsymbol{k}}$'s outside the triad. Because $\psi_{\boldsymbol{k}}$ is the Fourier transform of a real field, $\psi_{-\boldsymbol{k}} = \bar{\psi}_{\boldsymbol{k}}$. It is convenient to introduce $\Psi_{\boldsymbol{k}}(t) = \psi_{\boldsymbol{k}}(t)\mathrm{e}^{-\mathrm{i}\omega_{\boldsymbol{k}} t}$. In terms of $\Psi_{\boldsymbol{k}}$ we arrive at the following equations for the 3MT:

$$\left.\begin{aligned} \partial_t \Psi_{\boldsymbol{p}} &= T(\boldsymbol{p}, \boldsymbol{q}, \boldsymbol{p}_-)\, \Psi_{\boldsymbol{q}} \Psi_{\boldsymbol{p}_-} \mathrm{e}^{\mathrm{i}\Delta_- t}, \\ \partial_t \Psi_{\boldsymbol{q}} &= T(\boldsymbol{q}, \boldsymbol{p}, -\boldsymbol{p}_-) \Psi_{\boldsymbol{p}} \overline{\Psi}_{\boldsymbol{p}_-} \mathrm{e}^{-\mathrm{i}\Delta_- t}, \\ \partial_t \Psi_{\boldsymbol{p}_-} &= T(\boldsymbol{p}_-, \boldsymbol{p}, -\boldsymbol{q}) \Psi_{\boldsymbol{p}}\, \overline{\Psi}_{\boldsymbol{q}} \mathrm{e}^{-\mathrm{i}\Delta_- t}, \end{aligned}\right\} \tag{3.1}$$

where $\Delta_- = \omega_{\boldsymbol{p}} - \omega_{\boldsymbol{q}} - \omega_{\boldsymbol{p}_-}$. A similar set of equations can be derived for the other natural triad $(\boldsymbol{p}, -\boldsymbol{q}, \boldsymbol{p}_+)$, where $\boldsymbol{p}_+ = \boldsymbol{p} + \boldsymbol{q}$:

$$\left.\begin{aligned} \partial_t \Psi_{\boldsymbol{p}} &= T(\boldsymbol{p}, -\boldsymbol{q}, \boldsymbol{p}_+)\, \overline{\Psi}_{\boldsymbol{q}} \Psi_{\boldsymbol{p}_+} \mathrm{e}^{\mathrm{i}\Delta_+ t}, \\ \partial_t \Psi_{\boldsymbol{q}} &= T(\boldsymbol{q}, -\boldsymbol{p}, \boldsymbol{p}_+) \overline{\Psi}_{\boldsymbol{p}} \Psi_{\boldsymbol{p}_+} \mathrm{e}^{\mathrm{i}\Delta_+ t}, \\ \partial_t \Psi_{\boldsymbol{p}_+} &= T(\boldsymbol{p}_+, \boldsymbol{p}, \boldsymbol{q}) \Psi_{\boldsymbol{p}}\, \Psi_{\boldsymbol{q}} \mathrm{e}^{-\mathrm{i}\Delta_+ t}, \end{aligned}\right\} \tag{3.2}$$

where $\Delta_+ = \omega_{\boldsymbol{p}} + \omega_{\boldsymbol{q}} - \omega_{\boldsymbol{p}_+}$. If $\Delta_+ = 0$, the triad is exactly resonant. Then (3.2) form an exactly integrable set of equations that have been extensively studied (Kartashova & L'vov 2007; Bustamante & Kartashova 2009).

3.2. *Four-mode truncation*

The 4MT model retains both triads, $(\boldsymbol{p}, \boldsymbol{q}, \boldsymbol{p}_+)$ and $(\boldsymbol{p}, -\boldsymbol{q}, \boldsymbol{p}_-)$, where $\boldsymbol{p}_\pm = \boldsymbol{p} \pm \boldsymbol{q}$ mentioned above. The truncated equations combine (3.1) and (3.2):

$$\left.\begin{aligned} \partial_t \Psi_{\boldsymbol{p}} &= T(\boldsymbol{p}, \boldsymbol{q}, \boldsymbol{p}_-)\, \Psi_{\boldsymbol{q}} \Psi_{\boldsymbol{p}_-} \mathrm{e}^{\mathrm{i}\Delta_- t} + T(\boldsymbol{p}, -\boldsymbol{q}, \boldsymbol{p}_+)\, \overline{\Psi}_{\boldsymbol{q}} \Psi_{\boldsymbol{p}_-} \mathrm{e}^{\mathrm{i}\Delta_+ t}, \\ \partial_t \Psi_{\boldsymbol{q}} &= T(\boldsymbol{q}, \boldsymbol{p}, -\boldsymbol{p}_-) \Psi_{\boldsymbol{p}} \overline{\Psi}_{\boldsymbol{p}_-} \mathrm{e}^{-\mathrm{i}\Delta_- t} + T(\boldsymbol{q}, -\boldsymbol{p}, \boldsymbol{p}_+) \overline{\Psi}_{\boldsymbol{p}} \Psi_{\boldsymbol{p}_+} \mathrm{e}^{\mathrm{i}\Delta_+ t}, \\ \partial_t \Psi_{\boldsymbol{p}_-} &= T(\boldsymbol{p}_-, \boldsymbol{p}, -\boldsymbol{q}) \Psi_{\boldsymbol{p}}\, \overline{\Psi}_{\boldsymbol{q}} \mathrm{e}^{-\mathrm{i}\Delta_- t}, \\ \partial_t \Psi_{\boldsymbol{p}_+} &= T(\boldsymbol{p}_+, \boldsymbol{p}, \boldsymbol{q}) \Psi_{\boldsymbol{p}}\, \Psi_{\boldsymbol{q}} \mathrm{e}^{-\mathrm{i}\Delta_+ t}. \end{aligned}\right\} \tag{3.3}$$

Strictly speaking, the chosen four modes (ψ_0, $\psi_{\boldsymbol{q}}$, ψ_+ and ψ_-) are coupled to further modes and do not form a closed system. Indeed, even the linear problem closes only if all the satellites $\pm\boldsymbol{q}+m\boldsymbol{p}$ (m is a positive or negative integer) are included (Gill 1974). However, in considering the linear instability it is traditional to truncate the system to the four modes only with a justification that the higher-order satellites are less excited in the linear eigenvectors, which turns out to be a very good approximation for weak primary waves and quite reasonable for strong ones (Gill 1974). In this paper, we will test predictions of the 4MT system, both linear and nonlinear, against DNS of the full system.

3.3. *Nonlinearity parameter M*

The character of the instability is greatly influenced by the initial amplitude of the primary wave, $\Psi_0 = \Psi_{\boldsymbol{p}}|_{t=0}$ (Gill 1974). Following Gill (1974), we introduce the dimensionless amplitude

$$M = \frac{\Psi_0 p^3}{\beta}, \tag{3.4}$$

where M is a formal measure of the relative strength of the linear and nonlinear terms at the scale of the primary wave. We shall use M as an indicator of the typical level of nonlinearity. The actual level of nonlinearity, however, will clearly change during the time evolution and will not be uniform throughout the Fourier space. The case $M \gg 1$ corresponds to the case where the β-effect plays no role. For $F = 0$, this case reduces to the Euler equation limit and the well-studied instability of the plane parallel sinusoidal shear flow known as Kolmogorov flow (Arnold & Meshalkin 1960). Note that most papers on the modulational instability within the plasma context have dealt only with this limit (e.g. Smolyakov *et al.* 2000; Onishchenko *et al.* 2004). The limit $M \to \infty$ in Manin & Nazarenko (1994) and Smolyakov *et al.* (2000) was present 'implicitly' as a result of employing a Wentzel–Kramers–Brillouin (WKB) or scale separation approach, assuming that the modulation wavenumber q is much less than the carrier wavenumber p. Technically, this amounts to taking the limit $M = \Psi_0 p^3/\beta \to \infty$. Onishchenko *et al.* (2004) used the 4MT, as we do, but they also used a scale separation by expanding in small q, which leads to the same large-M results. The case $M \ll 1$ corresponds to the weak nonlinearity limit dominated by resonant wave triads. In this case, the four constituent modes (primary wave, modulation and two satellites) can be split into two coupled triads that produce independent contributions to the instability (Gill 1974). The instability associated with a single triad is known as the decay instability (Sagdeev & Galeev 1969). The condition $M \sim 1$ defines the Rhines scale k_r that marks a crossover from the hydrodynamic vortex to the wave behaviour (Rhines 1975).

4. Decay instability of a Rossby wave

In the decay instability, the primary wave decays into two secondary waves (see e.g. Sagdeev & Galeev 1969). We shall derive this instability from the 3MT in (3.1). Introducing the vector notation $\boldsymbol{\Psi} = (\Psi_{\boldsymbol{p}}, \Psi_{\boldsymbol{q}}, \Psi_{\boldsymbol{p}_-})$, a monochromatic primary wave is given by $\boldsymbol{\Psi}_0 = (\Psi_0, 0, 0)$, where Ψ_0 is a complex constant representing the amplitude of the initial primary wave. This is an exact solution of (3.1). We consider the stability of this solution to small perturbations involving the modes $\boldsymbol{q}$ and $\boldsymbol{p}_-$ by taking $\boldsymbol{\Psi} = \boldsymbol{\Psi}_0 + \epsilon\boldsymbol{\Psi}_1$ with the perturbation given by $\boldsymbol{\Psi}_1 = (0, \widetilde{\psi}_{\boldsymbol{q}}, \widetilde{\psi}_{\boldsymbol{p}_-})$. We linearize the equations and seek solutions for the evolution of the perturbation of the form $\widetilde{\psi}_{\boldsymbol{q}}(t) = A_{\boldsymbol{q}} \mathrm{e}^{-\mathrm{i}\,\Omega_{\boldsymbol{q}} t}$ and $\widetilde{\psi}_{\boldsymbol{p}_-}(t) = A_{\boldsymbol{p}_-} \mathrm{e}^{-\mathrm{i}\,\Omega_{\boldsymbol{p}_-} t}$. These solutions are harmonic if $\Omega_{\boldsymbol{q}}$ and

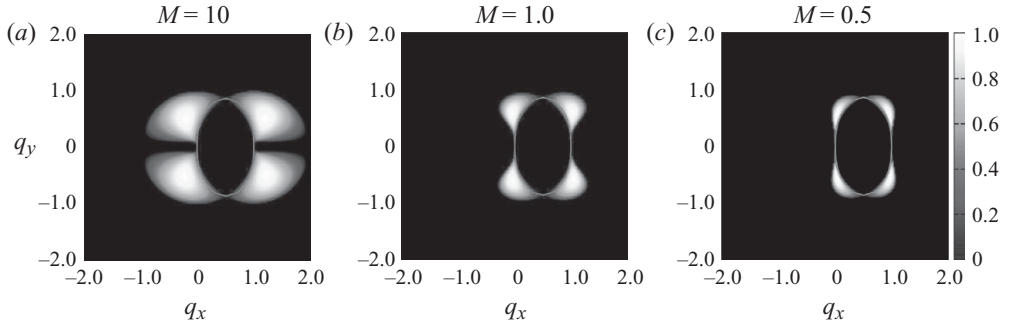

FIGURE 1. The growth rate of the decay instability (the negative imaginary part of the roots of (4.3)) is plotted as a function of $\boldsymbol{q}$ for a fixed meridional primary wave vector $\hat{\boldsymbol{p}} = (1, 0)$, for various values of the nonlinearity parameter M. The set of unstable perturbations becomes concentrated on the resonant manifolds (solid line) as the nonlinearity of the primary wave is decreased.

