Large-scale magnetohydrodynamic density-wave structures in the Andromeda nebula

Yu-Qing Lou* and Zuhui Fan

Department of Astronomy and Astrophysics, The University of Chicago, Chicago, IL 60637, USA

Accepted 2000 January 31. Received 2000 January 31; in original form 1998 January 12

ABSTRACT

The Andromeda galaxy (also referred to as M31 or NGC 224) shows approximately axisymmetric large-scale structures which are identified here with fast magnetohydrodynamic (MHD) density waves in a simple theoretical model characterized by a rotating composite system of a massive spherical halo (dark matter included), a stellar disc and a magnetized gas disc. For fast MHD density waves inside the Lindblad resonance, enhancements of stellar and gas surface mass densities as well as the azimuthal magnetic field are all roughly in phase. Based on this analysis and considerations of radiation processes in a magnetized interstellar medium, the observed overall correlation of annulus structures of M31 at many electromagnetic wavelengths can be naturally explained. In particular, the enhanced magnetic field (inferred from the synchrotron radio continuum) tends to align roughly along the magnetic torus of M31 as revealed by radio polarization studies. From observational estimates of physical parameters, the composite system of M31 is found to be quite stable—a fact that may account for the observed low global star formation rate in M31.

Key words: MHD – polarization – waves – stars: formation – galaxies: individual: Andromeda – galaxies: structure.

1 INTRODUCTION

The nearby Sb galaxy Andromeda nebula (also referred to as M31 or NGC 224) swirls like a maelstrom in visible light with an impressive central bulge and a number of luminous flocculent arms embedded with narrow dark dust lanes. At other electromagnetic wavelengths such as infrared bands (Habing et al. 1984; Battaner et al. 1986; Walterbos & Schwering 1987; Devereux et al. 1994), Hα (Arp 1964; Baade & Arp 1964; Devereux et al. 1994), ultraviolet (Deharveng et al. 1980), H II complexes (Baade & Arp 1964; Pooley 1969; Emerson 1974; Pellet et al. 1978), 21-cm emission from neutral hydrogen H I (Emerson 1974; Brinks & Shane 1984; Braun 1990a, 1991), synchrotron radio continuum (Pooley 1969; Berkhuysen & Wielebinski 1974; Berkhuysen 1977; Beck 1982; Berkhuysen, Wielebinski & Beck 1983; Beck, Berkhuysen & Hoernes 1998) and millimetre bands (Koper et al. 1991; Dame et al. 1993; Koper 1993; Devereux et al. 1994), a prominent yet largely axisymmetric annulus or torus1 with a radius of ~9–10 kpc and a slight warp surrounds the centre of M31 which is ~690 kpc away from us. Extensive comparison studies (e.g. Emerson 1974; Berkhuysen 1977; Deharveng et al. 1980; Habing et al. 1984; Walterbos & Kennicutt 1987; Dame et al. 1993; Koper 1993; Devereux et al. 1994) show that these ‘ring’ structures more or less coincide with salient optical arms 4 and 5 of M31 as numbered by Baade & Arp (1964), whereas ‘ring’ structures are barely detectable in the old stellar disc by red-light and near-infrared observations (Hiromoto, Maihara & Oda, 1983; Walterbos & Kennicutt 1988; Battaner et al. 1986; Walterbos & Schwering 1987). This latter fact indicates that large-scale density waves in the stellar disc are fairly weak in M31. Polarized synchrotron radio observations (Beck, Berkhuysen & Wielebinski 1980; Beck 1982; Beck et al. 1989; Braun 1990b; Beck et al. 1998) reveal a large-scale ordered magnetic field of a few μG in strength roughly aligned along the torus with some deviations in the inner disc (Ruzmaikin et al. 1990). Recently, local radio (at 73-, 20-, 11- and 6-cm wavelengths) and far-infrared (FIR; at 60- and 100-μm wavelengths) structural correlations within M31 have been clearly revealed by systematically examining Very Large Array (VLA)–Effelsberg and high-resolution Infrared Astronomical Satellite (IRAS) maps (Hoernes, Berkhuysen & Xu 1998). In particular, a non-thermal radio/cool dust correlation exists in addition to the thermal radio/warm dust correlation. All these empirical facts together indicate that a global process exists to orchestrate such types of large-scale structural

* E-mail: lou@oddjob.uchicago.edu; lou@asiaa.sinica.edu.tw

1 Systematic deviations from axisymmetry are mainly related to extremely tight spiral arms as seen in certain wavelengths with small arm pitch angles of <10°.
correlations. Meanwhile, different observational diagnostics involve specific radiation mechanisms that reflect the concurrence of various distinct sub-processes.

For theoretical interpretations of these observations there are two different perspectives, namely those of kinematic dynamo and those of magnetohydrodynamic (MHD) density waves. Given a specified disc rotation curve for M31, kinematic dynamo theories favour a dominant axisymmetric growth (Ruzmaikin, Sokoloff & Shukurov 1985; Beck et al. 1996; Moss et al. 1998) of seed magnetic fields in a gas disc alone; in this scenario, the stellar disc and self-gravity are not directly involved. By comparison, the large-scale dynamics of MHD density waves in a rotating composite system of stellar and magnetized gas discs coupled through the mutual gravitational interaction appears natural to explain the observed large-scale structural interrelations of M31. In this paper, we present theoretical analyses from the MHD density wave point of view (Fan & Lou 1996; Lou & Fan 1997, 1998a) using the standard MHD formalism and indicate the main differences in comparison with kinematic dynamo models. Even though a kinematic dynamo and the MHD density wave scenario are distinctly different in many respects, we feel that they are two important relevant aspects in galactic dynamics. If one thinks carefully about MHD equations, it is the intrinsic non-linearity of the problem that prevents the direct connection between these two complementary aspects at this stage. Finally, we also discuss the global star formation rate in M31 from the perspective of local instabilities in a magnetized composite disc system.

