Princeton Plasma Physics Laboratory

PPPL-4102
PPPL-4102

Bifurcation Theory of the Transition to Collisionless Ion-temperature-gradient-driven Plasma Turbulence

R.A. Kolesnikov and J.A. Krommes

September 2005

PPPL
PRINCETON PLASMA PHYSICS LABORATORY

Prepared for the U.S. Department of Energy under Contract DE-AC02-76CH03073.

PPPL Report Disclaimers

Full Legal Disclaimer

Trademark Disclaimer

PPPL Report Availability

This report is posted on the U.S. Department of Energy's Princeton Plasma Physics Laboratory Publications and Reports web site in Fiscal Year 2005. The home page for PPPL Reports and Publications is: http://www.pppl.gov/pub_report/

Office of Scientific and Technical Information (OSTI):

Available electronically at: http://www.osti.gov/bridge.

Available for a processing fee to U.S. Department of Energy and its contractors, in paper from:

U.S. Department of Energy
Office of Scientific and Technical Information
P.O. Box 62
Oak Ridge, TN 37831-0062

Telephone: (865) 576-8401

Fax: (865) 576-5728

E-mail: reports@adonis.osti.gov

National Technical Information Service (NTIS):

This report is available for sale to the general public from:

U.S. Department of Commerce
National Technical Information Service
5285 Port Royal Road
Springfield, VA 22161

Telephone: (800) 553-6847

Fax: (703) 605-6900

Email: orders@ntis.fedworld.gov

Online ordering: http://www.ntis.gov/ordering.htm

Bifurcation theory of the transition to collisionless ion-temperature-gradient-driven plasma turbulence

R. A. Kolesnikov and J. A. Krommes*

Plasma Physics Laboratory, Princeton University,

P.O. Box 451, Princeton, New Jersey 08543–0451

(Dated: September 13, 2005)

Abstract

The collisionless limit of the transition to ion-temperature-gradient-driven plasma turbulence is considered with a dynamical-systems approach. The importance of systematic analysis for understanding the differences in the bifurcations and dynamics of linearly damped and undamped systems is emphasized. A model with ten degrees of freedom is studied as a concrete example. A four-dimensional center manifold (CM) is analyzed, and fixed points of its dynamics are identified and used to predict a "Dimits shift" of the threshold for turbulence due to the excitation of zonal flows. The exact value of that shift in terms of physical parameters is established for the model; the effects of higher-order truncations on the dynamics are noted. Multiple-scale analysis of the CM equations is used to discuss possible effects of modulational instability on scenarios for the transition to turbulence in both collisional and collisionless cases.

PACS numbers: 52.35.Ra, 52.35.Qz, 05.45.-a

I. INTRODUCTION

Nonlinearly generated turbulence in hot inhomogeneous plasmas is the main reason for cross-field heat losses from fusion devices. Thus study of the regulatory mechanisms of turbulent transport is important for improvement of the magnetic confinement of plasmas.

It is believed that nonlinearly generated $\boldsymbol{E} \times \boldsymbol{B}$ poloidal (zonal) flows (ZF's) are major candidates for reducing unwanted radial transport; they act by shearing apart eddies associated with the underlying turbulence. It has been recognized both numerically[1] and experimentally[2] that ZF's play a central role in the process of low- to high-confinement transitions in the edge region, as well as in the formation of internal transport barriers. Since linear collisionless processes (Landau damping) do not damp ZF's in toroidal geometry,[3] ion–ion collisional dissipation[4] seems to be an important mechanism regulating the level of ZF's, so is one of the factors imposing limitations on confinement properties. In the hot core of tokamak plasmas, though, collisional dissipation is very small, raising the issue of the proper treatment of essentially collisionless nonlinear ZF dynamics. Many aspects of ZF physics have been recently reviewed by Diamond *et al.*[5]

An extreme limit that demonstrates the importance of ZF's is collisionless ion-temperature-gradient-driven (ITG) physics close to linear marginality. Large-scale collisionless gyrokinetic[6] and gyrofluid[7,8] simulations of ITG systems show that transport is virtually suppressed in a regime just above the threshold for marginal stability. That is, there is an effective nonlinear upshift of the critical temperature gradient for the onset of collisionless ITG turbulence. This is the so-called *Dimits shift*. In this paper, we consider the calculation of the Dimits shift with the aid of systematic bifurcation theory, using a simple but nontrivial model.

It has been realized that ZF's play an important role in the formation of the Dimits shift. In particular, Rogers *et al.*[9] argued that in this regime the plasma can generate undamped zonal flows that are sufficiently weak to remain stable yet are sufficiently strong to suppress the linear drift-wave (DW) mode. (Generally in this article we shall refer generically to DW's rather than the more specific ITG modes.) The boundary of the Dimits shift was asserted to be determined by the generation of tertiary modes that are observed to grow to nonlinear amplitudes and damp the zonal flows. This scenario involves three processes: primary instability of the DW's; secondary (Kelvin–Helmholtz-related) instability of the

ZF's (driven by DW's), which then (totally) suppress the DW's; and tertiary instability, which destabilizes the ZF's.

Since each of these processes has its own onset threshold, one might attempt to identify the above scenario with a sequence of bifurcations occurring in the system as some parameter ϵ (the temperature gradient for an ITG system) is increased. That works well for the transition to turbulence in collisional plasmas[10,11] or fully developed DW–ZF turbulence.[12] Consider, for example, the scenario for the transition to collisional resistive-g turbulence[10,11] in a truncated model. For that case, with the Rayleigh number serving as the bifurcation parameter, we would identify the above-described three instabilities with a sequence of three bifurcations occurring at $\epsilon^{(1)} \equiv \epsilon_c$, $\epsilon^{(2)}$, and $\epsilon^{(3)} \equiv \epsilon_*$. The first (Hopf) bifurcation happens at the linear threshold, producing steady convection without shear flow. The second bifurcation leads to steady "tilted-cell"[10] convection, an indication of the presence of shear flow. The third bifurcation results in "oscillatory convection," a signature of shear-flow destabilization.

However, an attempt to apply the scenario outlined above to a collisionless system encounters immediate problems. In particular, both secondary and tertiary instabilities, whether they are based on mode-coupling or modulational effects, rely on the nonlinear interaction DW + DW $\rightarrow$ ZF (two DW's beat together and drive a ZF). However, since the DW fluctuation level vanishes in the Dimits-shift regime, it is incorrect to talk about a steady-state loop DW $\rightarrow$ ZF $\rightarrow$ DW with no DW's but a nonzero level of ZF's. (Considering such a loop with nonzero values of both DW's and ZF's is relevant to both the transition to weak turbulence in collisional systems as well as the statistical theory of fully developed DW–ZF turbulence.[12]) Dastgeer *et al.*[7] seem to suggest that certain resonances enhance the ZF response, but even then ZF's cannot be driven if there are no DW's at all. Also, unlike the collisional scenario, no distinct $\epsilon^{(1)}$ and $\epsilon^{(2)}$ are observed in the collisionless simulations. Instead, as Rogers *et al.* noted, ZF's are excited by a burst of DW's [through a Kelvin–Helmholtz (KH) instability of radial streamers], which then die away leaving only the ZF's in steady state.

Although large simulations are important tools for the detailed modeling of complex behavior in modern tokamaks, they are frequently ill suited for the identification and detailed understanding of basic conceptual issues. Simplified analytical models are useful to study some aspects of the nonlinear evolution near the threshold for marginal stability. For example, Weiland and coworkers[7,8] have studied the generation of ZF's in certain collisionless fluid models with the aid of reductive perturbation theory.[13] However, those calculations did

not accomodate the presence of seed ZF modes, but rather allowed only nonlinear generation (as in standard collisional calculations). They thus missed the importance of sidebands and their role in the formation of the Dimits shift, as we will see. The present work was inspired, in part, by our desire to better understand those earlier calculations of Weiland and coworkers, particularly issues related to the ordering of the seed ZF's.

We shall describe a systematic dynamical systems analysis[14–16] of a simple low-order Galerkin-truncated model of an electrostatic, collisionless, curvature-driven, fluid ITG system near marginal stability and study the resulting sequence of bifurcations. (A briefer account of this work can be found in Ref. 17.) To be more specific, by "collisionless" ("collisional") we mean that the ZF's are taken to be linearly undamped (strongly damped). For both cases, we always assume a nonzero level of DW dissipation, $\eta_{\rm DW}$, passing only at the end of the calculation to the limit $\eta_{\rm DW} \to 0$. By keeping $\eta_{\rm DW} \neq 0$, we ensure that forced (background-gradient-driven), dissipative steady states are possible for the original system of partial differential equations (PDE's). This is in the same spirit as the usual model of high-Reynolds-number Navier–Stokes turbulence, for which it is well known[18] that the inertial-range Kolmogorov spectrum that arises in the presence of small but nonvanishing viscosity μ differs radically from the Gibbsian thermal-equilibrium solutions of the undamped, wave-number-truncated Euler equation — the spectra being distinguished by the interchange of the limits $t \to \infty$ and $\mu \to 0$.[19] The model we study can be seen as a generalization of the Navier–Stokes equation in which dissipation is applied selectively to some modes but not others. Not surprisingly, we will again encounter an interchange of limits, here related to the dissipation on the ZF's.

We have already mentioned the previous collisional bifurcation analysis of Horton *et al.*[10] It is noteworthy that such collisional calculations do not evince a Dimits shift. However, we will show that a properly formulated bifurcation theory does predict such a shift for a collisionless (ZF) model, in agreement with the numerical simulations. Although both collisional and collisionless models involve the generation of ZF's, the detailed physics of the bifurcations in the two cases differs because of the afore-mentioned interchange of limits; for the collisionless model, the limits must be taken in the order $\eta_{ZF} \to 0$, then $t \to \infty$. This results in a rather distinctive dynamics of the transition to collisionless turbulence, and the scenario we propose for the formation of the Dimits-shift regime differs in detail from those implied by the previous analytic analyses. Indeed, we will argue that there is just

one bifurcation point of interest. Also, for our particular model, we are able to predict the Dimits shift exactly as a function of physical parameters.

We stress that we do not attempt a quantitative calculation of the Dimits shift for realistic toroidal situations, which are best studied by large computer simulations. Our model does not capture a variety of physical processes practically important in both the production and termination of the Dimits-shift regime, so considerable further work is called for. The goal here is to emphasize certain qualitative issues, related to collisionless ZF dynamics and the proper formulation of collisionless bifurcation theory, which are best illustrated and understood by working with a relatively simple model that can be analyzed rather completely.

The paper is organized as follows. In Sec. II we briefly describe our curvature-driven ITG model and its lowest-order truncation relevant for collisionless dynamics. In Sec. III we review the basics of the dynamical-systems approach and emphasize the differences between its applications to collisional and to collisionless systems; we argue that standard methods that have been previously applied to collisional regimes cannot be used for collisionless ones. We apply a correct and consistent center-manifold calculation to the collisionless regime in Sec. IV. We exploit the simplicity of our ITG model to perform a nonperturbative calculation of an important fixed point in Sec. V. In Sec. VI we briefly discuss the possibility of modulational-instability effects on the transition to turbulence in both collisional and collisionless cases. In Sec. VII we comment on the addition of weak zonal-flow damping to the system as well as the effects of higher-order truncations. Discussion and conclusions are presented in Sec. VIII. We also include several appendices that provide detail omitted from Ref. 17 and the main body of this paper.

II. BASIC MODEL

Various ITG models have been discussed by Ottaviani *et al.*[20] and Beer.[21] We consider a simplified gyrofluid ITG system[21] for potential vorticity ϖ and pressure P driven by magnetic curvature. (The number of two coupled fields is the minimum required to produce a physics-based growth rate that is not just inserted "by hand"; the Hasegawa–Wakatani paradigm[18] is similar in this regard.) Let the unit vectors $\widehat{\boldsymbol{z}}$, $\widehat{\boldsymbol{x}}$, and $\widehat{\boldsymbol{y}}$ be associated with the magnetic field, radial, and (essentially) poloidal directions, respectively. Then in vector

form the system of PDE's [considered as two-dimensional (2D) in the plane perpendicular to $\widehat{\boldsymbol{z}}$] is

$$\partial_t \boldsymbol{u}(\boldsymbol{x},t) = \widehat{\mathsf{M}} \cdot \boldsymbol{u} + \widehat{\boldsymbol{N}}(\boldsymbol{u},\boldsymbol{u}), \tag{1}$$

where $\boldsymbol{u} \doteq (\varpi,\, P)^T$, T denotes transpose, hat denotes a differential operator, and the nonlinear term (bilinear in the field vector) describes simple $\boldsymbol{E}\times\boldsymbol{B}$ advection: $\widehat{\boldsymbol{N}}(\boldsymbol{u},\boldsymbol{u}) = -\widehat{\boldsymbol{z}}\times\nabla\varphi\cdot\nabla\boldsymbol{u}$. The electrostatic potential φ is obtained from $\varphi = \widehat{\mathcal{D}}^{-1}\varpi$, where $\widehat{\mathcal{D}} \doteq \widehat{\alpha} - \widehat{\nabla}^2$, $\widehat{\alpha}$ being zero for convective cells ($k_\parallel = 0$) and the identity operator otherwise.[12]

We have slightly modified the linear matrix $\widehat{\mathsf{M}}$ originally discussed in Ref. 21 in order to keep only the terms important for the dynamics considered; for further details, see Appendix A. The resulting matrix $\widehat{\mathsf{M}}$ has the form (A3),

$$\widehat{\mathsf{M}} = \begin{pmatrix} -i(\widehat{\Omega} - i\widehat{\eta}) & -i\widehat{b} \\ i\widehat{\epsilon} & -\widehat{d} \end{pmatrix}. \tag{2}$$

Here $\widehat{\Omega} \doteq -2i(\widehat{\mathcal{D}}^{-1} + \tau)\widehat{\partial}_y$, with τ being the ratio of ion and electron temperatures, is associated with the linear frequency of DW's; $\widehat{\eta} \doteq -\mu\widehat{\nabla}^2$ describes weak collisional damping on the DW's (only); $\widehat{b} \doteq -2i\widehat{\partial}_y$ provides the linear coupling between ϖ and P; $\widehat{d} \doteq \widehat{\nu} + \widehat{\eta}$, where $\widehat{\nu} = -\nu|\widehat{\partial}_y|$ represents the Landau damping effect in the gyrofluid closure[21]; and $\widehat{\epsilon} \doteq -i\tau\widehat{\mathcal{D}}^{-1}[R/L_T - (1+\tau\widehat{\mathcal{D}})]\widehat{\partial}_y$, where R/L_T is the ratio of the curvature and temperature-gradient scale lengths.

We consider a system of finite size, so we may use for the fields the representation $\boldsymbol{u} = \sum_{l,m} \boldsymbol{u}_{l,\widetilde{m}}(t)\sin(lk_x x)e^{i\widetilde{m}k_y y}$, where the summation is over all possible l's and m's with $\widetilde{m} \doteq 2\pi m/L_y k_y$. The wave numbers k_x and k_y correspond to the first linearly unstable DW mode. The quantities $2\pi/k_x$ and L_y are associated with the finite size of the system in the radial and poloidal directions due to magnetic shear and finite poloidal circumference, respectively. In choosing a standing wave in x, we subscribe to an argument from Ref. 10, which asserts this to be a crude representation of the localizing effect of magnetic shear. The rational number $\widetilde{m}$ accounts for the fact that poloidal wave numbers in the above representation are scaled to k_y instead of $2\pi/L_y$. We introduce it so that we can label modes by numbers of order unity (e.g., $\widetilde{m} = 1$) even though we are considering microturbulence ($k_y L_y \gg 1$).

A. Lowest-order Galerkin approximation

One of the methods useful for understanding the content of nonlinear systems of PDE's is projection of the dynamics onto a finite-dimensional subspace of modes. Some discussion of such truncations can be found in Ref. 16. We consider only truncations that are energetically consistent, i.e., that retain the conservation properties of the advective nonlinearities.

The general formula for the nonlinearities in Eq. (1) in terms of the $\boldsymbol{u}_{l,\widetilde{m}}$ variables is given by

$$\boldsymbol{N}_{l,\widetilde{m}}(\boldsymbol{u},\boldsymbol{u}) = \frac{i}{2}\Big(\sum_{m'=-\infty}^{+\infty} \sum_{l'=1}^{l} [\alpha^{l',l-l'}_{\widetilde{m}',\widetilde{m}-\widetilde{m}'} \boldsymbol{u}_{l-l',\widetilde{m}-\widetilde{m}'} - \beta^{l',l+l'}_{\widetilde{m}',\widetilde{m}-\widetilde{m}'} \boldsymbol{u}_{l+l',\widetilde{m}-\widetilde{m}'}] \varphi_{l',\widetilde{m}'}$$
$$+ \sum_{m'=-\infty}^{+\infty} \sum_{l'=l}^{\infty} [\beta^{l',l-l'}_{\widetilde{m}',\widetilde{m}-\widetilde{m}'} \boldsymbol{u}_{l-l',\widetilde{m}-\widetilde{m}'} - \beta^{l',l+l'}_{\widetilde{m}',\widetilde{m}-\widetilde{m}'} \boldsymbol{u}_{l+l',\widetilde{m}-\widetilde{m}'}] \varphi_{l',\widetilde{m}'} \Big) e^{i\widetilde{m}k_y y} \sin(lk_x x), \quad (3)$$

$$\alpha^{l',l''}_{\widetilde{m}',\widetilde{m}''} \doteq k_x k_y (\widetilde{m}' l'' - l' \widetilde{m}''), \quad (4a)$$

$$\beta^{l',l''}_{\widetilde{m}',\widetilde{m}''} \doteq k_x k_y (\widetilde{m}' l'' + l' \widetilde{m}''). \quad (4b)$$

For most of this work we consider the lowest energetically consistent truncation, which retains merely $\boldsymbol{u}_{1,1} \equiv \boldsymbol{u}_1$, $\boldsymbol{u}_{2,0} \equiv \boldsymbol{u}_2$, and $\boldsymbol{u}_{3,1} \equiv \boldsymbol{u}_3$ (as well as their complex conjugates). The DW mode $\boldsymbol{u}_1$ with wave numbers k_x and k_y represents both the first (as a function of increasing temperature gradient) bifurcating DW as well as a damped eigenmode (a two-field model has two eigenvalues for each Fourier amplitude). The amplitude $\boldsymbol{u}_3$ is a drift-wave sideband (SB); $\boldsymbol{u}_2$ represents zonal variation (ZF), present in the system as a result of the nonlinear interaction between the DW and the SB. This particular model does not retain streamers ($k_x = 0$), so it does not capture the KH mechanism of Ref. 9; however, it does permit ZF's to be generated from a DW transient.

In the Fourier representation, the $\widehat{\mathsf{M}}$ operator involves various $\boldsymbol{k}$-dependent coefficients Ω_i, η_i, b_i, ϵ_i, and d_i, where $i = 1, 2, 3$. By definition of the collisionless problem (no linear zonal damping), η_2 is taken to vanish. η_1 and η_3 are unfolding[14] parameters (the plane of the η_1 and η_3 variables contains a family of bifurcations). Since ϵ_1 is a function of the temperature gradient, we choose $\epsilon_1 \equiv \epsilon$ to be the bifurcation parameter of our problem. Note that although ϵ_3 is also a function of the temperature gradient, in the parameter regime of interest ϵ_3 does not become large enough to destabilize the SB mode and thus

is not treated as a bifurcation parameter; it is assumed to be constant in the perturbative expansion methods discussed later.

The vector equation (1) truncated to amplitudes $\boldsymbol{u}_1$, $\boldsymbol{u}_2$, and $\boldsymbol{u}_3$ becomes

$$\dot{\varpi}_1 = -i(\Omega_1 - i\eta_1)\varpi_1 - ib_1 P_1 + \frac{1}{2i}\left[\left(\frac{1}{\mathcal{D}_1} - \frac{1}{\mathcal{D}_2}\right)\varpi_1\varpi_2 - \left(\frac{1}{\mathcal{D}_3} - \frac{1}{\mathcal{D}_2}\right)\varpi_3\varpi_2\right], \tag{5a}$$

$$\dot{P}_1 = i\epsilon\varpi_1 - d_1 P_1 + \frac{1}{2i}\Big[\Big(\overbrace{\frac{\varpi_1}{\mathcal{D}_1}P_2}^{(1)} - \overbrace{\frac{\varpi_2}{\mathcal{D}_2}P_1}^{(0)}\Big) - \Big(\overbrace{\frac{\varpi_3}{\mathcal{D}_3}P_2}^{(2)} - \overbrace{\frac{\varpi_2}{\mathcal{D}_2}P_3}^{(3)}\Big)\Big], \tag{5b}$$

$$\dot{\varpi}_3 = -i(\Omega_3 - i\eta_3)\varpi_3 - ib_3 P_3 - \frac{1}{2i}\left(\frac{1}{\mathcal{D}_1} - \frac{1}{\mathcal{D}_2}\right)\varpi_1\varpi_2, \tag{5c}$$

$$\dot{P}_3 = i\epsilon_3\varpi_3 - d_3 P_3 - \frac{1}{2i}\Big(\overbrace{\frac{\varpi_1}{\mathcal{D}_1}P_2}^{(4)} - \overbrace{\frac{\varpi_2}{\mathcal{D}_2}P_1}^{(3)}\Big), \tag{5d}$$

$$\dot{\varpi}_2 = \left(\frac{1}{\mathcal{D}_1} - \frac{1}{\mathcal{D}_3}\right)\mathrm{Im}(\varpi_1\overline{\varpi}_3), \tag{5e}$$

$$\dot{P}_2 = \mathrm{Im}\left(\overbrace{\frac{\varpi_1}{\mathcal{D}_1}\overline{P}_3}^{(4)} + \overbrace{\frac{\varpi_3}{\mathcal{D}_3}\overline{P}_1}^{(2)}\right) - \mathrm{Im}\left(\overbrace{\frac{\varpi_1}{\mathcal{D}_1}\overline{P}_1}^{(1)}\right). \tag{5f}$$

Here an overbar denotes the complex conjugate. The zonal fields ϖ_2 and P_2 are real, so this system possesses ten real degrees of freedom. A consistency check on the truncation can be obtained by noting that the quantities $\mathcal{W} \doteq |\varpi_1|^2 + |\varpi_3|^2 + \varpi_2^2$ and $\mathcal{P} \doteq |P_1|^2 + |P_3|^2 + P_2^2$ are conserved by the nonlinearities. The cancellations of terms under $\dot{\mathcal{P}}$ are shown by the numbering; term 0 vanishes separately.