$\Omega_{\boldsymbol{p}_-}$ are purely real (stable case), and contain exponentially decaying and growing modes otherwise (unstable case). A straightforward calculation, the details of which are given in the online supplement (available at journals.cambridge.org/flm), leads to the following dispersion relation for the perturbation:

$$\Omega_{\boldsymbol{q}}(-\Omega_{\boldsymbol{q}} + \Delta_-) - T(\boldsymbol{q}, \boldsymbol{p}, -\boldsymbol{p}_-)\, T(\boldsymbol{p}_-, \boldsymbol{p}, -\boldsymbol{q})\, |\Psi_0|^2 = 0. \tag{4.1}$$

This has two roots. Instability occurs when $\Omega_{\boldsymbol{q}}$ has a non-zero imaginary part, so that $\widetilde{\psi}_{\boldsymbol{q}}(t) = A_{\boldsymbol{q}} \mathrm{e}^{\gamma_{\boldsymbol{q}} t}$ with $\gamma_{\boldsymbol{q}} = -\mathrm{Im}(\Omega_{\boldsymbol{q}})$. For an exactly resonant triad, $\Delta_- = 0$. For resonant triads, using (2.5), the roots of (4.1) are

$$\Omega = \pm \mathrm{i} \frac{|\Psi_0|\, |\boldsymbol{p} \times \boldsymbol{q}|}{\sqrt{(q^2 + F)(p_-^2 + F)}} \sqrt{(p^2 - q^2)(p_-^2 - p^2)}, \tag{4.2}$$

where we omitted the subscript $\boldsymbol{q}$ in $\Omega_{\boldsymbol{q}}$ for brevity. One can see that in this case instability occurs if $q^2 < p^2 < p_-^2$ or $q^2 > p^2 > p_-^2$.

Before investigating the non-resonant instability further, it is convenient to perform some rescalings. The dimensionless primary wave amplitude will be given by M defined in (3.4). We non-dimensionalize the other terms in (4.1) as follows: $\Omega \to \beta\Omega/p$, $F \to p^2 F$, $\boldsymbol{p} \to p\hat{\boldsymbol{p}}$ and $\boldsymbol{q} \to sp\hat{\boldsymbol{q}}$ where $\hat{\boldsymbol{p}} = (\hat{p}_x, \hat{p}_y)$ and $\hat{\boldsymbol{q}} = (\hat{q}_x, \hat{q}_y)$ are unit vectors pointing in the directions of $\boldsymbol{p}$ and $\boldsymbol{q}$, respectively. Equation (4.1) can then be re-arranged to the following form:

$$\Omega(-\Omega + \hat{\Delta}_-) - M^2 T(s\hat{\boldsymbol{q}}, \hat{\boldsymbol{p}}, -\hat{\boldsymbol{p}}_-)\, T(\hat{\boldsymbol{p}}_-, \hat{\boldsymbol{p}}, -\hat{\boldsymbol{q}}) = 0, \tag{4.3}$$

where $\hat{\boldsymbol{p}}_- = \hat{\boldsymbol{p}} - s\hat{\boldsymbol{q}}$ and $\hat{\Delta}_- = \omega_{\hat{\boldsymbol{p}}} - \omega_{s\hat{\boldsymbol{q}}} - \omega_{\hat{\boldsymbol{p}}_-}$. The two roots are

$$\Omega_\pm = \frac{1}{2}\left(\hat{\Delta}_- \pm \sqrt{(\hat{\Delta}_-)^2 - 4M^2 T(s\hat{\boldsymbol{q}}, \hat{\boldsymbol{p}}, -\hat{\boldsymbol{p}}_-)\, T(\hat{\boldsymbol{p}}_-, \hat{\boldsymbol{p}}, -\hat{\boldsymbol{q}})}\right). \tag{4.4}$$

To have an instability we require

$$\hat{\Delta}_- < 2M \sqrt{T(s\hat{\boldsymbol{q}}, \hat{\boldsymbol{p}}, -\hat{\boldsymbol{p}}_-)\, T(\hat{\boldsymbol{p}}_-, \hat{\boldsymbol{p}}, -\hat{\boldsymbol{q}})},$$

which demonstrates that the instability concentrates on the resonant manifold $\hat{\Delta}_- = 0$ as $M \to 0$. This is illustrated in figure 1. The corresponding analysis for the triad $(\boldsymbol{p}, -\boldsymbol{q}, \boldsymbol{p}_+)$ produces identical surfaces reflected about the vertical axis reflecting the

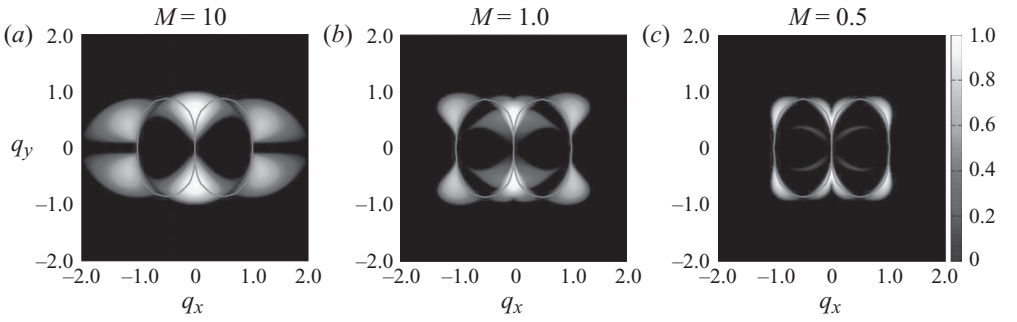

FIGURE 2. Growth rate of the modulational instability (the negative imaginary part of the roots of (5.2)) as a function of $\boldsymbol{q}$ for a fixed meridional primary wave vector, $\boldsymbol{p} = (1, 0)$ and $F = 0$. The values of the nonlinearity M for the initial primary wave are $M = 10$ (Euler limit), $M = 1$, $M = 3/4$, $M = 1/2$, $M = 1/4$ and $M = 1/10$ (weakly nonlinear limit). The set of unstable perturbations becomes concentrated on the resonant manifolds as the nonlinearity of the primary wave is decreased.

instability concentrating on the second resonant manifold $\hat{\Delta}_+ = 0$. As $M \to 0$, these two surfaces become disjoint from each other except near the origin $\boldsymbol{q} = 0$.

5. Modulational instability: linear analysis

Let us now derive the modulational instability in the same way as we have done for the decay instability. The modulational instability is studied using the 4MT. We begin by linearizing (3.3) about the pure primary wave solution, $\boldsymbol{\Psi}_0 = (\Psi_0, 0, 0, 0)$ where Ψ_0 is a complex constant representing the amplitude of the initial primary wave. We consider the stability of this solution to small perturbations involving the three modes $\boldsymbol{q}$, $\boldsymbol{p}_-$ and $\boldsymbol{p}_+$ by taking $\boldsymbol{\Psi} = \boldsymbol{\Psi}_0 + \epsilon \boldsymbol{\Psi}_1$ with the perturbation given by $\boldsymbol{\Psi}_1 = (0, \widetilde{\psi}_{\boldsymbol{q}}, \widetilde{\psi}_{\boldsymbol{p}_+}, \widetilde{\psi}_{\boldsymbol{p}_-})$. A standard linear stability analysis, which is detailed in the online supplement, yields the following dispersion relation for the perturbation (Gill 1974) (see also Lorentz 1972; Manin & Nazarenko 1994; Smolyakov *et al.* 2000; Onishchenko *et al.* 2004):

$$(q^2 + F)\Omega + \beta q_x + |\Psi_0|^2 \, |\boldsymbol{p} \times \boldsymbol{q}|^2 \, (p^2 - q^2) \times \left[\frac{p_+^2 - p^2}{(p_+^2 + F)(\Omega + \omega) + \beta p_{+x}} - \frac{p_-^2 - p^2}{(p_-^2 + F)(\Omega - \omega) + \beta p_{-x}} \right] = 0. \quad (5.1)$$

On non-dimensionalization, this reduces to

$$(s^2 + F)\Omega + s\hat{q}_x + M^2 s^2 (1 - s^2) \, |\hat{\boldsymbol{p}} \times \hat{\boldsymbol{q}}|^2 \, [T^+ - T^-] = 0, \quad (5.2)$$

where

$$T^{\pm} = \frac{|\hat{\boldsymbol{p}} \pm s\hat{\boldsymbol{q}}|^2 - 1}{(|\hat{\boldsymbol{p}} \pm s\hat{\boldsymbol{q}}| + F)\left(-\dfrac{\hat{p}_x}{1 + F} \pm \Omega \right) + \hat{p}_x \pm s\hat{q}_x}. \quad (5.3)$$

Equation (5.2) can be solved numerically, and sometimes analytically, for a given set of parameters to determine Ω, and where the parameters are M, F, s, θ_p and θ_q with θ_p and θ_q the angles between the x-axis and the primary wave vector and perturbation wave vector, respectively. The structure of the instability is strongly dependent on the value of M. This is shown in figure 2. We see that the modulational instability is, in

some sense, the nonlinear sum of the decay instabilities for the two triads, and we will clarify this issue in the next section.

