Role of Interaction between Magnetic Rossby Waves and Tachocline Differential Rotation in Producing Solar Seasons

Mausumi Dikpati\textsuperscript{1,} Scott W. McIntosh\textsuperscript{1,} Gregory Bothun\textsuperscript{2,} Paul S. Cally\textsuperscript{3,} Siddhartha S. Ghosh\textsuperscript{4,} Peter A. Gilman\textsuperscript{1,} and Orkan M. Umurhan\textsuperscript{5}

\textsuperscript{1}High Altitude Observatory, NCAR, 3080 Center Green Drive, Boulder, CO 80301, USA; dikpati@ucar.edu
\textsuperscript{2}Department of Physics, University of Oregon, 1274 University of Oregon, Eugene, OR 97403-1274, USA
\textsuperscript{3}School of Mathematical Sciences, 9 Rainforest Walk, Monash University, Clayton, Victoria 3800, Australia
\textsuperscript{4}Computational and Information System Laboratory, NCAR, 1850 Table Mesa Drive, Boulder, CO 80305, USA
\textsuperscript{5}SETI Institute and NASA Ames Research Center, Space Science and Astrobiology Division, Planetary Systems Branch, MS-245-3, Moffett Field, CA 94035, USA

Received 2017 November 20; revised 2017 December 29; accepted 2018 January 9; published 2018 February 1

Abstract

We present a nonlinear magnetohydrodynamic shallow-water model for the solar tachocline (MHD-SWT) that generates quasi-periodic tachocline nonlinear oscillations (TNOs) that can be identified with the recently discovered solar “seasons.” We discuss the properties of the hydrodynamic and magnetohydrodynamic Rossby waves that interact with the differential rotation and toroidal fields to sustain these oscillations, which occur due to back-and-forth energy exchanges among potential, kinetic, and magnetic energies. We perform model simulations for a few years, for selected example cases, in both hydrodynamic and magnetohydrodynamic regimes and show that the TNOs are robust features of the MHD-SWT model, occurring with periods of 2–20 months. We find that in certain cases multiple unstable shallow-water modes govern the dynamics, and TNO periods vary with time. In hydrodynamically governed TNOs, the energy exchange mechanism is simple, occurring between the Rossby waves and differential rotation. But in MHD cases, energy exchange becomes much more complex, involving energy flow among six energy reservoirs by means of eight different energy conversion processes. For toroidal magnetic bands of 5 and 35 kG peak amplitudes, both placed at 45° latitude and oppositely directed in north and south hemispheres, we show that the energy transfers responsible for TNO, as well as westward phase propagation, are evident in synoptic maps of the flow, magnetic field, and tachocline top-surface deformations. Nonlinear mode–mode interaction is particularly dramatic in the strong-field case. We also find that the TNO period increases with a decrease in rotation rate, implying that the younger Sun had more frequent seasons.

Key words: instabilities – magnetohydrodynamics (MHD) – Sun: activity – Sun: magnetic fields – Sun: rotation

1. Introduction

We live in the extended atmosphere of our star, the Sun. It is a variable star, and as a result, Earth is unavoidably impacted by its variability. Most of our star’s variability is associated with the solar cycle, the waxing and waning of sunspots and other solar magnetic features with an average period near 11 yr. This decadal-scale variability is readily seen in the Sun’s magnetic field, its radiative and particulate emission. Because the Sun rotates with a period near 27 days and has magnetic structure, radiation, and particle emissions that are functions of solar longitude as well as time, Earth also experiences strong solar variability on a 27-day timescale. The solar–terrestrial connection is now well established, as the interplanetary magnetic field serves as a conduit to deliver particle streams and magnetic fields generated by the solar magnetic activity to the atmosphere of Earth.

Recently it has been demonstrated that many measures of solar activity, in particular magnetic fields, particle emissions from flares, coronal mass ejections (CMEs), UV/EUV/X-ray radiation bursts, and solar wind, exhibit quasi-periodic “bursts” of activity with periods in the range of 6–18 months (McIntosh et al. 2015). These bursts of activity, called the “solar seasons,” have the same order-of-magnitude variability as the solar cycle itself and are the primary driver of energetic space weather phenomena. Closely related for long and short periods, respectively, are the observed “quasi-biennial oscillations (QBOs)” and Rieger periodicities (Rieger et al. 1984; Wolff 1992; Oliver et al. 1998; Ballester et al. 2002, 2004; Dimitropoulou et al. 2008). The solar seasons observed in the occurrence of auroras (Silverman 1990) clearly indicate the powerful solar–terrestrial connection. Gurgenashvili et al. (2017) discussed in a comprehensive manner the manifestations of solar seasons in varieties of solar and terrestrial data.

There is another property of solar activity that also influences space weather that reaches Earth, and that is the so-called “active longitudes” (Gyenge et al. 2016, 2017). These are intervals in longitude where for several rotations much of solar activity is concentrated. Because of their persistence, Earth feels periodic variations in solar outputs just from the fact that the Sun rotates. If there is just one active longitude, then this period is essentially the same as the Sun’s rotation period. But at any given time there can be more than one active longitude along a full circumference, leading to shorter periods, which differ according to the longitude spacing between the different longitudes.

Signals from active longitudes will contribute to the solar season signal, but the two phenomena are different. Solar seasons are defined from longitude-averaged data, while active longitude data are by definition longitude dependent. Solar season change requires time variation in these longitude averages, but the presence of active longitudes does not. But a change in active longitude amplitude will contribute to a change in seasons.

Solar seasons affect different layers of Earth’s atmosphere in different forms. Changes in the interplanetary magnetic field originating from the Sun interact with the geomagnetic field to
greatly affect the magnetosphere, leading to geomagnetic “storms” where charged particles enter the lower atmospheric layers by spiraling down geomagnetic fields near the poles. The variable UV/EUV and X-ray radiation causes changes in the thermosphere, mesosphere, and stratosphere where it is absorbed by molecular oxygen and ozone, causing the atmosphere to heat up and expand upward. The conductivity of the ionosphere is perturbed by all processes that contribute to its partial ionization. Smaller variations in visible radiation reach the lower atmosphere and have some effect on global circulation in the troposphere.

Recent reports to international governments have documented that these Sun-driven changes in the atmosphere have consequences for human activity (see, e.g., http://www.sciencedirect.com/science/article/pii/S0273117715002252). Satellite electronics can be damaged by bursts of energetic particles in flares and CMEs. Their orbital trajectories can change dramatically as a result of increased drag due to rapid radiative heating of the atmosphere. Energetic particles and radiation both pose significant hazards for astronauts in space and circumpolar commercial aircraft. Global positioning and navigation systems can be disrupted, costing private, public, and military agencies considerably. Radio communications can be scrambled. Closer to the ground, power grids can be seriously compromised by induced current surges that can lead to network-wide shutdowns and blackouts. For all these reasons and others, considerable effort is being made to forecast upcoming solar activity and its effects on Earth. But the current operational paradigm is that a forecast of event arrival and impact is made once a solar flare or CME has occurred and been observed, but while the event is upstream of Earth. No method exists to make a reliable prediction of when, or where, the next burst of solar activity will occur. To go beyond this paradigm for space weather, we need to forecast these seasonal bursts of activity by monitoring the dynamics of the subsurface spot-producing toroidal magnetic field systems that generate active regions long before they actually emerge.

Global organization of persistent longitude-dependent oscillating magnetic signals, such as those exhibited in the seasons of space weather, suggest an origin at the base of the convection zone—in the solar tachocline. Extensive studies of global hydrodynamics (HD) and MHD of the solar tachocline in quasi-3D shallow-water-type numerical models show persistent patterns of bulges and depressions of tachocline fluid with low longitudinal wavenumbers (Gilman 2000; Dikpati & Gilman 2001a, 2001b; Gilman & Dikpati 2002; Cally 2003; Cally et al. 2008; Dikpati 2012). Portions of the solar dynamo-generated, sunspot-producing, toroidal magnetic bands that coincide with bulging of the fluid are likely to emerge buoyantly to the surface through the convection zone. This bulging leads to magnetic flux emergence and the enhanced formation of active regions (Dikpati & Gilman 2005). The commonality of observation and global-scale organization of model-inferred magnetic activity supports assertions that the seasons of space weather originate in the solar tachocline.

Therefore, based on the paradigm that active-region-forming magnetic flux is originating in the solar tachocline, located at the base of the convection zone, we study the dynamics of the interaction between these spot-producing magnetic fields and tachocline differential rotation. This does not necessarily imply that the entire dynamo process has to operate there or that all bipolar magnetic regions originate there. Recent convective dynamo simulations suggest that flux emerging in solar active regions can also originate from the convection zone, in coherent toroidal bands that can coexist with the turbulent convection (Brown et al. 2010; Fan & Fang 2014). However, it has also been demonstrated that the addition of a tachocline in convective dynamo simulations can promote the generation of more persistent and coherent toroidal fields (Browning et al. 2006; Racine et al. 2011; Masada et al. 2013). Given our current state of knowledge, it is plausible to assume that magnetic flux is stored in the tachocline in the form of coherent toroidal magnetic bands and that these bands produce the largest active regions that are responsible for the most extreme space weather events and for the observed surges of activity.

Because the tachocline straddles the “bottom” of the convection zone, measured helioseismically at about 0.713 of the solar radius (Christensen-Dalsgaard et al. 1991), it includes both the “overshoot” part of the convection zone, where the stratification is probably slightly subadiabatic, and the uppermost part of the radiative interior, where the stratification is strongly subadiabatic. Being subadiabatic implies that toroidal fields generated by the solar dynamo may be stored there for some period of time, before they rise to the photosphere owing to magnetic buoyancy and advection by convection. This property, coupled with a lower level of turbulence there compared to the convection zone above, makes it likely that the tachocline will contain strong toroidal fields, perhaps 10–20 kG or even more, which have relatively smooth profiles with latitude and depth. These fields should be the source for emerging active regions that virtually always obey Hale’s polarity laws.