The choice of system (5) for the lowest truncation is not unique. Such bifurcations can be and have been studied using various sets of modes. For example, the transition from low-to high-confinement regimes in a resistive-g collisional system has been studied by Hu *et al.*[11]; however, certain aspects of a systematic center manifold calculation (discussed below in Sec. III) were not appreciated by those authors. The transition from Bohm to gyro-Bohm scalings for turbulent intensities was investigated for collisional dynamics by White *et al.*[22] with the familiar set of $\{\mathrm{DW}, \mathrm{ZF}, \mathrm{SB}\}$ modes. This set was chosen in such a way that there was no nonlinear interaction of DW's that produces a ZF mode. When applied to collisionless dynamics, that leads to an infinite-width Dimits-shift phenomenon for the system. The presence of the nonlinear interaction of DW's in our model system is a necessary element for the destabilization of the Dimits-shift regime, thus determining its

upper limit.

We will show in Sec. IV that our minimum set of modes ($\boldsymbol{u}_1$, $\boldsymbol{u}_2$, and $\boldsymbol{u}_3$) with ten real degrees of freedom is sufficient for understanding the basics of the collisionless dynamics in which we are interested. However, we also consider various ways in which additional modes influence the results. First, because we use a perturbative center manifold construction, our predictions cannot be extended all the way to the end of a nonzero-width Dimits-shift regime. Therefore, in Sec. V we perform a nonperturbative calculation to find the value of the Dimits shift. In Sec. VI we consider the addition of envelope-type perturbations to the basic modes in order to account for modulational-instability effects. In Sec. VII we discuss the effects of higher-order truncations.

B. Linear dynamics

The characteristic equation for the eigenvalues $\lambda^{(i)}$ of the model linearized around $\boldsymbol{u}_1 = \boldsymbol{u}_3 = \boldsymbol{0}$ and some chosen $\boldsymbol{u}_2 = \boldsymbol{z}_0 = (\varpi_2^0, P_2^0)^T$ and arbitrary ϵ is

$$\det[\mathcal{M} - \lambda^{(i)}(\boldsymbol{z}_0, \epsilon)\mathrm{I}] = 0, \tag{6}$$

with $\mathcal{M} \equiv \mathcal{M}(\boldsymbol{z}_0, \epsilon)$ being the 4×4 matrix describing the linear dynamics of the $\boldsymbol{u}_1$ and $\boldsymbol{u}_3$ amplitudes for the system parametrized by ϵ and linearized around $\boldsymbol{z}_0$:

$$\begin{aligned}\mathcal{M} \;=\; & \begin{pmatrix} -i(\Omega_1 - i\eta_1) & -ib_1 & 0 & 0 \\ i\epsilon & -d_1 & 0 & 0 \\ 0 & 0 & -i(\Omega_3 - i\eta_3) & -ib_3 \\ 0 & 0 & i\epsilon_3 & -d_3 \end{pmatrix} \\ & + \frac{1}{2i} \begin{pmatrix} \left(\frac{1}{\mathcal{D}_1} - \frac{1}{\mathcal{D}_2}\right)\varpi_2^0 & 0 & \left(\frac{1}{\mathcal{D}_3} - \frac{1}{\mathcal{D}_2}\right)\varpi_2^0 & 0 \\ \frac{1}{\mathcal{D}_1}P_2^0 & -\frac{1}{\mathcal{D}_2}\varpi_2^0 & -\frac{1}{\mathcal{D}_3}P_2^0 & \frac{1}{\mathcal{D}_2}\varpi_2^0 \\ -\left(\frac{1}{\mathcal{D}_1} - \frac{1}{\mathcal{D}_2}\right)\varpi_2^0 & 0 & 0 & 0 \\ -\frac{1}{\mathcal{D}_1}P_2^0 & \frac{1}{\mathcal{D}_2}\varpi_2^0 & 0 & 0 \end{pmatrix}. \end{aligned} \tag{7}$$

The left and right eigenvectors associated with these eigenvalues are distinct; both will be required for our calculations. The right eigenvectors $\boldsymbol{\mathcal{Q}}_i$ obey

$$\mathcal{M} \cdot \boldsymbol{\mathcal{Q}}_i(\boldsymbol{z}_0, \epsilon) = \lambda^{(i)}(\boldsymbol{z}_0, \epsilon)\boldsymbol{\mathcal{Q}}_i(\boldsymbol{z}_0, \epsilon). \tag{8}$$

For the left eigenvectors $\boldsymbol{\mathcal{P}}^i$, we have

$$\mathcal{M}^\dagger \cdot \boldsymbol{\mathcal{P}}^i(\boldsymbol{z}_0, \epsilon) = \overline{\lambda}^{(i)}(\boldsymbol{z}_0, \epsilon)\boldsymbol{\mathcal{P}}^i(\boldsymbol{z}_0, \epsilon), \tag{9}$$

where $\dagger$ denotes Hermitian conjugate and the index i stands for various combinations of the subscripts $(1, 3)$ and superscripts $(\pm)$ described below. These eigenvectors can be constrained to satisfy the orthonormality conditions

$$\boldsymbol{\mathcal{P}}^{i\dagger}(\boldsymbol{z}_0, \epsilon) \cdot \boldsymbol{\mathcal{Q}}_j(\boldsymbol{z}_0, \epsilon) = \delta^i_j. \tag{10}$$

Note that both $\boldsymbol{\mathcal{P}}^i$ and $\boldsymbol{\mathcal{Q}}_i$ are four-vectors. We can constrain the first element of each $\boldsymbol{\mathcal{Q}}$ to be equal to one.

In the limit $\boldsymbol{z}_0 \to \boldsymbol{0}$ the eigenmodes become pure DW and SB modes with eigenvalues λ_1^+ (the bifurcating DW, i.e., the one that as a function of ϵ first becomes linearly unstable), λ_1^- (a damped DW), and $\lambda_3^\pm$ (damped sidebands). For general $\boldsymbol{z}_0$, the eigenmodes are actually combinations of the pure DW and SB modes. Nevertheless, we name them using a notation that identifies their limit as $\boldsymbol{z}_0 \to \boldsymbol{0}$. Thus the eigenvalue $\lambda_1^+(\boldsymbol{z}_0, \epsilon)$ [associated with $\boldsymbol{\mathcal{P}}_1^+(\boldsymbol{z}_0, \epsilon)$ and $\boldsymbol{\mathcal{Q}}_1^+(\boldsymbol{z}_0, \epsilon)$] corresponds to the branch that is the first one to become linearly unstable as ϵ is increased and converges to the corresponding DW branch at $\boldsymbol{z}_0 = \boldsymbol{0}$. Similarly, we refer to $\lambda_1^-(\boldsymbol{z}_0, \epsilon)$ [associated with $\boldsymbol{\mathcal{P}}_1^-(\boldsymbol{z}_0, \epsilon)$ and $\boldsymbol{\mathcal{Q}}_1^-(\boldsymbol{z}_0, \epsilon)$] and $\lambda_3^\pm(\boldsymbol{z}_0, \epsilon)$ [associated with $\boldsymbol{\mathcal{P}}_3^\pm(\boldsymbol{z}_0, \epsilon)$ and $\boldsymbol{\mathcal{Q}}_3^\pm(\boldsymbol{z}_0, \epsilon)$].

For some given $\epsilon = \epsilon_0$, the dependence of $\mathrm{Re}[\lambda_1^+(\boldsymbol{z}, \epsilon_0)]$ on $\boldsymbol{z}$ is shown in Fig. 1. The $\boldsymbol{z}$ plane is divided into regions of positive and negative $\mathrm{Re}[\lambda_1^+(\boldsymbol{z}, \epsilon_0)]$. The curve $\boldsymbol{z} = \boldsymbol{z}(\epsilon_0)$ separating these regions is called the *marginal curve* and is given by $\mathrm{Re}[\lambda_1^+(\boldsymbol{z}, \epsilon_0)] = 0$. As ϵ_0 increases, numerical investigation reveals that the upper and lower regions of positive $\mathrm{Re}[\lambda_1^+(\boldsymbol{z}, \epsilon_0)]$ move towards each other [Fig. 1(a)], coalesce, then spread in the horizontal direction [Fig. 1(b)].

For the rest of the paper, we adopt the following notation for the linear quantities of the bifurcating mode. For the eigenvalue, we write $\lambda_+(\boldsymbol{z}_0, \epsilon) \equiv \lambda_1^+(\boldsymbol{z}_0, \epsilon)$. For the eigenvectors, we suppress the argument dependence and write $\boldsymbol{\mathcal{Q}}_+ \equiv \boldsymbol{\mathcal{Q}}_1^+(\boldsymbol{z}_0, \epsilon)$ and $\boldsymbol{\mathcal{P}}_+ \equiv \boldsymbol{\mathcal{P}}_1^+(\boldsymbol{z}_0, \epsilon)$ for the vectors calculated at arbitrary ϵ, and $\boldsymbol{\mathcal{Q}}_0 \equiv \boldsymbol{\mathcal{Q}}_1^+(\boldsymbol{z}_0, \epsilon_0)$ and $\boldsymbol{\mathcal{P}}_0 \equiv \boldsymbol{\mathcal{P}}_1^+(\boldsymbol{z}_0, \epsilon_0)$ for the ones calculated at the marginal $\epsilon = \epsilon_0$. Further details about the eigenvalues and eigenvectors can be found in Appendix C 1.

III. DYNAMICAL SYSTEMS APPROACH

Although a truncated system of ordinary differential equations (ODE's) provides only an approximate representation of the original PDE's, it may still contain qualitatively correct information about the complex dynamics of the original equation. Of course, one must be cautious. The famous Lorentz system of three coupled ODE's exhibits behavior at high Rayleigh number that probably has little to do with the Boussinesq PDE's from which it was derived. Also, higher-order Lorentz-like truncations manifest behavior quite different from that of the original Lorentz model. Nevertheless, certain kinds of qualitative questions may be answered successfully in low-order models.

A. Center manifolds

Even though our model system (5) represents a severe truncation of the original ITG dynamics, analytically it is still rather unmanageable with its $n = 10$ nonlinearly coupled degrees of freedom. Fortunately, dynamical systems theory provides a way of reducing the effective number of degrees of freedom. A systematic way of proceeding is to exploit the *Center Manifold Theorem*.[14,15] We consider only the case in which there are no positive linear eigenvalues at the point of first bifurcation. (A bifurcation is a change "in the topological structure of the solutions"[14] of a system of nonlinear equations as a parameter such as ϵ passes through certain values.) The theorem then states that at linear threshold the n-dimensional system dynamics are attracted to a smooth n_0-dimensional invariant subspace [the *center manifold* (CM)] as $t \to \infty$. At the point of bifurcation, the CM is tangent to the linear *center eigenspace* spanned by the eigenvectors of the n_0 modes whose eigenvalues λ ($\partial_t \to \lambda$) have zero real part. The other $n - n_0$ modes span the stable subspace. Because the stable modes are damped rapidly relative to the center modes, they achieve their asymptotic values quickly and thus become slaved to the center modes. Their effects are not negligible, in general, but the separation of time scales enables one to eliminate them analytically from the formalism. The n_0-dimensional dynamics on the CM represent the long-time asymptotic behavior of the original system. Essentially, the CM reduction eliminates transient dynamics and thereby simplifies study of the bifurcations that the system undergoes. (Asymptotically, different nonlinear systems may in fact have the same topology; possibilities are represented

by the simplest normal forms[14] of the nonlinear dynamics on the CM.) For the collisionless ITG regime, we will find $n_0 = 4$ (and by virtue of a symmetry, we can easily reduce that number to 3). The resulting 3D CM dynamics are substantially tractable.

To discuss the CM of Eqs. (5), we begin by decomposing the Fourier amplitudes $\boldsymbol{u}_1$ and $\boldsymbol{u}_3$ as

$$\boldsymbol{\mathcal{U}} \doteq \begin{pmatrix} \boldsymbol{u}_1 \\ \boldsymbol{u}_3 \end{pmatrix} = D\boldsymbol{\mathcal{Q}}_0 + \boldsymbol{\mathcal{Y}}. \tag{11}$$

Again, $\boldsymbol{\mathcal{Q}}_0$ and $\boldsymbol{\mathcal{P}}_0$ are the right and left eigenvectors [Eqs. (8) and (9)] of the linearly unstable DW branch, calculated at the marginal $\epsilon = \epsilon_0$ for given $\boldsymbol{z}_0$. Here $\boldsymbol{\mathcal{Y}} \doteq (\boldsymbol{y}_1, \boldsymbol{y}_3)^T$ is assumed to be orthogonal to the left eigenvector $\boldsymbol{\mathcal{P}}_0$ in the sense of the usual complex-valued scalar product: $\boldsymbol{\mathcal{P}}_0^\dagger \cdot \boldsymbol{\mathcal{Y}} = 0$. With the normalization (10), this implies that $D = \boldsymbol{\mathcal{P}}_0^\dagger \cdot \boldsymbol{\mathcal{U}}$. The real and imaginary parts of the complex number D (the amplitude of the bifurcating DW) provide two real (contravariant) coordinates on the center eigenspace. For further discussion of such decompositions, see Appendix B.

It remains to state how $\boldsymbol{u}_2$ is to be treated. To that end, we must distinguish between the collisional and collisionless problems because, as we will show, the dimensionality of the center eigenspace (and thus the CM) differs for those two cases. Although our interest is with the collisionless situation, insight is gained by first reviewing the collisional analysis.

B. Collisional center manifold

For the collisional problem in the regime of stable DW's, ZF modes are linearly damped; even if ZF's were initially present in the system at some finite level, they would be damped on a short time scale. Thus, in the collisional case the system (5) may only be expanded around $\boldsymbol{z}_0 = \boldsymbol{0}$. We call ϵ_c the bifurcation parameter at which the system expanded around $\boldsymbol{z}_0 = \boldsymbol{0}$ undergoes bifurcation. The entire collection of modes in the complex λ (eigenvalue) plane is shown in Fig. 2; the modes associated with the center eigenspace (and which thus determine the dynamics on the CM) are in the shaded area. For this case, the center eigenspace is 2D, described entirely by the complex D amplitude (for further discussion of such complex representations, see the last paragraph of Appendix B), and the complete decomposition for

the special case of $\boldsymbol{z}_0 = \boldsymbol{0}$ becomes

$$\boldsymbol{u}_1 = D\boldsymbol{q}_0 + \boldsymbol{y}_1, \tag{12a}$$

$$\boldsymbol{u}_2 = \boldsymbol{y}_2, \tag{12b}$$

$$\boldsymbol{u}_3 = \boldsymbol{y}_3. \tag{12c}$$

Here $D = \boldsymbol{p}_0^\dagger \cdot \boldsymbol{u}_1$ and the two-vectors $\boldsymbol{q}_0$ and $\boldsymbol{p}_0$ are composed of the first two elements of $\boldsymbol{\mathcal{Q}}_0$ and $\boldsymbol{\mathcal{P}}_0$, respectively. $\boldsymbol{y}_1$ describes the nonlinear curvature of the 2D CM with respect to the flat Cartesian center eigenspace.

If one pursues this collisional calculation (see Appendix C 2), one is led to a standard normal form [Eq. (C15)] for the intensity $I \doteq |D|^2$:

$$\dot{I} = 2\{\mathrm{Re}[\lambda_+(\boldsymbol{0}, \epsilon)]I - \mathrm{Re}(B_0)I^2\}, \tag{13}$$

where $\lambda_+(\boldsymbol{0}, \epsilon)$ is the eigenvalue of the unstable DW branch and B_0 is a known constant. Equation (13) (together with a phase equation that is not written here) describes a Hopf bifurcation (a complex-conjugate pair of eigenvalues crosses the imaginary λ axis) at the $\epsilon = \epsilon_c$ given by $\mathrm{Re}[\lambda_+(\boldsymbol{0}, \epsilon_c)] = 0$. Many physicists' intuitions about bifurcation phenomenology have been strongly influenced by the simple qualitative features of such a Hopf bifurcation (which is frequently relevant for systems with linear waves; see the last paragraph of Appendix B). If $\mathrm{Re}(B_0) > 0$ (an example of a supercritical bifurcation; see Ref. 18 for some introductory discussion), then slightly above linear threshold the DW's would saturate at a small intensity $I_0 \propto \mathrm{Re}[\lambda_+(\boldsymbol{0}, \epsilon)] \propto (\epsilon - \epsilon_c)^{1/2}$ [see Eq. (C16)]. Fig. 3 shows schematically the transition to such a "finite limit cycle" regime; for related discussion, see the calculations of collisionally damped ZF's in Refs. 10 and 11. If instead the bifurcation were subcritical[18] [$\mathrm{Re}(B_0) < 0$], then the DW's would jump to a finite level as ϵ is increased beyond ϵ_c. However, neither of these behaviors is observed in the collisionless simulations.

It is important to appreciate that each undamped zonal component contributes to the dimensionality of the CM. We included a single zonal mode $\boldsymbol{u}_2$ into our formalism (Sec. IV) to emphasize the importance of zonal modes being coordinates in the CM for collisionless systems and the roles of ZF's and SB's in the formation of the Dimits-shift regime. The simplicity of this collisionless model with just one (vector) ZF mode (i.e., two real ZF fields, resulting in a 4D CM) allows one to clearly identify the differences in the application of bifurcation theory and the resulting dynamics from their collisional counterparts. However,

a complete dynamical systems analysis would require inclusion of zonal flows with all possible radial wave numbers [not only $(2k_x, 0)$] into the CM, since all linearly undamped ZF's (with different wave numbers) would have to be included in the center eigenspace and would eventually become coordinates on the CM. Thus, the realistic dynamics of the original collisionless PDE (1) would include the effects of an infinite zonal mode spectrum and would in principle be described by an infinite-dimensional CM. In particular, the nontrivial fixed point $\mathcal{F}$ that later figures prominently in our analysis (Fig. 9) would have coordinates not just on a 2D plane but rather in an infinite-dimensional space of various ZF amplitudes. Nevertheless, we claim that our minimum model with only one ZF $\boldsymbol{u}_2$ mode is sufficient for understanding the basics of the collisionless dynamics in which we are interested.

1. *Collisionless center-manifold decomposition*

Since we assume here that the zonal modes are not linearly damped, the initial ZF level $\boldsymbol{z}_0$ may be finite; thus the system (5) needs to be linearized around an arbitrary $\boldsymbol{z}_0$. For the collisionless problem, we thus employ the decomposition

$$\boldsymbol{\mathcal{U}} = D\boldsymbol{\mathcal{Q}}_0 + \boldsymbol{\mathcal{Y}}, \tag{14a}$$

$$\boldsymbol{u}_2 = \boldsymbol{z}, \tag{14b}$$

with $\boldsymbol{\mathcal{U}} \doteq (\boldsymbol{u}_1, \boldsymbol{u}_3)^T$ and

$$\boldsymbol{z} = z_\varpi \begin{pmatrix} 1 \\ 0 \end{pmatrix} + z_P \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \tag{15}$$

The latter expansion is a special case of the contravariant representation (B5). That is, the vectors $(1, 0)^T$ and $(0, 1)^T$ are linearly independent eigenvectors of the 2×2 null matrix ($\eta_2 = 0$); with Eqs. (14b) and (15), we are stating that z_ϖ and z_P provide two real coordinates on the center eigenspace. (For the zonal variables, there is no stable subspace, so $\boldsymbol{y}_2 = \boldsymbol{0}$.) Thus $D \doteq \boldsymbol{\mathcal{P}}_0^\dagger \cdot \boldsymbol{\mathcal{U}}$ (one complex amplitude) and $\boldsymbol{z}$ (two real fields) provide four real coordinates on the 4D Cartesian center eigenspace, while $\boldsymbol{\mathcal{Y}} \doteq (\boldsymbol{y}_1, \boldsymbol{y}_3)^T$ describes the nonlinear curvature of the CM with respect to that space.

The collisionless system is significantly different from the collisional one. In particular, instead of having only one point (at $\epsilon = \epsilon_c$) where the collisional DW eigenvalues cross the imaginary axis, in the collisionless system at any given ϵ_0 (even $\epsilon_0 < \epsilon_c$) there are an infinite

number of points on the marginal curve (given by $\mathrm{Re}[\lambda_+(\boldsymbol{z},\epsilon_0)]=0$) where DW eigenvalues cross the imaginary axis. Alternatively, if one is given the initial conditions $\boldsymbol{z}=\boldsymbol{z}_0$ and ϵ is increased, in linear theory the initial level persists and one DW branch becomes unstable at some $\epsilon=\epsilon_0$ (found from $\mathrm{Re}[\lambda_+(\boldsymbol{z}_0,\epsilon_0)]=0$). It will be shown, however, that this does not mean that the system (initialized with $\boldsymbol{z}_0$) bifurcates at ϵ_0 since this transition does not necessarily change *topology*. This suggests that the bifurcation of a collisionless system (and thus the width of the Dimits shift regime) may depend on the initial conditions with which the system is started.

2. Numerical evidence for the collisionless center manifold

Numerical studies of our model system (5) of ODE's (ten real degrees of freedom) confirm the existence of a Dimits shift and the relaxation to a CM in that model. For $\epsilon_c<\epsilon<\epsilon_*$, it is found that for almost all initial conditions (IC's) an initial burst of fluctuations occurs that eventually dies away leaving only ZF's. This is illustrated in Fig. 5 for $k_\perp\rho_s=1$, $\tau=0.1$, $\mu=0.01$ (but $\eta_2=0$) and $\nu=1$; in terms of the temperature gradient, the linear threshold (associated with ϵ_c) is found to be $(R/L_T)_c\approx 1.8$. At $R/L_T\approx 2.2$ for many IC's, the final state is essentially unique; that is, many trajectories are attracted at least to the close vicinity of a nontrivial fixed point (at $\boldsymbol{z}=\boldsymbol{z}_{\mathcal{F}}\neq\mathbf{0}$, with all other fields vanishing). However, other IC's can lead to final states dependent on the IC's. For the chosen set of parameters, the Dimits shift terminates at $(R/L_T)_*\approx 2.4$ (associated with ϵ_*). For temperature gradients beyond the Dimits-shift regime ($\epsilon>\epsilon_*$), the model does not saturate in general. That is not important for a qualitative discussion of the Dimits shift since higher-order truncations (see Sec. VII) do saturate.