Generally speaking, large-scale structural interrelations of spiral galaxies as observed in many electromagnetic wavelengths must have a large-scale dynamic basis (cf. Beck et al. 1999 for observations of NGC 1097). Empirically, the neutral hydrogen Hı disc of a spiral galaxy is typically much larger than the optical disc (cf. Bureau et al. 1999 for NGC 2915 as an example). In particular, the Hı emission usually reveals a much larger somewhat fuzzy spiral pattern, which correlates well with the optical spiral pattern in the inner-disc region. As star formation tends to be more active along regions of high gas concentration, it is sensible to think that the Hı spiral density wave tends to be stronger towards the inner-disc region where numerous bright stars collectively give rise to the optical spiral pattern. In this scenario, the large-scale spiral structure of Hı gas emission is an inseparable component of the overall density-wave structure. Star formation activities gradually fade along spiral gas arms away from the galactic centre. As perturbation enhancements of magnetic fields and gas density can correlate with each other, it can happen that these two enhancements continue while optical features fade along spiral arms. The grand-design spiral galaxy NGC 2997 in the southern sky serves as such an example (Han et al. 1999; Lou, Han & Fan 1999). From the perspective of fast MHD density waves, we proposed recently (Lou et al. 1999) that the isolated arm of synchrotron emission observed in the south-east quadrant of NGC 2997 should be accompanied by an Hı arm. There is now preliminary evidence that this is indeed the case. In the case of M51, density waves are fairly strong such that large-scale Hı spiral structure is only moderately larger than the optical spiral structure (cf. figs 2 and 3 of Rots et al. 1990).

Nearly edge-on, with a high inclination angle of ~78° to the plane of the sky, M31 is a frequent target of observations at many electromagnetic wavelengths owing to its proximity (1 arcmin corresponds to ~200 pc). The disc rotation speed of M31 is generally ~230 km s^{-1} and the rotation curve appears more or less flat\(^2\) for \(r \geq 5\) kpc with a gradual decline towards the outer-disc portion (e.g. Braun 1991; Vallée 1994; Tenjes, Haud & Einasto 1994). The disc portion of M31 is fairly thin with an estimated thickness of ~1400 pc in radio continuum (Berkhuijsen 1977) and of ~500 pc in Hı emission (Emerson 1974). M31 may resemble our own Galaxy — the Milky Way — and large-scale annulus structures (in molecular bands) have also been detected in other nearby disc galaxies such as NGC 2841 and NGC 7331 (e.g. Young & Scoville 1982; Prieto et al. 1985). Therefore, the study of M31 is also significant for other similar galaxies.

Large-scale optical structures and dynamics of spiral galaxies have been extensively studied for more than three decades in the framework of density waves (Lin & Shu 1964; Toomre 1964; Lin 1967, 1987; Binney & Tremaine 1987; Bertin & Lin 1996), primarily in a thin rotating stellar disc. In order to decipher large-scale spiral structures observed at all available electromagnetic wavelengths and especially their well-organized overall interrelations, it is necessary and important for diagnostic as well as dynamic purposes to take into account large-scale dynamics of a magnetized gas disc together with density waves in a massive stellar disc. In the context of the global star formation rate (e.g. Kennicutt 1989; Elmegreen 1994; Silk 1997), the role of the galactic magnetic field cannot be overemphasized (Lou 1996; Lou & Fan 1998b). At the qualitative level, the density-wave scenario is conceptually consistent with the empirical findings that magnetic field strength and gas density are correlated in spiral galaxies (e.g. Berkhuijsen 1979; Niklas & Beck 1997). Prompted by optical and radio observations of the Whirlpool galaxy (M51 or NGC 5194; Neiningger 1992; Neiningger & Hollerrou 1996) and NGC 6946 (Beck & Hoernes 1996), we recently proposed the applications of fast and slow MHD density waves in a thin rotating magnetized gas disc for distinctly different interrelations of optical—radio spiral structures (Fan & Lou 1996; Lou & Fan 1998a). For the ring-like structures of M31 at various wavelengths, we now take the first step in analysing a theoretical model problem of axisymmetric MHD density waves in a rotating composite system of stellar and magnetized gas discs. Theoretically, this is a natural generalization of the analyses of Jog & Solomon (1984a,b) by including a magnetic field.

2 THEORETICAL MODEL ANALYSIS

M31 can be effectively modelled as a thin rotating composite disc system plus a central bulge and a massive spherical halo. While the inferred magnetic field based on polarized radio-continuum emissions (Beck et al. 1980, 1998) is more or less oriented along the torus, recent observations of Han, Beck & Berkhuijsen (1998) also tentatively indicate the presence of ring-like magnetic fields in M31 outside the annulus through rotation measures of 21 polarized background radio sources. The weak radio-continuum emission outside the annulus is attributed to a low cosmic-ray electron density. For a thin background axisymmetric gas disc in differential rotation, there threads an azimuthal magnetic field \(B_φ \propto r^{-1}\) from the divergence-free and force-free conditions, where the cylindrical coordinates \((r, \theta, z)\) are adopted. Such a \(B_φ\) configuration avoids the well-known winding dilemma in a disc system with a strong differential rotation (cf. Roberts & Yuan 1970). The curve of rotation speed \(V_φ(r) = rΩ(r)\) is taken to be