The presence of both differential rotation and toroidal fields in the tachocline raises the possibility that the latter are unstable to global disturbances with nonzero longitudinal wavenumbers (m) (Gilman & Fox 1997). Because of the subadiabatic stratification, these disturbances should have large latitudinal and longitudinal spatial scales compared to the thickness of the tachocline. In other words, the disturbances are nearly horizontal in spherical shells. It is well established that purely horizontal disturbances of differential rotation without toroidal fields are stable in the two-dimensional tachocline unless the tachocline latitudinal differential rotation is expressed including both the second- and fourth-order terms in sine latitude (Charbonneau et al. 1999; Arlt et al. 2005). Including the third dimension in a simplified way by allowing the top surface of the tachocline to deform, as in the so-called shallow-water models (Pedlosky 1987) long used in geophysical fluid dynamics, leads to vigorous instability in the tachocline (Dikpati & Gilman 2001b). For a magnetized tachocline, a vector-invariant MHD version of the shallow-water model was first built by Gilman (2000).

HD and MHD shallow-water models have been extensively applied in various astrophysical contexts over the past decade (Gilman & Dikpati 2002; Dikpati et al. 2003; Zaqarashvili et al. 2007, 2009, 2010a; Umurhan 2008, 2012; Dikpati 2012; Heng & Workman 2014; Klimchuk & Petrosyan 2017a, 2017b), and many details of MHD tachocline instability properties have been explored by many authors (Gilman & Dikpati 2002; Dikpati et al. 2003; Arlt et al. 2005; Dikpati & Gilman 2005). For toroidal field with latitude dependence included, the combination of differential rotation in latitude and toroidal
field is unstable for almost any magnitude of differential rotation and for any profile and amplitude of toroidal field, as well as any subadiabaticity. This is true even if the toroidal field energy is much larger than that present in the differential rotation. The $e$-folding growth times for unstable modes of this kind are very short compared to a sunspot cycle, of the order of a month to a year, so they are very relevant to magnetic patterns we observe on the Sun that evolve on timescales of months to a few years, such as the “seasonal” variability of the Sun.

These instabilities arise from the kinetic and magnetic energy stored in latitude gradients of rotation and toroidal fields. We expect the tachocline to have radial gradients in these quantities too, and so we ask about the role they might play in tachocline instabilities. For toroidal fields above about 4 kG, instabilities of radial gradients to global disturbances are suppressed by the toroidal field (Gilman 2015, 2017). Therefore, it is primarily the latitudinal gradients that should be responsible for global instabilities in the tachocline.

In an MHD shallow-water model of the solar tachocline the plasma fluid shell undergoes swellings and depressions at certain latitudes and longitudes. Therefore, certain portions of the dynamo-generated toroidal fields, frozen into tachocline plasma, fall into the swelling of the fluid shell, and certain portions fall into depressions. Those portions of the toroidal bands in swollen fluid are likely to enter the convection zone and rise to the surface. Hence, the surface patterns of magnetic activity contain an “imprint” of global tachocline MHD. These imprints were constructed by taking a linear combination of all plausible unstable longitudinal modes (see Dikpati & Gilman 2005), and the evolution of the latitude–longitude locations of active regions was simulated.

In the nonlinear evolution of shallow-water tachocline instabilities, Dikpati (2012) found that finite-amplitude disturbances exhibit amplitude oscillations between the disturbance- and reference-state differential rotation. Based on that hint, recently Dikpati et al. (2017) investigated the basic mechanism for these oscillations and established the physical foundation of these tachocline nonlinear oscillations, which they called TNOs. Dikpati et al. (2017) demonstrated that TNOs are essentially produced by the interaction between the Rossby waves and differential rotation in the form of back-and-forth exchange of energies, in a similar way to the nonlinear Orr mechanism (Orr 1907). TNOs can cause the seasonal variation in solar activity by producing enhanced bursts when the Rossby waves’ energy grows to its maximum by extracting energy from differential rotation. During this epoch, the tachocline top surface is maximally deformed, and hence nearly “frozen-in” toroidal fields can enter the convection zone from the tachocline, starting their buoyant rise to the surface to erupt as active regions. The bursty phase is followed by a relatively quiet phase, during which the differential rotation gets restored by extracting energy from Rossby waves and top-surface deformations subside. These TNOs were found to have periods between 2 and 20 months, for a wide range of effective gravity values, differential rotation amplitudes, strengths of toroidal magnetic bands, and their latitude locations.

We reason that bursty seasons on the Sun are most connected to the amplitude of the tachocline bulges that contain a toroidal field, while active longitudes are most related to the dominant persistent longitudinal wavenumber $m$ of the bulges that contain toroidal fields. Recent studies of active longitudes, such as Gyenge et al. (2016, 2017), indicate that $m = 1$ may be dominant, but $m = 2$ is also present in data for longitude distribution of solar flares and CMEs.

Traditional Rossby waves can occur in the Sun owing to vorticity conservation, very much in the way it occurs in Earth’s ocean and atmosphere. But in the solar tachocline, magnetic Rossby waves are equally important, for which vorticity is not conserved (Gilman & Fox 1997). Due to subadiabatic stratification, strong shear, and very low magnetic diffusivity, the solar tachocline is thought to be the seat of strong, spot-producing toroidal magnetic fields. Therefore, it is important to understand how the tachocline dynamics change owing to the interaction of magnetic Rossby waves. What follows in the next section is a brief description of traditional Rossby waves in the Sun. In Section 3 we present the mathematical formulation of a nonlinear MHD shallow-water tachocline model and solution methods. Section 4 presents results of complex interactions of hydrodynamic and magnetohydrodynamic Rossby waves with tachocline differential rotation and spot-producing toroidal magnetic field there and explains the physical mechanism involved in that. We close in Section 5 with concluding remarks.

2. HD and MHD Rossby Waves in the Sun

The existence of Rossby waves in Earth’s atmosphere has been known for a long time. Rossby waves are planetary waves that are naturally generated in the atmospheric fluid of any rotating planet. On Earth, Rossby waves are manifested in jet stream flow that is largely geostrophic with a wavenumber 3–5. Traditional Rossby waves can occur in the Sun in a very similar way, but there are bound to be differences due to the Sun being a magnetized plasma ball. Solar Rossby waves have been observed (McIntosh et al. 2017) not only in the Sun but also in many solar-like stars (Lanza et al. 2009).

The original theory of Rossby waves was based on assuming that the flow in them was nearly “geostrophic,” that is, being in near (but not exact) horizontal force balance between pressure gradients and Coriolis forces. In more modern theory, the flow is said to be at small “Rossby number,” $R_o$, which is just the dimensionless ratio between the typical rotation velocity and the rotational velocity of the interior of the spherical shell. For the solar tachocline, $R_o$ is well estimated just from the ratio of differential rotation over a significant fraction of the distance between equator and pole and the rotation of the interior below. This ratio is of order 0.1, small enough to yield Rossby waves as an important component of the global hydrodynamics. However, this does not mean that we have to assume small Rossby number to solve the equations we use. In fact, it is much better to solve more general equations that contain Rossby waves as a significant, hopefully observable, part of the dynamics, but not the only part. The basics of Rossby waves in thin spherical shells like planetary atmospheres and oceans and the solar tachocline are discussed in Pedlosky (1987) and Regev et al. (2016). On Earth, the location of Rossby waves is often an excellent predictor for changing weather patterns on a timescale of 3–5 days. In a similar vein, we would suggest that changes in solar Rossby wave pattern could be precursors of bursty seasons of the Sun.

Since Rossby waves are much less familiar in the solar context until recently, we describe schematically in Figure 1 how Rossby waves operate in terms of local vorticity about a local vertical axis. We use the well-known property that in the
simplest Rossby wave models the total vorticity about the local vertical axis of the spherical shell (component of rotation and vorticity of the flow in the local vertical direction) is conserved. An analogous argument can be made for the shallow-water system, with so-called “potential vorticity” (vorticity divided by shell thickness) conserved.

In Figure 1, we start from five patterns of plasma vorticity (top frame), initially at the same latitude in the northern hemisphere. If, randomly, two of them are moved poleward and one equatorward, then the two poleward-moving patterns will have increased relative vorticity and the equatorward-moving one will have decreased vorticity, due to the conservation of total vorticity (disturbance plus latitude gradient of the Coriolis parameter, which measures the vorticity of the rotating coordinate system). Thus, the poleward (equatorward) patterns will get an anticyclonic (cyclonic) relative vorticity. These anticyclonic/cyclonic motions will tend to move the other two undisturbed patterns and will change the vorticity in them. This vorticity will tend to restore the three patterns (first, third, and fifth patterns) back to their original position. Thus, a wave pattern will be formed and will move westward. These are depicted by patterns (1), (2), and (3) in panel (d).

Rossby waves are somewhat peculiar in that, even though an individual wave propagates westward in simple models, the group velocity of Rossby waves is directed eastward, so a “train” of Rossby waves can propagate eastward even when each individual wave in the “train” is propagating westward in phase (Regev et al. 2016).

When a toroidal magnetic field is added to the system, the HD Rossby waves become modified by the presence of perturbation magnetic forces. The degree of modification depends on the strength and profile of reference-state fields included. In this case there no longer is conservation of potential vorticity because of magnetic effects. Fluid elements in different locations are linked by magnetic field lines that can transmit vorticity from one element to another. Even for relatively weak toroidal fields the change from HD Rossby waves can be substantial; the MHD Rossby wave case has been studied extensively in references we cite below.