IV. CENTER MANIFOLD CALCULATION FOR COLLISIONLESS SYSTEM

As was mentioned earlier, the problem with the simple extrapolation of the collisional three-bifurcation scenario to a collisionless system is a misunderstanding of which modes must necessarily be included in the center eigenspace and in subsequent calculation of the CM. We will demonstrate that explicitly for our model (5) by both perturbative construction of the 4D CM and qualitative analysis of the dynamics on the CM (this section),

and by exact calculation of the relevant fixed point of the full nonlinearity (Sec. V). Both saturated collisionless dynamics and the Dimits shift $\Delta\epsilon$ can be predicted by analysis that involves the construction of the CM. It is critical to realize that when the zonal components are undamped [$\eta_2 = 0$; see Eqs. (5e) and (5f)], for any chosen bifurcation parameter ϵ_0 there is a marginal curve (Fig. 1) on the $\boldsymbol{z}$ plane, satisfying $\mathrm{Re}[\lambda_+(\boldsymbol{z}, \epsilon_0)] = 0$, at each point of which one of the DW branches becomes linearly unstable and the CM is 4D (Fig. 4; the modes in the CM are shaded). A significant difference from the collisional case, in which the expansion is done around the marginal point $\epsilon = \epsilon_c$, $\boldsymbol{z}_0 = \boldsymbol{0}$ where the nontrivial solution bifurcates, is that collisionless dynamics (for some given ϵ_0) requires expansion around this marginal curve.

To construct the CM, we write the decomposition (14), where $\{D, \boldsymbol{z}\}$ provides four real coordinates on the center eigenspace and $\boldsymbol{\mathcal{Y}}$ describes the nonlinear curvature of the CM with respect to that space.

It is also important to recall that we take the DW modes to be linearly damped ($\eta_{1,3}, d_{1,3} \neq 0$) in our collisionless model; "collisionless" refers only to linearly undamped ZF's. The dissipative terms in the linear dynamics of the DW's may be viewed as taking into account various small numerical dissipation effects always present in simulations. Eventually we will take the limit of zero DW dissipation and will show that it does not qualitatively change the dynamics resulting from our CM calculation.

Note that for understanding the form of the primary bifurcation in *collisional* dynamics,[11] it may be enough to identify the center eigenspace and only the ZF contribution to the nonlinear curvature part. However, such an approach, although it yields a correct result for the form of the collisional Hopf bifurcation (Appendix C 2), in general will not give correct quantitative or possibly even qualitative results a finite distance away from that bifurcation. (Perturbation theory provides only local information.) More importantly, it cannot give correct dynamics in the collisionless limit. Introduction of nonlinear curvature due to coupling to linearly damped sideband branches is a systematic way to proceed.

A. Center manifold calculation around the marginal curve $\mathrm{Re}[\boldsymbol{\lambda_+(z, \epsilon_0)}] = \boldsymbol{0}$

For a given $\boldsymbol{z}_0$, we choose ϵ_0 so that $\mathrm{Re}[\lambda_+(\boldsymbol{z}_0, \epsilon_0)] = 0$. To lowest order in an expansion around the marginal curve (Appendix C 3), we are led to the 4D system (C19) for the

dynamics on the CM. That can immediately be reduced to a 3D system for $I \doteq |D|^2$ and $\boldsymbol{z}$ by writing $D = \rho e^{i\theta}$ and noting that the θ dependence entirely decouples. Thus we find that the CM dynamics up to $O(|D, \overline{D}, \boldsymbol{z} - \boldsymbol{z}_0, \epsilon - \epsilon_0|^2)$ in the vicinity of $\boldsymbol{z}_0$ and ϵ_0 obey

$$\dot{I} = 2\,\mathrm{Re}[\Gamma(\boldsymbol{z}, \boldsymbol{z}_0, \epsilon, \epsilon_0)]I, \quad \dot{\boldsymbol{z}} = \boldsymbol{g}(\boldsymbol{z}_0, \epsilon_0)I. \tag{16}$$

The two-vector $\boldsymbol{g}(\boldsymbol{z}_0, \epsilon_0)$ and the function $\Gamma(\boldsymbol{z}, \boldsymbol{z}_0, \epsilon, \epsilon_0)$ are known. Note that $\Gamma(\boldsymbol{z}, \boldsymbol{z}_0, \epsilon, \epsilon_0)$ is effectively an expansion of $\lambda_+(\boldsymbol{z}, \epsilon)$ around $\boldsymbol{z} = \boldsymbol{z}_0$ and $\epsilon = \epsilon_0$. Supporting algebra and explicit formulas for these quantities are presented in Appendix C 3.

The origin $\mathcal{O}$ ($I = 0$, $\boldsymbol{z} = \boldsymbol{0}$) is a fixed point for this reduced dynamics; it is linearly stable for $\epsilon < \epsilon_c$ (the collisionless analog of ϵ_c here is found from $\mathrm{Re}[\lambda_+(\boldsymbol{0}, \epsilon_c)] = 0$) and unstable otherwise. However, the striking feature of this dynamics is that the entire $I = 0$ plane is invariant ($\dot{I} = 0$, $\dot{\boldsymbol{z}} = \boldsymbol{0}$). This unusual behavior is the first indication that for the collisionless situation the origin does not have the same preferred status as in other, more conventional problems.[10] Indeed, for any given ϵ the system also admits a *nontrivial fixed point* $\mathcal{F} = (I_{\mathcal{F}}, \boldsymbol{z}_{\mathcal{F}} \equiv \boldsymbol{z}_{\mathcal{F}}(\epsilon))$, which is the solution of

$$I_{\mathcal{F}} = 0, \qquad \boldsymbol{g}(\boldsymbol{z}_{\mathcal{F}}, \epsilon) = \boldsymbol{0}. \tag{17}$$

As was demonstrated in Fig. 5, such a fixed point is observed in numerical investigations of the original system (5). We will show that the stability of $\mathcal{F}$ determines the Dimits shift. A striking difference from the collisional Hopf bifurcation is that $\mathcal{F}$ is present even in the submarginal regime. Also, collisional analysis predicts that the normal form (13) is a function of the zonal dissipation η_2 ($B_0 \propto \eta_2^{-1}$) that does not converge to (16) in the $\eta_2 \to 0$ limit. Collisional analysis is equivalent to taking the $t \to \infty$ limit first, then taking the $\eta_2 \to 0$ limit (thereby calculating 2D CM dynamics). It is necessary to emphasize that collisionless dynamics requires that the $\eta_2 \to 0$ limit be taken first, thus ensuring that the following calculation of the CM dynamics (i.e., the $t \to \infty$ limit) is performed in 4D space.

Since the solution of $\boldsymbol{g}(\boldsymbol{0}, \epsilon_0) = \boldsymbol{0}$ [Eq. (17)] is $\epsilon_0 = 0$, $\mathcal{F}$ passes through the origin as ϵ passes through zero (Fig. 6). As ϵ increases, both the marginal curve and the fixed point move on the $\boldsymbol{z}$ plane, as schematically shown in Figs. 1 and 6.

The system (16) for describing the dynamics near the marginal curve was derived under the assumption that the coefficient $\boldsymbol{g}(\boldsymbol{z}_0, \epsilon_0)$ is nonzero. Remembering that the location of $\mathcal{F}$ is given by Eq. (17), we note that Eqs. (16) are valid only before $\mathcal{F}$ crosses the marginal curve

(i.e., for ϵ sufficiently small so that $\mathcal{F}$ is stable in the I direction). The dynamics predicted by Eqs. (16) are relatively trivial: depending on the sign of the initial $\mathrm{Re}[\Gamma(\boldsymbol{z}, \boldsymbol{z}_0, \epsilon, \epsilon_0)]$, all trajectories starting in the vicinity of the marginal curve are either attracted to the $I = 0$ plane (and end up close to the initial starting point) or are initially repelled but eventually attracted.

The point at which $\mathcal{F}$ crosses the marginal curve is obviously significant. We denote the ϵ at which that occurs by ϵ_*. That is, there is some $\epsilon = \epsilon_*$ such that $\mathcal{F} = (I_{\mathcal{F}} = 0, \boldsymbol{z}^*_{\mathcal{F}} \doteq \boldsymbol{z}_{\mathcal{F}}(\epsilon_*))$ simultaneously satisfies

$$\boldsymbol{g}(\boldsymbol{z}^*_{\mathcal{F}}, \epsilon_*) = \boldsymbol{0}, \quad \mathrm{Re}[\Gamma(\boldsymbol{z}^*_{\mathcal{F}}, \boldsymbol{z}^*_{\mathcal{F}}, \epsilon_*, \epsilon_*)] = 0. \tag{18}$$

To understand the $I, \boldsymbol{z}$ dynamics in the vicinity of ϵ_* and $\boldsymbol{z}^*_{\mathcal{F}}$, it is necessary to expand the $\boldsymbol{z}$ equation up to $O(|D, \overline{D}, \boldsymbol{z} - \boldsymbol{z}^*_{\mathcal{F}}, \epsilon - \epsilon_*|^3)$; we obtain [Eqs. (C24)]

$$\dot{D} = \Gamma(\boldsymbol{z}, \boldsymbol{z}^*_{\mathcal{F}}, \epsilon, \epsilon_*) D, \quad \dot{\boldsymbol{z}} = 2\,\mathrm{Re}[\mathcal{A}(\boldsymbol{z}^*_{\mathcal{F}}) \cdot \boldsymbol{\mathcal{Y}}\overline{D}], \tag{19}$$

where $\mathcal{A}$ is a known 2×4 matrix.

An expansion in terms of modes on the center eigenspace $\{D, \boldsymbol{z}, \epsilon\}$ (or symmetry considerations,[16] i.e., periodic boundary conditions applied in the poloidal direction, which require that $\boldsymbol{\mathcal{Y}}$ be independent of $\overline{D}$), dictates that $\boldsymbol{\mathcal{Y}} = \mathcal{W}^{\boldsymbol{z}D}(\boldsymbol{z}^*_{\mathcal{F}}) \cdot (\boldsymbol{z} - \boldsymbol{z}^*_{\mathcal{F}}) D + \boldsymbol{w}^{D\epsilon}(\boldsymbol{z}^*_{\mathcal{F}}) D(\epsilon - \epsilon_*) + \cdots$, where $\mathcal{W}^{\boldsymbol{z}D}(\boldsymbol{z}^*_{\mathcal{F}})$ is a constant 4×2 matrix and $\boldsymbol{w}^{D\epsilon}(\boldsymbol{z}^*_{\mathcal{F}})$ is a two-vector to be determined. That may be accomplished[14,15] by equating the time derivative of the power-series expansion of $\boldsymbol{\mathcal{Y}}$ to the right-hand side of the evolution equation that follows from the restriction of Eqs. (5) to the CM. In detail, we follow the projection method described by Kuznetsov,[15] which does not require a preliminary linear diagonalization. That procedure, essentially a determination of the contravariant representation of the shape of the CM in terms of the nonorthogonal right eigenvectors of the linear matrix (although it was not described in those terms by Kuznetsov), is reviewed in Appendix B.

Finally, we obtain [Eq. (C32)] the equations for the dynamics on the CM close to ϵ_* and $\boldsymbol{z}^*_{\mathcal{F}}$:

$$\dot{I} = 2\,\mathrm{Re}[\Gamma(\boldsymbol{z}, \boldsymbol{z}^*_{\mathcal{F}}, \epsilon, \epsilon_*)] I, \tag{20a}$$

$$\dot{\boldsymbol{z}} = [-(\epsilon - \epsilon_*)\boldsymbol{a}(\boldsymbol{z}^*_{\mathcal{F}}) + \mathsf{A}(\boldsymbol{z}^*_{\mathcal{F}}) \cdot (\boldsymbol{z} - \boldsymbol{z}^*_{\mathcal{F}})] I. \tag{20b}$$

The two-vector $\boldsymbol{a}(\boldsymbol{z}^*_{\mathcal{F}}) \doteq 2\,\mathrm{Re}[\mathcal{A}(\boldsymbol{z}^*_{\mathcal{F}}) \cdot \boldsymbol{w}^{D\epsilon}(\boldsymbol{z}^*_{\mathcal{F}})]$ and the 2×2 matrix $\mathsf{A}(\boldsymbol{z}^*_{\mathcal{F}}) \doteq 2\,\mathrm{Re}[\mathcal{A}(\boldsymbol{z}^*_{\mathcal{F}}) \cdot \mathcal{W}^{\boldsymbol{z}D}(\boldsymbol{z}^*_{\mathcal{F}})]$ are known.

At any given ϵ, there is an infinite number of points on the marginal curve $\mathrm{Re}[\lambda_+(\boldsymbol{z},\epsilon)]=0$ where one of the DW branches becomes linearly unstable and the CM variables evolve according to the normal form (16). But since for $\epsilon<\epsilon_*$ the topology of the system does not change, we do not think of these ϵ's as bifurcation points. It will be shown later that $\epsilon=\epsilon_*$ is the only point where topology changes, thus making the system undergo a bifurcation (Fig. 7) with the normal form given by Eqs. (20).

B. Local analysis around $\epsilon=0$ and $z_0=0$

If one is interested in dynamics close to the origin $\mathcal{O}$, or alternatively if $\boldsymbol{z}^*_{\mathcal{F}}$ is close to $\mathcal{O}$, one can consider a simplified version of (20). In particular, since $\mathcal{F}$ passes through $\mathcal{O}$ at $\epsilon=0$, one may set $\boldsymbol{z}^*_{\mathcal{F}}=\mathbf{0}$ and $\epsilon_*=0$ in Eqs. (20) to obtain

$$\dot{I}=2\,\mathrm{Re}[\Gamma(\boldsymbol{z},\epsilon)]I,\quad \dot{\boldsymbol{z}}=(-\boldsymbol{a}\epsilon+\mathsf{A}\cdot\boldsymbol{z})I, \tag{21}$$

with $\boldsymbol{a}\doteq\boldsymbol{a}(0)$, $\mathsf{A}\doteq\mathsf{A}(0)$, and $\Gamma(\boldsymbol{z},\epsilon)\doteq\Gamma(\boldsymbol{z},\mathbf{0},\epsilon,0)$ being known through $O(\boldsymbol{z}^2)$. All elements of A are positive and both of its eigenvalues are negative. Supporting algebra and explicit formulas for these quantities are presented in Appendix C 3. Considering the system (21) instead of (20) is equivalent to following the derivative of the $\mathcal{F}$ trajectory (while ϵ changes) at $\boldsymbol{z}_0=\mathbf{0}$ until it intersects the marginal curve (when $\mathrm{Re}[\lambda_+(\boldsymbol{z}_0,\epsilon)]=0$ is satisfied), instead of following the trajectory itself (Fig. 8).

Equations (21) give a simplified prediction for the fixed point $\mathcal{F}$ [Eqs. (18)]:

$$I_{\mathcal{F}}=0,\qquad \boldsymbol{z}_{\mathcal{F}}=\epsilon\mathsf{A}^{-1}\cdot\boldsymbol{a}. \tag{22}$$

This fixed point is *neutrally stable* against small variations in $\boldsymbol{z}$ because the variational equations are

$$\delta\dot{I} \;=\; 2\,\mathrm{Re}[\Gamma]\delta I+2\delta\,\mathrm{Re}[\Gamma]I_{\mathcal{F}}=2\,\mathrm{Re}[\Gamma]\delta I, \tag{23a}$$

$$\delta\dot{\boldsymbol{z}} \;=\; (-\boldsymbol{a}\epsilon+\mathsf{A}\cdot\boldsymbol{z}_{\mathcal{F}})\delta I+\mathsf{A}\cdot\delta\boldsymbol{z}I_{\mathcal{F}}=\mathbf{0}. \tag{23b}$$

This result can be confusing because a phase-plane analysis of Eqs. (21) predicts that the fixed point is absolutely stable on the invariant plane $I=0$. One may perform that analysis by noting that I cancels out under $\dot{P}_2/\dot{\varpi}_2=\dot{z}_P/\dot{z}_\varpi\doteq v(\boldsymbol{z})/u(\boldsymbol{z})$. A sketch of the phase trajectories is shown in Fig. 9. Note that these are not the actual trajectories of the 3D

dynamics but merely the projection of the dynamics onto the plane; once the system point reaches the $I = 0$ plane, it stays there, since that plane is "filled up with fixed points." All qualitative properties of Fig. 9 can be determined analytically. In particular, since both of its eigenvalues are negative (Appendix C 3), $\mathcal{F}$ is attracting in the $\boldsymbol{z}$ plane for all ϵ; in structure, it is a stable node.

In spite of the fact that $\mathcal{F}$ is absolutely stable on the $\boldsymbol{z}$ plane, the final $\boldsymbol{z}$ can be either $\mathcal{F}$ or may depend on IC's; the $\boldsymbol{z}$ position at $t = \infty$ cannot be obtained from the phase-plane analysis because the temporal I dynamics have been eliminated. A simple example that illustrates this unusual behavior is presented in Sec. IV C.

In the submarginal regime $\epsilon < \epsilon_c$, in which there are no linearly unstable modes, one has $\mathrm{Re}[\Gamma(\boldsymbol{z}, \epsilon)] < 0$ for all sufficiently small $\boldsymbol{z}$'s. This means that all trajectories starting in the vicinity of $\mathcal{O}$ are attracted to the $I = 0$ plane and end up close to the initial starting point.

In the supermarginal region $\epsilon_c < \epsilon < \epsilon_*$, the region in the vicinity of $\boldsymbol{z} = \boldsymbol{0}$ has $\mathrm{Re}[\Gamma(\boldsymbol{z}, \epsilon)] > 0$ [Eq. (21)], while the vicinity of $\mathcal{F}$ has $\mathrm{Re}[\Gamma(\boldsymbol{z}, \epsilon)] < 0$. Initially, most trajectories starting close to the origin $\mathcal{O}$ move away from it; eventually they end up either very close to $\mathcal{F}$ [for sufficiently large initial DW level $I(t = 0)$] or sometimes on the $I = 0$ plane at positions strongly depending on IC's. Such dynamics are consistent with the observed behavior (Fig. 5) above marginality: an initial burst of DW's generates ZF's, which then annihilate the DW's leaving only a steady ZF component as $t \to \infty$. This generation (secondary instability)/annihilation process is *transient*, so does not involve a distinct bifurcation point $\epsilon^{(2)}$. Also, note although $\mathcal{O}$ changes stability at $\epsilon = \epsilon_c$, $\mathcal{F}$ does not. Since the whole $I = 0$ plane is invariant, the global topology of the system does not change. Thus, ϵ_c does not serve as a distinct bifurcation point $\epsilon^{(1)}$ for the global system dynamics.

As ϵ is increased further through ϵ_*, $\mathrm{Re}[\Gamma(\boldsymbol{z}^*_{\mathcal{F}}, \epsilon)]$ becomes positive and many IC's are repelled from the $I = 0$ plane; numerically, we observe that the system does not saturate. This property of our specific model (5) is of little concern since simulations verify that higher truncations do saturate with nonzero levels of DW activity and characteristic chaotic behavior (Sec. VII).

Note that not all IC's explode for $\epsilon > \epsilon_*$; that is a property of only some of the trajectories starting in the region of positive $\mathrm{Re}[\Gamma(\boldsymbol{z}, \epsilon)]$. Many IC's with $\boldsymbol{z}$ sufficiently far away from the origin still have $\mathrm{Re}[\Gamma(\boldsymbol{z})] < 0$ and thus are attracted to the $I = 0$ plane. This makes the system started with this kind of IC become unstable at some ϵ larger than ϵ_*. We associate

$\Delta\epsilon \doteq \epsilon_* - \epsilon_c$ with the Dimits shift for the system started with a low level of initial ZF's. For more general IC's, the Dimits shift depends on the IC's and $\Delta\epsilon$ is the minimal possible width for the system.

It is useful to note that if we did not allow coupling to the linearly damped SB modes that contribute to the curvature of the CM (the way it is usually done for collisional systems), we would not get the sophisticated CM dynamics given by Eqs. (21). In particular, the $\mathsf{A}_3 \cdot \mathsf{W}_3$ term (Appendix C 3) in the equation would be absent and the dynamics would be trivially stable for any ϵ, thus giving an infinite Dimits shift.

Also, it is important to emphasize that using both the D and $\boldsymbol{z}$ amplitudes to describe the dynamics on the CM is necessary in order to recover the normal form (21). For example, White *et al.*[22] considered a collisional system of similar $\{\mathrm{DW}, \mathrm{ZF}, \mathrm{SB}\}$ modes. However, since their set was chosen so that the nonlinear interaction of DW's does not couple back to the ZF mode, a CM calculation of this system would lead to Eq. (21) without the $\boldsymbol{a}$ and $\mathsf{A}_1 \cdot \mathsf{W}_1^{\boldsymbol{z}D}$ terms (Appendix C 3). That would again lead to a trivially stable dynamics and prevent those authors from obtaining the finite-width Dimits-shift phenomenon in the collisionless limit of their model system. The presence of the nonlinear interaction of DW's in our model is critical for the destabilization of the Dimits-shift regime.

C. A simple illustration of the dependence of time-asymptotic dynamics on initial conditions

To understand how trajectories need *not* be attracted to $\mathcal{F}$ as $t \to \infty$ even though phase-plane analysis predicts that the fixed point is absolutely stable, consider the model

$$\dot{I} = -\lambda I, \qquad \dot{z} = (z_{\mathcal{F}} - z)I, \tag{24}$$

where the variables I and z are scalars and the parameters λ and $z_{\mathcal{F}}$ are real numbers with $\lambda \geq 0$. This system has a neutrally stable fixed point at $I = 0$, $z = z_{\mathcal{F}}$. The exact solution is

$$I(t) = I_0 e^{-\lambda t}, \quad z(t) = z_{\mathcal{F}} + (z_0 - z_{\mathcal{F}})S(t), \tag{25}$$

where

$$S(t) \doteq \exp\left[-I_0\left(\frac{1 - e^{-\lambda t}}{\lambda}\right)\right]. \tag{26}$$

For fixed $I = I_0$ ($\lambda = 0$), $z_{\mathcal{F}}$ is clearly stable by inspection of Eq. (24). [In more detail, for $\lambda = 0$ $S(t) \to \exp(-I_0 t)$, $\lim_{t\to\infty} S(t) = 0$, and $\lim_{t\to\infty} z(t) = z_{\mathcal{F}}$; this is the analog of the phase-plane analysis of $\boldsymbol{z}_{\mathcal{F}}$ described above.] However, for $\lambda \neq 0$ one has $\lim_{t\to\infty} S(t) = \exp(-\rho_0)$, where $\rho_0 \doteq I_0/\lambda$, so the time-asymptotic behavior depends on ρ_0. For large ρ_0, the final position is exponentially close to $z_{\mathcal{F}}$; for small ρ_0, however, $z(\infty)$ differs by a finite amount from $z_{\mathcal{F}}$. Only the single trajectory with $I_0 = \infty$ actually attains the fixed point as $t \to \infty$. This behavior arises because $\mathcal{F}$ is neutrally stable in the z direction; it is in qualitative agreement with numerical integration of the original model, as discussed in Sec. III B 2. Another way of explaining that the dynamics (25) need not necessarily end up at $z_{\mathcal{F}}$ as $t \to \infty$ is that any point $(I = 0, z)$ is a fixed point of Eqs. (24).