\(^2\)We are mainly concerned with gross features of the rotation curve averaged over a spatial scale compatible with that of relevant MHD density waves.
roughly the same for both stellar and gas discs, where $\Omega(r)$ is the sidereal disc angular rotation rate determined by the total mass contained in the entire galactic system, while the background stellar and gas surface mass densities, $\mu_S^0(r)$ and $\mu_g^0(r)$, are allowed to be different. In view of the large disc aspect ratio, the disc portion is assumed to be infinitely thin as an expedient approximation\(^3\) to the real disc system of M31. For the stellar disc part we adopt the more simple fluid description (cf. Binney & Tremaine 1987 and Bertin & Lin 1996), because the results derived from a fluid or a distribution function approach are expected to be more or less similar sufficiently far from the corotation and Lindblad resonances. In the magnetized gas disc part we invoke the MHD approximation (Lynden-Bell 1966; Roberts & Yuan 1970; Fan & Lou 1996; Lou & Fan 1998a) for large-scale perturbations, as magnetic field exerts a dynamic influence on neutral gas via sufficiently frequent collisions from charged particles. In this simple theoretical model, large-scale MHD density-wave perturbations in the stellar and magnetized gas discs are coupled through the mutual gravitational interaction as described by Poisson’s equation (Lin & Shu 1966; Jog & Solomon 1984a,b; Bertin & Romeo 1988; Elmegreen 1995; Jog 1996; Lou & Fan 1998b). In short, our procedure of analysing large-scale MHD density waves in a composite system of stellar and magnetized gas discs can be succinctly summarized as a fluid-MHD formalism.

Two parallel sets of linearized density-wave equations are involved for two-dimensional perturbations tangential to the composite disc plane, namely one set for hydrodynamic perturbations in the massive stellar disc (cf. equations 2.1–2.4) and the other set for MHD perturbations in the magnetized gas disc (cf. equations 2.6–2.10); the mutual gravitational coupling is dictated by Poisson’s equation (cf. equation 2.5). For the annulus structure of M31, we consider axisymmetric perturbations for the sake of simplicity. The current analysis can be readily extended to non-axisymmetric MHD perturbations if required. Note also that incompressible Alfvénic fluctuations involve perturbations perpendicular to the disc plane without disturbing the gravitational potential in the linear regime; thus, they can be consistently set to vanish in the present context. The radial and azimuthal velocity perturbations in the stellar disc are denoted by $v^\phi$ and $v^r$, the perturbation in $\mu_S^0$ by $\mu^S$, the enthalpy perturbation by $h^S$ and the associated gravitational potential perturbation by $\phi^S$, where the superscript $S$ indicates the association with a stellar disc (cf. Binney & Tremaine 1987). Likewise, $v_r$, $v_\phi$, $\rho$, $h$ and $\phi$ stand for the corresponding perturbation variables in the magnetized gas disc; in addition, the azimuthal magnetic field perturbation $b_\phi$ is dynamically involved in the momentum equation through the Lorentz force and its time evolution is governed by the magnetic induction equation.\(^4\) For large-scale density waves, the exp(\text{i}ot) dependence is assumed for all these perturbation variables, where $\omega$ is the angular frequency in the sidereal frame of reference.

The equations for density wave perturbations in the stellar disc are given by

$$i\omega v^S_\phi - 2\Omega v^S_\phi = -\frac{d}{dr} (\phi^S + \phi + h^S),$$

(2.1)

where $\kappa = (2\Omega/r)[d(r^2\Omega)/dr]$ defines the epicyclic frequency $\kappa$, $D_\phi$ is the radial velocity perturbation, $\rho$ is the mass density and $\mu^S$ is the azimuthal magnetic field perturbation.

The radial magnetic field perturbation vanishes for axisymmetric perturbations (Lou & Fan 1998a).

\(^3\)This restriction can be relaxed as required (Shu 1968; Vandervoort 1967, 1970a,b; Romeo 1992; Jog 1996).

\(^4\)The radial magnetic field perturbation $b_r$ vanishes for axisymmetric perturbations (Lou & Fan 1998a).

\(^6\)The radial magnetic field perturbation $b_r$ vanishes for axisymmetric perturbations (Lou & Fan 1998a).

\(^6\)The radial magnetic field perturbation $b_r$ vanishes for axisymmetric perturbations (Lou & Fan 1998a).
Fast MHD density waves of Andromeda

which generalizes the usual $Q$ parameter (cf. Safronov 1960; Toomre 1964) by incorporating the Alfvén wave speed $C_A$.

On the other hand, by allowing the second factor on the left-hand side of equation (2.18) to vanish, that is

$$\omega^2 = \kappa^2 + \left(D_S^2 - \frac{2\pi G\mu_0^5}{|k|}\right)k^2, \tag{2.21}$$

one obtains the well-known dispersion relation for hydrodynamic density waves in a thin stellar disc alone (Lin & Shu 1964). For the stellar disc alone to be locally stable for all $|k|$, one returns to the Toomre stability criterion $Q_S^2 = \kappa^2 D_S^2/(|\pi G\mu_0^5|^2) > 1$ where $Q_S$ is the stellar $Q$ parameter (Safronov 1960; Toomre 1964).

In a compact form, the necessary and sufficient condition for equation (2.18) to have two positive roots of $\omega^2$ is then

$$(D_S - 1)(D_G - 1) > 1, \tag{2.22}$$

where

$$D_S = \frac{\kappa^2 + k^2 C_S^2}{2\pi G\mu_0^5 |k|}, \tag{2.23}$$

and

$$D_G = \frac{\kappa^2 + k^2 (C_A^2 + C_L^2)}{2\pi G\mu_0^5 |k|}.$$  \tag{2.24}

Condition (2.22) necessarily implies $D_S > 1$ and $D_G > 1$; the former guarantees that the magnetized gas disc alone is locally stable against fast MHD density wave disturbances, while the latter guarantees that the stellar disc alone is locally stable against hydrodynamic density wave perturbations. One important observation is that while both $D_S > 1$ and $D_G > 1$, the inequality (2.22) may not be satisfied. In other words, the composite disc system is more prone to instability as a result of the mutual gravitational coupling between the two discs. This point was first brought out by Jog & Solomon (1984a,b) in the case without magnetic field.