3. MHD Shallow-water Model Equations and Solution Technique

3.1. Description of MHD Shallow-water Model and Assumptions

We summarize the discussion on the concept of applying hydrodynamic (HD) “shallow-water” (SW) equations to global flows in thin layers of stars, such as their tachoclines (Dikpati & Gilman 2001b), and their generalization to the MHD case (Gilman 2000) briefly here.
Hydrodynamic shallow-water models have a long history of useful applications to problems in atmospheric, oceanic, and planetary fluid dynamics (Stoker 1957; Pedlosky 1987). They were first developed for oceanic tides and surface water waves. The idea was to capture the effects of a gravity-restoring force acting at the interface of an incompressible fluid (such as the ocean) with its surroundings, such as the atmosphere above. It can be applied to many fluid systems in which the flow to be modeled has much larger horizontal than vertical scales. This allows the fluid to be treated as in hydrostatic balance in the local vertical, even though vertical motions are allowed. Being hydrostatic, the horizontal pressure gradient force can be written in terms of horizontal gradients in the thickness of the shallow layer. The simplest version of a shallow-water system contains a single layer with a rigid bottom and a top that is free to deform. This is the version we will consider here. For the tachocline, the rigid bottom corresponds to the very subadiabatic radiative interior, and the top the convection zone above, which, being adiabatically stratified, offers little resistance to vertical deformation from below. Shallow-water systems with multiple layers in the vertical also exist, particularly as applied to the ocean. They could also be used for the tachocline, if we wish to include separate layers for the “overshoot” and “radiative” tachoclines, which have quite different degrees of subadiabaticity.

In one-layer shallow-water systems the horizontal flows are independent of height, while the vertical velocity is a linear function of height. Therefore, the fluid moves horizontally in vertical columns, which can stretch or compress in the vertical, according to mass conservation, as the top surface deforms. Shallow-water systems usually include no diffusion and so represent “ideal fluids.” Therefore, they conserve their total energy, kinetic plus potential, when integrated over the whole volume of the shell. This property is very useful when testing the accuracy of numerical schemes used to solve the equations and for interpreting solutions found.

Since the tachocline is very likely to contain strong magnetic fields, in an MHD generalization for the shallow-water equations, magnetic fields are assumed to be confined to the layer even when it deforms (Gilman 2000), so field lines on the top surface remain there always. Horizontal fields are independent of height, and the vertical field is a linear function of height, exactly the same as their velocity counterparts. In this system it is the total pressure (hydrostatic plus magnetic) gradient that is proportional to the horizontal gradient of shell thickness. Total magnetic flux is conserved, so there are no magnetic monopoles, which is captured by a modified divergence-free field condition, namely, that the horizontal divergence of the product of the height and horizontal field is satisfied everywhere. In the MHD shallow-water system total energy is still conserved, but this includes kinetic, potential, and magnetic energies.

### 3.2. MHD Shallow-water Equations in Spherical Geometry

Here we present the full set of MHD shallow-water equations in spherical geometry and in the rotating frame of reference (rotating with core rotation rate \( \omega_c \), which is equivalent to the rotation rate at 32° at the tachocline depth). These equations are very similar to that presented in Gilman & Dikpati (2002), except that theirs were in inertial frame. We denote \( h \) the height change relative to the undisturbed height, \( u, v \) the horizontal velocities in longitude \( \lambda \) and latitude \( \phi \) directions, respectively, and \( a, b \) the corresponding magnetic field components and write the nonlinear one-layer, dimensionless MHD shallow-water equations as

\[
\frac{\partial u}{\partial t} = \nu \cos \phi \left[ \frac{\partial v}{\partial \lambda} - \frac{\partial}{\partial \phi} (u \cos \phi) \right] - \frac{1}{\cos \phi} \frac{\partial}{\partial \phi} \left( \frac{u^2 + v^2}{2} \right) - G \frac{1}{\cos \phi} \frac{\partial h}{\partial \lambda} + 2\omega_c v \sin \phi
\]

\[
- \frac{1}{\cos \phi} \frac{\partial}{\partial \phi} \left( \frac{u^2 + v^2}{2} \right) - G \frac{1}{\cos \phi} \frac{\partial h}{\partial \lambda} - 2\omega_c u \sin \phi
\]

\[
+ a \cos \phi \left[ \frac{\partial b}{\partial \lambda} - \frac{\partial}{\partial \phi} (a \cos \phi) \right] + \frac{\partial}{\partial \phi} \left( \frac{a^2 + b^2}{2} \right),
\]

(1)

\[
\frac{\partial v}{\partial t} = -u \cos \phi \left[ \frac{\partial v}{\partial \lambda} - \frac{\partial}{\partial \phi} (u \cos \phi) \right] - \frac{1}{\cos \phi} \frac{\partial}{\partial \phi} \left( \frac{u^2 + v^2}{2} \right) - G \frac{1}{\cos \phi} \frac{\partial h}{\partial \lambda} - 2\omega_c u \sin \phi
\]

\[
+ a \cos \phi \left[ \frac{\partial b}{\partial \lambda} - \frac{\partial}{\partial \phi} (a \cos \phi) \right] + \frac{\partial}{\partial \phi} \left( \frac{a^2 + b^2}{2} \right).
\]

(2)

\[
\frac{\partial}{\partial t} \left( 1 + h \right) = \frac{1}{\cos \phi} \frac{\partial}{\partial \phi} \left[ (1 + h) u \right] - \frac{1}{\cos \phi} \frac{\partial}{\partial \phi} \left[ (1 + h) v \cos \phi \right],
\]

(3)

\[
\frac{\partial a}{\partial t} = \frac{\partial}{\partial \phi} (ub - va) + \frac{a}{\cos \phi} \left[ \frac{\partial u}{\partial \lambda} + \frac{\partial}{\partial \phi} (v \cos \phi) \right] - \frac{u}{\cos \phi} \left[ \frac{\partial a}{\partial \lambda} + \frac{\partial}{\partial \phi} (b \cos \phi) \right],
\]

(4)

\[
\frac{\partial b}{\partial t} = -\frac{1}{\cos \phi} \frac{\partial}{\partial \phi} (ub - va) + \frac{b}{\cos \phi} \left[ \frac{\partial u}{\partial \lambda} + \frac{\partial}{\partial \phi} (v \cos \phi) \right] - \frac{v}{\cos \phi} \left[ \frac{\partial b}{\partial \lambda} + \frac{\partial}{\partial \phi} (b \cos \phi) \right],
\]

(5)

\[
\frac{\partial}{\partial \lambda} \left[ (1 + h) a \right] + \frac{\partial}{\partial \phi} \left[ (1 + h) b \cos \phi \right] = 0.
\]

(6)

The length and time are scaled respectively by the radius \( r_0 \) of the fluid shell and by the inverse of the interior rotation rate \( \omega_c^{-1} \); \( G = gH / r_0^2 \omega_c^2 \) is a dimensionless parameter, in which \( g \) is the “effective” or “reduced” gravity of the stratified layer of undisturbed dimensional thickness \( H \). \( G \sim 10^3 \times |\nabla - \nabla_{ad}| \), the fractional departure of the actual temperature gradient from the adiabatic gradient. Dikpati et al. (2017) used \( \delta \) for \( |\nabla - \nabla_{ad}| \); hence, we will use \( \delta \) from now onward. The factor \( (4\pi \rho)^{-1/2} \) has been absorbed into the magnetic components (\( \rho \) is the fluid density). In the overshoot part of the tachocline (located between 0.7 and 0.72 \( R_\odot \)), the value of the subadiabatic temperature gradient is \( 10^{-5} \leq \delta \leq 10^{-4} \), and in the radiative tachocline (between 0.68 and 0.7 \( R_\odot \)) \( 10^{-2} \leq \delta \leq 10^{-1} \).
Helioseismically determined differential rotation can be expressed in the rotating frame as
\[
\omega = s_0 - s_2 \mu^2 - s_4 \mu^4 - \omega_c ,
\]
in which \( \mu \) is the sine latitude (\( \sin \phi \)) and \( s_0, s_2, s_4 \) are the numerical coefficients. The interior rotation rate, \( \omega_c \), approximately matches the rotation rate at 32° latitude at the tachocline. If the \( s_2 \) and \( s_4 \) are 0, there is no differential rotation, and \( s_0 \) is required to be 1 in order to have \( \omega = 0 \) in the rotating frame. But for a finite differential rotation amplitude, \( s_0 \) is not simply 1 and is chosen so that \( \omega = 0 \) at 32° latitude. Then with the choice of values of \( s_2 \) and \( s_4 \), the differential rotation amplitude becomes \( (s_2 + s_4)/s_0 \).

Following Dikpati & Gilman (1999), we use the same Gaussian expression to represent the spot-producing toroidal magnetic band. For the sake of completeness, we write the expression of the reference-state toroidal magnetic band \((a_0 = \alpha_0 \cos \phi; \alpha_0 \) is the angular measure) as
\[
\alpha_0 = p_0 (e^{-\beta(\mu - d_0)^2} - e^{-\beta(\mu + d_0)^2}).
\]

In expression (8), \( p_0 \) is the field strength, \( \beta \) controls the width of the Gaussian band, \( d_0 \) is the sine latitude of the center of the band, and \( p \) is the prefactor to scale the peak field strength with the change in \( \beta \) in the Gaussian profile so that the value of \( p_0 \) denotes the peak field strength.