6. Comparison of the 3MT and the 4MT models with DNS of the full CHM system

There is sometimes confusion in the literature, perhaps partially semantic, on whether the modulational instability of the Rossby and drift waves is governed by three-wave or four-wave interactions. Here we will clarify this issue. Gill (1974) showed that as $M \to 0$, the modulational instability obtained within the 4MT model localizes on the two resonant manifolds for the two triads $\Delta_+ = 0$ and $\Delta_- = 0$. Because these two curves are mostly disjoint from each other (except for the origin), in the weakly nonlinear limit, the modulational instability is just a simple sum of the two decay instabilities. Namely, the two unstable eigenvectors of the instability of the 4MT will coincide with the eigenvectors of the two respective branches of the decay instability (i.e. the fourth mode in such 4MT eigenvectors will have zero amplitude). In particular, the maximum growth rates of the 3MT and 4MT instabilities become identical. For larger values of M, the growth rate of the modulational instability is typically larger than that of the corresponding decay instability.

However, for the typical set-up where the primary wave is purely meridional and the modulation is purely zonal, the wave vector $\boldsymbol{q}$ is equally close to both branches of the three-wave resonant manifold. This is because these resonant manifolds cross zero of the $\boldsymbol{q}$-space in the direction of the q_y-axis, i.e. in the zonal direction. Thus, the above speculations about the equivalence of the 3MT and the 4MT for weak waves may not apply to such a set-up. Therefore, let us consider the weakly nonlinear case ($M = 0.1$) and examine (numerically obtained) predictions of the 3MT and the 4MT models and compare them to DNS of the full CHM system, in the following two cases:

(*i*) The primary wave is purely meridional, $\boldsymbol{p} = (10, 0)$, and the modulation is purely zonal, $\boldsymbol{q} = (0, 1)$; and

(*ii*) The primary wave is purely meridional, $\boldsymbol{p} = (10, 0)$, and the modulation is off-zonal. We take $\boldsymbol{q} = (9, 6)$. This is close to the maximum of the most unstable mode on the resonant curve. We cannot select the exact value of the maximum because the discrete wavenumbers in the periodic box do not typically lie exactly on the resonant manifolds. This is a subtlety that can have strong implications for numerical simulations of very weakly nonlinear regimes (Connaughton, Nazarenko & Pushkarev 2001), which we have been careful to avoid here.

Let us first consider case (*i*) when the modulation is purely zonal. Figure 3(*a*) compares $|\Psi_{\boldsymbol{q}}|$ obtained from the solution of (2.3) with that obtained from numerical solutions of 3MT in (3.1), and 4MT in (3.3). The initial condition was constructed from the unstable eigenvector for the decay instability (see the online supplement). We see that the growth rate predicted for the decay instability is not observed. The DNS instead seems to follow the growth rate for the modulational instability. From the zoomed-in plot of the early time evolution shown in figure 3(*b*), we see that the full dynamics very quickly generates the mode $\boldsymbol{p}_+$ that is absent from (3.1). The full system then quickly deviates from the solution of the 3MT in a time of the order of the inverse of the instability growth rate. However, the set of four modes takes much longer to generate any further modes. Thus, in this set-up, the 4MT (3.3) provide a much better description of the full dynamics for times up to 10 instability times.

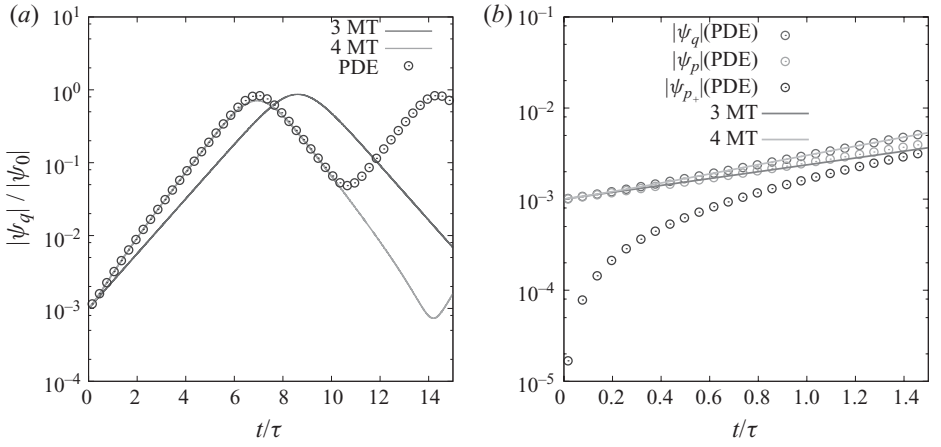

FIGURE 3. Amplitude of the zonal mode with wavenumber $\boldsymbol{q}$ for $M = 0.1$, $\boldsymbol{p} = (10, 0)$ obtained from DNS and from solutions of 3MT and 4MT models. Case (i): purely zonal modulations, $\boldsymbol{q} = (0, 1)$. (*a*) Long-time evolution. (*b*) Zoomed view of early time evolution.

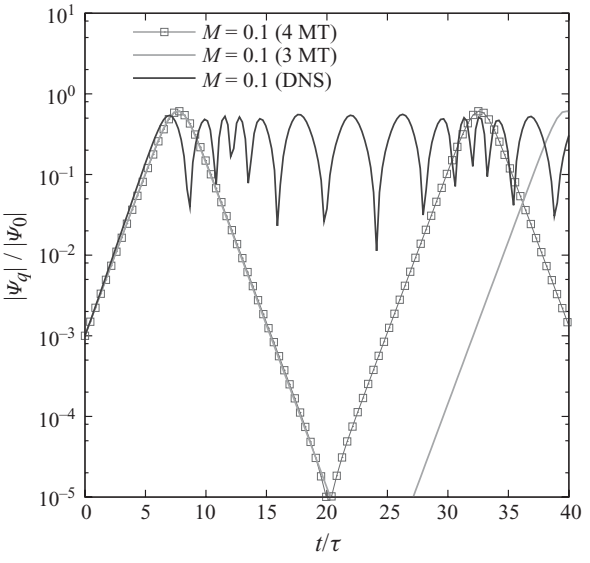

FIGURE 4. Same as in figure 3(*a*) but for case (ii): off-zonal modulations, $\boldsymbol{q} = (9, 6)$.

Let us now consider the case (*ii*) when the modulation is off-zonal. Figure 4 compares $|\Psi_{\boldsymbol{q}}|$ obtained from the solution of (2.3) with that obtained from solutions of 3MT (see (3.1)) and 4MT (see (3.3)) for an initial condition being the unstable eigenvector for the decay instability (see the online supplement). As expected, now the 3MT and 4MT models give practically identical results, and both of these models agree well with DNS up to the time equal to seven inverse growth rates. They predict well the maximum of the zonal jet amplitude, although the subsequent stage of decrease is not described as well as in case (*i*) by the 4MT model.

From these results we conclude that the three-wave interaction is indeed the basic nonlinear process when $M \ll 1$ provided the triad is not degenerate, in the sense that it does not contain quasi-resonant modes that are equidistant from two different

resonant manifolds as happens when the vector $\boldsymbol{q}$ is zonal. In these cases, the 3MT system is just as good as the 4MT and it describes well the full CHM system for over several characteristic times (i.e. the inverse instability growth rates). On the other hand, the most relevant configuration with $\boldsymbol{q}$ zonal is, in fact, degenerate. In this case, however, the 4MT model works well over many characteristic times, whereas the 3MT fails almost immediately. Thus, to have a wider range of applicability, we will study the 4MT model and abandon the 3MT model in the remainder of this paper.

Next, let us study the modulational instability arising from the 4MT in greater detail.

7. Instability for purely meridional primary wave and purely zonal modulation

The case of a purely zonal primary wave ($\hat{\boldsymbol{p}} = (1, 0)$) and purely meridional perturbation ($\hat{\boldsymbol{q}} = (0, 1)$) is of physical interest and simpler. The dispersion relation then reduces to

$$\Omega^3 + \frac{s^4\left[2M^2(1-s^2)(1+F)^2(s^2+F+1) - (s^2+F)\right]}{(1+F)^2(s^2+1+F)^2(s^2+F)}\Omega = 0, \tag{7.1}$$

which has roots

$$\Omega = 0, \tag{7.2}$$

$$\Omega = \frac{\pm \mathrm{i}s^2}{(1+F)(s^2+1+F)}\left[\frac{2M^2(1-s^2)(1+F)^2(s^2+F+1) - (s^2+F)}{s^2+F}\right]^{1/2}. \tag{7.3}$$

The instability question reduces to finding when the quantity under the square root is positive. Recall that s is the ratio q/p of the modulus of the modulation wave vector to that of the primary wave vector. Letting $s^2 = y$, one obtains a quadratic for the quantity under the square root that is positive in the range $s \in (-s_{max}, s_{max})$, where

$$s_{max}^2 = \frac{1+2M^2F(1+F)^2}{2M^2(1+F)^2}\left[-1+\left(1+4\frac{(2M^2(1+F)^3-F)(2M^2(1+F)^2)}{(1+2M^2F(1+F)^2)^2}\right)^{1/2}\right]. \tag{7.4}$$

When $F = 0$ the analysis becomes particularly simple. For any M, there is always a range, $(0, s_{max})$, of unstable long-wavelength perturbations. Note that s_{max} is given by

$$s_{max} = \sqrt{\frac{-1+\sqrt{1+16M^4}}{4M^2}}. \tag{7.5}$$

Within this range, the growth rate is

$$\Omega = \frac{s\left(2M^2(1-s^4) - s^2\right)^{1/2}}{(s^2+1)}. \tag{7.6}$$

The growth rate has a single maximum at $s_0 = \sqrt{y_0}$, where y_0 is the positive root of

$$y^3 + 3y^2 + \left(1 + \frac{1}{M^2}\right)y - 1 = 0. \tag{7.7}$$

One can show that $s_0 \to \sqrt{\sqrt{2}-1}$ as $M \to \infty$ and $s_0 = M + O(M^2)$ as $M \to 0$. One would be interested to know when the maximally unstable meridional perturbation is a local maximum with respect to nearby non-meridional perturbations. To ascertain