V. NONPERTURBATIVE CALCULATION OF THE FIXED POINT $\mathcal{F}$

The center manifold calculation for collisionless systems presented in the previous section is perturbative around $\epsilon = 0$, with the amplitudes of the undamped modes in the center eigenspace being $O(\epsilon)$ and the other modal amplitudes being at least $O(\epsilon^2)$. Perturbative CM calculations provide only approximations to the value of ϵ_* and the location $\boldsymbol{z}_{\mathcal{F}}$ of $\mathcal{F}$. Performing an expansion up to only $O(\epsilon^3)$ cannot guarantee that the obtained *global* dynamics are qualitatively the same as for the original unperturbed system, since perturbation expansion is local.[23] For example, there may be other fixed points that do not intersect the origin at $\epsilon = 0$. Fortunately, the present model is simple enough that $\mathcal{F}$ and ϵ_* can be calculated exactly. By doing so, we will shown that the perturbative CM calculation captures the basic dynamics in the Dimits-shift regime.

Rigorous equations for fixed points are motivated by the observation that the dynamics ought to relax rapidly to the CM. Since the DW vorticity ϖ_1 always has a component in the CM (since we are discussing the bifurcating DW), we define the normalized variables

$$P_1' \doteq P_1/\varpi_1 = \zeta_1' e^{i\phi_1'}, \tag{27a}$$

$$\varpi_3' \doteq \varpi_3/\varpi_1 = \rho_3' e^{i\theta_3'}, \tag{27b}$$

$$P_3' \doteq P_3/\varpi_1 = \zeta_3' e^{i\phi_3'}; \tag{27c}$$

$\boldsymbol{z}$ is not normalized, being already entirely in the CM. Although for $\epsilon < \epsilon_*$ all original variables (except for $\boldsymbol{z}$) are dynamically driven to zero, the normalized variables remain

nonzero as $t \to \infty$. This expedites tracking the fixed point $\boldsymbol{z}_{\mathcal{F}}$. Upon deriving evolution equations for the normalized variables from Eqs. (5) and passing to an amplitude–phase representation, we are led after tedious algebra to evolution equations for ζ_1', ζ_3', ρ_3', ϕ_1', ϕ_3', and θ_3'. For example, when written in normalized variables, the zonal vorticity (ϖ_2) equation (5e) becomes

$$\begin{aligned} \dot{\varpi}_2 &= |\varpi_1|^2 \left(\frac{1}{\mathcal{D}_1} \operatorname{Im}(\overline{\varpi}_3') + \frac{1}{\mathcal{D}_3} \operatorname{Im} \varpi_3' \right) \\ &= -I \left(\frac{1}{\mathcal{D}_1} - \frac{1}{\mathcal{D}_3} \right) \rho_3' \sin\theta_3', \end{aligned} \tag{28}$$

where $I \doteq |\varpi_1|^2$ is the DW intensity. Similarly, the zonal pressure (P_2) equation (5f) transforms to

$$\begin{aligned} \dot{P}_2 &= I \operatorname{Im} \left(\frac{1}{\mathcal{D}_1} \zeta_3' e^{-i\phi_3'} + \frac{1}{\mathcal{D}_3} \rho_3' \zeta_1' e^{i(\theta_3' - \phi_1')} \right) \\ &\quad - I \operatorname{Im} \left(\frac{1}{\mathcal{D}_1} \zeta_1' e^{-i\phi_1'} \right). \end{aligned} \tag{29}$$

Demanding the steady state $\dot{\varpi}_2 = 0$ in Eq. (28) leads to the constraint $\sin\theta_3' = 0$ or $\theta_3' = n\pi$. Similarly, requiring $\dot{P}_2 = 0$ gives from (29) the constraint

$$Y_3 = \left[1 - (-1)^n \left(\frac{\mathcal{D}_1}{\mathcal{D}_3} \right) \rho_3' \right] Y_1, \tag{30}$$

where we use the notation

$$X_i \doteq \zeta_i' \cos\phi_i', \quad Y_i \doteq \zeta_i' \sin\phi_i', \quad i = 1, 3. \tag{31}$$

Similar steady-state constraints on the other variables can also be obtained; they determine the locations of the fixed points on the CM. For $\theta_3' = n\pi$, the steady-state conditions for P_1, ϖ_3, and P_3 can be transformed to

$$\Omega_1 (X_1'^2 + Y_1'^2) + \left(\epsilon \rho_3^{-1} - \frac{P_2'}{2\mathcal{D}_1} \right) X_1' = 0, \tag{32a}$$

$$-C_1 (X_1'^2 + Y_1'^2) + \left(\epsilon \rho_3^{-1} - \frac{P_2'}{2\mathcal{D}_1} \right) Y_1' = 0, \tag{32b}$$

$$\frac{1}{2} \left(\frac{1}{\mathcal{D}_1} - \frac{1}{\mathcal{D}_2} \right) \varpi_2' - b_3 X_3' - (-1)^n \Delta\Omega = 0, \tag{32c}$$

$$(-1)^n b_3 Y_3' - \Delta\eta = 0, \tag{32d}$$

$$\Omega_1 (X_3'^2 + Y_3'^2) + \left(\frac{P_2'}{2\mathcal{D}_1} + (-1)^n \epsilon_3 \right) X_3' = 0, \tag{32e}$$

$$-C_3 (X_3'^2 + Y_3'^2) + \left(\epsilon \rho_3^{-1} - \frac{P_2'}{2\mathcal{D}_1} \right) Y_3' = 0, \tag{32f}$$

where

$$X_i' \doteq X_i/\rho_3', \quad Y_i' \doteq Y_i/\rho_3', \quad i = 1, 3. \tag{33}$$

$$\varpi_2' \doteq \varpi_2/\rho_3', \quad P_2' \doteq P_2/\rho_3', \quad i = 1, 3. \tag{34}$$

$$C_1 \doteq d_1 - \eta_1 = \nu, \tag{35a}$$

$$C_3 \doteq d_3 - \eta_1 = \nu + \Delta\eta, \tag{35b}$$

$$\Delta\eta \doteq \eta_3 - \eta_1. \tag{35c}$$

We seek solutions for ρ_3' that scale with ϵ. For $n = 0$ there is a consistent solution with that ordering:

$$Y_1'(\epsilon) = \frac{\Delta\eta}{b_3}\rho_3'(\epsilon), \tag{36}$$

$$\rho_3'(\epsilon) = \epsilon \left\{ \left[\left(\frac{\Omega_1^2 + \nu^2}{\nu} \right) + \left(\frac{\Omega_1^2 + C_3^2}{C_3} \right) \right] \frac{\Delta\eta}{b_3} - \epsilon_3 \right\}^{-1}. \tag{37}$$

The position $\boldsymbol{z}_{\mathcal{F}} = (\varpi_2(\epsilon), P_2(\epsilon))$ of the fixed point on the $I = 0$ plane is then

$$P_2(\epsilon) = -2\mathcal{D}_1 \left[\left(\frac{\Omega_1^2 + C_3^2}{C_3} \right) \left(\frac{\Delta\eta}{b_3} \right) - \epsilon_3 \right] \rho_3'(\epsilon), \tag{38a}$$

$$\varpi_2(\epsilon) = -2 \left(\frac{1}{\mathcal{D}_1} - \frac{1}{\mathcal{D}_2} \right)^{-1} \times \left[\Delta\Omega - \left(\frac{\Omega_1}{C_3} \right) \Delta\eta \right] \rho_3'(\epsilon). \tag{38b}$$

Here $\Delta\Omega \doteq \Omega_3 - \Omega_1 < 0$. This fixed point passes through the origin at $\epsilon = 0$. Further nontrivial algebra also shows that to lowest order in ϵ the result (38) agrees with the location of $\mathcal{F}$ [Eq. (22)] found from the perturbative CM construction. For any ϵ, numerical solution of the fixed-point equations demonstrates agreement with the numerically observed $\mathcal{F}$ through six decimal places, providing an important check on the tedious algebraic details.

Although solutions of these nonperturbative equations capture all (possibly global) fixed points of the original system (the perturbative CM calculation is local), we have found only the $\mathcal{F}$ described above. We have no categorical proof that no other fixed points exist, although no other stable ones have emerged from an admittedly very incomplete numerical search of the phase space. We believe that if they do exist they are all saddle points, which would not modify the qualitative asymptotics we have described.

With nonperturbative results in hand, we can formulate an exact equation for the Dimits shift. From Eq. (5a) for $\varpi_1 = \rho_1 e^{i\theta_1}$ and the steady-state conditions, one finds

$$\dot{\rho}_1 = (-\eta_1 + b_1 Y_1)\rho_1, \tag{39}$$

Thus the fixed point $\mathcal{F}$ is destabilized in the I direction when

$$Y_1(\epsilon_*) = \eta_1/b_1. \tag{40}$$

Upon combining this with Eqs. (37) and (40), we obtain

$$\epsilon_* = \frac{\eta_1}{\Delta\eta}\frac{b_3}{b_1}\left\{\left[\left(\frac{\Omega_1^2+\nu^2}{\nu}\right)+\left(\frac{\Omega_1^2+C_3^2}{C_3}\right)\right]\frac{\Delta\eta}{b_3}-\epsilon_3\right\}. \tag{41}$$

The condition (41) is equivalent to the requirement $\mathrm{Re}[\Gamma(\boldsymbol{z}^*_{\mathcal{F}}, \epsilon_*)] = 0$ for the destabilization of $\mathcal{F}$ in the CM calculation. Numerical work demonstrates agreement with our simulation value of ϵ_*; for the set of parameters used in Sec. III to obtain Fig. 5, Eq. (41) yields $(R/L_T)_* \approx 2.4$, which agrees with the simulation result. The destabilization process is not a KH instability but rather an ITG instability modified by stabilizing ZF shear.

So far in the calculation we have kept nonzero DW collisional dissipation $\eta_{1,3}$. The reason is that the right-hand sides of Eqs. (36) and (40) vanish with the DW dissipation, and this singular situation would not allow us to obtain any tractable results if we set $\eta_{1,3}$ to zero in the very beginning. Instead, we take the limit $\eta_{1,3} \to 0$ in Eq. (41). An important observation emerges in this limit, for which it is easy to show that Eq. (41) transforms to

$$\epsilon_* = -\epsilon_3(b_3/b_1)[(k_x^2+k_y^2)/8k_x^2] > \epsilon_c, \tag{42}$$

the latter inequality following straightforwardly from the fact that $\epsilon_c = 0$ at $\eta_1 = 0$. Thus a nonzero Dimits shift $\Delta\epsilon = \epsilon_* - \epsilon_c$ arises even in the limit of vanishing DW collisional dissipation, which substantially enhances the relevance of our model to the large-scale collisionless simulations.

Using the expressions for the coefficients in Eq. (2), we can rewrite formula (42) in terms of physical quantities to obtain

$$\left(\frac{R}{L_T}\right)_* = 1+\tau(1+k_x^2+k_y^2)+8\tau k_x^2\left[\left(\frac{1+9k_x^2+k_y^2}{1+k_x^2+k_y^2}\right)\left(\frac{8k_x^2}{k_x^2+k_y^2}\right)+1\right]^{-1}, \tag{43}$$

where k_x and k_y are normalized to ρ_s^{-1}. The temperature gradient at marginality is $(R/L_T)_c = 1+\tau(1+k_x^2+k_y^2)$, so the last term in Eq. (43) gives the Dimits shift. To

give some comparison with the results of gyrofluid simulations,[24] we take equal electron and ion temperatures ($\tau = 1$), $k_\perp \rho_s = 1$ for the fastest growing modes, and no drift wave dissipation ($\mu = 0$). Equation (43) yields $(R/L_T)_c \approx 4$ and $(R/L_T)_* \approx 4.5$ for the values of the temperature gradient at marginality and the Dimits shift termination, respectively. (Flux tube toroidal gyrofluid simulations predict these quantities to be roughly 4 and 5, with only half the shift accounted for according to the gyrokinetic results of Dimits *et al.*[6]) Note that although our model is obviously too simplistic to give precise quantitative results, it still can capture important qualitative aspects of the dynamics.

VI. MULTIPLE-SCALE ANALYSIS AND MODULATIONAL INSTABILITY

The possibility of modulational instability must also be addressed when discussing the mechanisms of the transition to turbulence. Envelope-type long-wavelength modulations[25] imposed on the low-order truncated system (5) may be unstable under certain circumstances. In particular, the steady-state solutions that exist between the bifurcation points found in Appendixes C 2 and C 3 may become modulationally unstable, thus complicating the picture of the transition to turbulence. Our work in this area is incomplete; however, we include some discussion in order to both make contact with the reductive perturbation approach used by Dastgeer *et al.*[7] and to emphasize important qualitative differences between collisional and collisionless systems. Further exploration of the inhomogeneous equations derived here would be an interesting area for future research.

We assume that the characteristic spatial scale of envelope-type modulations $\boldsymbol{R}$ of the original pump (linearly unstable) DW mode is much larger than the wavelength of the DW itself, and also that the evolution of the modulational instability (with characteristic time scale T) is much slower than the linear DW growth rate. This permits expansion of the time and space derivatives with respect to a smallness parameter δ. Upon letting t_0 and ∇_0^{-1} be the typical time and space scales of the original DW mode in the system and introducing

multiple slow scales, we expand

$$\begin{aligned}\partial_t &= \partial_{t_0} + \delta\partial_T \\ &= \partial_{t_0} + \delta\partial_{T_1} + \delta^2\partial_{T_2} + \cdots, \qquad (44a)\\ \partial_{\boldsymbol{x}} &= \nabla_0 + \delta\partial_{\boldsymbol{R}} = \nabla_0 + \delta\left(\partial_X,\, \partial_Y\right)^T \\ &= \nabla_0 + \delta\partial_{\boldsymbol{R}_1} + \delta^2\partial_{\boldsymbol{R}_2} + \cdots. \qquad (44b)\end{aligned}$$

The parameter δ is chosen so that a minimum number of expansions is needed to arrive at the dynamics on the time scale of interest, while at the same time not losing any important fast dynamics. For a collisional system the parameter may be conveniently chosen to be $\delta \doteq (\epsilon - \epsilon_c)^{1/2}$, which measures the distance of the temperature gradient away from the marginal point. For a collisionless (or weakly collisional) system, $\delta \doteq \epsilon - \epsilon_0$ (ϵ_0 being found from $\mathrm{Re}[\lambda_+(\boldsymbol{z}_0, \epsilon_0)] = 0$ with $\boldsymbol{z}_0$ being some chosen point on the $\boldsymbol{z}$ plane) scales with the distance of the temperature gradient away from the *marginal curve* (Sec. II B).

A. Collisional system

In a collisional system [where zonal flow modes are linearly damped by $\eta_2 = O(1)$], the $O(\delta)$ expansion will generate the original DW mode with marginal eigenfrequency $\Omega_D = i\lambda_+(\boldsymbol{0}, \epsilon_c)$, which constitutes the center eigenspace. Zonal modes do not enter at this order, since if introduced they would be linearly damped. They do enter at $O(\delta^2)$ in a balance between (i) nonlinear generation by DW–DW interaction, and (ii) linear damping. These DW and ZF modes are enough to obtain a nontrivial equation for an envelope. The evolution of the envelope modulation $\widetilde{D}(T, \boldsymbol{R})$ of the underlying DW harmonic is described by [Eq. (D15)]

$$\begin{aligned}\partial_{\xi_D}\widetilde{D} &\doteq \left(\partial_T + \boldsymbol{V}_D\cdot\partial_{\boldsymbol{R}}\right)\widetilde{D} \qquad (45a)\\ &= \breve{\lambda}_+^{[1]}\widetilde{D} + \widehat{\Lambda}_{\boldsymbol{R}}^{(2)}\widetilde{D} - B_0|\widetilde{D}|^2\widetilde{D}, \qquad (45b)\end{aligned}$$

where the coefficients $\breve{\lambda}_+^{[1]}$ and B_0 describe the basic linear and nonlinear interactions while $\boldsymbol{V}_D$ gives the group velocity of the ITG wave packet. The notation $\breve{\lambda}^{[n]}$ denotes the expansion of the ϵ-dependent eigenvalue through n orders in $\epsilon - \epsilon_c$, excluding the zeroth-order term [which is accounted for in the underlying, rapidly varying harmonic dependence; see Eq. (D2)]. Clearly the $\breve{\lambda}_+^{[1]}$ in Eq. (45b) is an approximation to $\breve{\lambda}_+^{[\infty]} \doteq \lambda_+(\boldsymbol{0}, \epsilon) - \lambda_+(\boldsymbol{0}, \epsilon_c)$.

The nonuniform operator $\widehat{\Lambda}^{(2)}_{\boldsymbol{R}} = (\partial_{\boldsymbol{R}})^T \cdot \boldsymbol{\Lambda} \cdot \partial_{\boldsymbol{R}}$ (where $\boldsymbol{\Lambda}$ is a known matrix), determines the form of the "neutral surface" [defined by $\sigma(\Delta\boldsymbol{k}_{\rm neutr}, \epsilon) = 0$ with $\sigma(\Delta\boldsymbol{k}, \epsilon) \doteq \mathrm{Re}\,\breve{\lambda}^{[1]}_{+} - \Delta\boldsymbol{k} \cdot \boldsymbol{\Lambda} \cdot \Delta\boldsymbol{k}$] for the first bifurcating mode [which destabilizes at $\boldsymbol{k}_c = (k_x, k_y)^T$]. Note that sidebands are generated only at third order, so do not contribute to Eq. (45b). That equation may be reduced to Eq. (13) with $\lambda_+(\boldsymbol{0}, \epsilon) \approx -i\Omega_D + \breve{\lambda}^{[1]}_{+}$ *via* the addition of high-frequency DW dynamics by the transformation

$$\dot{D} = (\partial_{t_0} + \partial_{\xi_D})\, \widetilde{D} e^{-i\Omega_D t} \tag{46}$$

and the neglect of nonuniform terms.

Equation (45b) is a two-dimensional version of the standard Ginzburg–Landau equation,[26] with expansion performed in both the X (radial) and Y (poloidal) directions. It allows solutions of the form

$$\widetilde{D} = D_0 e^{i[(\boldsymbol{k}_c + \Delta\boldsymbol{k})\cdot\boldsymbol{R} + \omega\xi_D]} + \delta\widetilde{D}(\boldsymbol{R}, \xi_D), \tag{47}$$

where the first term is the steady-state solution of the uniform equation with an envelope type of modulation $(\boldsymbol{k}_c + \Delta\boldsymbol{k})\cdot\boldsymbol{R} + \omega\xi_D$, where $\omega \doteq -\Delta\boldsymbol{k}\cdot\mathrm{Im}\,\boldsymbol{\Lambda}\cdot\Delta\boldsymbol{k} - \mathrm{Im}\,B_0\sigma(\Delta\boldsymbol{k}, \epsilon)/\,\mathrm{Re}\,B_0$. The amplitude $|D_0|^2 = \sigma(\Delta\boldsymbol{k}, \epsilon)/\,\mathrm{Re}\,B_0$ is related to the corresponding fixed point (C16). The second term

$$\begin{aligned} \delta\widetilde{D}(\boldsymbol{R}, \xi_D) \;=\; & a(\boldsymbol{R}) e^{i[(\boldsymbol{k}_c + \Delta\boldsymbol{k}')\cdot\boldsymbol{R} + \omega'\xi_D]} \\ & + b(\boldsymbol{R}) e^{i[(\boldsymbol{k}_c + \Delta\boldsymbol{k}'')\cdot\boldsymbol{R} + \omega''\xi_D]} \end{aligned} \tag{48}$$

is its perturbation with some $\boldsymbol{k}_c + \Delta\boldsymbol{k}'$ and $\boldsymbol{k}_c + \Delta\boldsymbol{k}''$ wave numbers, which may become unstable by feeding from the base pump DW. The relationships between the wave numbers and growth rates are given by

$$\Delta\boldsymbol{k}' + \Delta\boldsymbol{k}'' = 2\Delta\boldsymbol{k}, \quad \omega' + \omega'' = 2\omega. \tag{49}$$

Analysis of the Ginzburg–Landau equation shows[27] that if $|\boldsymbol{k}_c + \Delta\boldsymbol{k}| < f(\Delta\boldsymbol{k}_{\rm neutr})|\boldsymbol{k}_c + \Delta\boldsymbol{k}_{\rm neutr}|$ with some $f(\Delta\boldsymbol{k}_{\rm neutr}) < 1$, then the wave numbers $\boldsymbol{k}_c + \Delta\boldsymbol{k}'$ and $\boldsymbol{k}_c + \Delta\boldsymbol{k}''$ are stable. In a range $f(\Delta\boldsymbol{k}_{\rm neutr})|\boldsymbol{k}_c + \Delta\boldsymbol{k}_{\rm neutr}| < |\boldsymbol{k}_c + \Delta\boldsymbol{k}| < |\boldsymbol{k}_c + \Delta\boldsymbol{k}_{\rm neutr}|$, the wave number $\boldsymbol{k}_c + \Delta\boldsymbol{k}$ is unstable due to feeding from the base pump wave $\boldsymbol{k}_c + \Delta\boldsymbol{k}$. This effect is called the Benjamin–Feir resonance mechanism.[27] Note that the nature of toroidal configurations allows only a discrete number of modes in both the poloidal and the radial directions. In particular,

the minimum poloidal wave number is determined by the poloidal circumference, while the minimum radial wave number is set by the properties of linear (or possibly nonlinear) eigenfunctions in the presence of magnetic shear. Thus, collisional analysis predicts that as the temperature gradient increases away from marginality, the spread of the neutral surface grows and at some $\epsilon_m \gtrsim \epsilon_c$ one or more modes enter the range unstable to the Benjamin–Feir mechanism.