### 3.3. Energy Equations and Energy Flow Diagrams

The governing nonlinear Equations (1)–(6) have no dissipation in them, so they conserve the total energy (potential + kinetic + magnetic) of the system. Energies, energy conversions, and energy equations can be useful as diagnostic tools for understanding the nonlinear dynamics of the system, including nonlinear oscillations, even if we do not actually integrate these equations with time. It is useful to subdivide each form of energy into its axisymmetric part, denoted by an overbar, and the longitudinal average of all departures from axisymmetry, variables in which are denoted by primes. With this subdivision, there are six energy “reservoirs,” two for each of the three energy types, denoted by \( \overline{P}, P', K, K', M, M' \), respectively, for potential, kinetic, and magnetic energies. The total energy of the system, \( T_{en} = \overline{P} + P' + \overline{K} + K' + \overline{M} + M' \), is conserved. We write down below the energy integrals for the total system, followed by the energy integrals for the six reservoirs, integrated over the entire spherical shell. The total energy \( (T_{en}) \) of the system can be given by
\[
T_{en} = \int_0^{2\pi} \int_{-\pi/2}^{\pi/2} \cos \phi d\phi d\lambda \cdot \left[ (1 + h) \left( \frac{u^2 + v^2}{2} + \frac{a^2 + b^2}{2} \right) + \frac{G \cdot h(h + 2)}{2} \right].
\]

or, in a \((\mu, \lambda)\) coordinate system, in which \( \mu = \sin \phi \), the total energy integral can be written as
\[
T_{en} = \int_0^{2\pi} \int_{-\pi/2}^{\pi/2} d\mu d\lambda \cdot \left[ (1 + h) \left( \frac{u^2 + v^2}{2} + \frac{a^2 + b^2}{2} \right) + \frac{G \cdot h(h + 2)}{2} \right].
\]

The background potential energy, \( G/2 \), of the system is taken out, since it is unavailable for extraction, because it can do no work. In Equations (9a) and (9b), the total energy \( (T_{en}) \) consists of total potential energy \( (P_{tot}) \), total kinetic energy \( (K_{tot}) \), and total magnetic energy \( (M_{tot}) \), which are
\[
P_{tot} = \int_0^{2\pi} \int_{-\pi/2}^{\pi/2} \cos \phi d\phi d\lambda \cdot \left[ \frac{G \cdot h(h + 2)}{2} \right].
\]
\[
K_{tot} = \int_0^{2\pi} \int_{-\pi/2}^{\pi/2} \cos \phi d\phi d\lambda \cdot \left[ (1 + h) \left( \frac{u^2 + v^2}{2} \right) \right].
\]
\[
M_{tot} = \int_0^{2\pi} \int_{-\pi/2}^{\pi/2} \cos \phi d\phi d\lambda \cdot \left[ (1 + h) \left( \frac{a^2 + b^2}{2} \right) \right].
\]

Since each of these total energies contains both the reference-state and perturbation energy reservoirs, we perform the integration over longitude for each of \( P_{tot}, K_{tot}, \) and \( M_{tot} \) and obtain the corresponding energy integrals for the reference-state and perturbation energy reservoirs:
\[
\overline{P} = \int_{-\pi/2}^{\pi/2} \cos \phi d\phi \cdot \left[ \frac{G \cdot \overline{h}(\overline{h} + 2)}{2} \right].
\]
\[
P' = P_{tot} - \overline{P} = \int_{-\pi/2}^{\pi/2} \cos \phi d\phi \cdot \left[ \frac{G \cdot h^2}{2} \right].
\]
\[
\overline{K} = \int_{-\pi/2}^{\pi/2} \cos \phi d\phi \cdot \left[ (1 + \overline{h}) (\overline{a^2 + v^2}) \right].
\]
\[
K' = K_{tot} - \overline{K} = \int_{-\pi/2}^{\pi/2} \cos \phi d\phi \cdot \left[ (1 + \overline{h}) \left( \frac{u^2 + v^2}{2} \right) + \overline{h}^2 u' v' + \overline{h}^2 \left( a^2 + b^2 \right) \right].
\]
\[
\overline{M} = \int_{-\pi/2}^{\pi/2} \cos \phi d\phi \cdot \left[ (1 + \overline{h}) \left( \frac{a^2 + b^2}{2} \right) \right].
\]
\[
M' = M_{tot} - \overline{M} = \int_{-\pi/2}^{\pi/2} \cos \phi d\phi \cdot \left[ \frac{G \cdot \overline{h}^2 (\overline{a^2 + b^2})}{2} \right].
\]

Note that the last two terms in each of Equations (16) and (18) contain third-order correlations and therefore can be ignored in linear calculations. However, these terms may be important in certain situations during the nonlinear evolution of a
shallow-water system, particularly for low $G$, for which $\eta' \neq 0$ is not necessarily $\ll 1$.

All of these energies are linked to each other by energy conversion processes that appear with opposite signs in the energy equations. We give mathematical detail for the energy equations, including the form of energy conversions from one form to another, in the Appendix. Physically, each of the energy conversion integrals can be described as a form of work done to convert energy from one form to another within the system. The energy conversions of the system can be conveniently summarized in a graphical way in the energy diagram presented in Figure 2 (see the Appendix). Each of the energies and energy conversion processes represented in Figure 2 is time dependent in an actual simulation. Since the total energy of the system is conserved, an increase of one energy form must be accompanied by a compensating decrease in another energy form. Thus, it is easy to see that the system is likely to support the so-called TNOs of Dikpati et al. (2017) among the various energies shown. We also should note that not all energy reservoirs are linked to each other. In general, this is because all energy conversions must involve work done on or by the fluid. Therefore, there are no direct energy conversion links between potential and magnetic energy; they must all go through a kinetic energy reservoir. Analogously, there is no direct conversion of energy between $\mathcal{M}$ and $\mathcal{M}'$. The energy flow must be through a kinetic energy reservoir to get from one magnetic energy form to another (see, e.g., Figure 1 and related text in Gilman & Fox 1997). Somewhat analogously, there are no direct conversion links of the form $(\mathcal{P}, \mathcal{K}')$ or $(\mathcal{P}', \mathcal{K})$ (Dikpati & Gilman 2001a).

In a linear calculation, the energy flows from the reference state to the perturbation (see, e.g., Dikpati & Gilman 2001b), but in the nonlinear evolution, energy can flow back and forth between reference-state and perturbation energy reservoirs. In Figure 2 $(\mathcal{P}, \mathcal{K})$ represents the energy flow from the reference-state potential energy reservoirs to the reference-state kinetic energy reservoir; this happens primarily through the work done by meridional circulation against or with meridional pressure gradients to change the axisymmetric mass distribution. $(\mathcal{P}', \mathcal{K}')$ represents changes in potential energy due to rearranging the mass distribution from axisymmetric to nonaxisymmetric by perturbation velocities. $(\mathcal{P}', \mathcal{K}')$ is the work to create perturbation kinetic energy from potential energy by perturbation pressure forces. $(\mathcal{K}, \mathcal{M}')$ is the work done by Reynolds stresses to convert mainly differential rotation energy into perturbation kinetic energy. $(\mathcal{M}', \mathcal{K}')$ is the work done by the Maxwell stress to convert mainly the differential rotation energy into perturbation magnetic energy. $(\mathcal{M}', \mathcal{K}')$ is the work done by the “mixed stress” to convert mainly the toroidal magnetic field energy into perturbation kinetic energy. $(\mathcal{K}', \mathcal{M}')$ is the work done by the perturbation $j \times b$ force to convert perturbation kinetic into perturbation magnetic energy.

From the energy diagram it is clear that there are multiple possible paths of energy flow that can reverse periodically to support the TNOs. The dominant path will likely be determined by the initial amplitudes of the various energy reservoirs. If initial toroidal fields are weak, the oscillation will be driven primarily by hydrodynamic energy conversions with a relatively passive, but still periodic, response from magnetic energies. For stronger initial magnetic energies, the...
dominant path for oscillation will involve periodic changes in both magnetic and kinetic energies.

\subsection*{3.4. Range of Physical Phenomena Contained in Governing Equations}

The possible range of behavior included in Equations (1)–(6) is very broad. There are steady equilibrium solutions for differential rotation, toroidal field (in the MHD case), and tachocline thickness variations with latitude, previously described in Dikpati & Gilman (2001a, 2001b). These are limited only by the requirement that the tachocline have finite thickness at all latitudes. For lower effective gravity and stronger magnetic fields the deformation of the tachocline top surface can be large, because larger thickness variations are needed in that case to generate the latitudinal hydrostatic pressure gradients that can balance the Coriolis force of differential rotation and the magnetic curvature stress from the toroidal field. Therefore, there is a limit for the combination of low effective gravity and strong magnetic field we can consider so that the global tachocline fluid remains global and does not split into two or multiple, disconnected fluid regions.

Perturbation equations linearized about the steady differential rotation, toroidal field, and layer thickness contain HD and MHD Rossby and gravity waves, as well as Alfven waves (De Sterck 2001; Schecter et al. 2001; Zaqarashvili et al. 2007; Heng & Spitzkovsky 2009; Zaqarashvili et al. 2009, 2010a, 2010b, 2015; Gurgensashvili et al. 2017; Zaqarashvili 2017). They also contain unstable modes that extract kinetic, magnetic, and/or potential energy from the reference states they are perturbing (Gilman & Fox 1997; Charbonneau et al. 1999; Dikpati & Gilman 1999, 2001b; Gilman & Dikpati 2000, 2002; Dikpati et al. 2003). Multiple modes of different longitudinal wavenumber can be excited that can “beat” with each other, producing additional apparent oscillations and amplitude variations in the combined signal (Dikpati & Gilman 2005; Zaqarashvili et al. 2015). Linear solutions, if the reference state satisfies certain symmetry conditions, cleanly separate into two opposite symmetries about the equator, leading to neutral wave and unstable modes that are completely independent of each other for opposite symmetries. This property is common to many of the linear studies cited above.

From previous studies in the nonlinear realm, it is known that there are all the same waves and unstable modes possible, but with finite amplitude (De Sterck 2001; Schecter et al. 2001), as well as so-called “solitary waves” (London 2017), and nonlinear evolution of energies through exchanges among reference-state and perturbation energy reservoirs (Cally et al. 2003; Dikpati 2012; Balk 2014; Raphaldini & Raupp 2015). There can also be nonlinear interactions between unstable, growing modes and coexisting neutral MHD modes (MHD Rossby waves in various forms) from which even resonance can occur. In the nonlinear case, in general there no longer are separate solutions for opposite symmetries; instead, they are coupled together by the nonlinear terms, unless initial conditions are chosen with special symmetry characteristics. Even then, the symmetries may mix owing to truncation-level departures from symmetry that grow over time. Both linear and nonlinear phenomena manifest themselves in different forms at different latitudes, because of the rotation and curvature of the spherical shell.