Comparing with earlier analyses\(^5\) (Jog & Solomon 1984a,b; Bertin & Romeo 1988; Elmegreen 1995; Jog 1996; Lou & Fan 1998b), the modification in (2.18) is to simply replace the sound speed $C_S$ by the magnetosonic speed $(C_S^2 + C_L^2)^{1/2}$. This can be readily understood physically as the presence of $B_\theta$ giving rise to an additional pressure effect. With this key parameter adjustment, the detailed stability analyses (cf. Elmegreen 1995; Jog 1996; Lou & Fan 1998b) to determine $Q_{\text{eff}}$ – the effective Toomre $Q$ parameter for a composite disc system – are directly applicable here. In other words, an effective magnetic $Q$ parameter, namely $Q_{\text{M,eff}}$, can be readily determined. The main conclusion is that an azimuthal background magnetic field $B_\theta$ tends to stabilize a composite disc system against axisymmetric perturbations as compared with the corresponding hydrodynamic counterpart (without a magnetic field). As this type of theoretical analysis has been invoked in recent years to understand global star formation rates in spiral galaxies, the inclusion of a magnetic field here is natural and important. In particular, our analysis here is relevant to the fact that star formation activities are somewhat low in M31 as compared to other disc galaxies.

Dispersion relation (2.18) actually contains two MHD

\(^5\) Lin & Shu (1966, 1968) first derived the dispersion relation for density waves in a composite system of a stellar disc (described by a stellar distribution function) and a gas disc (described by a fluid approach).
density-wave modes for $\omega^2$, namely
\[ \omega^2 \equiv [2\kappa^2 + k^2(D^2_S + C^2_S + C^2_\Lambda) - 2\pi G(\mu_0 + \mu^S_0)]/2 \]
\[ \pm (k^2/2)[(D^2_S - C^2_S - C^2_\Lambda - 2\pi G(\mu^S_0 - \mu_0)/|k|)]^2 \]
\[ + 16\pi^2G^2\mu^S_0\mu_0/k^2]^{1/2}, \tag{2.25} \]
where the plus- and minus-sign solutions correspond to MHD density waves outside and inside the Lindblad resonance, respectively. This mode number doubling is not surprising and can be readily understood by the close analogy of plasma waves in a fully ionized plasma in which electrons and ions are coupled through the electric force. From dispersion relation (2.25), the radial group speeds of MHD density waves can be derived, namely
\[ \frac{d\omega}{dk} \equiv \frac{|k|(D^2_S + C^2_S + C^2_\Lambda - \pi G(\mu_0 + \mu^S_0))}{2\omega} \]
\[ \pm \frac{1}{2\omega} \left[ (D^2_S - C^2_S - C^2_\Lambda)\|k^2 - 2\pi G(\mu^S_0 - \mu_0)\|k^2 \right] \]
\[ + 16\pi^2G^2\mu^S_0\mu_0\|k^2\|^{-1/2} \times \left[ (D^2_S - C^2_S - C^2_\Lambda)\|k^2 - 2\pi G(\mu^S_0 - \mu_0)\| \right] \]
\[ \times \left[ (D^2_S - C^2_S - C^2_\Lambda)\|k - \pi G(\mu^S_0 - \mu_0)\| \right] \]
\[ + 8\pi^2G^2\mu^S_0\mu_0\|k\|. \tag{2.26} \]
where $\omega$ is given by equation (2.25).

By direct substitution of solutions (2.25) into equation (2.16), one derives
\[ \frac{-2\pi G\mu_0\mu^S}{|k|} = \frac{\mu^S}{2} \left\{ \left[ D^2_S - 2\pi G\mu^S_0 \|k\| \right] + \left[ C^2_S + C^2_\Lambda - 2\pi G\mu_0 \|k\| \right] \right\} \]
\[ \pm \frac{\mu^S}{2} \left\{ \left( D^2_S - 2\pi G\mu^S_0 \|k\| \right) - \left( C^2_S + C^2_\Lambda - 2\pi G\mu_0 \|k\| \right) \right\}^2 \]
\[ + \frac{16\pi^2G^2\mu^S_0\mu_0}{k^2} \right]^{1/2}, \tag{2.27} \]
where $\pm$ signs correspond to $\mu \propto \pm \mu^S$. We have thus shown explicitly that within the Lindblad resonance enhancements of $\mu$ and $\mu^S$ track each other for the minus-sign solution, while outside the Lindblad resonance enhancements of $\mu$ and $\mu^S$ actually anti-correlate with each other for the plus-sign solution. The physical interpretation is as follows. When $\mu$ and $\mu^S$ track each other, the effect of self-gravity is enhanced and the MHD density wave speed becomes slower. In contrast, when $\mu$ and $\mu^S$ anti-correlate with each other, the effect of self-gravity reduces and the MHD density wave speed becomes faster (cf. Lou & Fan 1998b for the hydrodynamic counterparts of these effects). These perturbation mass density correlations also remain valid during the initial growth (that is, $\omega^2 < 0$ and $|\omega| < \kappa$ in the regime of small perturbation magnitudes) of instabilities associated with MHD density waves. Note that for realistic situations, $\mu^S$ and $\mu$ can be comparable in magnitude even though $\mu^S_0$ is typically an order of magnitude larger than $\mu_0$. As long as $\mu^S_0/\mu_0$ is sufficiently small such that $\mu/\mu_0$ remains small, the linear approximation can be consistently valid (see estimates for M31 in Section 4).