B. Collisionless system

Collisionless systems differ substantially from collisional ones. Applying multiple-scale analysis to a weakly collisional system requires the introduction of zonal flow modes at $O(\delta)$ [as opposed to a collisional system, in which they enter the problem at $O(\delta^2)$]. This is another way of stating that the collisionless center eigenspace contains both DW's and ZF's. This presence of ZF's at $O(\delta)$ forces sidebands to be generated at $O(\delta^2)$ [as opposed to a collisional system, in which the sideband modes are generated at $O(\delta^3)$]. The resulting system of nonuniform equations up to $O(\delta^3)$ (Appendix D 2) for the envelope is [Eq. (D35)]

$$\partial_{\xi_D} \widetilde{D} = [\Gamma^{(1)}(\boldsymbol{z}, \epsilon) + \widehat{\Lambda}^{(2)}_{\boldsymbol{R}}]\widetilde{D} + \widehat{\boldsymbol{\Lambda}}^{(1)}_{\boldsymbol{R}}(\boldsymbol{z}, \widetilde{D}), \tag{50a}$$

$$\partial_T \boldsymbol{z} = (-\epsilon \boldsymbol{a} + \mathsf{A} \cdot \boldsymbol{z})|\widetilde{D}|^2 + \widehat{\boldsymbol{\Sigma}}^{(1)}_{\boldsymbol{R}}(\widetilde{D}, \overline{\widetilde{D}}). \tag{50b}$$

There is no ZF linear dissipation term, since we assume a purely collisionless regime ($\eta_2 = 0$). [More precisely, collisionless means $\eta_2 = O(\delta^n)$ with $n \geq 4$, since we carried out the expansion only up to $O(\delta^3)$.] Also note that without the nonuniform terms, the system (50) reduces to Eqs. (21) under the transformation (46). Thus, it captures both the dynamics of Eqs. (5) in the Dimits-shift regime as well as modulational effects.

1. *Modulational stability in the Dimits-shift regime*

Obviously the stability properties of Eqs. (C35) in the Dimits-shift regime Δ ($\epsilon < \epsilon_*$), particularly the stability of the fixed point $\mathcal{F}(I_{\mathcal{F}} = 0, \boldsymbol{z}_{\mathcal{F}})$ pertain to Eqs. (50) as well except that the location of $\mathcal{F}$ is slightly perturbed to

$$\boldsymbol{z}_{\mathcal{F}} = \mathsf{A}^{-1} \cdot (\epsilon \boldsymbol{a} - \boldsymbol{\Sigma}^{(1)}_{\boldsymbol{K}}) \tag{51}$$

and the criterion $\mathrm{Re}[\Gamma^{(1)}(\boldsymbol{z}^*_{\mathcal{F}} = \boldsymbol{z}_{\mathcal{F}}(\epsilon_*), \epsilon_*)] = 0$ for finding the critical ϵ_* is slightly changed to

$$\mathrm{Re}[\Gamma^{(1)}(\boldsymbol{z}^*_{\mathcal{F}}, \epsilon_*) + \Lambda^{(2)}_{\boldsymbol{K}} + \boldsymbol{\Lambda}^{(1)}_{\boldsymbol{K}} \boldsymbol{z}^*_{\mathcal{F}}] = 0. \tag{52}$$

Here $\Lambda^{(2)}_{\boldsymbol{K}}$, $\boldsymbol{\Lambda}^{(1)}_{\boldsymbol{K}}$ and $\boldsymbol{\Sigma}^{(1)}_{\boldsymbol{K}}$ are the $\boldsymbol{k}$-space versions of the operators $\widehat{\Lambda}^{(2)}_{\boldsymbol{R}}$, *etc.* Thus, the CM dynamics of the collisionless system (50) are modulationally stable in the Dimits-shift regime. This differs from the collisional CM dynamics (45b), which become modulationally unstable near marginality $\epsilon = \epsilon_c$.

Calculations similar to our multiple-scale analysis have been attempted with regard to the dynamics of the Dimits-shift regime. Weiland *et al.*[8] used the reductive perturbation method[13] to approach the generation of ZF modes in a collisionless reactive model. That calculation did not allow the presence of ZF modes at first order, but rather only allowed them to be generated at second order above linear marginality threshold as in the standard collisional calculations. Such an approach misses the importance of sidebands and their role in the formation of the Dimits shift. In particular, in that analysis generation of ZF's is found from the $O((\epsilon - \epsilon_c)^3)$ equation $\partial_T \boldsymbol{z} = \widehat{\boldsymbol{\Sigma}}^{(1)}_{\boldsymbol{R}}(\widetilde{D}, \overline{\widetilde{D}})$, which does not capture the saturation mechanism for ZF's in the collisionless system or allow one to calculate the destabilization (upper limit) of the Dimits-shift regime.

2. Transition to collisionless ITG turbulence

At $\epsilon = \epsilon^{(3)} = \epsilon_*$ satisfying Eq. (52), the Dimits-shift regime $\Delta\epsilon = \epsilon_* - \epsilon_c$ of no DW activity is terminated. Above this instability limit, the system (50) cannot saturate for some IC's. As we remark in the next section, higher-order modal truncations do saturate. In essence, the multiple-scale analysis also adds more modes, although ones ordered to be of long wavelength relative to the primary modes retained in the homogeneous theory. Ultimately, we cannot make definitive statements about the transition to collisionless turbulence until we combine multiple-scale analysis with higher-order truncations; that remains to be done. However, we will comment briefly on the general structure of the theory and derive the general form of a collisionless Ginzburg–Landau equation.

To find the dynamics of the original system above ϵ_*, one needs to expand Eqs. (1) up to $O(\delta^4)$. Following the formalism presented in Appendix D 2, we collect the $O(\delta^4)$ terms and calculate their contribution to the $\boldsymbol{z}$ part of the constraint equations (50); we will see later

that corrections to the $\widetilde{D}$ part of (50) are unnecessary. This constraint includes corrections to already existing uniform terms as well as various new nonuniform ones. The resulting system for the envelope amplitudes is

$$\partial_{\xi_D}\widetilde{D} = [\Gamma^{(1)}(\boldsymbol{z},\epsilon) + \widehat{\Lambda}^{(2)}_{\boldsymbol{R}}]\widetilde{D} + \widehat{\boldsymbol{\Lambda}}^{(1)}_{\boldsymbol{R}}(\boldsymbol{z},\widetilde{D}), \tag{53a}$$

$$\begin{aligned}\partial_T \boldsymbol{z} = {} & (-\boldsymbol{a}\epsilon + \mathsf{A}(\boldsymbol{z})\cdot\boldsymbol{z})|\widetilde{D}|^2 + \widehat{\boldsymbol{\Sigma}}^{(1)}_{\boldsymbol{R}}(\widetilde{D},\overline{\widetilde{D}}) \\ & + \widehat{\boldsymbol{\Sigma}}^{(2)}_{\boldsymbol{R}}(\widetilde{D},\overline{\widetilde{D}}) + \widehat{\Sigma}^{(3)}_{\boldsymbol{R}}\cdot\boldsymbol{z}.\end{aligned} \tag{53b}$$

As one can see, the equation for the $\widetilde{D}$ amplitude is unchanged, whereas the uniform part of the $\boldsymbol{z}$ equation now involves an amplitude-dependent matrix $\mathsf{A}(\boldsymbol{z})$ and its nonuniform part has two new terms involving the $\widehat{\boldsymbol{\Sigma}}^{(2)}_{\boldsymbol{R}}$ and $\widehat{\Sigma}^{(3)}_{\boldsymbol{R}}$ operators.

For $\epsilon < \epsilon_*$, the system (53) is stable for any IC's. The nonuniform terms slightly change the stability and the location of the special fixed point, but otherwise do not play any interesting role in the dynamics.

For $\epsilon \geq \epsilon_*$, some solutions (with IC's that generate trajectories ending up close to $\mathcal{F}$) become unstable (cannot saturate via just uniform terms). After some transient dynamics, these solutions satisfy $\mathrm{Re}[\Gamma(\boldsymbol{z}_\mathcal{F},\epsilon)] \geq 0$ and $-\boldsymbol{a}\epsilon + \mathsf{A}(\boldsymbol{z}_\mathcal{F})\cdot\boldsymbol{z}_\mathcal{F} = \boldsymbol{0}$. Thus, to study the evolution of these solutions for $\epsilon \geq \epsilon_*$ we use the system

$$\partial_{\xi_D}\widetilde{D} = [\Gamma^{(1)}(\boldsymbol{z}_\mathcal{F},\epsilon) + \widehat{\Lambda}^{(2)}_{\boldsymbol{R}}]\widetilde{D} + \widehat{\boldsymbol{\Lambda}}^{(1)}_{\boldsymbol{R}}(\boldsymbol{z},\widetilde{D}), \tag{54a}$$

$$\partial_T \boldsymbol{z} = \widehat{\boldsymbol{\Sigma}}^{(2)}_{\boldsymbol{R}}(\widetilde{D},\overline{\widetilde{D}}) + \widehat{\Sigma}^{(3)}_{\boldsymbol{R}}\cdot\boldsymbol{z}. \tag{54b}$$

In a regime where the $\boldsymbol{z}$ amplitude variation is slow compared to that of $\widetilde{D}$, the system (54) converges to[5]

$$\partial_{\xi_D}\widetilde{D} = [\Gamma^{(1)}(\boldsymbol{z}_\mathcal{F},\epsilon) + \widehat{\Lambda}^{(2)}_{\boldsymbol{R}}]\widetilde{D} + \widehat{\boldsymbol{\Lambda}}^{(1)}_{\boldsymbol{R}}(\boldsymbol{z},\widetilde{D}), \tag{55a}$$

$$\boldsymbol{z} = -(\widehat{\Sigma}^{(3)}_{\boldsymbol{R}})^{-1}\cdot\widehat{\boldsymbol{\Sigma}}^{(2)}_{\boldsymbol{R}}(\widetilde{D},\overline{\widetilde{D}}). \tag{55b}$$

Since in the $\partial/\partial Y = 0$ limit the last operator in the first equation turns out to be a constant (independent of $\boldsymbol{R}$)[5] multiplied by $|\widetilde{D}|^2\widetilde{D}$, Eqs. (55) provide a collisionless version of the Ginzburg–Landau equation that describes the usual Benjamin–Feir resonance mechanism just above the ϵ_* of the collisionless system. In particular, since the poloidal wave numbers are discrete (determined by the poloidal circumference), as the temperature gradient increases away from $\epsilon = \epsilon_*$ the spread of the neutral surface grows; at some $\epsilon_m > \epsilon_*$, we predict that one or more modes should enter the range unstable to the Benjamin–Feir mechanism for the transition to collisionless turbulence.

VII. EFFECTS OF WEAK COLLISIONALITY AND HIGHER-ORDER TRUNCATIONS

Now consider the addition of very *weak* zonal damping: $\dot{\boldsymbol{z}} = -\eta_2 \boldsymbol{z} + \cdots$. ($\eta_2$ can also be considered to be an unfolding parameter.) By weak, we mean that the zonal modes are still placed into the 4D center eigenspace. Alternatively, the weakness of zonal damping means that $\eta_2 = O(\delta^n)$ ($n \geq 2$) and thus enters no sooner than third order in the multiple-scale expansion described in Sec. VI.

If one assumes that $\eta_2 = O(\delta^3)$, expansion up to $O(\delta^4)$ (with the neglect of nonuniform terms) yields

$$\dot{D} = \underbrace{\Gamma(\boldsymbol{z}, \epsilon) D}_{O(\delta^3)}, \tag{56a}$$

$$\dot{\boldsymbol{z}} = \underbrace{(-\boldsymbol{a}\epsilon + \mathsf{A} \cdot \boldsymbol{z})|D|^2}_{O(\delta^3)} - \underbrace{\eta_2 \boldsymbol{z}}_{O(\delta^4)}. \tag{56b}$$

Here we changed back to the D variable (which includes the DW harmonic dependence) by performing the transformation (46) and neglecting nonuniform terms.

The collisionless dynamics involves different time scales: t_0 is associated with the DW linear damping rate, T_1 characterizes motion of the wave packet with the group velocity, and T_2 describes evolution toward the $\mathcal{F}$ determined by the $O(\delta^3)$ terms in (56). The presence of the small damping parameter η_2 introduces a new, very long time scale T_3 and perturbs the position of $\mathcal{F}$. For $\epsilon < \epsilon_*$, arbitrary initial conditions typically move rapidly to the vicinity of the original fixed point (of the undamped dynamics), then slowly relax to the final steady state determined by all of the underbraced terms. That state involves a small, nonturbulent DW component.

This disparity of time scales underlies the bursting behavior observed in Ref. 4 for weakly collisional runs. That does not occur in the lowest-order truncation studied here, but does occur in higher-order ones, whose additional degrees of freedom allow $\mathcal{F}$ to be destabilized in other directions and, thus, the trajectories to be ejected from its vicinity after the slow relaxation, producing a bursting type of dynamics with a characteristic collision frequency. Preliminary long-time (many-burst) integrations of such truncations show relaxation to a quasiregular state (Fig. 11). We believe that limited computational resources precluded the authors of Ref. 4 from integrating more than a few bursts and seeing this quasiregular state

on a long time scale.

In real experiments one observes, and one is interested in, plasma turbulence occurring on the transport timescale T_n ($n \gg 3$), which is of the order of a second, much longer than the characteristic damping time of zonal flows (of the order of the ion–ion collision time). Strictly speaking, study of this regime requires higher-order expansions and truncations than we have considered here. Nonetheless, as is easy to see from our treatment of weakly collisional systems, the smallness of the damping parameter results in a sophisticated mechanism of the transition to turbulence that includes the Dimits-shift phenomenon.

VIII. DISCUSSION

In summary, we have considered a simple yet instructive model for the transition to collisionless ion-temperature-gradient-driven plasma turbulence. The excitation of zonal flows, important in a variety of physics contexts, plays a crucial role in the dynamics of that transition. A quantitatively accurate calculation of the critical temperature gradient for the onset of turbulence (known as the Dimits shift) with complete toroidal physics is best left to large simulations. Here we focused on the detailed understanding of basic conceptual issues and clarified some subtle asymptotic behavior. Unlike some other models of recent interest,[11,22] our model captures a zonal-flow-driven Dimits shift even with a very simplified set of modes. By using tools from dynamical systems theory, we have shown how the Dimits shift is related to a certain fixed point of the nonlinear system (with just one bifurcation point of interest) and how that shift can be calculated in terms of the physical parameters of the model.

As opposed to systems with linearly damped (weakly damped in reality) zonal modes with a 2D center manifold, the systems with linearly undamped zonal modes have at least a 4D CM, thus leading to nontrivial dynamics in the Dimits-shift regime. The differing dimensionality of the CM's for the two cases is the result of the interchange of the limits $t \to \infty$ and zonal damping $\eta_2 \to 0$. The Dimits shift occurs when the $\eta_2 \to 0$ limit is taken first. This important conceptual issue has been previously overlooked when trying to apply standard collisional dynamical-systems techniques to collisionless models. We wish to stress that this asymptotic interchange of limits, related to very weak zonal damping, underlies the very existence of the Dimits shift. The significance of such subtle asymptotology has been

previously discussed in the related context of the collisionless simulation of drift waves.[19]

We employed multiple-scale analysis to study the possibility of the fixed point becoming unstable to modulational (envelope) perturbations. For collisionless systems, we derived the system of nonuniform equations (D35) as a generalization of the Ginzburg–Landau equation. We concluded that the Dimits-shift regime is not destroyed by modulational instability.

The low-degree-of-freedom methods that work well for our Galerkin-truncated model cannot be practically applied to the original collisionless ITG system (1) of PDE's, even close to marginality. To do so, one would have to add an infinite number of undamped zonal modes into the CM calculation (Appendix C 3) or into the solution of the $O(\delta)$ equation in multiple-scale analysis (Appendix D). The resulting CM would be infinite-dimensional, and one would have to calculate the fixed point in infinite-dimensional space. (In practice, not all zonal modes can be considered to be collisionless, but the relevant space would still be very large.) A related point is that the Dimits shift depends on the order of truncation. Together, these observations point to the desirability of alternate methods of calculation that apply directly to the PDE's.

In any event, we do not claim to have calculated the Dimits shift for a realistic system of PDE's. We do, however, believe that the insights gained from this calculation remain relevant for more complicated and physically complete models. Systematic dynamical systems analysis has been little used for problems of many-mode plasma microturbulence, and indeed the difficulties are severe for situations of fully developed turbulence. However, for qualitative discussion of the dynamical mechanisms involved in the transition to turbulence, the present calculation demonstrates the utility of such analysis and emphasizes the importance of the center manifold. The detailed study of models with a small number of degrees of freedom has important pedagogical worth and, for qualitative understanding, provides a viable alternative to large, brute-force, and expensive numerical simulations.

Acknowledgments

We would like to thank J. Weiland for sending us a preprint of Ref. 8. We have profited from valuable discussions with P. Holmes. This work was supported by U. S. Dept. of Energy Contract No. DE-AC02-76-CHO-3073.

APPENDIX A: APPROXIMATIONS TO THE LINEAR MATRIX

As a basis for the linear matrix M in our Eqs. (1), we begin with the original linear matrix from Ref. 21 for the description of reactive ITG dynamics of the $(\varphi, P)^T$ fields:

$$\widehat{\mathsf{M}} = \begin{pmatrix} i(-\omega_* + 2\omega_d)/(\tau + b) & 2i\omega_d/(\tau + b) \\ -i(1+\eta_i)\omega_* + 3i\omega_d[1-(\tau+b)] & 6i\omega_d \end{pmatrix}. \tag{A1}$$

Here τ is the ratio of ion and electron temperatures, $b \doteq -\widehat{\nabla}^2\rho_i^2$, and ρ_i is the ion gyroradius. The ion diamagnetic drift frequency ω_* is given by $\omega_* \doteq -k_y\rho_i v_{ti}/L_n$, and the toroidal drift frequency ω_d for a large-aspect-ratio tokamak in the outer midplane of the torus is $\omega_d \doteq \omega_* L_n/R$. Here v_{ti} is the ion thermal velocity, L_n is the ion density-gradient scale length, and R is the magnetic curvature. Hats denote integro-differential operators; with our choice of boundary conditions, those become multiplicative in $\boldsymbol{k}$ space.

We shall modify the original matrix (A1) in order to retain only terms essential for the dynamics in which we are interested. In particular, we consider the flat-density limit and neglect all terms involving ω_* except for the $\eta_i\omega_*$ driving term (which is proportional to L_T^{-1}).

One significant feature of such reactive models (which contain no classical collisional dissipation) is the need to somehow include the Landau damping effect. That can be done through a gyrofluid closure.[21] In particular, the presence of Landau damping terms in the pressure equation enforces physically correct properties like the dependence of the dispersion relation on τ or the finite frequency of the first bifurcating mode.

Upon introducing $\omega_{ds} \doteq \omega_d/\tau$ and incorporating the above corrections, we transform the matrix (A1) in terms of the $(\varpi \doteq \widehat{\mathcal{D}}^{-1}\varphi, P)^T$ fields to obtain

$$\widehat{\mathsf{M}}/\omega_{ds} = \begin{pmatrix} 2i\widehat{\mathcal{D}}^{-1} & 2i \\ \left(-i\tau\dfrac{R}{L_T}\widehat{\mathcal{D}}^{-1} + 3i\tau(\widehat{\mathcal{D}}^{-1} - \tau) - \dfrac{|\omega_{ds}|}{\omega_{ds}}\tau^2\nu\right) & \left(6i\tau + \dfrac{|\omega_{ds}|}{\omega_{ds}}\tau\nu\right) \end{pmatrix}. \tag{A2}$$

Here we used $\widehat{\mathcal{D}} \doteq (\widehat{\alpha} - \widehat{\nabla}^2)$ [with $\widehat{\alpha}$ being zero for convective cells ($k_\parallel = 0$) and the identity operator otherwise.[12]] and $\eta_i = L_n/L_T$. The parameter ν accounts for Landau damping.

In the limit of small τ, the system $(\varpi, P)^T$ with the linear matrix (A2) becomes linearly unstable at $R/L_T = (1 + \tau\widehat{\mathcal{D}})$. Thus, by neglecting terms of higher order in τ while simultaneously appropriately modifying the definitions of the driving term and the frequency $\widehat{\Omega}$, we are able to obtain a system that both is as simplified as possible and also possesses the correct linear threshold as a function of various parameters.

The resulting matrix $\widehat{\mathsf{M}}$ has the form

$$\widehat{\mathsf{M}} = \begin{pmatrix} -i(\widehat{\Omega} - i\widehat{\eta}) & -i\widehat{b} \\ i\widehat{\epsilon} & -\widehat{d} \end{pmatrix}. \tag{A3}$$

Here $\widehat{\Omega} \doteq -2i(\widehat{\mathcal{D}}^{-1} + \tau)\widehat{\partial}_y$ and $\widehat{\epsilon} \doteq -i\tau\widehat{\mathcal{D}}^{-1}[R/L_T - (1 + \tau\widehat{\mathcal{D}})]\widehat{\partial}_y$. The coefficient $\widehat{\eta} \doteq -\mu\widehat{\nabla}^2$ describes weak collisional damping and is included here to extend Eq. (A1) to the dissipative case; $\widehat{b} \doteq -2i\widehat{\partial}_y$; and $\widehat{d} \doteq \widehat{\overline{\nu}} + \widehat{\eta}$, where $\widehat{\overline{\nu}} = -\nu|\widehat{\partial}_y|$. Note that the coefficients in Eq. (A3) are dimensionless and normalized to ω_{ds}.

APPENDIX B: PROJECTION METHOD FOR PERTURBATIVE CENTER MANIFOLD CALCULATIONS

We review the methodology for perturbative calculation of the center manifold of the n-dimensional system

$$\partial_t \boldsymbol{u} = \mathsf{M}(\epsilon) \cdot \boldsymbol{u} + \boldsymbol{N}[\boldsymbol{u}], \tag{B1}$$

where $\boldsymbol{N} = O(u^2)$. We assume that we can write $\mathsf{M}(\epsilon) = \mathsf{M}^{(0)} + \epsilon\mathsf{M}^{(1)}$; generalization to arbitrary ϵ dependence is straightforward. $\mathsf{M}(\epsilon)$ is an arbitrary $n \times n$ matrix, not necessarily Hermitian, parametrized by ϵ. It possesses eigenvalues $\lambda_i(\epsilon)$ $(i = 1, \ldots, n)$ with left eigenvectors $\boldsymbol{p}^i$ and right eigenvectors $\boldsymbol{q}_i$, assumed to obey $\boldsymbol{p}^{i\dagger} \cdot \boldsymbol{q}_j = \delta^i_j$. It is assumed that at $\epsilon = 0$ n_0 eigenvalues lie on the imaginary axis. The associated eigenvectors span an n_0-dimensional subspace, the *center eigenspace*; all other eigenvalues are assumed to be stable. The Center Manifold Theorem (CMT) then states that there exists an n_0-dimensional invariant manifold, the *center manifold*, that is tangent to the center eigenspace at $\epsilon = 0$.

This basic version of the theorem characterizes the CM for fixed $\epsilon = 0$. In order to deal with systems parametrized by ϵ, and thus to consider dynamics at $\epsilon \neq 0$, one may adjoin to Eq. (B1) the equation $\dot{\epsilon} = 0$, thus including ϵ as a dynamical variable and augmenting the dimensionality by one.[14,15] The CMT applies to this new $(n+1)$-D "suspended" system, and

predicts an (n_0+1)-D CM tangent at the origin to an (n_0+1)-D center eigenspace. The goal is to calculate the shape of that surface perturbatively.[28] (Note that in this paper we quote the dimensionality of the CM as n_0, not n_0+1.)