In the limit of high $G$ the governing Equations (1)–(6) approach those of the corresponding 2D problem, such as first solved in Gilman & Fox (1997). In this limit, terms in Equations (1), (2), (4), and (5) that involve the horizontal divergence of either horizontal velocity or magnetic fields become small compared to other terms. In the limit of weak but not vanishing magnetic fields, the terms in Equations (1) and (2) containing products of magnetic field components are small compared to the hydrodynamic terms, but all terms in the full induction Equations (4) and (5) are still important. In this regime the magnetic fields are passive, not affecting the hydrodynamics of the system. For somewhat higher magnetic fields, the magnetic stress terms in the equations of motion are important, but not dominant. They can destabilize differential rotation that is stable to perturbations when no magnetic fields are present. For the highest field strengths, the electromagnetic stresses dominate over purely hydrodynamic stresses. Then the system behaves very differently than in the hydrodynamic case. We present here the solutions from the full nonlinear MHD-SWT model for both the weak toroidal magnetic band of peak field strength of 5 kG and the strong magnetic band of peak field strength of 35 kG for a low effective gravity value ($G = 0.5$).

\subsection*{3.5. Solution Technique and Initial Condition}

The numerical algorithm and the technique for solving nonlinear shallow-water equations for the solar tachocline have been discussed in detail in Dikpati (2012). Here we briefly discuss the major steps taken in developing the nonlinear shallow-water code:

\begin{enumerate}
\item The variable $h$ is decomposed in scalar spherical harmonics and $u, v, a, b$ in vector spherical harmonics to deal with the pole problem.
\item Nonlinear terms in the equations are computed following pseudo-spectral implementation given in Swarztrauber (1996).
\item A fourth-order Runge–Kutta time integration scheme is implemented for time evolution; semi-implicit dynamics is included following Hack & Jakob (1992) in order to integrate out high-frequency gravity waves. This allows the use of larger time steps.
\item Momentum is checked and balanced in every few thousand steps, in which the model evolves for about 3 days. As discussed in the literature, in order to take care of aliasing error (which is called Gibb’s phenomenon) in the pseudo-spectral formalism, a small numerical viscosity is added as a standard technique. This results in introducing small errors in conserved quantities, for example, in angular momentum integrated over the whole system. In order to deal with this issue, we evaluate the excess or deficit of angular momentum once in every few thousand time steps, and respectively subtract (in case of excess) or add back (in case of deficit) by distributing to the grid points (see, e.g., Gottlieb & Shu 1997; Gelb & Gleeson 2001; Kattelans & Heimrichs 2009).
\item Computer-intensive synthesis and analysis steps are run concurrently in multiple parallel threads on modern many-core processors.
\end{enumerate}

Initial conditions are chosen by setting the differential rotation and toroidal magnetic field profiles in latitude. The reference states are independent of longitude. In order to
construct the initial state of the total differential rotation of the system (and total magnetic field in the case of MHD), first the unstable modes with various longitudinal wavenumbers for the aforementioned reference states are calculated from the corresponding linear instability problem. Then a perturbation with an amplitude of, say, approximately 40% with respect to the reference-state differential rotation amplitude is added to the unperturbed differential rotation. In terms of energy, it is about 16% perturbation energy with respect to the reference-state kinetic energy of differential rotation. The perturbation is constructed in one of two ways: (i) by assigning a random combination of amplitudes to latitude–longitude disturbances of several unstable eigenmodes of the linear system so that the peak amplitudes add to 40%, or (ii) by assigning the 40% amplitude to a single unstable mode that has the highest growth rate.

In previous calculations by some of the authors (Cally 2003; Cally et al. 2003; Dikpati 2012; Dikpati et al. 2017), the nonlinear evolution of the system was found to converge to a similar quasi-periodic exchange among energies of the system in about 1 month, during which the system evolved to a full nonlinear state from the above two initial conditions. However, a more systematic survey of the effect of other initial conditions, such as starting from purely random number perturbations, which may take a much longer time to settle down and sort out the most unstable modes, will be explored in a forthcoming paper.

4. Results

4.1. TNO Generated by Hydrodynamic Rossby Waves’ Interaction with Tachocline Differential Rotation

Extensive linear analysis of shallow-water instability of solar tachocline differential rotation revealed that the latitudinal differential rotation can be unstable to 2D and quasi-3D shallow-water-type HD and MHD disturbances (Gilman & Fox 1997; Dikpati & Gilman 1999; Gilman & Dikpati 2002; Dikpati et al. 2003; Zaqarashvili et al. 2009, 2010a, 2010b). The latitude–longitude disturbance plannorms of the eigenfunction of the unstable modes obtained in linear studies show the eastward tilts (see, e.g., Figure 7 of Dikpati & Gilman 2001a) when it is hydrodynamically dominated, but they show westward tilts when magnetic fields dominate the dynamics. The angular momentum gets transported poleward by the Reynolds stress or Maxwell stress to drive the instability respectively in the hydrodynamic and magnetohydrodynamic regimes.

However, in the nonlinear evolution of a nearly dissipationless system of tachocline latitudinal differential rotation and Rossby waves, Dikpati et al. (2017) found a back-and-forth exchange of angular momentum between the reference-state differential rotation and the disturbances (which are essentially the Rossby waves in tachocline), and hence an oscillation between between the reference-state and perturbation kinetic energies. Their Figure 3 showed that the reference-state kinetic energy ($K$) from tachocline latitudinal differential rotation and perturbation kinetic energy ($K'$) from Rossby waves undergo an out-of-phase oscillation, $K'$ being in its maximum when $K$ is minimum and vice versa. The pole-to-equator differential rotation, which is the unperturbed, reference-state differential rotation, was chosen to be 21% (with $s_2/s_0 = 0.15$ and $s_4/s_0 = 0.06$ in the expression of $\omega$ given in Equation (7)), and the effective gravity parameter was $G = 0.5$.

The example case of Dikpati et al. (2017) produced an oscillation between $K$ and $K'$ of about 6 months. The oscillations of the tilts of the plasma patterns from eastward to westward and back are simple, because that case was dominated by the evolution of $m = 1$ symmetric mode (i.e., symmetric flow and antisymmetric height deformation about the equator), which had faster growth rate (with $e$-folding growth time of about a year) than the other possible unstable mode, namely, the $m = 2$ antisymmetric mode (see, e.g., Table 1). For a different choice of parameters, if several unstable modes can be excited with similar growth rates, interaction among modes makes the situation more complex.

We show in Figure 3 the evolution of the disturbance pattern for a pole-to-equator differential rotation of 27% ($s_2/s_0 = 0.27$ and $s_4/s_0 = 0.0$) and $G = 0.5$. For these choices of parameters, $m = 1$ symmetric and antisymmetric modes and $m = 2$ antisymmetric mode have similar growth rates, with $e$-folding growth occurring approximately in 1–2 yr. Three panels reveal that the tilts of the disturbance pattern, no matter how complex, are oscillating from eastward to westward through neutral tilts in the middle panel and hence causing the back-and-forth angular momentum exchange between the reference-state differential rotation and Rossby waves.

Figure 4 shows the corresponding evolution of $K$ and $K'$. The oscillation between $K$ of the reference-state differential rotation and $K'$ of Rossby waves is quasi-periodic, instead of perfectly periodic, due to interactions among all possible unstable modes. If the dynamics is dominated by these parameters, we should expect quasi-periodic seasons of solar activity. In fact, observations indicate quasi-periodic bursts rather than strictly periodic events. However, in order to be able to simulate these quasi-periodic bursts for a longer time interval of a few years within a solar cycle, it is necessary to estimate the initial condition at the tachocline that is producing the imprints of bursty events seen at the surface. This is possible through the implementation of data assimilation methods, which are being developed for solar models with good progress (see, e.g., Jouve et al. 2011; Dikpati et al. 2014, 2016; Hickmann et al. 2015).

Dikpati et al. (2017) showed how the period of TNOs varies with differential rotation amplitude, effective gravity, and latitude. For a given rotation rate of the star’s interior. We know that there are solar-like stars with both slower and faster rotations. Figure 5 depicts the period of a TNO for different stellar rotation rates with the same absolute

<table>
<thead>
<tr>
<th>$s_2/s_0$</th>
<th>$s_4/s_0$</th>
<th>$G$</th>
<th>Longitudinal Wavenumber</th>
<th>Mode</th>
<th>$e$-folding Growth</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.15</td>
<td>0.06</td>
<td>0.5</td>
<td>$m = 1$</td>
<td>Symmetric</td>
<td>0.008</td>
</tr>
<tr>
<td>0.27</td>
<td>0.0</td>
<td>0.5</td>
<td>$m = 1$</td>
<td>Symmetric</td>
<td>0.004</td>
</tr>
</tbody>
</table>

Table 1: Linear Growth Rate of Various Unstable Modes for Selected Differential Rotation Amplitudes and $G$ Values
differential rotation between equator and pole. We performed two sets of simulation experiments, one with 5% initial perturbation with respect to the kinetic energy of differential rotation and the other with 16% initial perturbation. TNO periods as a function of rotation rate of the star are respectively plotted with filled circles and diamonds. The curve for the simulation experiment with 5% initial perturbation in energy lies above that obtained with 16% initial perturbation, because the TNO is expected to slow down with the decrease in initial perturbation (see, e.g., Figure 3(b) of Dikpati 2012). We find from both simulation experiments that, as the rotation rate is increased, for the same dimensional differential rotation the TNO period gets shorter. Stars with twice the rotation rate of the Sun have oscillations whose period is shorter, about two-thirds that of the Sun. This is presumably because, as a percentage of the interior rotation rate, the differential rotation is only half of that of the solar case. This means that there is only one-fourth the kinetic energy of differential rotation available to feed an oscillation, and hence it is quickly spent and then gets rebuilt, thus completing a full TNO cycle quickly.