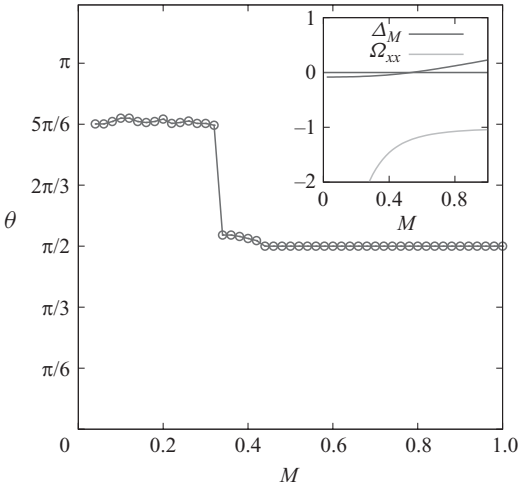

FIGURE 5. Angle θ between the $\boldsymbol{q}$ wave vector of the maximally unstable perturbation and the x-axis as a function of M. Inset, Δ_M and Ω_{xx} as a function of M illustrating the transition of the maximum growth rate for on-axis perturbations from a local maximum to a saddle point at $M \approx 0.53$.

this one should perform the 'second derivative test' for finding critical points of functions of two variables. To do this, we first need to calculate the sign of the determinant

$$\Delta_M(\hat{q}_x, \hat{q}_y) = \begin{vmatrix} \dfrac{\partial^2 \Omega}{\partial^2 \hat{q}_x} & \dfrac{\partial^2 \Omega}{\partial \hat{q}_x \partial \hat{q}_y} \\ \dfrac{\partial^2 \Omega}{\partial \hat{q}_x \partial \hat{q}_y} & \dfrac{\partial^2 \Omega}{\partial^2 \hat{q}_y} \end{vmatrix} \tag{7.8}$$

evaluated at the critical point $(\hat{q}_x, \hat{q}_y) = (0, s_0)$. This can be done semi-analytically using *Mathematica* and is plotted in the inset of figure 5. We find that $\Delta_M > 0$ with $(\partial^2 \Omega)/(\partial^2 \hat{q}_x) < 0$ (the criterion for a local maximum) for $M > M_c$. Also, $\Delta_M < 0$ with $(\partial^2 \Omega)/(\partial^2 \hat{q}_x) < 0$ (the criterion for a saddle) for $M > M_c$. The critical value of M is found numerically to be $M_c \approx 0.534734$. Numerical explorations show that the local maximum found for $M > M_c$ is actually global. For $M > M_c$, therefore, the fastest growing perturbation is indeed zonal. As M decreases below M_c the most unstable perturbation moves to a point with a finite value of q_x. The maximally unstable perturbation for $M < 0.53$ tends to a point on the resonant manifold making an angle of $5\pi/6$ with the x-axis. The dependence of this angle on M is shown in figure 5. A clear transition from an axial maximum to an off-axis maximum is clearly visible. Some additional analysis for the case of finite F is available in the online supplement.

8. Role of the primary wave amplitude

We have discussed the role of the nonlinearity parameter M for the most unstable modulation (i.e. zonal for big M and inclined for small M) as well as for the different regimes in the finite F case. Figure 2 shows plots of the instability growth rate as a function of $\boldsymbol{q} = (q_x, q_y)$ for several different values of M for fixed meridional $\boldsymbol{p}$ and $F = 0$. In particular, we see how the maximum growth rate flips from zonal to off-zonal $\boldsymbol{q}$ when M is reduced (below see also about the collapse of the unstable region to the resonant curve).

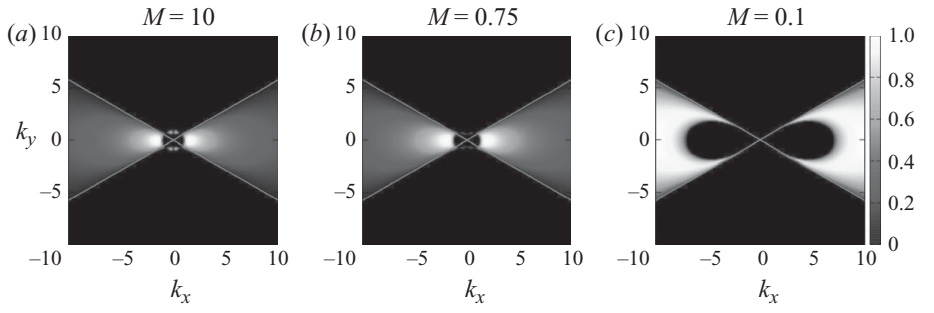

FIGURE 6. Growth rate of the modulational instability (found from (5.2)) as a function of $\boldsymbol{p}$ for a fixed zonal modulation wave vector, $\boldsymbol{q} = (0, 1)$ and $F = 0$. The values of M for the initial primary wave are $M = 10$ (Euler limit), $M = 1$, $M = 3/4$, $M = 1/2$, $M = 1/4$ and $M = 1/10$ (weakly nonlinear limit).

Another natural way to visualize the structure of the set of unstable perturbations is to fix the wave vector of the perturbation mode, $\boldsymbol{q}$, and plot the instability growth rate as a function of the primary wave vector, $\boldsymbol{p}$. Figure 6 does this, plotting the instability growth rate as a function of $\boldsymbol{p} = (p_x, p_y)$ for several different values of M with fixed zonal $\boldsymbol{q}$ and $F = 0$. We see that nonlinearity reduction 'eats into' the instability cone, i.e. makes some wavenumbers inside the cone stable. At the same time, the nonlinearity makes some wavenumbers outside the cone unstable. Note that even for large M, the maximum growth rate occurs outside of the cone, for the primary wave orientations closer to zonal than to the meridional direction; see figure 6 for $M = 10$. This fact is easy to overlook if one considers only the limit $M \to \infty$ (as it is common in the plasma literature) because, in this limit, the growth rate maximum is for the meridional primary waves. On the other hand, the choice of the primary wave direction is often dictated not by the maximum growth rate of the modulational instability, but by the structure of the primary instability creating the Rossy and drift waves (ITG instability in plasmas and the baroclinic instability in GFD).

Let us now specially consider the limits $M \gg 1$ and $M \ll 1$.

8.1. *Limit* $M \gg 1$

The limit of large nonlinearity $M \gg 1$ is a particularly simple and well-studied one (Arnold & Meshalkin 1960; Lorentz 1972; Gill 1974; Manin & Nazarenko 1994; Smolyakov *et al.* 2000; Onishchenko *et al.* 2004). As we mentioned before, the β-effect becomes unimportant and, for $F = 0$, this case reduces to instability of Kolmogorov flow in the Euler equations (i.e. sinusoidal plane-parallel shear). In this case, the most unstable modulation is perpendicular to the primary wave. The instability criterion reduces to (Gill 1974)

$$\cos^2\phi < \left(1 + \frac{q^2}{p^2}\right) \Big/ 4,$$

where ϕ is the angle between $\boldsymbol{p}$ and $\boldsymbol{q}$. For the scale-separated case $q \ll p$, this condition describes an 'instability cone' (Manin & Nazarenko 1994; Smolyakov *et al.* 2000; Onishchenko *et al.* 2004)

$$|\phi| < \pi/6.$$

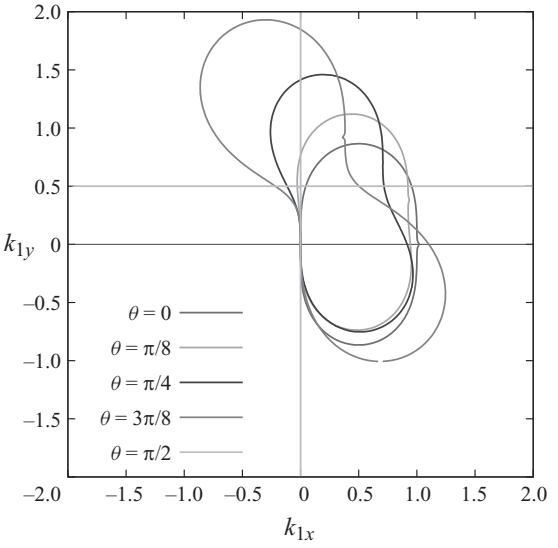

FIGURE 7. Shape of the resonant manifold determined by (8.3) with $\boldsymbol{k}_3 = (\cos(\theta), \sin(\theta))$ for several values of θ for the $F = 0$ case.

A finite deformation radius modifies this cone to a larger instability area (Manin & Nazarenko 1994; Smolyakov *et al.* 2000)

$$F + p_x^2 - 3p_y^2 > 0.$$

We repeat that one has to use the results obtained in the limit $M \to \infty$ with great caution, because even for large M's the most unstable primary wave is not predicted correctly in this limit.

8.2. *Limit* $M \ll 1$

In the limit of weak nonlinearity, $M \ll 1$, the dynamics is completely wave dominated (Gill 1974). The nonlinear terms allow waves to interact weakly and exchange energy. Because the nonlinearity is quadratic, wave interactions are triadic (three-wave resonances are allowed by the dispersion relation in (2.4)). Any triad of waves having wave vectors $\boldsymbol{k}_1$, $\boldsymbol{k}_2$ and $\boldsymbol{k}_3$ interact only if they satisfy the resonance conditions:

$$\boldsymbol{k}_3 = \boldsymbol{k}_1 + \boldsymbol{k}_2, \tag{8.1}$$

$$\omega(\boldsymbol{k}_3) = \omega(\boldsymbol{k}_1) + \omega(\boldsymbol{k}_2). \tag{8.2}$$

Using (2.4), (8.2) gives an implicit equation for the resonant manifold of a given $\boldsymbol{k}_3 = (k_{3x}, k_{3y})$:

$$\frac{k_{1x}}{k_{1x}^2 + k_{1y}^2 + F} + \frac{k_{3x} - k_{1x}}{(k_{3x} - k_{1x})^2 + (k_{3y} - k_{1y})^2 + F} - \frac{k_{3x}}{k_{3x}^2 + k_{3y}^2 + F} = 0. \tag{8.3}$$

Because the system is anisotropic, the shape of resonant manifold depends on the direction of $\boldsymbol{k}_3$ as shown in figure 7.

These resonant manifolds are relevant even for finite nonlinearity because the support of the instability concentrates close to the resonant curves as M is decreased as shown in figure 2. Even for $M = 1$, there is a strong connection between the resonant curves and the shape of the set of modulationally unstable perturbations.