In early accounts of the perturbative calculation,[14] a preliminary linear transformation was performed to bring M to diagonal form. That is unnecessary, however; projection methods[15] based on general contravariant representations can be used instead. We shall discuss two variants of this approach, Method A and Method B. In both, we exploit the fact that the $\boldsymbol{q}$'s form a nonorthogonal basis, with the $\boldsymbol{p}$'s providing a dual or reciprocal basis. In Method A, we expand $\boldsymbol{u}$ contravariantly in terms of the $\boldsymbol{q}_0$'s, where $\boldsymbol{q}_0 \equiv \boldsymbol{q}(\epsilon = 0)$ and $\boldsymbol{p}_0 \equiv \boldsymbol{p}(\epsilon = 0)$. Thus

$$\boldsymbol{u} = \sum_{i \in n_0} D_0^i \boldsymbol{q}_{0i} + \boldsymbol{y}_0, \tag{B2}$$

where $\boldsymbol{y}_0 \doteq \sum_{k \notin n_0} y_0^k \boldsymbol{q}_{0,k}$ lies in the stable subspace. The contravariant coordinates D^i and y^k are interpreted in Fig. 12 for the case $n = 2$, $n_0 = 1$. The modal amplitudes are obtained by noting from Eq. (B2) that $D_0^i = \boldsymbol{p}_0^{i\,\dagger} \cdot \boldsymbol{u}$. Then, by applying the operator $\mathsf{P}_0 = \sum_{i \in n_0} \boldsymbol{q}_{0i}\, \boldsymbol{p}_0^{i\dagger}$ (which projects onto the center eigenspace) to Eq. (B1) and recalling that the $\boldsymbol{p}_0^i$'s are associated with the eigenvectors of $\mathsf{M}^{(0)}$ that lie on the imaginary axis, one obtains

$$\dot{D}_0^i = -i\Omega_{0,j}^i D_0^j + \boldsymbol{p}_0^{i\dagger} \cdot (\boldsymbol{N} + \boldsymbol{N}') \tag{B3}$$

(summation convention), where $\Omega_{0,j}^i \doteq i\boldsymbol{p}_0^{i\dagger} \cdot \mathsf{M}^{(0)} \cdot \boldsymbol{q}_{0,j}$ gives the marginal eigenfrequencies and $\boldsymbol{N}' \doteq \epsilon \mathsf{M}^{(1)} \cdot (D^j \boldsymbol{q}_{0,j} + \boldsymbol{y}_0)$ is a contribution to the effective nonlinearity that arises because we choose to expand around marginality. The term $\epsilon \boldsymbol{p}_0^{i\dagger} \cdot \mathsf{M}^{(1)} \cdot \boldsymbol{q}_{0,j}$ provides an $O(\epsilon)$ correction to the unstable eigenvalue(s), while the $O(\boldsymbol{y}_0)$ part of $\boldsymbol{N}'$ will generate eigenvalue corrections of $O(\epsilon^2)$. On the CM, the CMT guarantees that the y^k's are at least quadratic in any combination of ϵ and D^i. They can be calculated perturbatively[28] by assuming a Taylor expansion for $\boldsymbol{y}_0$ (beginning at second order), then equating its time derivative to the right-hand side of the dynamical equation for $\boldsymbol{y}_0$, obtained by applying the orthogonal projection operator $\mathsf{Q}_0 \doteq \mathsf{I} - \mathsf{P}_0$ to Eq. (B1):

$$\dot{\boldsymbol{y}}_0 = \mathsf{M}^{(0)} \cdot \boldsymbol{y}_0 + \mathsf{Q}_0 \cdot (\boldsymbol{N} + \boldsymbol{N}'). \tag{B4}$$

The resulting solution for $\boldsymbol{y}_0$ can then be used to evaluate the nonlinear terms in Eq. (B3) to any desired order.

Although this method is completely systematic and furnishes the shape of the CM directly in terms of the original variables, one disadvantage is that the unstable eigenvalue is produced perturbatively, cluttering the algebra. It is therefore sometimes convenient (Method B) to expand in terms of the $\boldsymbol{q} \equiv \boldsymbol{q}(\epsilon)$ basis:

$$\boldsymbol{u} = D^i \boldsymbol{q}_i + \boldsymbol{y}, \tag{B5}$$

where $\boldsymbol{p}^{i\dagger} \cdot \boldsymbol{y} = 0$. One has (having utilized $\mathsf{Q} \doteq \mathsf{I} - \mathsf{P}$, $\mathsf{P} \doteq \sum_{i \in n_0} \boldsymbol{q}_i \, \boldsymbol{p}^{i\dagger}$)

$$\dot{D}^i \;=\; \lambda_i(\epsilon) D^i + \boldsymbol{p}^{i\dagger} \cdot \boldsymbol{N}, \tag{B6a}$$

$$\dot{\boldsymbol{y}} \;=\; \mathsf{M} \cdot \boldsymbol{y} + \mathsf{Q} \cdot \boldsymbol{N}. \tag{B6b}$$

Here there is no $O(\epsilon)$ correction to the nonlinearity; the full $\lambda_i(\epsilon)$ has appeared directly. The basic perturbation expansion proceeds as before. However, it is important to note that, because the eigenvectors rotate as a function of ϵ, the geometric shape of the coordinate-independent CM is now represented by different contravariant coordinates; that is, $D^i \neq D_0^i$ and the final equation for $\dot{D}^i$ on the CM differs by small terms involving ϵ from the equation for $\dot{D}_0^i$. Effectively, Method B incorporates a nonlinear transformation of variables that must be calculated explicitly if connection needs to be made to the original D_0's. However, topological properties, including the stability of fixed points, are invariant to this transformation.

In the present work, we use Method A exclusively.

In Sec. III B we discuss the 2D Hopf bifurcation, yet write Eq. (12a), $\boldsymbol{u}_1 = D\boldsymbol{q}_0 + \boldsymbol{y}_1$ [an example of the Method A decomposition (B2)], in terms of a single basis vector $\boldsymbol{q}_0$. This is correct because we do not explicitly consider the complex conjugates of the $\boldsymbol{u}$ amplitudes (which are required for reality of the original $\boldsymbol{x}$-space fields). A basic example is furnished by the harmonic oscillator equations $\dot{x} = \Omega y$ and $\dot{y} = -\Omega x$ (where we consider x and y to be real), or $\dot{\boldsymbol{x}} = \mathsf{M} \cdot \boldsymbol{x}$ with $\mathsf{M} \doteq \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \Omega$. The eigenvalues of M are $\lambda_\pm = \mp i\Omega$, and the eigenvectors are $\boldsymbol{q}_\pm = (1, \mp i)^T$ and $\boldsymbol{p}^\pm = \frac{1}{2}\boldsymbol{q}_\pm$. $\boldsymbol{q}_+$ and $\boldsymbol{q}_-$ span a 2D space of complex-valued vectors. The representation of a general complex vector $\boldsymbol{z}$, $\boldsymbol{z} = a^+\boldsymbol{q}_+ + a^-\boldsymbol{q}_-$, reduces for real $\boldsymbol{z} = \boldsymbol{x}$ to $\boldsymbol{x} = a^+\boldsymbol{q}_+ + \text{c.c.}$; here $\boldsymbol{q}_+$ is analogous to the $\boldsymbol{q}_0$ in Eq. (12a). Furthermore, the oscillator amplitude $D \equiv D^+ \doteq \boldsymbol{p}^{+\dagger} \cdot \boldsymbol{x} = \frac{1}{2}(x + iy)$ obeys $D = -i\Omega D$. In this representation, the complex-conjugate pair of eigenvalues discussed in the Hopf bifurcation arises because the equation $\overline{D} = i\Omega\overline{D}$ is also satisfied.

APPENDIX C: CENTER MANIFOLD CALCULATIONS

Here we provide details of the perturbative construction of the CM for both the collisional and collisionless cases.

1. Linear dynamics

The eigenvalues $\lambda_\pm(\mathbf{0},\epsilon)$ for the two DW branches are

$$\lambda_\pm(\mathbf{0},\epsilon) = \frac{1}{2}\{-(\nu_1 + 2\eta_1 + i\Omega_1) \pm [(\nu_1 - i\Omega_1)^2 + 4b_1\epsilon]^{1/2}\}. \tag{C1}$$

The branch with $\lambda_+(\mathbf{0},\epsilon)$ is the first to become linearly unstable; it does so at the ϵ_c found from $\mathrm{Re}[\lambda_+(\mathbf{0},\epsilon_c)] = 0$. At marginality $\lambda_+(\mathbf{0},\epsilon_c) = -i\Omega_D$ is purely imaginary with $\Omega_D = d_1\Omega_1/(d_1+\eta_1)$. This result is most easily obtained by decomposing the characteristic equation into real and imaginary parts, rather than by manipulating Eq. (C1) directly.

Because M has no special symmetry, the left and right eigenvectors associated with these eigenvalues are distinct; as explained in Appendix B, both will be required. The right eigenvector $\boldsymbol{q}$ obeys $\mathsf{M}_1 \cdot \boldsymbol{q}_\pm = \lambda_\pm(\mathbf{0},\epsilon)\boldsymbol{q}_\pm$. Because the normalization of $\boldsymbol{q}$ is arbitrary, we may fix its first (vorticity) component to one. Then

$$\boldsymbol{q}_\pm = \begin{pmatrix} 1 \\ [-i(\Omega_1 - i\eta_1) - \lambda_\pm(\mathbf{0},\epsilon)]/(ib_1) \end{pmatrix}. \tag{C2}$$

For the left eigenvector $\boldsymbol{p}$, we have $\mathsf{M}_1^\dagger \cdot \boldsymbol{p}_\pm = \overline{\lambda}_\pm(\mathbf{0},\epsilon)\boldsymbol{p}_\pm$ and

$$\boldsymbol{p}_\pm = \frac{1}{N_\pm}\begin{pmatrix} 1 \\ ib_1/[d_1 + \overline{\lambda}_\pm(\mathbf{0},\epsilon)] \end{pmatrix}. \tag{C3}$$

Here we choose $N_\pm$ to satisfy the arbitrary normalization $\boldsymbol{p}_\pm^\dagger \cdot \boldsymbol{q}_\pm = 1$. Also note that the orthogonality conditions $\boldsymbol{p}_\pm^\dagger \cdot \boldsymbol{q}_\mp = 0$ are satisfied. Nontrivial consequences of this orthonormality are that the ϖ and P components of the $\boldsymbol{p}$'s satisfy $p_{-,\varpi} = 1 - p_{+,\varpi}$ and $p_{-,P} = -p_{+,P}$.

When $\eta_1 = 0$, the DW threshold is $\epsilon_c = 0$. At that threshold, the DW eigenvalues are $\lambda_+(\mathbf{0},0) = -i\Omega_1$ and $\lambda_-(\mathbf{0},0) = -\overline{\nu}$. For this special case, the bifurcating ITG mode has no pressure component at the point of bifurcation: $\boldsymbol{q}_+ = (1,\,0)^T$.

2. Center manifold calculation for collisional system

Here we consider the CM calculation of the collisional [$\eta_2 = O(1)$] system (5) in order to allow comparison with the weakly collisional calculations presented in Sec. IV and Appendix C 3. The overall collection of modes in this low-order truncated model in the eigenvalue plane is shown in Fig. 2. Shaded on the figure are the modes included in the center eigenspace and which therefore determine the dynamical constraints on the CM. Note that the two damped zonal modes with linear matrix $\mathsf{M}_2 = -\eta_2 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ are not coordinates on the CM for $\eta_2 = O(1)$.

We apply the projection method A (see Appendix B) to the suspended system, yielding CM dynamics described by the $D = \boldsymbol{p}_0^\dagger \cdot \boldsymbol{u}_1 \in T^c$ variable on the center eigenspace. Here $\boldsymbol{p}_0 \equiv \boldsymbol{p}_+(\epsilon = \epsilon_c)$ and $\boldsymbol{q}_0 \equiv \boldsymbol{q}_+(\epsilon = \epsilon_c) \equiv (1, q^{(2)})^T$. The vectors $\{\boldsymbol{y}_1, \boldsymbol{y}_2, \boldsymbol{y}_3\} \in T^s$ span the stable eigenspace.

Application of the projection operator P_0 to the $\boldsymbol{u}_1$ equations of the system (5) gives for the dynamics on the CM

$$\dot{D} = \boldsymbol{p}_0^\dagger \cdot \mathsf{M}_1^{(0)} \cdot \boldsymbol{q}_0 D + \boldsymbol{p}_0^\dagger \cdot (\boldsymbol{N}_1 + \boldsymbol{N}_1'), \tag{C4a}$$

$$\dot{\epsilon} = 0, \tag{C4b}$$

where $\mathsf{M}_i^{(0)} \doteq \mathsf{M}_i(\epsilon = \epsilon_c)$ and $\boldsymbol{N}_1' \doteq (\epsilon - \epsilon_c)\mathsf{M}_1^{(1)} \cdot (\boldsymbol{q}_0 D + \boldsymbol{y}_1)$ with $\mathsf{M}_1^{(1)} \doteq \partial_\epsilon \mathsf{M}_{1|\epsilon_c}$. Note that in this method only $\mathsf{M}_1^{(0)}$ is treated as a linear term; the other terms of the expansion of M_1 are considered to be nonlinear (since they involve at least one power each of $\epsilon - \epsilon_c$ and D).

The vectors $\{\boldsymbol{y}_1, \boldsymbol{y}_2, \boldsymbol{y}_3\} \in T^s$ are slaved to the center modes; they are at least $O(|D, \overline{D}, \epsilon - \epsilon_c|^2)$ and thus contribute to the nonlinear curvature of the CM. Application of the orthogonal projection operator $\mathsf{Q}_0 \doteq \mathsf{I} - \boldsymbol{q}_0 \boldsymbol{p}_0^\dagger$ to Eqs. (5) gives, upon using Eqs. (12),

$$\dot{\boldsymbol{y}}_1 = \mathsf{M}_1^{(0)} \cdot \boldsymbol{y}_1 + \mathsf{Q}_0 \cdot (\boldsymbol{N}_1 + \boldsymbol{N}_1'), \tag{C5a}$$

$$\dot{\boldsymbol{y}}_2 = \mathsf{M}_2^{(0)} \cdot \boldsymbol{y}_2 + \boldsymbol{N}_2, \tag{C5b}$$

$$\dot{\boldsymbol{y}}_3 = \mathsf{M}_3^{(0)} \cdot \boldsymbol{y}_3 + \boldsymbol{N}_3. \tag{C5c}$$

Keeping only terms up to $O(|D, \overline{D}, \epsilon - \epsilon_c|^3)$ in Eq. (C4a) gives

$$\dot{D} = [\lambda_+^{[1]}(\boldsymbol{0}, \epsilon) + \boldsymbol{G}_0 \cdot \boldsymbol{y}_2]D + (\epsilon - \epsilon_c)\boldsymbol{p}_0^\dagger \cdot \mathsf{M}^{(1)} \cdot \boldsymbol{y}_1, \tag{C6}$$

where $\lambda^{[n]}$ denotes the expansion of λ through order n in $\epsilon - \epsilon_c$ and

$$\boldsymbol{G}_0 \doteq \frac{1}{2i}\boldsymbol{p}_0^\dagger \cdot \begin{pmatrix} \left(\dfrac{1}{\mathcal{D}_2} - \dfrac{1}{\mathcal{D}_1}\right) & 0 \\ \dfrac{1}{\mathcal{D}_2} q^{(2)} & -\dfrac{1}{\mathcal{D}_1} \end{pmatrix}. \tag{C7}$$

Analogously, keeping only terms up to $O(|D, \overline{D}, \epsilon - \epsilon_c|^2)$ in Eqs. (C5) gives

$$\dot{\boldsymbol{y}}_1 = \mathsf{M}_1^{(0)} \cdot \boldsymbol{y}_1 + \boldsymbol{K}_0 D(\epsilon - \epsilon_c), \tag{C8a}$$

$$\dot{\boldsymbol{y}}_2 = \mathsf{M}_2^{(0)} \cdot \boldsymbol{y}_2 + \mathrm{Re}(\boldsymbol{J}_0) D\overline{D}, \tag{C8b}$$

$$\dot{\boldsymbol{y}}_3 = \mathsf{M}_3^{(0)} \cdot \boldsymbol{y}_3. \tag{C8c}$$

Here $\boldsymbol{K}_0 \doteq i(\mathsf{I} - \boldsymbol{q}_0 \boldsymbol{p}_0^\dagger) \cdot \boldsymbol{e}_\perp$ [the specific form $\mathsf{M}_1^{(1)} \propto \boldsymbol{e}_\perp \boldsymbol{e}^T$ and the constraint $\boldsymbol{e}^T \cdot \boldsymbol{q}_0 = 1$ were used, where $\boldsymbol{e} \doteq (1,0)^T$ and $\boldsymbol{e}_\perp \doteq (0,1)^T$] and

$$\boldsymbol{J}_0 \doteq \frac{1}{2i}\left(0,\ \frac{1}{\mathcal{D}_1}\overline{q}^{(2)}\right)^T. \tag{C9}$$

We seek solutions for the $\boldsymbol{y}_i$ $(i = 1, 2, 3)$ curvatures in terms of the coordinates on the CM up to $O(|D, \overline{D}, \epsilon - \epsilon_c|^2)$ in the form

$$\begin{aligned} \boldsymbol{y}_i = {} & \boldsymbol{w}_i^{DD} D^2 + \boldsymbol{w}_i^{\overline{D}\,\overline{D}} \overline{D}\,\overline{D} + \boldsymbol{w}_i^{D\overline{D}} D\overline{D} + \boldsymbol{w}_i^{D\epsilon} D(\epsilon - \epsilon_c) \\ & + \boldsymbol{w}_i^{\overline{D}\epsilon} \overline{D}(\epsilon - \epsilon_c) + \boldsymbol{w}_i^{\epsilon\epsilon} (\epsilon - \epsilon_c)^2 + \cdots . \end{aligned} \tag{C10}$$

Time-differentiating $\boldsymbol{y}_i$ yields

$$\dot{\boldsymbol{y}}_i = \left(\frac{\partial \boldsymbol{y}_i}{\partial D}, \frac{\partial \boldsymbol{y}_i}{\partial \overline{D}}, \frac{\partial \boldsymbol{y}_i}{\partial \epsilon}\right) \cdot \begin{pmatrix} \lambda_+(\boldsymbol{0}, \epsilon_c) D + \cdots \\ \overline{\lambda}_+(\boldsymbol{0}, \epsilon_c) \overline{D} + \cdots \\ 0 \end{pmatrix}, \tag{C11}$$

where the various derivatives are obtained from the ansatz (C10). Upon replacing the $\dot{\boldsymbol{y}}_i$ in (C8) by the right-hand sides of Eqs. (C8), we obtain

$$-\mathsf{M}_2^{(0)} \cdot \boldsymbol{w}_2^{D\overline{D}} = \mathrm{Re}(\boldsymbol{J}_0), \tag{C12a}$$

$$[\lambda_+(\boldsymbol{0}, \epsilon_c)\mathsf{I} - \mathsf{M}_1^{(0)}] \cdot \boldsymbol{w}_1^{D\epsilon} = \boldsymbol{K}_0 D(\epsilon - \epsilon_c). \tag{C12b}$$

The solution of Eq. (C12a) is straightforwardly $\boldsymbol{w}_2^{D\overline{D}} = \eta_2^{-1}\,\mathrm{Re}(\boldsymbol{J}_0)$. Solution of Eq. (C12b) for $\boldsymbol{w}_1^{D\epsilon}$ is complicated by the fact that the matrix $\mathsf{M}_1^{(0)} - \lambda_+(\boldsymbol{0}, \epsilon_c)\mathsf{I}$ has a null eigenvector, so it is not invertible in the conventional sense. However, upon recalling the Fredholm

Alternative Theorem, we note that the solvability condition $\boldsymbol{p}_0^\dagger \cdot \boldsymbol{K}_0 = 0$ is satisfied and that also $\boldsymbol{y}$ is constrained by definition to satisfy $\boldsymbol{p}_0^\dagger \cdot \boldsymbol{y} = 0$. Because the two right eigenvectors $\boldsymbol{q}_0$ and $\boldsymbol{q}_0^-$ [where $\boldsymbol{q}_0^- \doteq \boldsymbol{q}_-(\epsilon = \epsilon_c)$] are linearly independent and span the 2D DW subspace, we conclude that $\boldsymbol{w}_1^{D\epsilon} \propto \boldsymbol{q}_0^-$ and find

$$\boldsymbol{w}_1^{D\epsilon} = i \left(\frac{\boldsymbol{p}_-^\dagger(\epsilon = \epsilon_c) \cdot \boldsymbol{e}_\perp}{\lambda_+(\boldsymbol{0}, \epsilon_c) - \lambda_-(\boldsymbol{0}, \epsilon_c)} \right) \boldsymbol{q}_0^- . \tag{C13}$$

The rest of the coefficients in Eq. (C10) vanish. Thus the curvature contributions to the CM up to $O(|D, \overline{D}, \epsilon - \epsilon_c|^2)$ are given by

$$\boldsymbol{y}_1 = \boldsymbol{w}_1^{D\epsilon} D(\epsilon - \epsilon_c), \quad \boldsymbol{y}_2 = \eta_2^{-1} \operatorname{Re}(\boldsymbol{J}_0) D\overline{D}. \tag{C14}$$

The dynamics on the CM can now be obtained by substituting Eqs. (C14) into Eq. (C6):

$$\dot{D} = \lambda_+^{[2]}(\boldsymbol{0}, \epsilon) D - B_0 |D|^2 D, \tag{C15}$$

with $B_0 \doteq -\boldsymbol{G}_0 \cdot \operatorname{Re}(\boldsymbol{J}_0)/\eta_2$. $\lambda_+^{[2]}$ is obviously an approximation to $\lambda_+^{[\infty]} = \lambda_+$, so we drop the superscript. This equation has the normal form of a system undergoing a Hopf bifurcation at $\operatorname{Re}[\lambda_+(\boldsymbol{0}, \epsilon)] = 0$. That is, with $I \doteq |D|^2$, the trivial fixed point of Eq. (C15), $I = 0$, is stable for $\operatorname{Re}[\lambda_+(\boldsymbol{0}, \epsilon)] \leq 0$ and unstable otherwise.