4.2. Physics of Magnetic Rossby Waves' Interaction with Tachocline Differential Rotation

Dikpati et al. (2017) described schematically the mechanics of interaction between the hydrodynamic Rossby waves and differential rotation in producing TNOs, in terms of how the periodically reversing “tilted” structure of velocity perturbations exchanges their kinetic energy with the differential rotation. As we have seen in Figure 2 for the energy reservoirs and transfers in the MHD case, with magnetic fields actively participating in the dynamics, there are two additional energy reservoirs and several additional paths over which energy can be transferred between reservoirs. To understand the physics
behind the mechanism of interaction among HD and MHD Rossby waves, differential rotation, and magnetic fields, we present in Figure 6 a schematic diagram, which is similar to Figure 2 of Dikpati et al. (2017), but with magnetic perturbations included. In Figure 6 red denotes hydrodynamic variables and structures, and blue represents magnetic variables and structures. We start from a system of differential rotation that is maximum at the equator and a banded toroidal field of opposite sign in the northern and southern hemispheres.

From experience in studying the structures of growing velocity and magnetic field perturbations, we know that there is a tendency for the velocity and magnetic perturbations to have similar patterns in longitude and latitude, but with a significant phase difference in longitude between them. In Figure 6 the uppermost pattern depicts a starting point in time, with velocity and magnetic patterns tilted “upstream” away from the equator. In this configuration the differential rotation is stable to the hydrodynamic perturbation, because the tilt shown implies angular momentum transport by the perturbation Reynolds stress toward the equator in both hemispheres, which would increase the differential rotation. But the same sense of tilt in the magnetic patterns implies a transfer of angular momentum away from the equator by the Maxwell stress. Therefore, in this representation there will be net extraction of energy from differential rotation to grow perturbations only if the Maxwell stress is large enough to overcome the Reynolds stress. From linear instability studies we know that this is true even when the toroidal field is rather weak, leading to instability of the combination of differential rotation and toroidal field that is stable if the toroidal field is absent.

From this starting point the nonlinear oscillation proceeds, with extraction of kinetic energy from differential rotation until it is no longer unstable. At this point both velocity and magnetic perturbations lose their tilts and temporarily become “energetically neutral” (middle panel). But then the system overshoots, leading to opposite disturbance tilts (bottom panel), implying that the differential rotation is rebuilding its amplitude.

As shown in the schematic, there is another energy exchange process going, involving energy stored in the toroidal field itself. Changes in toroidal field energy are driven by the so-called “mixed stress,” or correlation between perturbation velocity and magnetic field patterns. In the energy Equations (10)–(15), this is represented by \( \mathbf{F}_t, K' \). Qualitatively, it has the effect of changing the toroidal field profile through the sequence of patterns shown on the extreme right in the schematic.

The schematic does not show the relative strengths of the various transport and energy conversion processes. Fundamentally that is determined by the relative size of the energy reservoirs \( \mathcal{K} \) and \( \mathcal{M} \). If \( \mathcal{K} \gg \mathcal{M} \), then the Reynolds and Maxwell stresses will dominate in driving the instability and nonlinear oscillation, while if \( \mathcal{K} \ll \mathcal{M} \), the mixed stress will dominate in driving the instability. The resulting nonlinear oscillations in these two cases will be somewhat different. We will examine these differences below and in papers to follow.

4.3. Exchanges among Energy Reservoirs for an Example Case of a 10° Band at 45° Latitude

4.3.1. Case Study for Weak Toroidal Band with 5 kG Peak Field

Here we give an example of a nonlinear oscillation when magnetic fields are present in the shallow-water system. For this case we have chosen the amplitude of differential rotation as to be again 21%, and the toroidal field as to have 5 kG peak field strength. Initially the perturbation energies are 16% of the initial differential rotation energy.

The evolution of the various energy reservoirs for about 90 time units is shown in Figure 7. This is for a case for which the initial magnetic energy is about 4 times smaller than the kinetic energy of differential rotation, when integrated over the whole volume of the tachocline fluid shell. But in the latitude range of the toroidal band, magnetic field and differential rotation energies are of the same order. The toroidal field peak is about 5 kG. We clearly see the nonlinear oscillation in all six reservoirs, for which there is a clear period near 60 units, or 7.2 months. Furthermore, all reference-state energies \( (\mathcal{K}, \mathcal{P}, \mathcal{M}) \), fall and rise together, as do all perturbation energies \( (K', P', M') \), but out of phase with the reference states.
In all simulations, the perturbation potential energy $P'$ is a good measure of the amount of deformation of the top boundary of the tachocline. As argued in Dikpati & Gilman (2005), the maximum deformation is where we should expect toroidal flux to be injected into the convection zone and rise to the surface to form active regions. In this simulation, as well as the HD simulations here and in Dikpati et al. (2017), $P'$ is maximum near where both $K'$ and $M'$ are maximum. Therefore, the maxima in both horizontal and vertical magnetic perturbations occur when the deformations of the top surface of the tachocline fluid are maximized, which increases the likelihood of a new bursty season.

It is instructive to also look at longitude–latitude plots of the field and flow for this case. Figure 8 shows an example, a sequence of four frames, containing perturbation velocities (black arrows), perturbation magnetic fields (white arrows), and tachocline top-surface deformations (blue is negative, hence representing depression, and red swelling). Velocity and magnetic arrows are scaled to a single maximum value, so the relative lengths of velocity and magnetic arrows are accurate. In Figure 7 the perturbation magnetic energy is much smaller than the perturbation kinetic energy, but this is due primarily to the fact that perturbation velocities occur at all latitudes, while perturbation magnetic fields are confined to the neighborhood of the latitude of the undisturbed toroidal band, in this case at 45°. The four frames shown are from approximately 15, 30, 45, and 60 time units in the simulation, corresponding to rising, peaking, and falling and minimum perturbation kinetic energy, respectively. Note that 100 time units correspond to approximately 1 yr. The fact that near the toroidal field band maximum velocity and magnetic arrows are of similar length says that there the dynamical effect of the field is strong.

![Figure 6](image)

**Figure 6.** Schematic diagram illustrating time-varying interactions among differential rotation, toroidal field, and tilts in perturbation velocity and magnetic patterns through one period of an MHD tachocline nonlinear oscillation. Velocity profiles and patterns are plotted in red; magnetic profiles and patterns are shown in blue. The dashed profiles on the right are repeats of those on the left, added to show evolution more clearly. The details of the evolution are discussed in the text.

![Figure 7](image)

**Figure 7.** Time evolution of the six energy reservoirs (kinetic in red, potential in blue, magnetic in purple) for an example MHD simulation with a toroidal band of 10° width and 5 kG peak field strength, placed at 45° latitude in each hemisphere. The black line near the top is the total energy of the system, which is conserved over the whole length of the simulation.
Figure 8. Sequence of synoptic maps of perturbation velocities (black arrows), magnetic fields (white arrows), and top boundary deformations (color maps in which red represents bulges; blue, depressions), at times 15, 30, 45, and 60 units (arranged top to bottom) for the evolution of the system of HD and MHD Rossby waves along with reference states of 21% differential rotation and 5 kG toroidal magnetic bands oppositely directed in the northern and southern hemispheres.
It is immediately evident that the chart for the active participation of a magnetic band is significantly more complicated than the typical HD patterns shown in Figure 3. Even a modest toroidal band actively participating in the dynamics has a significant influence on the velocity patterns, particularly in the cross-equatorial latitude range from 40°N to 40°S. Perturbation velocities there are much larger than in the corresponding case without toroidal bands present. In effect, the toroidal magnetic bands are “channeling” some of the flow in the low-latitude belt. Strong cross-equatorial flow is also evident in all four frames. But the sense of tilt of the velocity vectors in these low latitudes reverses with longitude, so there is no obvious net angular momentum transport either toward or away from the equator, when the velocity correlation is averaged in longitude. To first order, the velocities look neutral energetically, with the toroidal magnetic bands at 45° acting as a wave guide. Clearly, with the toroidal magnetic bands participating actively in the dynamics, there are stronger dynamical links between northern and southern hemispheres.

When we focus on the patterns in the latitudes where the toroidal magnetic bands reside, we can pick out tilting velocity and magnetic field vectors that look qualitatively similar to those in the schematic seen in Figure 6, corresponding to poleward or equatorward transport of angular momentum by Reynolds and Maxwell stresses. We can also detect phase differences in longitude, implying that mixed stresses are present that can extract energy from the toroidal field. When we compare typical features in Figure 8 with the energy variations with time in Figure 7, unlike in the HD case (Figure 3) where a rise in $K'$ energy must come from a decline in $K$ energy, in the MHD case $K'$ could be rising also owing to transfers from both $\mathbf{M}$ and $M'$, $\mathbf{M}$ is declining and $M'$ is rising, so in this case $K'$ is getting energy both from $\mathbf{K}$ through Reynolds stress and from $\mathbf{M}$ through mixed stress; $M'$ is getting energy from $\mathbf{K}$ through Maxwell stress. Unlike the idealized schematics in Figure 6, there may be anomalous tilts in velocity or magnetic perturbations at different times. Nevertheless, we see the least magnitude of tilt in most features in panel (d), which from Figure 7 is the time of minimum energy for perturbation kinetic energy and toroidal field, but where perturbation magnetic energy and differential rotation energies are maximized. We infer that at this time the Maxwell stress, coming from the relatively large magnetic perturbations near 45°, has rebuilt the differential rotation by the maximum amount that it is able to achieve.

If we focus on the height deformations, we can see that there are clearly only certain longitude intervals where the toroidal fields coincide with the outward bulges in the top surface. It is these sites where we should expect new bursts of solar activity to be appearing when magnetic fields coincide there. We can also see that both the color and arrow patterns, mostly $m = 1$ patterns, propagate retrograde relative to the rotation of the system, completing a 360° movement relative to the rotating frame in about 30 time units, or about 4 months (compare panel (a) with panel (c), or panel (b) with panel (d)). All latitudes seem to move at about the same rate. Two effects are at work here: without differential rotation present a pure HD Rossby wave has the largest retrograde propagation with $m = 1$. Second, the global disturbance patterns, being MHD, are at least somewhat tied to the rotation of the latitude, in this case 45°, of the peak of the toroidal field. This adds an additional retrograde rotation. We anticipate that for a toroidal band placed in the middle of the zone of sunspot occurrence, say, at 20°, this retrograde speed may be substantially smaller, or can be prograde.