In fact, figure 2 shows two resonant curves corresponding to two resonant triads:

$$\boldsymbol{k}_1 = \boldsymbol{p}, \quad \boldsymbol{k}_2 = \boldsymbol{q} \quad \text{and} \quad \boldsymbol{k}_3 = \boldsymbol{p}_+ = \boldsymbol{p} + \boldsymbol{q}$$

and

$$\boldsymbol{k}_1 = \boldsymbol{p}_- = \boldsymbol{p} - \boldsymbol{q}, \quad \boldsymbol{k}_2 = \boldsymbol{q} \quad \text{and} \quad \boldsymbol{k}_3 = \boldsymbol{p}.$$

Out of the four wavenumbers in our truncated system, $\boldsymbol{p}, \boldsymbol{q}, \boldsymbol{p}_-$ and $\boldsymbol{p}_+$, three are resonant (or nearly resonant) and the remaining one is non-resonant ($\boldsymbol{p}_-$ or $\boldsymbol{p}_+$). As we mentioned before, this picture is correct in non-degenerate situations, when $\boldsymbol{q}$ is not zonal. Then, for $M \to 0$ the amplitude of this non-resonant mode in the instability eigenvector tends to zero, so effectively there are only three active modes, and one can use the results obtained above for the 3MT model. In particular, (4.2) gives the instability growth rate:

$$\gamma = \frac{|\psi_0||\boldsymbol{q} \times \boldsymbol{p}|\sqrt{(p^2 - q^2)(p_+^2 - p^2)}}{\sqrt{(p_+^2 + F)(q^2 + F)}}. \tag{8.4}$$

This expression was previously obtained in the case $F = 0$ by Gill (1974) based on the 4MT model. One can see that instability of the primary wave occurs if its wavenumber length lies in between the wavenumber lengths of the waves it decays into, $q < p < p_+$. This condition has a nice dual-cascade interpretation: to decay the wave must be able to transfer its energy to a large scale and its enstrophy to a smaller scale. For $F = 0$, the typical instability growth rate is $\gamma \sim U_0 p$, where $U_0 = p\psi_0$ is the velocity amplitude of the primary wave (Gill 1974). In the large F case, the instability is slowed by the factor F/p^2 (but not arrested).

Another interesting feature of the instability for $M \ll 1$ is seen in figure 6, where we can see that (for fixed zonal q) the unstable region becomes narrow and collapses onto the sides of the 'cone', i.e. onto the lines $p_y = \pm p_x/\sqrt{3}$. This fact can be explained by considering the resonant curve for $q \ll p$ where it behaves as $q_x = -2(p_x p_y/p^4)q_y^3$. For instability, this curve has to pass as close as possible to the vertical (zonal) axis (where we have chosen our $\boldsymbol{q}$). Thus, we need to minimize the above coefficient $(p_x p_y/p^4)$ (e.g. with respect to p_y for fixed p_x), which immediately gives $p_y = \pm p_x/\sqrt{3}$.

For small M, the maximally unstable modulation $\boldsymbol{q}$ is off-zonal, which may be important for determining the final statistical state of the nonlinear evolution. As we will see later, this state appears to have a predominantly off-zonal component even if the initial modulation is chosen to be zonal.

9. Nonlinear evolution

From now on, we study only systems with infinite deformation radius, $F = 0$. To test the linear predictions and to study the nonlinear evolution, we have performed DNS of the CHM system (2.1), using a standard pseudo-spectral method with resolution up to 1024^2 and hyperviscosity parameters $\nu_n = 4.5\text{e}^{-30}$. We solve numerically the 4MT system (3.3), and compare it with DNS. Although the 4MT was used as the departure point for the linear stability analysis, it is a fully nonlinear set of equations in its own right. In addition to checking the linear instability predictions against DNS, we will also explore the extent to which the nonlinear dynamics of the 4MT capture the behaviour of the full PDE. In all cases, we choose the initial condition to be along the unstable eigenvector of the 4MT.

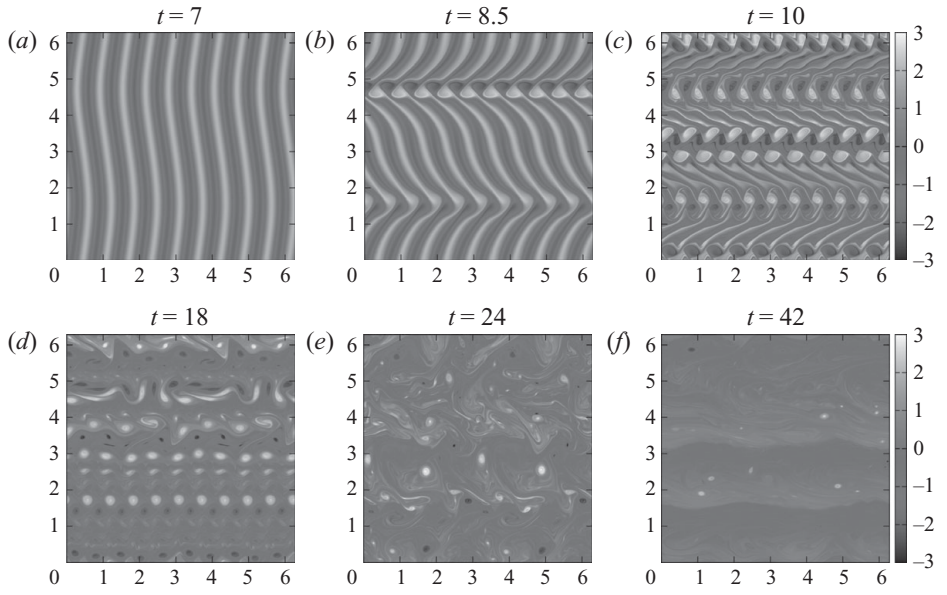

FIGURE 8. Vorticity snapshots showing the growth, saturation and transition to turbulence of a zonal perturbation of a meridional primary wave having $M = 10$.

9.1. *Case of meridional primary wave and zonal modulation*

Let us consider the geometry that we dealt with most: purely meridional primary wave and purely zonal modulation. Some results for non-zonal modulations are available in the online supplement. We choose $\boldsymbol{p} = (10, 0)$ and $\boldsymbol{q} = (0, 1)$. A series of frames of the vorticity field for the cases of strong ($M = 10$), medium ($M = 1$) and weak ($M = 0.1$) nonlinearities obtained by DNS are shown in figures 8, 9 and 10, respectively. The evolution of the mean zonal velocity $\overline{u}(y)$, averaged over x, obtained from DNS for the same set of nonlinearities is shown in figure 11 for times close to the formation of the jet. Finally, evolution of the amplitude, $|\psi_{\boldsymbol{q}}|$, of the zonal mode for the same runs is shown in figure 12. For comparison, we also put the corresponding values of $|\psi_{\boldsymbol{q}}|$ obtained from the 4MT.

Immediately, one can see that the initial stage of evolution agrees very well with the predictions of the linear theory obtained from the 4MT. Moreover, the 4MT works rather well beyond the linear stage, particularly in the $M = 0.1$ case, where the initial growth reverses in agreement with the (periodic) behaviour of the four-mode system. For $M = 1$, the system's growth does not reverse, but rather experiences a saturation at the level where the four-mode system reaches maximum and reverses. The most surprising behaviour is seen for $M = 10$ where the linear exponential growth continues well beyond the point of reversal of the four-wave system, even though the system is clearly nonlinear at these times and follows a self-similar evolution, see below.

9.1.1. *Self-similar jet pinching*

A common feature of the nonlinear saturation stage of the jet growth is self-focusing of the zonal jets that become very narrow with respect to the initial modulation wavelength. This self-focusing cannot be described by the 4MT because such an harmonic jet shape involve strong contributions from higher harmonics $\boldsymbol{p} \pm n\boldsymbol{q}$. For large M and $q \ll k$, such jet 'pinching' was predicted theoretically by Manin & Nazarenko (1994), where self-similar solutions were obtained describing a collapse of the jet width. These strong narrow zonal jets are expected to produce transport

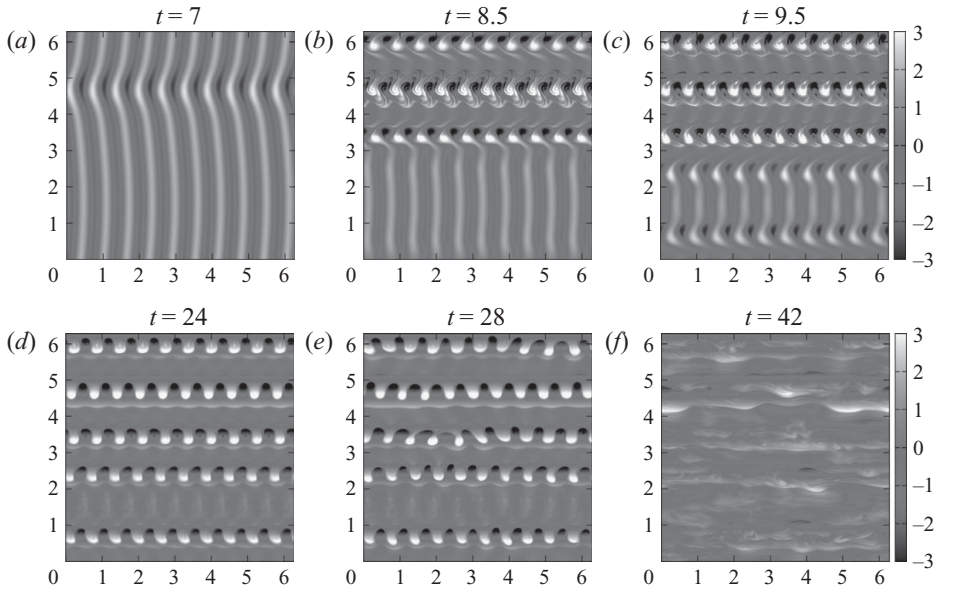

FIGURE 9. Vorticity snapshots showing the growth, saturation and transition to turbulence of a zonal perturbation of a meridional primary wave having $M = 1$.