From equations (2.6)–(2.10), it follows that
\[ \frac{b_v/B_0}{b_0} = \frac{k^2[2\pi G(\mu^S + \mu)/|k| - C^2_\Lambda\mu_0]}{k^2 + k^2C^2_\Lambda - \omega^2}. \tag{2.28} \]
From the magnetic field perturbation relation (2.28), one can further show, for both solutions of (2.25), that
\[ b_v/B_0 \equiv \mu/\mu_0, \tag{2.29} \]
which indicates gross in-phase enhancements of $b_v$ and $\mu$. As $v_S^v \equiv -(\omega/k)(\mu^S/\mu_0)$ from equation (2.3) and $v_v \equiv -(\omega/k)(\mu/\mu_0)$ from equation (2.9) for large $k$, it follows that $v_S^v$ and $v_v$ are out of and in phase for the plus- and minus-sign solutions of (2.25), respectively. Likewise, as $v_S^v = i\omega v_S^v/(2\omega k)$ from equation (2.2) and $v_v = i\omega v_v/(2\omega k)$ from equation (2.8), it then follows that $v_S^v$ and $v_v$ are out of and in phase for the plus- and minus-sign solutions of (2.25), respectively. Velocity perturbations in the stellar disc are typically smaller in magnitude than velocity perturbations in the magnetized gas disc. While $v_S^v$ and $v_v$ (or $v_S^v$ and $v_v'$) have a phase difference of $\pi/2$, their magnitudes can be roughly comparable. We specifically note these phase relationships of velocity perturbations as they can be determined observationally via Doppler shifts (e.g. Visser 1980a,b).

In the development of a hydrodynamic density wave theory (without magnetic field), one important aspect is to specifically relate the theory to various available galactic observations. Except for observations in red light and near-infrared bands that reveal large-scale spiral structures in the stellar disc, most other available observational diagnostics involve directly and indirectly the magnetized interstellar gas medium, i.e. the gas disc. There have been conscious efforts to include interstellar gas effects (either in a dynamically consistent manner or in an approximate manner by treating the gas response passively) in the past. The main motivation was to establish the relation between velocity deviations and column density perturbations of neutral hydrogen $H_1$ gas associated with hydrodynamic density waves theoretically. It is then possible to compare theoretical results with $H_1$ observations of Doppler shifts and column density (cf. Lin, Yuan & Shu 1969; Visser 1980a,b). This research programme was successful in supporting the basic theoretical idea of large-scale density waves in spiral galaxies. The main contribution of our MHD density wave analysis is the inclusion of large-scale magnetic field effects which have been detected through synchrotron radio-continuum emissions. As perturbations of stellar density, gas density, velocity and magnetic field are related to each other in the theoretical framework of MHD density waves in a composite disc system, one can test these specific interrelations observationally.

In order to relate these results to the observed annulus of M31, we identify the roughly ring-like structures of M31 with large-scale fast MHD density waves within the Lindblad resonance (i.e. the minus-sign solution of equation 2.25) from the following considerations and qualifications. First, all Population I signatures more or less coincide (Arp 1964; Baade & Arp 1964; Emerson 1974; Walterbos & Kennicutt 1988; Braun 1990a; Devereux et al. 1994; Hoernes et al. 1998), although the $H_1$ gas ‘ring’ (Emerson 1974) appears somewhat extended. Secondly, the ‘ring’ of radio continuum (Pooley 1969; Berkhuysen 1977; Beck et al. 1998) (related to overall enhancements of both large-scale regular and small-scale random magnetic fields plus cosmic-ray electrons) and the gas ‘ring’ overlap remarkably well. Finally, in addition to random magnetic fields, polarization studies (Beck et al. 1980; Beck 1982; Beck et al. 1998) of synchrotron radio-continuum emission do indicate an enhancement of large-scale ordered magnetic fields at the ‘ring’ with orientations approximately tangential to the torus (Sofue & Takano 1981; Han et al. 1998). The fact that a ‘ring’ structure is hardly detectable in the old stellar
disc may be explained by a considerably smaller $\mu^2/\mu_0^2$ compared with $\mu/\mu_0$ (see specific estimates in Section 3).

3 MHD DENSITY WAVES VERSUS KINEMATIC DYNAMOS

The fully self-consistent MHD dynamo theory as formally formulated is non-linear in general because of the inclusion of the Lorentz-force feedback on the flow velocity field (Lou 1993). This intrinsic non-linearity makes the problem extremely difficult. In practice, a time-dependent, three-dimensional non-linear simulation remains a challenging numerical task. The mean-field kinematic dynamo theory was thus introduced to simplify relevant theoretical analyses.

The magnetic field structure of M31 was studied in the past using mean-field kinematic dynamo models (Ruzmaikin & Shukurov 1981; Ruzmaikin et al. 1985; Deinzer, Grosser & Schmitt 1993; Moss et al. 1998). Basic premises of a turbulent mean-field kinematic dynamo include (i) an appropriate separation of small-scale turbulence and large-scale mean-field behaviour and (ii) the assumption that the initial large-scale magnetic field in a prescribed flow profile is extremely weak, such that the non-linear dynamic feedback of the Lorentz force on the flow can be ignored. Within this framework of simplification, the dynamo problem becomes essentially linear for the mean magnetic field evolution. For the turbulent coefficients involved, it is then crucial to determine their forms and estimate their magnitudes. For example, the inclusion of quenching effects in the turbulent coefficients reflects some interesting ideas about non-linear feedback and it is worthwhile to perform numerical experiments with them to some extent (e.g. Moss et al. 1998). Nevertheless, the specific forms and magnitudes of these coefficients are very much uncertain and have been, in fact, quite controversial in recent years. Given these qualifications, modern, considerably sophisticated, dynamo models can produce specific results for a prescribed rotation curve. For M31, it appears that the quantitative dynamo results are somehow sensitive to details of the adopted rotation curve (cf. Moss et al. 1998 and references therein).