The multiple pathways possible for energy conversion of one form to another is bound to add more variability to the nonlinear oscillations, in both period and amplitude, but TNOs are definitely still present. A future paper will present results for a detailed energetic analysis for solutions for a wide range of placements of toroidal bands and field strengths, as well as different effective gravities.

4.3.2. Case Study for Strong Toroidal Band with 35 kG Peak Field

Results for the high toroidal field case are shown in Figures 9 and 10. Figure 9 contains the time variations in all the energy reservoirs, analogous to those for the low-field case seen in Figure 7. In the strong-field case, the reference-state magnetic energy $\mathbf{M}$ initially is large compared to that in the differential rotation $\mathbf{K}$. As the perturbations grow, they are being supplied much more energy from the toroidal field than from differential rotation; $\mathbf{M}$ declines much more in both absolute and percentage terms than does $\mathbf{K}$. Both $M'$ and $K'$ are growing rapidly. We know from the energy flow diagram in Figure 2, as well as from the Appendix, that $K'$ grows primarily by action of the mixed stress arising from the phase difference in longitude between velocity and magnetic field perturbations. Here the relatively weak differential rotation is acting as a catalyst. The strong toroidal field implies a strong perturbation $j \times B$ body force, which quickly converts the growing perturbation kinetic energy into perturbation magnetic energy. The strong Maxwell stress acting on the differential rotation adds to the perturbation magnetic energy. Together these two transfer processes lead to perturbation magnetic energy being larger than perturbation kinetic energy. For strong fields, potential energy plays very little role in energy transfers. However, tachocline bulges associated with variations in $P'$ are still very much present. It is not surprising that in the strong-field case $\mathbf{M}$ and $\mathbf{K}$ rise and fall together, because both changes are driven by magnetic stresses and $j \times B$ forces, which dominate over Reynolds stresses and all other hydrodynamic processes.
The differences between energy variations in strong- and weak-field cases is dramatic. For the strong-field case, the system is so nonlinear that perturbation kinetic and magnetic energies are for long stretches of time larger than the respective reference states. That is virtually never true in the weak-field case. Also, in the strong-field case, once the nonlinear solutions

Figure 10. Same as in Figure 8, but for toroidal magnetic bands with 35 kG peak field strength, oppositely directed in the northern and southern hemispheres.
are fully developed, perturbation kinetic and magnetic energies are similar in magnitude. This is a form of “equipartition of energy” made possible by the strong linkage between the two forms by the strong magnetic fields. It is interesting to note that such an equipartition is characteristic of linear Alfvén waves, no matter what the field strength. The perturbations here are not Alfvén waves, but rather a form of MHD Rossby waves that have qualitatively similar equipartition properties, while propagating in longitude at rates governed by the local rotation rate and Coriolis forces. Because the toroidal field is strong, this speed is close to the local rotation rate at the latitude (in this case 45°) where the toroidal field peaks.

In Figure 10 we display a sequence of planforms of the perturbations in velocity (black arrows), magnetic fields (white arrows), and tachocline thickness (color shades) in the same way as shown in Figure 8 for the weak-field case, but the snapshots are spaced nonuniformly in time according to the variation in dynamics. In frame 10a, seven time units after the start of the integration, we see the growing velocity and magnetic perturbations in a relatively simple m = 1 pattern.

There are clearly tilts and phase differences in the patterns that are causing energy to be extracted from the reference state, particularly the toroidal field. By t = 22 units, when from Figure 9 a substantial amount of energy has been extracted from the system, the patterns have become much more complex, with higher wavenumbers and both longitude and latitude now present. In subsequent frames shown, for t = 35 and t = 45, the patterns wax and wane somewhat in complexity (t = 35 has less complex pattern structure than t = 45). Clearly there is far more going on here than a simple “tipping” of toroidal field rings, as seen in earlier 2D nonlinear MHD simulations (Cally et al. 2003). Another strong feature of the sequence of frames is that the perturbation magnetic field clearly spreads wider in latitude with time. This is a symptom of the reference-state toroidal field being widened, which it must be in order to supply energy to the perturbations. With total magnetic flux being conserved, the peak toroidal field must decline. But from Figure 9, a minimum in the field is reached, beyond which the toroidal field begins to rebuild, by taking back energy from the perturbations. Energetically this happens by the phase difference between magnetic and velocity perturbations becoming smaller, even temporarily reversing. At the same time, the tilts in field lines also reverse, allowing a rebuild of the differential rotation. In all frames shown one can identify examples of tilted flows and fields, implying Reynolds and Maxwell stresses at work, as well as phase differences between velocity and magnetic perturbations that imply that mixed stresses are present. There is enough complexity in the planforms that it is possible to have tilts of opposite sign in different locations at the same time, and both signs of phase differences as well. It is only the net integrated effect of all the local features that determines how the various energy reservoirs evolve with time.

5. Concluding Remarks

We have developed and presented here a full nonlinear magnetohydrodynamic shallow-water model for the solar tachocline that can simulate nonlinear dynamics for a wide range of differential rotations and toroidal fields that are plausible for the tachocline, and for a wide range of possible effective gravities, including both weak (for the overshoot tachocline) and strong (for the radiative tachocline).

We describe the full range of energy reservoirs and energy conversions that are included in this model, and we illustrate schematically the dominant transfer processes that can oscillate in direction within the overall system for which total (potential + kinetic + magnetic) energies are conserved for extended numerical integrations. These processes generally involve Reynolds and Maxwell stresses interacting with differential rotation and MHD Rossby waves, and “mixed stresses” that directly affect the toroidal field. Because in the shallow-water system the thickness of the tachocline fluid shell varies with longitude, latitude, and time, there are additional energy transfers between reservoirs that involve correlations of thickness variations with perturbation velocities, magnetic fields, or both. These may be particularly important for cases of low effective gravity, for which the thickness variation amplitudes will be largest. The possible oscillatory behavior, and nonlinear behavior generally, is inherently more complex when magnetic fields actively participate in the dynamics, since it has six energy reservoirs and eight possible pathways among reservoirs, compared to four reservoirs and four pathways in a purely hydrodynamic system, when tachocline magnetic fields are only passively participating.

We simulate examples of TNOs for both HD and MHD cases, which both involve interactions with Rossby waves with multiple longitudinal wavenumber m. In the HD case we show that even with multiple m values having similar amplitude, the oscillation is coherent in time and selects a dominant symmetry about the equator. The oscillation period is also relatively insensitive to an amplitude range of differential rotation plausible for the tachocline. In the MHD case we should expect nonlinear oscillatory behavior of the system to depend significantly on the profile and amplitude of the toroidal field present, the latitude of the magnetic band, its width as well as amplitude, and whether it has a dominant symmetry about the equator or not. In the initial example, for a band centered at 45° latitude, of width 10°, and dimensional amplitude of about 5 kG, we find that the TNO period still falls within the range reported in Dikpati et al. (2017).

We are also able to verify with disturbance planform plots the perturbation velocity and magnetic pattern tilts and phase differences in longitude needed to effect the energy transfers necessary to sustain the oscillations. In particular, we find that the Reynolds and Maxwell stresses tend to oppose each other, which, other properties being equal, should tend to lengthen the oscillation period compared to an HD case with the same differential rotation, as does the fact that growth of perturbation velocities tends to be opposed by the perturbation j × B force. But with the increase of field strength beyond 20 kG, periods decrease again, as seen in Dikpati et al. (2017). This is most likely due to the mixed stress dominating over Maxwell’s stress.

For both weak and strong toroidal fields, we find that the perturbation energies (P, K, and M) are all nearly in phase, but out of phase with all reference-state energies (K, P, M). When the peak total (toroidal plus perturbation) magnetic field is found at a latitude and longitude where the deformation of the top boundary is also at or near a maximum, is where and when we might expect an emerging bursty season to be of the greatest amplitude. It is these events that should lead to the most extreme space weather events that impact Earth. The degree to which this happens can only be determined by assimilating surface magnetogram data into the MHD-SWT model to find the
latitude–longitude locations of peaks of top-surface height deformation, velocity, and magnetic perturbations.

Forthcoming papers will answer the following basic questions about TNOs: (i) how the energy conversion processes vary with the change in the latitude location of the toroidal band, and (ii) how the TNO properties differ when there appears the opposite-polarity new cycle’s toroidal bands at high latitudes along with the old cycle’s bands present at low latitudes near the equator.

We thank Phil Judge for thoroughly reviewing the entire manuscript, Yuhong Fan for helpful discussion on energy conversion through the work of the perturbation $j \times b$ force, and Prof. Eyal Heifetz for helpful discussions on the nonlinear Or mechanism. We extend our thanks to an anonymous referee of this paper for his/her helpful and constructive comments, which have helped significantly improve our paper. The simulation runs have been performed on the Cheyenne Supercomputer of NWSC2/NCAR under project No. P22104000 and NCAR Strategic Capability project with award No. NHA00013. The National Center for Atmospheric Research is sponsored by the National Science Foundation.

**Appendix**

Here we give more detail on energy equations and energy conversions for the MHD-SWT system defined by Equations (1)–(6) in the main text. We begin with equations for the total potential, kinetic, and magnetic energies of the system that are not separated into energies for longitudinally averaged variables and departure therefrom. We first derive equations for the time rate of change of each energy at a longitude and latitude point of the spherical shell, by multiplying Equations (1)–(5), respectively, by the quantities $(1 + h)u$, $(1 + h)v$, $(1 + h)G$, $(1 + h)a$, $(1 + h)b$ and then adding, respectively, to Equations (1), (2), (4), (5), and (3) multiplied, respectively, by $a^2/2$, $v^2/2$, $a^2/2$, $b^2/2$. Following these steps, we arrive at equations for, respectively, the time rates of change of the potential energy $G(1 + h)^2/2$, the kinetic energy $(1 + h)(u^2 + v^2)/2$, and the magnetic energy $(1 + h)(a^2 + b^2)/2$ at each longitude and latitude point in the shallow-water shell. Then, if we integrate over all longitudes and latitudes, employing differentiation and integration by parts, the fact that all functions are single valued in longitude, and Prof. Eyal Heifetz for helpful discussions on the nonlinear Or mechanism. We extend our thanks to an anonymous referee of this paper for his/her helpful and constructive comments, which have helped significantly improve our paper. The simulation runs have been performed on the Cheyenne Supercomputer of NWSC2/NCAR under project No. P22104000 and NCAR Strategic Capability project with award No. NHA00013. The National Center for Atmospheric Research is sponsored by the National Science Foundation.