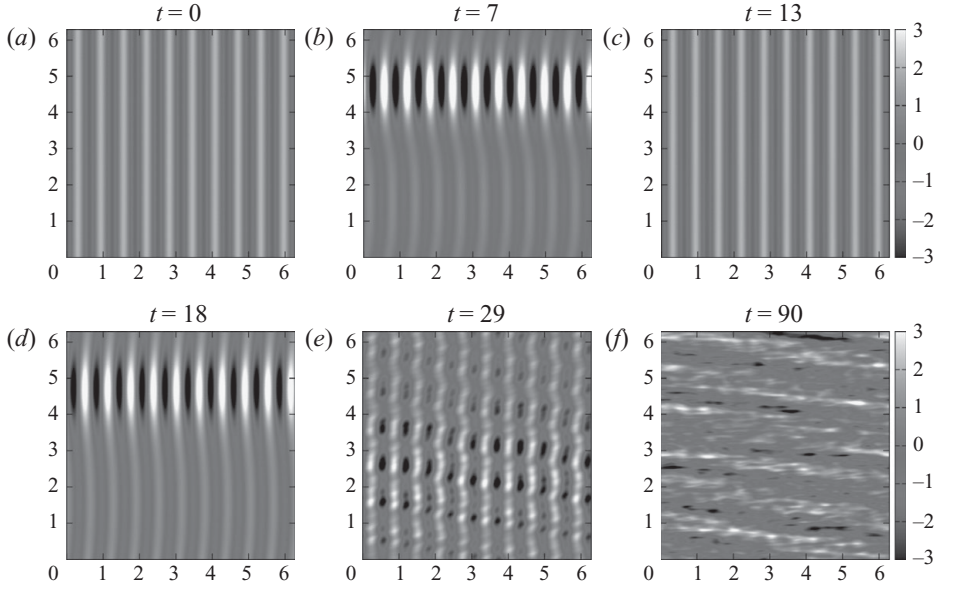

FIGURE 10. Vorticity snapshots showing the growth, saturation and transition to turbulence of a zonal perturbation of a meridional primary wave having $M = 0.1$.

barriers in the transverse (y) direction, which is important in both fusion plasma and the geophysical contexts.

Figure 13 shows the zonal velocity $\overline{u}$ rescaled with self-similar variables as $\overline{u}(y,t) = a(t)\, f(b(t)y)$ in the run with $M = 10$. The self-similar stage occurs in the time interval corresponding to the overshoot in figure 12(a), i.e. after the 4MT has reached its maximum but before DNS saturated at a plateau. Empirically, we obtain $a(t) = u_0\, \mathrm{e}^{\gamma t}$ and $b(t) = \mathrm{e}^{1.85t}$. At least during the decade of amplitude growth

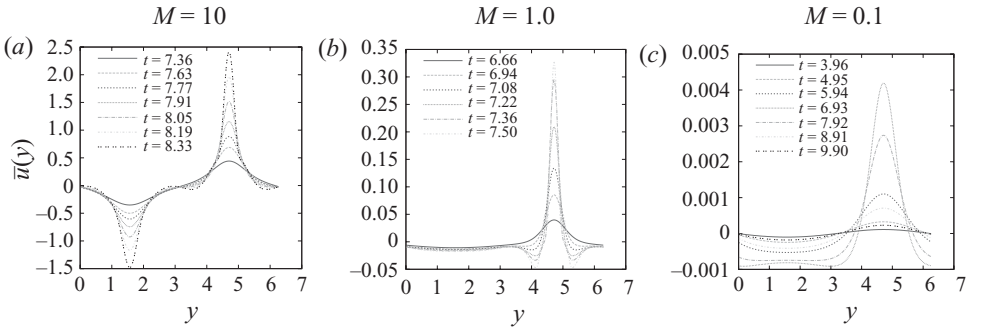

FIGURE 11. Mean zonal velocity profiles for (*a*) $M = 10$, (*b*) $M = 1.0$ and (*c*) $M = 0.1$.

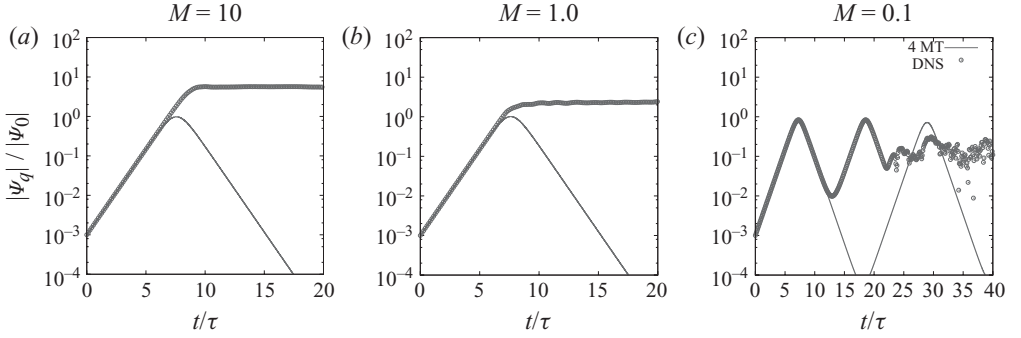

FIGURE 12. Comparison of the growth of the zonal mode $\boldsymbol{q}$ obtained by DNS versus solving 4MT system. In each case, the primary wavenumber is $\boldsymbol{p} = (10, 0)$ and the modulation wavenumber is $\boldsymbol{q} = (0, 1)$. The nonlinearity levels are (*a*) $M = 10$, (*b*) $M = 1.0$ and (*c*) $M = 0.1$. In each case, time has been scaled by τ (the inverse of the instability growth rate).

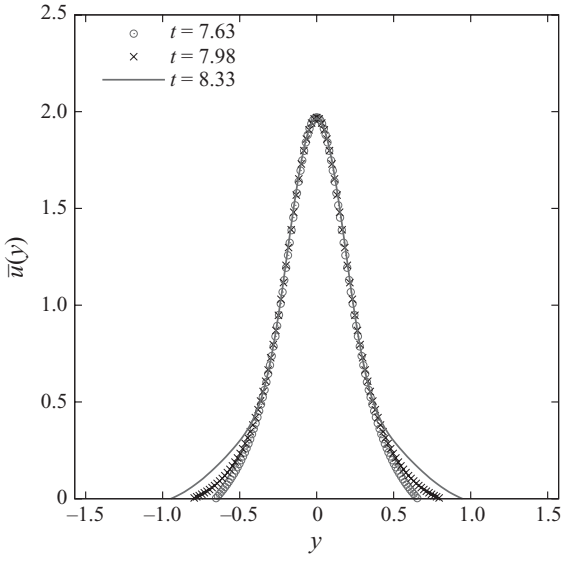

FIGURE 13. Zonal velocity $\overline{u}(y)$ obtained for $M = 10$ rescaled with self-similar variables: $\overline{u}(y, t) = a(t)\, f(b(t)y)$ with $a(t) = u_0 \mathrm{e}^{\gamma t}$ and $b(t) = \mathrm{e}^{1.85t}$ for $M = 10$.

observed before pinching occurs, one can see a good evidence of self-similar behaviour and, remarkably, the nonlinear growth at the self-similar stage follows the same exponential law with growth rate γ as on the linear instability stage. Note that the self-similar solutions were obtained by Manin & Nazarenko (1994) based on the scale separated description and, therefore, the self-similar pinching must stop when the scale separation property breaks down because of the jet narrowing (at which point a roll-up into vortices occurs, see below). In the smaller M runs, the overshoot is absent and the amplitude of the zonal mode decreases after reaching a maximum in correspondence with the solution of the 4MT. The self-focusing is thereby much reduced and the self-similar stage is not clearly observed.

9.1.2. *The role of nonlinearity M*

One can also see some qualitative behaviour differences for different degrees of nonlinearity M. First of all, we see that the east–west asymmetry is larger for weaker waves, which is seen as asymmetry of the top and bottom halves of the vorticity distributions in figures 8–10. This is natural considering that for large nonlinearities the beta-effect, which is the cause of the east–west asymmetry, is less important. Furthermore, we see that for large M the nonlinear evolution is vortex dominated and that the vorticity of the initial primary wave rolls into vortices organized into Kármán-like vortex streets. This corresponds to the moment when the jet velocity reaches a plateau in figures 12(a) and 12(b). On the other hand, in the weak wave case $M \ll 1$, one cannot see vortex roll-ups and the dynamics remain wave dominated.

For large nonlinearities, at the final stages the vortex streets break up because of a vortex pairing instability, which is followed by a transition to turbulence. Such turbulence is anisotropic with a pronounced zonal jet component. At $t = 42$ in figure 8 we can see a well-formed potential vorticity staircase structure as described by Dritschel & McIntyre (2008).

For small nonlinearities, $M \ll 1$, the system's nonlinear evolution starts with self-focusing of the primary wave, but this is followed by a quasi-oscillatory behaviour where the system returns close to the initial state. This is very well reproduced by the four-wave truncation. The same effect was also noted for the generalized Hasegawa–Mima model by Manfredi, Roach & Dendy (2001) and in the atmospheric dynamics context by Mahanti (1981). However, the periodic behaviour is not sustained and a transition to an anisotropic turbulent state occurs. Interestingly, the dominant jet structures observed in such a turbulent state are off-zonal. This effect may be connected to the off-zonal 'striations' reported for the ocean observations by Maximenko *et al.* (2008). We currently regard this connection with some caution since we have not performed any time averaging, whereas the ocean striations are sufficiently weak that they only become evident in the averaged data. We hope to investigate this further in future work.

For $M \sim 1$ or greater, the vortex streets represent the two-dimensional fine structure of the saturated zonal jets (i.e. at the plateau part of figures 12a and 12b). Such vortex street configurations are more stable than the plane parallel (x-independent) flows with the same zonal velocity profiles. This can be understood heuristically (see e.g. McWilliams 2006, chapter 3) by considering the vortices to impart some eddy viscosity to the mean zonal velocity profile. Recall that stability of the latter is determined by the Rayleigh–Kuo necessary instability condition (Kuo 1949),

$$\partial_{yy}\overline{u}(y) - \beta > 0. \tag{9.1}$$

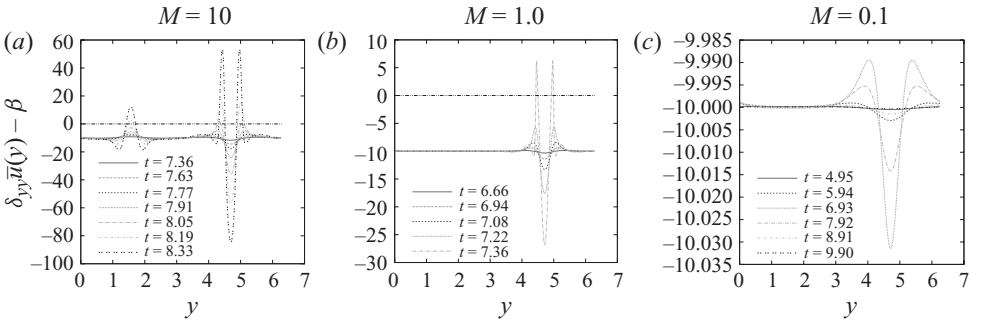

FIGURE 14. The Rayleigh–Kuo profiles for (*a*) $M = 10$, (*b*) $M = 1.0$ and (*c*) $M = 0.1$.