The earlier $\alpha\omega$-dynamo model (Ruzmaikin & Shukurov 1981; Ruzmaikin et al. 1985) was based on a rotation curve with a pronounced double-peaked form, and the calculated magnetic field enhancements were confined to $0 \leq r \leq 2$ and $7 \leq r \leq 20$ kpc. By adopting a newly determined rotation curve of M31 and a set of parameters involved in a more advanced $\alpha^2\omega$-dynamo model, Moss et al. (1998) produced an extended belt of enhanced magnetic field in the radial range $7 \leq r \leq 12$ kpc. Meanwhile, their model also predicts a belt of maximum magnetic field between 2- and 6-kpc radii. While the former result is largely consistent with observations of synchrotron radio emission from M31, the latter is apparently not. They speculated that the absence of the inner ring ($2 \leq r \leq 6$ kpc) might be caused by a non-equipartition of cosmic ray and magnetic energy densities there. In other words, the cosmic ray energy density is significantly lower than the magnetic energy density at the inner ring such that the resulting synchrotron emission is weak.

At this stage of theoretical development, our analysis of MHD density waves described in Section 2 is a local one as stipulated by the WKBJ approximation, and the background disc system is assumed to be very thin. A full-blown comparison with M31 observations and with numerical results of dynamo model calculations is premature. Nevertheless, it is useful to note several main differences between the MHD density-wave approach and kinematic dynamo models. First, MHD density-wave formulation involves self-gravity which is absent in kinematic dynamo models. The long-range gravitational force plays a key role in organizing large-scale density-wave structures and in transporting angular momentum and energy outward for trailing spirals (Lynden-Bell & Kalnajs 1972; Fan & Lou 1999). Secondly, one important issue addressed by kinematic dynamo models is the initial seed-field amplification which is bypassed in the density-wave scenario by simply stipulating the existence of a mean background ring-like magnetic field. Conceptually, dynamo and MHD density-wave processes are most likely concurrent in spiral galaxies (Mestel 1999, private communication) and the separation of the two processes is somewhat artificial. The main obstacle here is the intrinsic non-linearity of the problem. Thirdly, MHD density waves involve large-scale dynamic interplay between a massive stellar disc, a gas disc and magnetic field, which is ignored in dynamo models. The coupling between the discs is the mutual gravitation and the coupling between the gas and magnetic field is the Lorentz force. In principle, one should be also concerned with the cosmic-ray gas as its energy density is typically comparable to that of the magnetic field (Lou & Fan 2000a, in preparation). Fourthly, as a result of dynamic interaction, there exist two distinctly different MHD density wave modes, namely fast and slow (Fan & Lou 1996), information about which is completely lost in kinematic dynamo treatments (Moss 1998; Shukurov 1998). In general, the two MHD density wave modes may coexist in a spiral galaxy (cf. Fan & Lou 1997). Depending on specific situations, either one may become more prominent than the other (Fan & Lou 1997). In the cases of M51 (Roberts & Yuan 1970; Mathewson, van der Kruit & Brouw 1972; Neininger 1992; Fan & Lou 1996; Lou & Fan 1998a) and NGC 2997 (Han et al. 1999; Lou et al. 1999), fast MHD density waves are dominant. In the case of NGC 6946 (Beck & Hoernes 1996; Fan & Lou 1996; Lou & Fan 1998a), slow MHD density waves overwhelm fast ones. Finally, for MHD density waves, it is possible to examine observationally the large-scale interrelations between mass density, magnetic field and velocity perturbations – an important testing procedure that is precluded by the assumption of kinematic dynamos. For conventional density-wave theory (without magnetic field), the large-scale interrelation between density and velocity perturbations provides a valuable link between theory and observations (Lin, Yuan & Shu 1969; Visser 1980a,b).

While we emphasize the new perspective of MHD density waves for magnetized spiral galaxies (including M31 specifically discussed in this paper), the known limitations of a local analysis within a thin disc system need eventually to be systematically removed. As our current theoretical analysis invokes the local WKBJ approximation (in the sense that the radial wavelength is sufficiently small compared with the radial spatial scale of the background variation, that is $kr \gg 1$), these MHD density waves propagate in the radial direction. As $m = 0$ under the assumed axisymmetry, there is no azimuthal propagation. For non-axisymmetric spiral MHD density waves in general, there always exists an azimuthal propagation with a pattern speed $\omega_\phi = \omega/m$. Over more than three decades of theoretical development, there has been a long debate about whether large-scale structures of spiral galaxies as we observe them are quasi-stationary (cf. Lin & Shu 1964; Lin 1967, 1987; Bertin & Lin 1996) or transient (cf. Toomre 1969, 1977). Strictly speaking, this issue cannot be answered by a local WKBJ-type analysis, as (MHD) density waves always propagate locally in the absence of relevant boundary conditions. One main reason for introducing the local
approximation is to simplify Poisson’s equation (equation 2.5), which relates mass density and gravitational potential (cf. Shu 1970), to perform basic analytical analyses, and to identify the physical nature of large-scale galactic structures. It is possible to include some higher order terms in the simplified Poisson equation and push analytical analyses further (cf. Bertin et al. 1989b and extensive references therein).