**Appendix**

Here we give more detail on energy equations and energy conversions for the MHD-SWT system defined by Equations (1)–(6) in the main text. We begin with equations for the total potential, kinetic, and magnetic energies of the system that are not separated into energies for longitudinally averaged variables and departure therefrom. We first derive equations for the time rate of change of each energy at a longitude and latitude point of the spherical shell, by multiplying Equations (1)–(5), respectively, by the quantities $(1 + h)u$, $(1 + h)v$, $(1 + h)G$, $(1 + h)a$, $(1 + h)b$ and then adding, respectively, to Equations (1), (2), (4), (5), and (3) multiplied, respectively, by $a^2/2$, $v^2/2$, $a^2/2$, $b^2/2$. Following these steps, we arrive at equations for, respectively, the time rates of change of the potential energy $G(1 + h)^2/2$, the kinetic energy $(1 + h)(u^2 + v^2)/2$, and the magnetic energy $(1 + h)(a^2 + b^2)/2$ at each longitude and latitude point in the shallow-water shell. Then, if we integrate over all longitudes and latitudes, employing differentiation and integration by parts, the fact that all functions are single valued in longitude, and

\[
\frac{\partial P_{\text{tot}}}{\partial t} = (P, K) \tag{19}
\]

\[
\frac{\partial K_{\text{tot}}}{\partial t} = -(P, K) + (K, M) \tag{20}
\]

\[
\frac{\partial M_{\text{tot}}}{\partial t} = -(K, M), \tag{21}
\]

in which $P_{\text{tot}}$, $K_{\text{tot}}$, $M_{\text{tot}}$ are as written in Equations (10)–(12),

\[
(P, K) = G \int_0^{2\pi} \int_{-\pi/2}^{\pi/2} \left\{ u \frac{\partial}{\partial \lambda} [(1 + h)^2/2] + v \cos \phi \frac{\partial}{\partial \phi} [(1 + h)^2/2] \right\} d\lambda d\phi \tag{22}
\]

and

\[
(K, M) = \int_0^{2\pi} \int_{-\pi/2}^{\pi/2} (1 + h)u
\]

\[
\times \left[ \frac{\partial}{\partial \lambda} (a^2/2) + b \frac{\partial}{\partial \phi} (a \cos \phi) \right] d\lambda d\phi
\]

\[
+ \int_0^{2\pi} \int_{-\pi/2}^{\pi/2} (1 + h)v
\]

\[
\times \left[ a^2 \sin \phi + b \frac{\partial}{\partial \lambda} + \cos \phi \frac{\partial}{\partial \phi} (b^2/2) \right] d\lambda d\phi. \tag{23}
\]

Then in Equations (19)–(21) above, the “energy conversion integral” $(P, K)$ clearly represents work done by or against (depending on sign) the horizontal pressure gradients to exchange energy between $P$ and $K$, while $(K, M)$ represents work done by or against the electromagnetic body force, to exchange energy between $K_{\text{tot}}$ and $M_{\text{tot}}$. By inspection, summing Equations (19)–(21) yields zero on the right-hand side, so the rate of change of total (potential+kinetic +magnetic) is zero, proving that the MHD shallow-water system indeed conserves total energy.

This conservation property remains when we divide all energies into reference-state (denoted by overbars) and perturbation (denoted by primes) reservoirs, as defined in Equations (13)–(18), but now energy conversions are allowed between reference-state and perturbation energies. We can work all this out using the same methods as above for the total energies, but the process is much more complicated, since there are twice as many energy reservoirs and four times as many categories of energy conversions as seen in Equations (19)–(21). In addition, each energy conversion integral has many more terms in it, especially those involving conversions between perturbation kinetic and magnetic energies. By inspection, we see that energy integrals in Equations (19)–(21) involve triple products of variables, and energy conversion integrals, quadruple products. Thus, when each variable is split into overbar and prime parts, there will result terms without primes and terms with one, two, three, or four primes. When integrated over longitude, all terms with single primes average out, so they need not be considered further. Terms with no primes represent energies and energy conversions strictly for and among reference-state variables. It is not difficult to show that energy conversion terms taking energy from or adding to a reference-state reservoir to or from a perturbation reservoir involve only quadratic products of primed variables. The total number of these is relatively modest. By contrast, energy conversion integrals between two perturbation reservoirs will include terms with quadratic, triple, and quadruple products, of which in the full system there are potentially hundreds, depending on how they are written. In our judgment, it is not productive to derive, display, or calculate all these terms individually. Instead, here we limit ourselves to those quadratic terms that define the energy conversions between reference-state and perturbation reservoirs and show how in an actual simulation the rest of the terms, in the aggregate, can be inferred without directly calculating them.

In practical terms, we arrive at energy equations for the reference-state energies $\bar{P}, \bar{K}, \bar{M}$ by averaging Equations (1)–(6) over longitude after expanding each variable into a longitude-averaged part (overbars) and departures (primes). Then, we follow exactly the same
procedure as resulted in Equations (19)–(21). In symbolic form, the result is

$$\frac{\partial F}{\partial t} = (\overline{F}, \overline{K}) + (\overline{F}, P')$$

(24)

$$\frac{\partial K}{\partial t} = - (\overline{F}, \overline{K}) + (\overline{K}, \overline{M}) + (\overline{K}, K') + (\overline{K}, M')$$

(25)

$$\frac{\partial M}{\partial t} = -(\overline{K}, \overline{M}) + (\overline{M}, K'),$$

(26)

in which the energy conversion integrals now become

$$(\overline{F}, \overline{K}) = \int_{-\pi/2}^{\pi/2} G(1 + \hbar)\overline{\nabla}_h \cos \phi d\phi;$$

(27)

$$(\overline{F}, P') = \int_{-\pi/2}^{\pi/2} \overline{Gh} \overline{\nabla}_h \cos \phi d\phi;$$

(28)

$$(\overline{K}, \overline{M}) = \int_{-\pi/2}^{\pi/2} \left(1 + \hbar \right) \left[ \overline{\nabla}_h \overline{\nu} \overline{\nabla}_h (\overline{a} \cos \phi) + \overline{\nu} \overline{\nabla}_h \overline{\nu} \overline{\nabla}_h (\overline{b}^2/2) \right] d\phi;$$

(29)

$$(\overline{K}, K') = -\int_{-\pi/2}^{\pi/2} \left(1 + \hbar \right) \left\{ \overline{\nabla}_h \overline{\nabla}_h (u' \cos \phi) + \overline{\nu} \left[ \overline{u}' \cos \phi + \overline{u} \frac{\partial \overline{u}}{\partial \lambda} \cos \phi \overline{\nu} \right] \frac{\partial \overline{\nu}^2/2}{\partial \phi} \right\} d\phi;$$

(30)

$$(\overline{M}, \overline{M}') = \int_{-\pi/2}^{\pi/2} \left\{ \overline{nh} \overline{\nabla}_h (a' \cos \phi) \right.$$  

$$+ \overline{\nu} \left[ \overline{d}' \overline{d} + \overline{d}^2 \sin \phi + \overline{\nu} \frac{\partial \overline{\nu}^2/2}{\partial \phi} \right] \right\} d\phi;$$

(31)

Given the energy conversion integrals defined in Equations (27)–(32) above, we can write down the energy equations for perturbation energies in symbolic form as

$$\frac{\partial P'}{\partial t} = -(\overline{F}, P') + (\overline{F}, P');$$

(33)

$$\frac{\partial K'}{\partial t} = -(\overline{P'}, K') - (\overline{K}, K') - (\overline{M}, K') + (K', M');$$

(34)

$$\frac{\partial M'}{\partial t} = -(K', M') - (\overline{K}, M').$$

(35)

We can use these three equations for perturbation energy changes in sequence to infer the signs and magnitudes of the conversion integrals ($P', K'$) and ($K', M'$), since with an actual integration of Equations (1)–(6) we can calculate all the energies defined by Equations (13)–(18) as frequently as we wish, along with the integrals defined in Equations (27)–(32), by separating axisymmetric from perturbation parts of each variable. Then once ($P', K'$) is inferred from Equation (33), since all other terms are known, we can infer ($K', M'$) from Equation (34). Then Equation (35) can be used as a check, since all terms in it are then known.

Finally, we note that Figure 2 is a diagrammatic representation of the six energy Equations (24)–(26) and (33)–(35).

### ORCID iDs

Mausumi Dikpati @ https://orcid.org/0000-0002-2227-0488  
Scott W. McIntosh @ https://orcid.org/0000-0002-7369-1776  
Gregory Bothun @ https://orcid.org/0000-0002-9515-6360  
Paul S. Cally @ https://orcid.org/0000-0001-5794-8810  
Peter A. Gilman @ https://orcid.org/0000-0002-1639-6252

### References

Cally, P. S. 2003, MNRA, 339, 957  
Cally, P. S., Dikpati, M., & Gilman, P. A. 2008, MNRA, 391, 891  
De Sterck, H. 2001, PhPl, 8, 3293  
Dikpati, M., Cally, P. S., McIntosh, S. W., & Heifetz, E. 2017, NatSci, 7, 14750  
Dimitropoulou, M., Moussas, X., & Strintzi, D. 2008, MNRA, 386, 2278  
Gelb, A., & Gleeson, J. P. 2001, MWRv, 129, 2346  