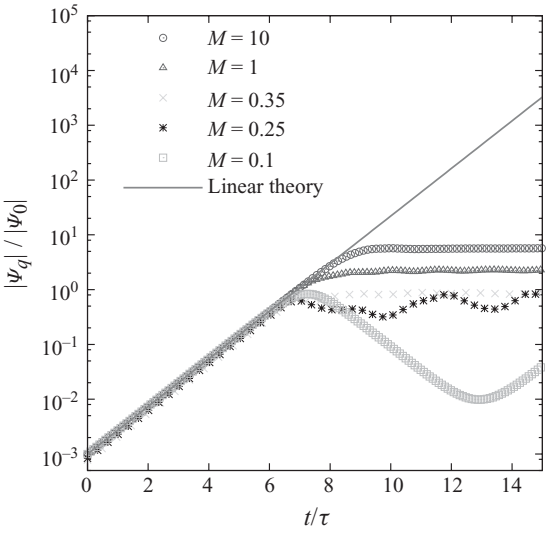

FIGURE 15. Growth of zonal perturbations due to modulational instability of a meridional primary wave having $\boldsymbol{p} = (10, 0)$ for several different values of M. The amplitude of the zonal mode has been scaled by Ψ_0 and time has been scaled by τ (the inverse of the instability growth rate).

Figure 14 plots the profiles of $\partial_{yy}\overline{u}(y) - \beta$ at different moments in time corresponding to runs with $M = 10$, $M = 1$ and $M = 0.1$, respectively. In figures 14(*a*) and 14(*b*) these profiles cross the x-axis (especially far in the $M = 10$ case), which implies that the zonal flows get stronger than the limiting values implied by the Rayleigh–Kuo condition. This is a result of a competition between the instability and the jet pinching process. For large M, the pinching is self-accelerating (self-similar), and it manages to significantly compress/amplify the unstable jet in the finite time needed for the instability to develop (i.e. the inverse growth rate). On the other hand, in the case $M = 0.1$ the jet strength reaches a maximum and then decreases, remaining in the stable range according to the criterion (9.1).

9.1.3. *Critical nonlinearity M_* and breakdown of 4MT*

These results allow us to draw conclusions about the critical value of nonlinearity, $M = M_*$, which separates the two qualitatively different types of behaviours: vortex roll-up and saturation versus the oscillatory dynamics; see figure 15. If the jet strength

maximum, as predicted by the 4MT, exceeds the values of Rayleigh–Kuo necessary instability condition (9.1), then the vortex roll-up occurs and the jet enters into a saturated, relatively long-lived plateau stage. At this moment, the system's behaviour starts to depart from the 4MT model. On the other hand, if the jet strength maximum, as predicted by the 4MT, remains below the Rayleigh–Kuo threshold, then the system's growth reverses and it follows the 4MT dynamics for longer time.

This simple picture leads to a qualitative physical estimate for M_* and the saturated velocity of the jet. Because the x-periodicity is preserved, the step of the vortex street is equal to the wavelength of the original primary wave. The vortices in the stable vortex street are approximately round and the y-spacing between the vortices is approximately the same as the x-spacing. Thus, the saturation width of the pinched jet is of the order of the wavelength of the initial primary wave. Lagrangian conservation of the potential vorticity determines the final saturated amplitude of the jet. Indeed, the same vorticity as in the initial primary wave rolls into the vortices (the βy part of the potential vorticity does not play much role here because the fluid parcels remain at about the same y's) and the rest of the vorticity is shed in between the vortex streets, shredded by shearing and dissipated. Thus, the jet saturation velocity is of the order of the velocity amplitude of the initial primary wave

$$\overline{u}_{max} \sim \frac{M\beta}{p^2}. \tag{9.2}$$

This estimate is well confirmed by our numerical results for $M = 1$ and $M = 10$. Indeed, taking values of u_{max} from figures 11(a) and 11(b) we get

$$\overline{u}_{max} \approx 3\frac{M\beta}{p^2}. \tag{9.3}$$

Now, estimating $\partial_{yy}\overline{u}(y)$ as $p^2 u_{max}$ and using (9.3), we can rewrite the instability condition (9.1) in a very simple form as

$$M > M_* \sim 1/3.$$

Our numerics show that $M_* \approx 0.25$–0.35; see figure 15. Note that the boundary is not sharp. For $M = 0.25$, the dynamics are definitely wave dominated; however, some elongated fuzzy vortices are still apparent, whereas for $M = 0.35$ streets of round vortices are clearly formed with some wave-like oscillations still present.

10. Stable case

We have considered in detail various situations where the linear theory based on the 4MT model predicts instability. We have also investigated the linearly stable case.

For small M, the zonal mode in the modulationally stable case behaves as expected, following the 4MT theory without growth of the mode. In this case, deviations from the 4MT are tiny; hence, we omitted the corresponding graph. For $M \gg 1$, the situation is more interesting. Figure 16 shows the evolution of the zonal mode for the run with $\boldsymbol{p} = (8, 6)$ and $\boldsymbol{q} = (0, 1)$ and for $M = 10$, which corresponds to a linearly stable configuration within the 4MT model. We see agreement with the 4MT stability prediction at early times, i.e. the zonal mode is not growing in figure 16 for $t \lesssim 1$. However, after about one time scale the zonal mode quickly breaks into growth, increasing (more or less exponentially) by two orders of magnitude. Hence, the 4MT instability criterion must be used with caution if $M \gg 1$. Furthermore, for the $M \gg 1$ stable case, Manin & Nazarenko (1994) predicted steepening of the zonal velocity

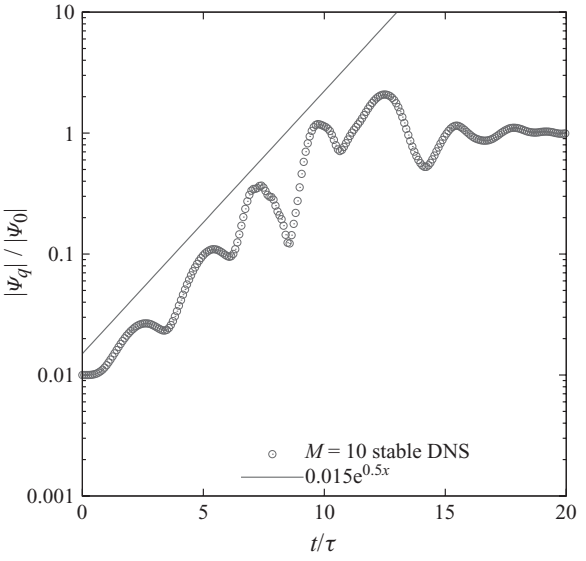

FIGURE 16. Growth of the zonal mode $\boldsymbol{q}$ obtained by DNS for $\boldsymbol{p} = (8, 6)$ and $\boldsymbol{q} = (0, 1)$ and for $M = 10$.

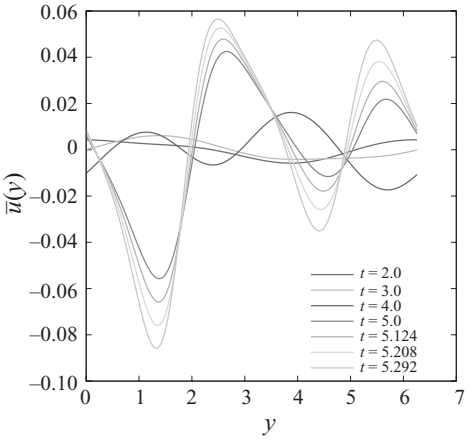

FIGURE 17. Mean zonal velocity profile for the stable configuration $\boldsymbol{p} = (8, 6)$, $\boldsymbol{q} = (0, 1)$ and $M = 10$.

profile, and this is evident in figure 17, where the initial sinusoidal profile develops into a triangular Burgers shock-type profile.

11. Summary and conclusions

In this paper, we dealt with the theory and numerical simulations of the modulational instability of the Rossby/drift waves described by the CHM model. We have revisited the linear theory of Gill (1974) using the four-mode truncation and further elaborated on the role of the primary wave amplitude/nonlinearity, the role of the deformation radius and the role of resonant wave interactions in the case of a weakly nonlinear primary wave. We found a change of the most unstable modulation from zonal to off-zonal when the primary wave nonlinearity parameter M falls below

a critical value, $M > 0.53$. This latter effect may be important for understanding the recent ocean observation of off-zonal jet striations (Maximenko *et al.* 2008). It is also a likely mechanism for generation of off-zonal random jets in our numerical simulations at the late development stages of the modulational instability for the $M = 0.1$ case.

We established how the modulational instability relates to the decay instability obtained within the 3MT in order to clarify the question of whether the dominant nonlinear mechanism of the modulational instability is three-wave or four-wave. By comparison with DNS, we found that 3MT works very well for low nonlinearities M when the primary wave and the modulation belong to the same resonant triad that is non-degenerate, i.e. when it does not include wave vectors too close to $k = 0$ point where the two branches of the resonant curve intersect. However, this excludes the most popular choice of purely zonal modulations, for which 3MT appears to be a bad model. On the other hand, 4MT is more general and it works very well for small nonlinearities M including the case of the purely zonal modulations. Moreover, 4MT also works well for the initial evolution in the strongly nonlinear cases, $M \gtrsim 1$, including the linear growth stage and the prediction of the critical nonlinearity M_* for which transition from the saturated to the oscillatory nonlinear regimes is observed. For $M > M_*$, 4MT description breaks down when the jet breaks up into a vortex street, which is natural because hydrodynamic vortices behave very differently from waves.

After the roll-up the full system enters into a saturated quasi-stable state that persists for a relatively long time but eventually decays because of the presence of hyper-viscosity. On the other hand, the corresponding 4MT system keeps going through an infinite sequence of nonlinear oscillations. If M is small and the roll-ups do not occur (or are delayed) the full system may follow its 4MT counterpart for much longer: its initial growth can reverse and may exhibit the nonlinear oscillations associated with the 4MT.