Within the past decade or so, numerical codes (e.g. Pannatoni 1983; Bertin et al. 1989a) have been developed to solve the set of linear ordinary integro-differential equations for hydrodynamic density waves in a single stellar disc, that is, the Poisson equation (2.5) can be solved globally without the restrictive local approximation. Meanwhile, one needs to formulate appropriate boundary conditions near the disc centre (e.g. the so-called Q-barrier that shields the inner Lindblad resonance) and radiation conditions outside the outer Lindblad resonance; in numerical integrations one should also be careful about corotation and Lindblad resonances in general. This is exactly the case in the so-called modal formulation (cf. Bertin et al. 1989a,b; Bertin & Lin 1996). In this modal formulation of global density waves (i.e. without magnetic field), large-scale spiral structures are superpositions of several relevant modes (i.e. those with maximum growth); while propagating azimuthally (for \( m \neq 0 \)), these density-wave structures are quasi-stationary in the radial direction. It is the combined effect of disc differential rotation and gravitational instability that sustains these spiral structures. The Q-barrier around the centre and the corotation radius act as reflection ‘mirrors’ to trap density waves. Meanwhile, one must allow a partial leakage of density waves at corotation. As (MHD) density waves are negative-energy waves (e.g. Bertin & Lin 1996; Fan & Lou 1999), a partial radial leakage of waves leads to growth (ultimately at the cost of disc differential rotation).

In view of the historical and recent developments of density-wave theory (without magnetic field), a global analysis of MHD density waves by numerically integrating a set of ordinary integro-differential equations\(^6\) (cf. Bertin et al. 1989a,b; Bertin & Lin 1996) and an analysis on the effect of finite disc thickness (e.g. Jog 1996) can be carried out in an orderly manner. These are just two examples; there are other aspects that one can pursue in analogy to hydrodynamic density-wave analysis. As it is widely accepted that the overall density-wave scenario forms the basic framework for studying large-scale structures of spiral galaxies, it is almost inevitable that the theoretical concept of MHD density waves will play a key role in understanding the large-scale behaviour of galactic magnetic fields, especially in view of the large-scale structural correlations in optical and radio-continuum bands.

4 DISCUSSION

In order to account for various radiation signatures from annulus structures of M31, one must go beyond the fluid-MHD calculation of density waves in a composite disc system and consider the following processes in the magnetized interstellar medium. First, at the intuitive level, one would expect large-scale concentrations of dust grains and atomic/molecular gas to be roughly coincident as a result of entrainment by the gravitational potential trough, effective grain/magnetic-field interaction or sufficiently frequent collisions between gas and dust (Hoernes et al. 1998). Secondly, young massive stars continuously form in zonal regions of high gas concentrations where conglomeration of molecular clouds and clumps (Koper et al. 1991; Dame et al. 1993; Koper 1993) is triggered by various instabilities, especially in view of the magnetic field presence (Field 1965; Parker 1966; Elmegreen 1994). Thirdly, enhanced thermal infrared emissions are expected from regions of high concentrations of dust grains, which interact with profuse ultraviolet photons from young massive stars (cf. Xu & Helou 1996 and references therein). Fourthly, various instabilities and activities on smaller scales associated with large-scale high-density gas arms and enhanced magnetic field tend to disrupt the order of magnetic fields to various degrees. Finally, there is growing evidence that, on large scales, cosmic-ray production must be somehow linked to global star-formation activities in a galaxy; for M31 it would be very difficult for cosmic-ray electrons to escape from the central bulge and to form a ‘ring’ \(\approx 9–10\) kpc away. In other words, the presence of enhanced cosmic-ray electrons in the ‘radio-continuum belt’ should be produced locally. In this respect, the MHD-density-wave scenario provides a natural conceptual link among enhancements of HI gas, galactic magnetic fields, star formation and cosmic-ray production.

When all these aspects are taken together, it is not surprising that large-scale ‘ring’ structures of M31 in H\(\alpha\), young massive stars, H\(\alpha\), H\(\gamma\) complexes, OB associations (van den Bergh 1964), ultraviolet, thermal infrared emissions, molecular clouds (Battaner et al. 1986; Koper 1993) and synchrotron radio continuum (Beck et al. 1980, 1998; Hoernes et al. 1998) appear more or less coincident. It also comes as no surprise that synchrotron radio emission reveals both ordered and random magnetic field components along the annulus. The faint ‘ring’ in diffuse X-ray emission (West, Barber & Folgerer 1997) may result from enhanced stellar X-ray coronal contributions as well as from enhanced magnetic fields with associated heating in a tenuous magnetized gas. We further anticipate the existence of a gamma-ray ‘ring’ structure as a result of interaction between cosmic-ray nuclei and high-density annulus components of the interstellar medium.\(^7\)

From available observations (Emerson 1974; Berkhuijsen 1977; Beck et al. 1980; Nolthenius & Ford 1987; Koper 1993; Braun 1991; Tenjes et al. 1994), one can estimate that \(V_g \approx 250\,\text{km}\,\text{s}^{-1}\), \(\bar{\Omega} \approx 8.3 \times 10^{-16}\,\text{s}^{-1}\) and \(\kappa \approx 1.18 \times 10^{15}\,\text{s}^{-1}\) at the annulus radius of \(R \approx 10\,\text{kpc}\); \(D_{\odot} = 5 \times 10^5\,\text{cm}\,\text{s}^{-1}\), \(C_{\odot} \approx 10^6\,\text{cm}\,\text{s}^{-1}\) and \(C_{\odot} \approx 9 \times 10^5\,\text{cm}\,\text{s}^{-1}\); \(\mu_0^3 = 45\,\text{M}_\odot\,\text{pc}^{-2}\) (here disc mass is taken to be \(\approx 40\%\) of the total mass) and \(\mu_1 \approx 6\,\text{M}_\odot\,\text{pc}^{-2}\). It then follows that the stellar disc \(Q\)-parameter \(Q_\star = D_\star \kappa / (\pi g \mu_0^3) \approx 2.8\), the magnetized gas disc \(Q\)-parameter \(Q_{\text{gas}} = (C_{\odot}^2 + C_{\odot}^2)^{1/2} / (\pi g \mu_0^3) \approx 5.7\) and the disc surface mass density ratio \(\mu_1 / (\mu_0 + \mu_2) \approx 0.12\). Based on these estimates of \(Q_\star, Q_{\text{gas}}\) and \(\mu_1 / (\mu_0 + \mu_2)\), the effective magnetic \(Q\)-parameter \(Q_{\text{M eff}}\) for the rotating composite disc system can be readily determined (cf. Elmegreen 1995; Jog 1996), and it is found that the annulus region