The other fixed point is the nontrivial one

$$I = \operatorname{Re}[\lambda_+(\boldsymbol{0}, \epsilon)] / \operatorname{Re}(B_0). \tag{C16}$$

This exists only above the linear threshold, i.e., $\operatorname{Re}[\lambda_+(\boldsymbol{0}, \epsilon)] > 0$. Since $\operatorname{Re}(\boldsymbol{G}_0) \cdot \operatorname{Re}(\boldsymbol{J}_0) < 0$, one has $B_0 > 0$; this solution is supercritical (Fig. 3).

3. Center manifold calculation for collisionless system

Here a system with no zonal damping is considered and a systematic CM calculation is performed to identify the collisionless dynamics. To proceed, we set $\eta_2, d_2 = 0$ and note that the zonal modes, now with $\mathsf{M}_2 = \mathsf{0}$, possess trivial linear dynamics.

Since in the collisionless system the whole plane $I = 0$ is invariant, it is necessary to consider the system in the vicinity of an arbitrary $\boldsymbol{z}_0$ (Sec. II B) on that plane. In particular, for any given ϵ_0, there is on that plane a curve $\operatorname{Re}[\lambda_+(\boldsymbol{z}, \epsilon_0)] = 0$ on which one of the DW branches is marginally stable (Fig. 1).

For any point on this marginal curve, the overall collection of modes in our low-order truncated model (5) in the eigenvalue plane is shown in Fig. 4. The shaded modes are the ones in the CM. Note that undamped zonal modes are part of the CM, which is now 4D. Collisionless dynamics requires expansion around the marginal curve.

a. Center manifold calculation: Around the marginal curve $\mathrm{Re}[\boldsymbol{\lambda_+(z,\epsilon_0)}] = \mathbf{0}$

Again, we follow the projection method A (Appendix B) for a suspended CM calculation. Here, for some given $\boldsymbol{z}_0$, we choose ϵ_0 so that it is on the marginal curve $\mathrm{Re}[\lambda_+(\boldsymbol{z}_0,\epsilon_0)] = 0$. Utilizing the P_0 and Q_0 operators, which serve to project the dynamics onto the center and stable eigenspaces respectively, we perform the decomposition (14). The eigenvectors $\boldsymbol{\mathcal{Q}}_0 = \boldsymbol{\mathcal{Q}}_1^+(\boldsymbol{z}_0,\epsilon_0)$ and $\boldsymbol{\mathcal{P}}_0 = \boldsymbol{\mathcal{P}}_1^+(\boldsymbol{z}_0,\epsilon_0)$ are calculated on the marginal curve.

Application of the P_0 operator to the original DW–SB set $\boldsymbol{\mathcal{U}} = (\boldsymbol{u}_1, \boldsymbol{u}_3)^T$ gives the new (complex) variable $D = \boldsymbol{\mathcal{P}}_0^\dagger \cdot \boldsymbol{\mathcal{U}}$. Since the zonal mode $\boldsymbol{u}_2$ is linearly undamped, the $\boldsymbol{z}$ variable also spans part of the center eigenspace. Hence the $\{D, \boldsymbol{z}\} \in T^c$ set of variables serves as coordinates on the 4D CM.

The equations for the D and $\boldsymbol{z}$ coordinates on the CM are obtained from the original equations (5):

$$\dot{D} = \boldsymbol{\mathcal{P}}_0 \cdot \mathcal{M}^{(0)} \cdot \boldsymbol{\mathcal{Q}}_0 D + \boldsymbol{\mathcal{P}}_0^\dagger \cdot (\boldsymbol{\mathcal{N}} + \boldsymbol{\mathcal{N}}'), \tag{C17a}$$

$$\dot{\boldsymbol{z}} = \boldsymbol{N}_2, \tag{C17b}$$

$$\dot{\epsilon} = 0. \tag{C17c}$$

Here $\boldsymbol{\mathcal{N}} \doteq (\boldsymbol{N}_1, \boldsymbol{N}_3)^T$, $\mathcal{M}^{(0)} \doteq \mathcal{M}(\epsilon = \epsilon_0)$, and $\boldsymbol{\mathcal{N}}' \doteq \mathcal{M}^{(1)} \cdot (\boldsymbol{\mathcal{Q}}_0 D + \boldsymbol{\mathcal{Y}})$ with $\mathcal{M}^{(1)} \doteq \partial_\epsilon \mathcal{M}_{|\epsilon_0}$. Again, note that in Method A only $\mathcal{M}^{(0)}$ is treated as linear, while the other terms of the $\mathcal{M}$ expansion are considered to be nonlinear.

The equation for the $\boldsymbol{\mathcal{Y}}$ variable, which describes the dynamics on the stable manifold, is obtained by applying the orthogonal projection operator $\mathsf{Q}_0 \doteq \mathsf{I} - \boldsymbol{\mathcal{Q}}_0 \boldsymbol{\mathcal{P}}_0^\dagger$ to Eq. (5):

$$\dot{\boldsymbol{\mathcal{Y}}} = \mathcal{M}^{(0)} \cdot \boldsymbol{\mathcal{Y}} + \mathsf{Q}_0 \cdot (\boldsymbol{\mathcal{N}} + \boldsymbol{\mathcal{N}}'). \tag{C18}$$

Since $\boldsymbol{\mathcal{Y}} \in T^s$ lies on the stable manifold, it describes the nonlinear curvature of the CM and thus is at least $O(|D, \overline{D}, \boldsymbol{z} - \boldsymbol{z}_0, \epsilon - \epsilon_0|^2)$.

Keeping only terms up to $O(|D, \overline{D}, \boldsymbol{z} - \boldsymbol{z}_0, \epsilon - \epsilon_0|^2)$ in the equations for the D and $\boldsymbol{z}$ variables gives

$$\dot{D} = \Gamma(\boldsymbol{z}, \boldsymbol{z}_0, \epsilon, \epsilon_0) D, \quad \dot{\boldsymbol{z}} = \boldsymbol{g}(\boldsymbol{z}_0, \epsilon_0) D \overline{D}. \tag{C19}$$

Here $\boldsymbol{z}$ is assumed to be in the vicinity of $\boldsymbol{z}_0$. The function $\Gamma(\boldsymbol{z}, \boldsymbol{z}_0, \epsilon, \epsilon_0)$ is an expansion of the complete eigenvalue $\lambda_+(\boldsymbol{z}, \epsilon)$ around $\boldsymbol{z}_0$ and ϵ_0:

$$\Gamma(\boldsymbol{z}, \boldsymbol{z}_0, \epsilon, \epsilon_0) = \lambda_+^{[1]}(\boldsymbol{z}_0, \epsilon) + \boldsymbol{\mathcal{P}}_0^\dagger \cdot \mathsf{G}(\boldsymbol{z}_0, \epsilon_0) \cdot (\boldsymbol{z} - \boldsymbol{z}_0). \tag{C20}$$

The functions $\mathsf{G}(\boldsymbol{z}_0, \epsilon_0)$ and $\boldsymbol{g}(\boldsymbol{z}_0, \epsilon_0)$ are given by

$$\mathsf{G}(\boldsymbol{z}_0, \epsilon_0) = \frac{1}{2i} \begin{pmatrix} \left(\frac{1}{\mathcal{D}_1} - \frac{1}{\mathcal{D}_2}\right) + \left(\frac{1}{\mathcal{D}_3} - \frac{1}{\mathcal{D}_2}\right) q^{(3)} & 0 \\ -\frac{1}{\mathcal{D}_2} q^{(2)} + \frac{1}{\mathcal{D}_2} q^{(4)} & \frac{1}{\mathcal{D}_1} - \frac{1}{\mathcal{D}_3} q^{(3)} \\ -\left(\frac{1}{\mathcal{D}_1} - \frac{1}{\mathcal{D}_2}\right) & 0 \\ \frac{1}{\mathcal{D}_2} q^{(3)} & -\frac{1}{\mathcal{D}_1} \end{pmatrix}, \tag{C21a}$$

$$\boldsymbol{g}(\boldsymbol{z}_0, \epsilon_0) = \mathrm{Im} \begin{pmatrix} \left(\frac{1}{\mathcal{D}_1} - \frac{1}{\mathcal{D}_3}\right) \overline{q}^{(3)} \\ \frac{1}{\mathcal{D}_1}(\overline{q}^{(4)} - q^{(3)} \overline{q}^{(2)}) + \frac{1}{\mathcal{D}_3} q^{(2)} \overline{q}^{(3)} \end{pmatrix}. \tag{C21b}$$

Here we have used the notation $\boldsymbol{\mathcal{Q}}_0 = (1, q^{(2)}, q^{(3)}, q^{(4)})^T$ and $\boldsymbol{\mathcal{P}}_0 = (p^{(1)}, p^{(2)}, p^{(3)}, p^{(4)})^T$.

Keeping terms up to only $O(|D, \overline{D}, \boldsymbol{z} - \boldsymbol{z}_0, \epsilon - \epsilon_0|^2)$ in Eq. (C18) for $\boldsymbol{\mathcal{Y}}$ gives

$$\dot{\boldsymbol{\mathcal{Y}}} = \mathcal{M}^{(0)} \cdot \boldsymbol{\mathcal{Y}} + \mathsf{H} \cdot (\boldsymbol{z} - \boldsymbol{z}_0) D + \boldsymbol{h} D (\epsilon - \epsilon_0), \tag{C22}$$

where $\boldsymbol{h} \doteq i[(0,1,0,0)^T - \overline{p}^{(2)} \boldsymbol{\mathcal{Q}}_0]$ and the 4×2 matrix H is given by $\mathsf{H} \doteq (\mathsf{I} - \boldsymbol{\mathcal{Q}}_0 \boldsymbol{\mathcal{P}}_0^\dagger) \cdot \mathsf{G}$.

At some $\epsilon = \epsilon_*$, there is a $\boldsymbol{z}_{\mathcal{F}}^* = \boldsymbol{z}_{\mathcal{F}}(\epsilon_*)$ such that

$$\mathrm{Re}[\Gamma(\boldsymbol{z}_{\mathcal{F}}^*, \boldsymbol{z}_{\mathcal{F}}^*, \epsilon_*, \epsilon_*)] = 0, \tag{C23a}$$

$$\boldsymbol{g}(\boldsymbol{z}_{\mathcal{F}}^*, \epsilon_*) = \boldsymbol{0}. \tag{C23b}$$

To understand the $\boldsymbol{z}$ dynamics in the vicinity of $\epsilon = \epsilon_*$ and $\boldsymbol{z} = \boldsymbol{z}_{\mathcal{F}}^*$, it is necessary to expand Eq. (C17b) up to $O(|D, \overline{D}, \boldsymbol{z} - \boldsymbol{z}_{\mathcal{F}}^*, \epsilon - \epsilon_*|^3)$; we obtain

$$\dot{D} = \Gamma(\boldsymbol{z}, \boldsymbol{z}_{\mathcal{F}}^*, \epsilon, \epsilon_*) D, \quad \dot{\boldsymbol{z}} = 2\,\mathrm{Re}[\mathcal{A}(\boldsymbol{z}_{\mathcal{F}}^*) \cdot \boldsymbol{\mathcal{Y}} \overline{D}], \tag{C24}$$

with the 4×2 matrix $\mathcal{A}(\boldsymbol{z}^*_{\mathcal{F}}) \doteq (2i)^{-1}(\mathsf{A}_1(\boldsymbol{z}^*_{\mathcal{F}}), \mathsf{A}_3(\boldsymbol{z}^*_{\mathcal{F}}))$,

$$\mathsf{A}_1(\boldsymbol{z}_0) \doteq \begin{pmatrix} \left(\frac{1}{\mathcal{D}_1} - \frac{1}{\mathcal{D}_3}\right)\overline{q}^{(3)} & 0 \\ \frac{1}{\mathcal{D}_1}(\overline{q}^{(4)} - \overline{q}^{(2)}) & -\frac{1}{\mathcal{D}_3}\overline{q}^{(3)} + \frac{1}{\mathcal{D}_1} \end{pmatrix}, \tag{C25a}$$

$$\mathsf{A}_3(\boldsymbol{z}_0) \doteq \begin{pmatrix} -\left(\frac{1}{\mathcal{D}_1} - \frac{1}{\mathcal{D}_3}\right) & 0 \\ \frac{1}{\mathcal{D}_3}\overline{q}^{(2)} & -\frac{1}{\mathcal{D}_1} \end{pmatrix}. \tag{C25b}$$

We seek solutions of Eq. (C22) for the $\boldsymbol{\mathcal{Y}}$ curvature in terms of coordinates on the CM up to $O(|D, \overline{D}, \boldsymbol{z} - \boldsymbol{z}^*_{\mathcal{F}}, \epsilon - \epsilon_*|^2)$ in the form

$$\boldsymbol{\mathcal{Y}} = \mathcal{W}^{\boldsymbol{z}D}(\boldsymbol{z}^*_{\mathcal{F}}) \cdot (\boldsymbol{z} - \boldsymbol{z}^*_{\mathcal{F}})D + \boldsymbol{w}^{D\epsilon}(\boldsymbol{z}^*_{\mathcal{F}})D(\epsilon - \epsilon_*) + \cdots. \tag{C26}$$

Various terms representing other quadratic combinations of the set $\{D, \overline{D}, \boldsymbol{z} - \boldsymbol{z}^*_{\mathcal{F}}, \epsilon - \epsilon_*\}$ are not written explicitly in Eq. (C26); it can be easily shown that those have zero coefficients.

Time-differentiating Eq. (C26) yields

$$\dot{\boldsymbol{\mathcal{Y}}} = \left(\frac{\partial \boldsymbol{\mathcal{Y}}}{\partial D}, \frac{\partial \boldsymbol{\mathcal{Y}}}{\partial \overline{D}}, \frac{\partial \boldsymbol{\mathcal{Y}}}{\partial \boldsymbol{z}}, \frac{\partial \boldsymbol{\mathcal{Y}}}{\partial \epsilon}\right) \cdot \begin{pmatrix} \lambda_+(\boldsymbol{z}^*_{\mathcal{F}}, \epsilon_*)D + \cdots \\ \overline{\lambda}_+(\boldsymbol{z}^*_{\mathcal{F}}, \epsilon_*)\overline{D} + \cdots \\ \mathbf{0} \\ 0 \end{pmatrix}, \tag{C27}$$

where the various derivatives are estimated according to Eq. (C26). Upon equating the representations (C27) and (C22), we obtain

$$[\lambda_+(\boldsymbol{z}^*_{\mathcal{F}}, \epsilon_*)\mathsf{I} - \mathcal{M}^{(0)}] \cdot \mathcal{W}^{\boldsymbol{z}D}(\boldsymbol{z}^*_{\mathcal{F}}) = \mathsf{H}, \tag{C28a}$$

$$[\lambda_+(\boldsymbol{z}^*_{\mathcal{F}}, \epsilon_*)\mathsf{I} - \mathcal{M}^{(0)}] \cdot \boldsymbol{w}^{D\epsilon}(\boldsymbol{z}^*_{\mathcal{F}}) = \boldsymbol{h}. \tag{C28b}$$

To obtain $\mathcal{W}^{\boldsymbol{z}D}(\boldsymbol{z}^*_{\mathcal{F}})$ and $\boldsymbol{w}^{D\epsilon}(\boldsymbol{z}^*_{\mathcal{F}})$, we note that the matrix $\lambda_+(\boldsymbol{z}^*_{\mathcal{F}}, \epsilon_*)\mathsf{I} - \mathcal{M}^{(0)}$ is not invertible; upon using the Fredholm Alternative Theorem, we find that $\boldsymbol{w}^{D\epsilon}(\boldsymbol{z}^*_{\mathcal{F}})$ is a linear combination of the eigenvectors that span the stable eigenspace:

$$\begin{aligned} \boldsymbol{w}^{D\epsilon}(\boldsymbol{z}^*_{\mathcal{F}}) = [&\alpha_1^- \boldsymbol{\mathcal{Q}}_1^-(\boldsymbol{z}_0, \epsilon_0) \\ &+\alpha_3^+ \boldsymbol{\mathcal{Q}}_3^+(\boldsymbol{z}_0, \epsilon_0) + \alpha_3^- \boldsymbol{\mathcal{Q}}_3^-(\boldsymbol{z}_0, \epsilon_0)]\Big|_{\substack{\boldsymbol{z}_0 = \boldsymbol{z}^*_{\mathcal{F}} \\ \epsilon_0 = \epsilon_*}}, \end{aligned} \tag{C29}$$

where the coefficients α_1^-, α_3^+, and α_3^- are to be found from Eq. (C28b). We do not write them explicitly because the approximation in Sec. C 3 b is much simpler.

The coefficient $\mathcal{W}^{zD}(\boldsymbol{z}^*_{\mathcal{F}})$ is a 4×2 matrix:

$$\mathcal{W}^{zD}(\boldsymbol{z}^*_{\mathcal{F}}) = \sum_i \boldsymbol{\mathcal{Q}}_i(\boldsymbol{z}_0,\epsilon_0)\cdot(\beta_i,\gamma_i)|_{\substack{\boldsymbol{z}_0=\boldsymbol{z}^*_{\mathcal{F}}\\ \epsilon_0=\epsilon_*}}\ , \tag{C30}$$

where the summation parameter i stands for the following combinations of superscripts and subscripts: $(-,1)$, $(+,3)$, and $(-,3)$. The various coefficients are to be found from Eq. (C28a).

Thus the curvature contribution to the CM up to $O(|D,\overline{D},\boldsymbol{z}-\boldsymbol{z}^*_{\mathcal{F}},\epsilon-\epsilon_*|^2)$ is given by

$$\boldsymbol{\mathcal{Y}} = \mathcal{W}^{zD}(\boldsymbol{z}^*_{\mathcal{F}})\cdot(\boldsymbol{z}-\boldsymbol{z}^*_{\mathcal{F}})D + \boldsymbol{w}^{D\epsilon}(\boldsymbol{z}^*_{\mathcal{F}})D(\epsilon-\epsilon_*). \tag{C31}$$

Substituting this expression into equation (C24) yields

$$\dot{D} \;=\; \Gamma(\boldsymbol{z},\boldsymbol{z}^*_{\mathcal{F}},\epsilon,\epsilon_*)D, \tag{C32a}$$

$$\begin{aligned}\dot{\boldsymbol{z}} \;=\;& 2\,\mathrm{Re}[-(\epsilon-\epsilon_*)\boldsymbol{b}(\boldsymbol{z}^*_{\mathcal{F}})\\ &+\mathsf{A}(\boldsymbol{z}^*_{\mathcal{F}})\cdot(\boldsymbol{z}-\boldsymbol{z}^*_{\mathcal{F}})]D\overline{D},\end{aligned} \tag{C32b}$$

with $\mathsf{A}(\boldsymbol{z}^*_{\mathcal{F}}) \doteq \mathcal{A}(\boldsymbol{z}^*_{\mathcal{F}})\cdot\mathcal{W}^{zD}(\boldsymbol{z}^*_{\mathcal{F}})$ and $\boldsymbol{b}(\boldsymbol{z}^*_{\mathcal{F}}) \doteq \mathcal{A}(\boldsymbol{z}^*_{\mathcal{F}})\cdot\boldsymbol{w}^{D\epsilon}(\boldsymbol{z}^*_{\mathcal{F}})$.

b. Local analysis around $\boldsymbol{z_0}=\boldsymbol{0}$ and $\boldsymbol{\epsilon}=\boldsymbol{0}$.

In the general case, the expressions (C23) for the location of $\mathcal{F}$ and its nearby dynamics [Eqs. (C19) and (C24)] are quite tedious to analyze. However, in the case when ϵ_* is sufficiently close to zero [and upon noting from Eq. (C23) that $\mathcal{F}$ crosses the origin at $\epsilon=0$], those expressions may be simplified by expansion around $\boldsymbol{z}_0=\boldsymbol{0}$ and $\epsilon=0$.