We would like to emphasize two physical effects that can be important for both plasma and GFD systems. For $M \gtrsim 1$, we observe the formation of stable, narrow zonal jets, in agreement with earlier theoretical predictions of Manin & Nazarenko (1994). As we mentioned, these jets are more stable than one would expect based on the Rayleigh–Kuo criterion alone because their two-dimensional structure consists of stable vortex streets. Such narrow jets represent very effective transport barriers. The second physical effect occurs at low nonlinearities M. This is the fact that the system tends to select the states with somewhat off-zonal jet structures. This tendency to favour the off-zonal structures is seen already on the level of the linear analysis, where as we showed the most unstable modulation changes from zonal to off-zonal when the nonlinearity is reduced. Possibly, this mechanism can explain recent ocean observations of off-zonal jet striations (Maximenko *et al.* 2008) (note that the nonlinearity of the ocean Rossby waves in these situations is likely to be rather low). While the presence of baroclinic dynamics is likely to modify the dynamics, numerical simulations of the two-layer case with subcritical baroclinicity confirm the instability of a baroclinic Rossby wave leading to the formation of off-zonal jets with the orientation predicted by the linear analysis in this setting as well. A detailed discussion of this case is clearly beyond the scope of this paper, other than noting that in this latter baroclinic setting, it would be interesting to establish the relationship of modulational instability to the mechanisms recently proposed by Berloff, Kamenkovich & Pedlosky (2009). Furthermore, the presence of lateral boundaries (continents) is likely to modify the dynamics of modulational instability.

Thus, an extension of the present analysis to the case of a closed basin in terms of basin Rossby modes should clarify the possibility of forming alternating zonal jets (see e.g. Nadiga 2006) through this instability mechanism.

In plasmas, the CHM model used in this paper is also an oversimplification, and one should introduce further modifications in order to be able to make realistic predictions. A widely discussed modification is the introduction to CHM of an extra term subtracting from the field variable its zonal average, in order to model the effect of averaging over the magnetic surface (Dorland *et al.* 1990). Manfredi, Roach & Dendy (2001) and Dewar & Abdullatif (2007) showed that such a modification may have a profound effect on MI. It would be interesting to apply our approach to such a modified CHM system and to quantify the similarities and differences with the original CHM model. In future, it would also be important to study in more detail how the transport properties are affected by both zonal and off-zonal jets that arise in the strongly and weakly nonlinear cases respectively. It would also be interesting to study situations where a broad spectrum of modulations is present initially and, in particular, to verify that the system selects the most unstable one. If the primary wave spectrum is not narrow, it would be of interest to study when the modulational instability wins over the inverse cascade mechanism.

Supplementary material is available at journals.cambridge.org/flm.

REFERENCES

Arnold, V. I. & Meshalkin, L. D. 1960 Seminar led by A. N. Kolmogorov on selected problems of analysis (1958–1959). *Usp. Mat. Nauk* **15**, 247.

Balk, A. M. 1991 A new invariant for Rossby wave systems. *Phys. Lett.* A **155**, 20–24.

Balk, A. M. 1997 New conservation laws for the interaction of nonlinear waves. *SIAM Rev.* **39** (1), 68–94.

Balk, A. M., Nazarenko, S. V. & Zakharov, V. E. 1990*a* Non-local drift wave turbulence. *Sov. Phys. JETP* **71**, 249–260.

Balk, A. M., Nazarenko, S. V. & Zakharov, V. E. 1990*b* On the non-local turbulence of drift type waves. *Phys. Lett.* A **146**, 217–221.

Balk, A. M., Nazarenko, S. V. & Zakharov, V. E. 1991 A new invariant for drift turbulence. *Phys. Lett.* A **152**, 276–280.

Benjamin, T. & Feir, J. 1967 The disintegration of wave trains on deep water. Part 1. Theory. *J. Fluid Mech.* **27**, 417.

Berloff, P., Kamenkovich, I. & Pedlosky, J. 2009 A mechanism of formation of multiple zonal jets in the oceans. *J. Fluid Mech.* **628**, 395.

Bustamante, M. D. & Kartashova, E. 2009 Effect of the dynamical phases on the nonlinear amplitudes' evolution. *Europhys. Lett.* **85** (3), 34002.

Charney, J. G. 1949 On a physical basis for numerical prediction of large-scale motions in the atmosphere. *J. Meteorol.* **6**, 371–85.

Connaughton, C., Nazarenko, S. V. & Pushkarev, A. N. 2001 Discreteness and quasi–resonances in weak turbulence of capillary waves. *Phys. Rev.* E **63** (4), 046306.

Dewar, R. L. & Abdullatif, R. F. 2007 Zonal flow generation by modulational instability. In *Frontiers in Turbulence and Coherent Structures: Proceedings of the CSIRO/COSNet Workshop on Turbulence and Coherent Structures* (ed. J. Denier & J. S. Frederiksen), *World Scientific Lecture Notes in Complex Systems*, vol. 6, pp. 415–430. World Scientific.

Diamond, P. H., Itoh, S.-I., Itoh, K. & Hahm, T. S. 2005 Zonal flows in plasma: a review. *Plasma Phys. Control. Fusion* **47** (5), R35–R161.

Dorland, W., Hammett, G. W., Chen, L., Park, W., Cowley, S. C., Hamaguchi, S. & Horton, W. 1990 Numerical simulations of nonlinear 3-D ITG fluid turbulence with an improved Landau damping model. *Bull. Am. Phys. Soc.* **35**, 2005.

Dritschel, D. G. & McIntyre, M. E. 2008 Multiple jets as PV staircases: the Phillips effect and the resilience of eddy-transport barriers. *J. Atmos. Sci.* **65**, 855–874.

Galperin, B., Nakano, H., Huang, H.-P. & Sukoriansky, S. 2004 The ubiquitous zonal jets in the atmospheres of giant planets and Earth's oceans. *Geophys. Res. Lett.* **31**, L13303.

Gill, A. E. 1974 The stability on planetary waves on an infinite beta-plane. *Geophys. Fluid Dyn.* **6**, 29–47.

Hasegawa, A. & Mima, K. 1978 Pseudo-three-dimensional turbulence in magnetized non-uniform plasma. *Phys. Fluids* **21**, 87.

Horton, W. & Ichikawa, Y.-H. 1996 *Chaos and Structures in Nonlinear Plasmas*. World Scientific.

James, I. N. 1987 Suppression of baroclinic instability in horizontally sheared flows. *J. Atmos. Sci.* **44** (24), 3710.

Janssen, P. A. E. M. 2003 Nonlinear four-wave interactions and freak waves. *J. Phys. Ocean.* **33**, 863–884.

Kartashova, E. & L'vov, V. S. 2007 Model of intraseasonal oscillations in Earth's atmosphere. *Phys. Rev. Lett.* **98** (19), 198501.

Kraichnan, R. H. 1967 Inertial ranges in two-dimensional turbulence. *Phys. Fluids* **10**, 1417–1423.

Kuo, H. L. 1949 Dynamic instability of two-dimensional nondivergent flow in a barotropic atmosphere. *J. Meteorol.* **6**, 105–122.

Lewis, J. M. 1988 Clarifying the dynamics of the general circulation: Phillips's 1956 experiment. *Bull. Am. Meteorol. Soc.* **79** (1), 39–60.

Lorentz, E. N. 1972 Barotropic instability of Rossby wave motion. *J. Atmos. Sci.* **29**, 258–269.

Mahanti, A. C. 1981 The oscillation between Rossby wave and zonal flow in a barotropic fluid. *Arch. Met. Geoph. Biokl.* A **30**, 211–225.

Manfredi, G., Roach, C. M. & Dendy, R. O. 2001 Zonal flow and streamer generation in drift turbulence. *Plasma Phys. Control. Fusion* **43**, 825–837.

Manin, D. Yu. & Nazarenko, S. V. 1994 Nonlinear interaction of small-scale Rossby waves with an intense large-scale zonal flow. *Phys. Fluids* **6** (3), 1158–1167.

Maximenko, N. A., Melnichenko, O. V., Niiler, P. P. & Sasaki, H. 2008 Stationary mesoscale jet-like features in the ocean. *Geophys. Res. Lett.* **35**, L08603.

McWilliams, J. C. 2006 *Fundamentals of Geophysical Fluid Dynamics*. Cambridge University Press.

Mima, K. & Lee, Y. C. 1980 Modulational instability of strongly dispersive drift waves and formation of convective cells. *Phys. Fluids* **23**, 105.

Nadiga, B. T. 2006 On zonal jets in oceans. *Geophys. Res. Lett.* **33** (10), L10601.

Nazarenko, S. & Quinn, B. 2009 Triple cascade behaviour in quasigeostrophic and drift turbulence and generation of zonal jets. *Phys. Rev. Lett.* **103** (11), 118501.

Onishchenko, O. G., Pokhotelov, O. A., Sagdeev, R. Z., Shukla, P. K. & Stenflo, L. 2004 Generation of zonal flows by Rossby waves in the atmosphere. *Nonlinear Proc. Geophys.* **11**, 241–244.

Onorato, M., Osborne, A. R., Serio, M. & Bertone, S. 2001 Freak waves in random oceanic sea states. *Phys. Rev. Lett.* **86**, 5831.

Rhines, P. 1975 Waves and turbulence on a betaplane. *J. Fluid Mech.* **69**, 417–443.

Rudakov, L. I. & Sagdeev, R. Z. 1961 On the instability of a non-uniform rarefied plasma in a strong magnetic field. *Sov. Phys. Dokl.* **6**, 415.

Sagdeev, Z. & Galeev, A. A. 1969 *Nonlinear Plasma Theory*. Benjamin.

Sanchez-Lavega, A., Rojas, J. F. & Sada, P. V. 2000 Saturn's zonal winds at cloud level. *Icarus* **147**, 405.

Simon, A. A. 1999 The structure and temporal stability of Jupiter's zonal winds: a study of the north tropical region. *Icarus* **141**, 29.

Smolyakov, A. I., Diamond, P. H. & Shevchenko, V. I. 2000 Zonal flow generation by parametric instability in magnetized plasmas and geostrophic fluids. *Phys. Plasmas* **7**, 1349.

Wagner, F., Becker, G., Behringer, K., Campbell, D., Eberhagen, A., Engelhardt, W., Fussmann, G., Gehre, O., Gernhardt, J., Gierke, G. V., Haas, G., Huang, M., Karger, F., Keilhacker, M., Klüber, O., Kornherr, M., Lackner, K., Lisitano, G., Lister, G. G., Mayer, H. M., Meisel, D., Müller, E. R., Murmann, H., Niedermeyer, H., Poschenrieder, W., Rapp, H. & Röhr, H. 1982 Regime of improved confinement and high beta in neutral-beam-heated divertor discharges of the ASDEX tokamak. *Phys. Rev. Lett.* **49** (19), 1408–1412.