\(^6\)As a result of the long-range nature of the self-gravity force, the surface mass density perturbation \(\mu^0 + \mu\) and gravitational potential perturbation \(\phi + \phi\) are connected by the integral relation equation (2.11), which remains valid whether the gas is magnetized or not. Local relation (2.15) is a result of the WKBJ approximation (Shu 1970).

\(^7\)There was a keen interest in a possible enhancement of \(\gamma\)-ray flux from Andromeda for a completely different reason, namely the debate on the physical origin of \(\gamma\)-ray bursts in the past several years. Our interest here is in the correlation of a \(\gamma\)-ray enhancement with the ‘ring’ structure, even though the overall emission level of \(\gamma\) rays from Andromeda would be very weak as a result of the distance.
of M31 with Population I signatures (van den Bergh 1964; Pellet et al. 1978; Berkhuijsen 1977) is quite stable at the present epoch. Away from the annulus, $\mu_0$ can reach values (Berkhuijsen 1977) less than the peak value at $\sim 10$ kpc by as much as a factor $\approx 5$; under these conditions the composite disc system is expected to be generally stable on large scales. This appears natural to explain the observed low star formation rate in M31.

Dispersion relation (2.18) was derived for axisymmetric MHD density waves at a single $\omega$. For a group of waves with different $\omega$ (and thus different $k$), an MHD density-wave packet can be constructed by integration (or summation) in the frequency domain. For the chosen parameters (in the preceding paragraph) around the annulus with an estimated $k \sim (1.5 \text{kpc})^{-1}$, $\omega \sim 0.7 \times 10^{-15} \text{s}^{-1}$ by the minus-sign solution of (2.25) and the corresponding radial group speed is estimated to be $\sim 3 \times 10^8 \text{cm s}^{-1}$. One can also show that $\mu^5/\mu_0^5 \sim 0.1/\mu_0$. For $\mu/\mu_0 \sim 0.1$, one would then have $\mu^5/\mu_0^5 \sim 0.01$. This result appears qualitatively consistent with extremely weak density-wave ‘ring’ fluctuations in the old stellar disc (e.g. Walterbos & Kennicutt 1988).

It is expected that various non-linear effects associated with MHD density waves are involved in the actual ‘ring’ structures of M31, especially in the magnetized gas disc. In particular, MHD shocks in the gas disc (Roberts & Yuan 1970) have long been suspected to trigger a chain of active star formation activities (e.g. Elmegreen 1994). As a result of the typical lifetime of $\sim 10^7$ yr for young massive stars, the narrow dark gas/dust lanes are expected to be slightly shifted relative to the bright ridges of young stars and H II complexes. We emphasize that as stars form in gas, the galactic magnetic field must influence the global star formation rate in a non-trivial manner.

While the overall ‘ring’ structures of M31 at many electromagnetic wavelengths do appear grossly axisymmetric about the galactic centre, deviations can be seen on all discernible scales. In fact, optical as well as H I images for the disc portion of M31 are sometimes interpreted as extremely tight-winding spiral arms, although not without some ambiguities owing to the high inclination of M31. At any rate, our fluid-MHD calculations on the assumption of axisymmetry are quite indicative of the underlying large-scale dynamics and can be readily extended to the non-axisymmetric case when necessary (Lou & Fan 1998a,b; 2000b, in preparation).

Finally, we note the presence of two types of MHD density waves given by equation (2.25); our main emphasis in this paper has been on the minus-sign solution in the application to the annulus of M31. Generally speaking, at this level of fluid-MHD calculations, there is no a priori reason to exclude one mode from the other. One way of determining mode selection processes is to adopt the modal formalism (Bertin et al. 1989a,b; Bertin & Lin 1996) in which the potential–density integral (2.11) is computed numerically (Pannatoni 1983) and the fluid-MHD density-wave equations in the composite disc system are integrated with a prescribed rotation curve and properly specified boundary conditions. It remains to be shown that, in a modal formalism, global MHD density waves inside the Lindblad resonance are indeed preferable to the ones outside.

ACKNOWLEDGMENTS

We are grateful to Dr R. Beck and Director R. Wielebinski for their arrangements and hospitality during our visits at the Max-Planck-Institut für Radioastronomie, and thank Drs R. Beck, E. M. Berkhuysen, M. Krause and N. Neininger for extensive discussions. This research was supported in part by grants from US NSF (ATM-9302037 and AST-9731623), NASA Space Physics Theory Program and NASA SR&T grants to the University of Chicago, by funds from the Institute of Astronomy & Astrophysics, Academia Sinica (NSC-88-2816-M-001-0010-6) and the Institute of Astronomy, National Central University, by the Visiting Scientist Programs at the International Space Science Institute, Bern, Switzerland and at the Mathematical Institute, University of St Andrews, Fife, Scotland.

REFERENCES

Brinks E., Shane W. V., 1984, A&AS, 55, 179
Fan Z. H., Lou Y.-Q., 1996, Nat, 383, 800
This paper has been typeset from a \TeX/LaTeX file prepared by the author.