Linear quantities are obtained from Eqs. (6), (C2), and (C3) at $\boldsymbol{z}_0=\boldsymbol{0}$ and expanded around $\epsilon=0$:

$$\lambda_+(\boldsymbol{0},\epsilon) \;\approx\; -i(\Omega_1-i\eta_1)+\frac{b_1\epsilon}{\nu_1-i\Omega_1}, \tag{C33a}$$

$$\lambda_-(\boldsymbol{0},\epsilon) \;=\; \lambda_1^-(\boldsymbol{0},\epsilon)\approx -(\nu_1+\eta_1), \tag{C33b}$$

$\boldsymbol{\mathcal{Q}}_1^+(\boldsymbol{0},0)=(\boldsymbol{q}_0,\boldsymbol{0})^T$, $\boldsymbol{\mathcal{P}}_1^+(\boldsymbol{0},0)=(\boldsymbol{p}_0,\boldsymbol{0})^T$, and $\boldsymbol{\mathcal{Q}}_1^-(\boldsymbol{0},0)=(\boldsymbol{q}_0^-,\boldsymbol{0})^T$, with $\boldsymbol{q}_0\doteq(1,0)^T$,

$$\boldsymbol{q}_0^- \;\doteq\; \begin{pmatrix}1\\ (\nu_1-i\Omega_1)/(ib_1)\end{pmatrix}, \tag{C34a}$$

$$\boldsymbol{p}_0 \;\doteq\; \begin{pmatrix}1\\ ib_1/(\nu_1+i\Omega_1)\end{pmatrix}. \tag{C34b}$$

Upon substituting these quantities into Eqs. (C32), we obtain

$$\dot{D} = \Gamma(\boldsymbol{z}, \epsilon) D, \quad \dot{\boldsymbol{z}} = (-\boldsymbol{a}\epsilon + \mathsf{A} \cdot \boldsymbol{z})|D|^2 \tag{C35}$$

with $\boldsymbol{a} \doteq 2\,\mathrm{Re}[\boldsymbol{b}(\boldsymbol{0})] = 2\,\mathrm{Re}[\mathcal{A}(\boldsymbol{0}) \cdot \boldsymbol{w}^{D\epsilon}(\boldsymbol{0})]$, $\mathsf{A} \doteq 2\,\mathrm{Re}[\mathsf{A}(\boldsymbol{0})] = 2\,\mathrm{Re}[\mathcal{A}(\boldsymbol{0}) \cdot \mathcal{W}^{\boldsymbol{z}D}(\boldsymbol{0})]$, and

$$\begin{aligned} \Gamma(\boldsymbol{z}, \epsilon) &\doteq \Gamma(\boldsymbol{z}, \boldsymbol{0}, \epsilon, 0) \\ &= \lambda_+(\boldsymbol{0}, \epsilon) + \boldsymbol{p}_0^\dagger \cdot \mathsf{G}(\boldsymbol{0}, 0) \cdot \boldsymbol{z}, \end{aligned} \tag{C36a}$$

$$\mathsf{G}(\boldsymbol{0}, 0) \doteq \frac{1}{2i} \begin{pmatrix} \left(\frac{1}{\mathcal{D}_1} - \frac{1}{\mathcal{D}_2}\right) & 0 \\ 0 & \frac{1}{\mathcal{D}_1} \\ -\left(\frac{1}{\mathcal{D}_1} - \frac{1}{\mathcal{D}_2}\right) & 0 \\ 0 & -\frac{1}{\mathcal{D}_1} \end{pmatrix}. \tag{C36b}$$

The matrix $\mathcal{A}(\boldsymbol{0}) \doteq (2i)^{-1}(\mathsf{A}_1(\boldsymbol{0}), \mathsf{A}_3(\boldsymbol{0}))$ is given by

$$\mathsf{A}_1(\boldsymbol{0}) \doteq \begin{pmatrix} 0 & 0 \\ 0 & \frac{1}{\mathcal{D}_1} \end{pmatrix}, \tag{C37a}$$

$$\mathsf{A}_3(\boldsymbol{0}) \doteq \begin{pmatrix} -\left(\frac{1}{\mathcal{D}_1} - \frac{1}{\mathcal{D}_3}\right) & 0 \\ 0 & -\frac{1}{\mathcal{D}_1} \end{pmatrix}. \tag{C37b}$$

Since local analysis performs an expansion around $\boldsymbol{z}_0$ and $\epsilon = 0$, the last two elements of the various eigenvectors vanish, meaning that the CM variable $D = \boldsymbol{\mathcal{P}}_0^\dagger \cdot \boldsymbol{\mathcal{U}} = \boldsymbol{p}_0^\dagger \cdot \boldsymbol{u}_1 \in T^c$ is a combination of the DW components only and is decoupled from the SB dynamics, which contribute only to the stable eigenspace. Hence Eqs. (C26) for the nonlinear curvature in the local approximation transform to

$$\boldsymbol{y}_{1,3} = \mathcal{W}_{1,3}^{\boldsymbol{z}D}(\boldsymbol{0}) \cdot \boldsymbol{z} D + \boldsymbol{w}_{1,3}^{D\epsilon}(\boldsymbol{0}) D\epsilon, \tag{C38}$$

and the equations for the coefficients (C28) are given by $\mathcal{W}^{\boldsymbol{z}D}(\boldsymbol{0}) \doteq (\mathcal{W}_1^{\boldsymbol{z}D}(\boldsymbol{0}), \mathcal{W}_3^{\boldsymbol{z}D}(\boldsymbol{0}))^T$ and $\boldsymbol{w}^{D\epsilon}(\boldsymbol{0}) \doteq (\boldsymbol{w}_1^{D\epsilon}(\boldsymbol{0}), \boldsymbol{w}_3^{D\epsilon}(\boldsymbol{0}))^T$, with

$$[\lambda_+(\boldsymbol{0}, 0)\mathsf{I} - \mathsf{M}_1^{(0)}] \cdot \mathcal{W}_1^{\boldsymbol{z}D}(\boldsymbol{0}) = \mathsf{H}_1, \tag{C39a}$$

$$[\lambda_+(\boldsymbol{0}, 0)\mathsf{I} - \mathsf{M}_3^{(0)}] \cdot \mathcal{W}_3^{\boldsymbol{z}D}(\boldsymbol{0}) = \mathsf{H}_3, \tag{C39b}$$

$$[\lambda_+(\boldsymbol{0}, 0)\mathsf{I} - \mathsf{M}_1^{(0)}] \cdot \boldsymbol{w}_1^{D\epsilon}(\boldsymbol{0}) = \boldsymbol{h}_1, \tag{C39c}$$

$$[\lambda_+(\boldsymbol{0}, 0)\mathsf{I} - \mathsf{M}_3^{(0)}] \cdot \boldsymbol{w}_3^{D\epsilon}(\boldsymbol{0}) = \boldsymbol{0}, \tag{C39d}$$

$\boldsymbol{h}_1 \doteq i(-\overline{p}^{(2)}, 1)^T$,

$$\mathsf{H}_1 \doteq \frac{1}{2i\mathcal{D}_1}\begin{pmatrix} 0 & -\overline{p}^{(2)} \\ 0 & 1 \end{pmatrix}, \tag{C40a}$$

$$\mathsf{H}_3 \doteq -\frac{1}{2i\mathcal{D}_1\mathcal{D}_2}\begin{pmatrix} \mathcal{D}_2 - \mathcal{D}_1 & 0 \\ 0 & \mathcal{D}_2 \end{pmatrix}. \tag{C40b}$$

The solution of the system (C39) is

$$\mathcal{W}_1^{\boldsymbol{z}D}(\mathbf{0}) = \boldsymbol{q}_0^- \cdot (\alpha, \beta), \tag{C41a}$$

$$\mathcal{W}_3^{\boldsymbol{z}D}(\mathbf{0}) = [\lambda_+(\mathbf{0}, 0)\mathsf{I} - \mathsf{M}_3^{(0)}]^{-1} \cdot \mathsf{H}_3, \tag{C41b}$$

$$\boldsymbol{w}_1^{D\epsilon}(\mathbf{0}) = -i\frac{\overline{p}^{(2)}}{\lambda_+(\mathbf{0}, 0) - \lambda_-(\mathbf{0}, 0)}\boldsymbol{q}_0^-, \tag{C41c}$$

$$\boldsymbol{w}_3^{D\epsilon}(\mathbf{0}) = \mathbf{0}, \tag{C41d}$$

where $\alpha = 0$ and

$$\beta = -\frac{1}{\mathcal{D}_1}\frac{\overline{p}^{(2)}}{\lambda_+(\mathbf{0}, \epsilon_0) - \lambda_-(\mathbf{0}, \epsilon_0)}. \tag{C42}$$

The vector $\boldsymbol{a}$ is easily calculated to give $\boldsymbol{a} = \big(0, -2\Omega_1/\mathcal{D}_1(\nu^2 + \Omega_1^2)\big)^T$.

It is easy to see that all elements of A are positive. Some nontrivial algebra shows that all of its eigenvalues are negative if the product of $|\det(\lambda_+(\mathbf{0}, 0)\mathsf{I} - \mathsf{M}_3^{(0)})|^2$ and the $(2, 2)$ element of $\mathrm{Re}[\mathsf{A}_1(\mathbf{0}) \cdot \mathsf{W}_1^{\boldsymbol{z}D}(\mathbf{0})]$ is positive, which is always satisfied.

APPENDIX D: MULTIPLE-SCALE ANALYSIS

We record here some details of the multiple-scale analysis of both collisional and collisionless systems. The procedure can be seen as a generalization of the CM calculation to include the essentially infinite number of long-wavelength modes that may be excited in a spatially extended system. When such modes are ignored (i.e., when weak envelope inhomogeneity is neglected), the method reduces to an alternate algorithm for constructing the dynamics on the CM. In that case, the distinction between the projection method described in Appendix B and the multiple-time-scale analysis presented here is that the former works directly with the "final" amplitude D on the center eigenspace, whereas the latter constructs D order by order in the slow time variations, determining different dynamical equations at each order that are eventually added together to obtain the final result. A

useful introduction to the envelope formalism for spatially extended systems is given by Manneville.[29]

1. Collisional system

We perform the time and space expansion (44) of various derivatives in the original equation (1) with the smallness parameter for time and space expansion δ chosen to be $\delta \doteq (\epsilon - \epsilon_c)^{1/2}$; this is a measure of the distance (in terms of the temperature gradient) from the linear threshold. We seek solutions of (1) of the form $\boldsymbol{u} = \delta \boldsymbol{u}^{(1)} + \delta^2 \boldsymbol{u}^{(2)} + \cdots$. Note that the assumption of a collisional system implies that $\eta_2 = O(1)$.

***O*($\boldsymbol{\delta}$):** Upon collecting the $O(\delta)$ terms, we have

$$(\partial_{t_0} \mathsf{I} - \widehat{\mathsf{M}}^{(0)}) \cdot \boldsymbol{u}^{(1)} = \boldsymbol{0}. \tag{D1}$$

Here $\widehat{\mathsf{M}}^{(n)}$ is the nth-order term in the δ expansion of the linear $\widehat{\mathsf{M}}$ operator.

If $\boldsymbol{u}^{(1)}$ is expanded in the complete set of right eigenvectors $\boldsymbol{q}_0 = \boldsymbol{q}_+(\epsilon = \epsilon_c)$, the components in the stable subspace (including both ZF's and SB's) will rapidly damp to zero on the t_0 time scale, leaving one with the marginally stable drift wave

$$\boldsymbol{u}^{(1)} = D^{(1)} \boldsymbol{q}_D + \text{c.c.}, \tag{D2}$$

where $\boldsymbol{q}_D \doteq \boldsymbol{q}_0 e^{ik_y y} \sin(k_x x) e^{-i\Omega_D t}$, and $D^{(1)} = D^{(1)}(T, \boldsymbol{R})$ is the modulation (envelope) function of the drift-wave carrier. Note that the various $\boldsymbol{u}$ modes used here differ from the ones used in previous sections by the explicit retention of harmonic dependences. For the DW frequency $\Omega_D = i\lambda_+(\boldsymbol{0}, \epsilon_c)$ and the eigenvectors $\boldsymbol{p}_0$ and $\boldsymbol{q}_0$, refer to Appendix C 2.

***O*($\boldsymbol{\delta^2}$):** At second order, we have

$$\begin{aligned}(\partial_{t_0} \mathsf{I} - \widehat{\mathsf{M}}^{(0)}) \cdot \boldsymbol{u}^{(2)} \;=\; & -\partial_{T_1} \boldsymbol{u}^{(1)} + \widehat{\mathsf{M}}^{(1)} \cdot \boldsymbol{u}^{(1)} \\ & + \widehat{\boldsymbol{N}}^{(0)}(\boldsymbol{u}^{(1)}, \boldsymbol{u}^{(1)}).\end{aligned} \tag{D3}$$

In order that this equation be solvable, the right-hand side must be orthogonal (in the sense of the usual complex-valued scalar product) to any left null eigenvectors of the lowest-order operator $\partial_{t_0} \mathsf{I} - \widehat{\mathsf{M}}^{(0)}$. Those have the form of $\boldsymbol{q}_D$ except that $\boldsymbol{q}_0$ is replaced by $\boldsymbol{p}_0$. Upon applying the $\widehat{\mathsf{P}}_0$ operator

$$\widehat{\mathsf{P}}_0 = \lim_{T \to \infty} \int_0^T \frac{dt}{T} \int_0^{L_x} \frac{dx}{L_x} \int_0^{L_y} \frac{dy}{L_y} \, \boldsymbol{q}_0 \, \boldsymbol{p}_D^\dagger \tag{D4}$$

(i.e., projecting onto the center eigenspace) to Eq. (D3), we obtain a constraint on the $D^{(1)}$ amplitude involving the T_1 time scale:

$$\partial_{T_1} D^{(1)} = \boldsymbol{p}_0^\dagger \cdot \mathsf{M}_1^{(1)} \cdot \boldsymbol{q}_0 D^{(1)} = -\boldsymbol{V}_D^{(1)} \cdot \partial_{\boldsymbol{R}_1} D^{(1)}. \tag{D5}$$

It can be shown that $\boldsymbol{V}_D^{(1)}$ is just the DW group velocity: $\boldsymbol{V}_D^{(1)} = \partial\Omega_D/\partial\boldsymbol{k}$.

Having satisfied the solvability equation (D5), we may write Eq. (D3) in the form

$$\begin{aligned}(\partial_{t_0}\mathsf{I} - \widehat{\mathsf{M}}^{(0)}) \cdot \boldsymbol{u}^{(2)} \;=\; & \widehat{\mathsf{Q}}_0 \cdot \widehat{\mathsf{M}}^{(1)} \cdot \boldsymbol{u}^{(1)} \\ & + \widehat{\boldsymbol{N}}^{(0)}(\boldsymbol{u}^{(1)}, \boldsymbol{u}^{(1)}).\end{aligned} \tag{D6}$$

One can construct the most general solution of Eq. (D6) as a sum of homogeneous ($D^{(2)}\boldsymbol{q}_D$) and inhomogeneous ($\boldsymbol{y}^{(2)}$) parts:

$$\boldsymbol{u}^{(2)} = D^{(2)}\boldsymbol{q}_D + \boldsymbol{y}^{(2)} + \text{c.c.}, \tag{D7}$$

where $\widehat{\mathsf{P}}_0 \cdot \boldsymbol{y}^{(2)} = 0$ and $\boldsymbol{y}^{(2)} = \boldsymbol{y}_1^{(2)} + \boldsymbol{y}_2^{(2)}$. [As usual, the subscripts denote wave-number dependence; the $(2,0)$ zonal harmonic arises from the nonlinearity.] Here the methodology has led us again to the contravariant construction of the CM, with $\boldsymbol{y}^{(2)}$ describing the second-order curvature; cf. Eq. (B2).

The homogeneous amplitude $D^{(2)}$ is undetermined at this order.

The inhomogeneous solution $\boldsymbol{y}_2^{(2)} = \boldsymbol{w}_2 D^{(1)}\overline{D}^{(1)}\sin(2k_x x)$ describes the nonlinearly generated zonal modes with

$$\boldsymbol{w}_2 = -\left(\mathsf{M}_2^{(0)}\right)^{-1} \cdot \boldsymbol{N}_2^{(0)}(\boldsymbol{u}^{(1)}, \boldsymbol{u}^{(1)}). \tag{D8}$$

Here we introduced the notation

$$\begin{aligned}\boldsymbol{N}_{\mathcal{K}}^{(n)}(\boldsymbol{a},\boldsymbol{b}) = (1 - \tfrac{1}{2}\delta_{\boldsymbol{a},\boldsymbol{b}}) \int_0^T \frac{dt}{T} \int_0^{L_x} \frac{dx}{L_x} \int_0^{L_y} \frac{dy}{L_y} \\ \times \left[\widehat{\boldsymbol{N}}^{(n)}(\boldsymbol{a},\boldsymbol{b}) e^{im_{\mathcal{K}} k_y y} \sin(l_{\mathcal{K}} k_x x) e^{-i\Omega_{\mathcal{K}} t}\right]\end{aligned} \tag{D9}$$

for the nonlinear term(s) averaged over the harmonic dependence of mode $\mathcal{K} = (l_{\mathcal{K}}, m_{\mathcal{K}})$ with frequency $\Omega_{\mathcal{K}}$. Note that since $\eta_2 = O(1)$, $\mathsf{M}_2^{(0)} = -\eta_2 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ is a nontrivial, invertible, order-unity matrix. Thus $\boldsymbol{w}_2$ may be easily calculated to give Eq. (C12a).

The other inhomogeneous term can be written in the form

$$\boldsymbol{y}_1^{(2)} = S^{(1)} D^{(1)} \boldsymbol{q}_D^-. \tag{D10}$$

With $\boldsymbol{q}_D^- \doteq \boldsymbol{q}_0^- e^{ik_y y}\sin(k_x x)e^{-i\Omega_D t}$ and $\boldsymbol{q}_0^- = \boldsymbol{q}_-(\epsilon=\epsilon_c)$, $S^{(1)}$ can be calculated to give

$$S^{(1)} = \frac{\boldsymbol{p}_-^\dagger(\epsilon=\epsilon_c)\cdot \mathsf{M}_1^{(1)}\cdot \boldsymbol{q}_0}{\lambda_+(\mathbf{0},\epsilon_c)-\lambda_-(\mathbf{0},\epsilon_c)}. \tag{D11}$$

Both $\boldsymbol{y}_1^{(2)}$ and $\boldsymbol{y}_2^{(2)}$ will be used in the third-order constraint.

$\boldsymbol{O(\delta^3)}$**:** Upon collecting the $O(\delta^3)$ terms, we have

$$\begin{aligned}(\partial_{t_0}\mathsf{I}-\widehat{\mathsf{M}}^{(0)})\cdot\boldsymbol{u}^{(3)} &= -\partial_{T_1}\boldsymbol{u}^{(2)}-\partial_{T_2}\boldsymbol{u}^{(1)}\\ &+\widehat{\mathsf{M}}^{(1)}\cdot\boldsymbol{u}^{(2)}+\widehat{\mathsf{M}}^{(2)}\cdot\boldsymbol{u}^{(1)}\\ &+\widehat{\boldsymbol{N}}^{(0)}(\boldsymbol{u}^{(1)},\boldsymbol{u}^{(2)})+\widehat{\boldsymbol{N}}^{(1)}(\boldsymbol{u}^{(1)},\boldsymbol{u}^{(1)}).\end{aligned} \tag{D12}$$

Upon applying the $\widehat{\mathsf{P}}_0$ operator to Eq. (D12), we get a constraint on the $D^{(1)}$ amplitude involving the T_2 time scale:

$$\begin{aligned}\partial_{T_2}D^{(1)} \;+\; &(\partial_{T_1}+\boldsymbol{V}_D\cdot\partial_{\boldsymbol{R}_1})D^{(2)}\\ =\;&\boldsymbol{p}_0^\dagger\cdot\mathsf{M}_1^{(1)}\cdot\boldsymbol{q}_0^-\widehat{\mathsf{S}}^{(1)}D^{(1)}+\boldsymbol{p}_0^\dagger\cdot\mathsf{M}_1^{(2)}\cdot\boldsymbol{q}_0D^{(1)}\\ &+\boldsymbol{p}_0^\dagger\cdot\boldsymbol{N}_1^{(0)}(\boldsymbol{u}^{(1)},\boldsymbol{y}_2^{(2)}).\end{aligned} \tag{D13}$$

There is sufficient freedom to constrain $D^{(2)}$ to propagate with the group velocity; thus $D^{(2)}$ just provides a trivial correction to the basic amplitude. Upon introducing a new variable according to $\widetilde{D}=\delta D^{(1)}+\delta^2 D^{(2)}$ and adding the constraints (D5) and (D13), we obtain a constraint on the envelope $\widetilde{D}$:

$$\begin{aligned}\partial_{\xi_D}\widetilde{D} \;=\;&\boldsymbol{p}_0^\dagger\cdot\mathsf{M}_1^{(1)}\cdot\boldsymbol{q}_0^-\widehat{\mathsf{S}}^{(1)}\widetilde{D}+\boldsymbol{p}_0^\dagger\cdot\mathsf{M}_1^{(2)}\cdot\boldsymbol{q}_0\widetilde{D}\\ &+\boldsymbol{p}_0^\dagger\cdot\boldsymbol{N}_1^{(0)}(\boldsymbol{u}_1,\boldsymbol{u}_2),\end{aligned} \tag{D14}$$

with $\boldsymbol{u}_1 \doteq \widetilde{D}\boldsymbol{q}_0+\delta^2\boldsymbol{y}_1^{(2)}$, $\boldsymbol{u}_2\doteq\delta^2\boldsymbol{y}_2^{(2)}$, and $\partial_{\xi_D}\doteq\partial_T+\boldsymbol{V}_D\cdot\partial_{\boldsymbol{R}}$. Here the notation has been changed back to the T and $\boldsymbol{R}$ variables according to Eq. (44). The group velocity $\boldsymbol{V}_D\doteq\boldsymbol{V}_D^{(1)}+\delta\boldsymbol{V}_D^{(2)}$ includes a second-order correction; see Eq. (D16).

After computing the various coefficients in the above constraint, we have up to $O(\delta^3)$

$$\partial_{\xi_D}\widetilde{D}=\breve{\lambda}_+^{[2]}\widetilde{D}+\widehat{\Lambda}_{\boldsymbol{R}}^{(2)}\widetilde{D}-B_0|\widetilde{D}|^2\widetilde{D}. \tag{D15}$$

Here we used

$$\begin{aligned}\boldsymbol{p}_0^\dagger\cdot\mathsf{M}_1^{(1)}\cdot\boldsymbol{q}_0^-\widehat{\mathsf{S}}^{(1)} \;+\;&\boldsymbol{p}_0^\dagger\cdot\mathsf{M}_1^{(2)}\cdot\boldsymbol{q}_0\\ =\;&\breve{\lambda}_+^{[2]}+\boldsymbol{V}_D^{(2)}\cdot\partial_{\boldsymbol{R}}+\widehat{\Lambda}_{\boldsymbol{R}}^{(2)}.\end{aligned} \tag{D16}$$

$\breve{\lambda}_+^{[2]}$ is an approximation to $\breve{\lambda}_+^{[\infty]} = \lambda_+(\mathbf{0},\epsilon) - \lambda_+(\mathbf{0},\epsilon_c)$. Refer to Appendix C 1 for that quantity and the mode-coupling coefficient B_0. The linear dispersive nonuniform operator $\widehat{\Lambda}_{\boldsymbol{R}}^{(2)}$ is of second order in the long-scale gradient operator $\partial_{\boldsymbol{R}}$; we do not write its complicated explicit form here. It describes the form of the nonneutral surface close to the bifurcation point of the first linearly unstable mode, thus incorporating a continuous unstable band of modes into the formation of unstable envelope modulations.

2. Collisionless system ($\epsilon < \epsilon_*$)

Now we apply the expansion (44) of various derivatives to the original equation (1) with no ZF damping ($\eta_2 = 0$). We seek solutions of (1) of the form $\boldsymbol{u} = \delta \boldsymbol{u}^{(1)} + \delta^2 \boldsymbol{u}^{(2)} + \cdots$. The smallness parameter δ is now chosen to be $\delta = \epsilon - \epsilon_0$, with ϵ_0 found from $\mathrm{Re}[\lambda_+(\boldsymbol{z}_0,\epsilon_0)] = 0$ ($\boldsymbol{z}_0$ is some given point on $\boldsymbol{z}$ plane). This is a condition for the system to be close to the marginal curve (Sec. II B).

***O*($\boldsymbol{\delta}$):** Upon collecting the $O(\delta)$ terms, we have

$$(\partial_{t_0}\mathsf{I} - \mathcal{M}^{(0)}) \cdot \boldsymbol{\mathcal{U}}^{(1)} = \mathbf{0}, \tag{D17a}$$

$$(\partial_{t_0}\mathsf{I} - \mathsf{M}_2^{(0)}) \cdot \boldsymbol{u}_2^{(1)} = \partial_{t_0}\boldsymbol{u}_2^{(1)} = \mathbf{0}. \tag{D17b}$$

The quantity $\boldsymbol{\mathcal{U}}^{(1)}$ describes the marginally stable DW, while the linearly undamped ZF $\boldsymbol{u}_2^{(1)}$ is initially introduced into the system. The solution of Eq. (D17a) describes two types of DW's: $\boldsymbol{u}_1^{(1)}$, with harmonic dependence $e^{ik_y y}\sin(k_x x)e^{-i\Omega_D t}$; and $\boldsymbol{u}_{3,\Omega_S}^{(1)}$, with $e^{ik_y y}\sin(3k_x x)e^{-i\Omega_S t}$. Here the DW frequencies obey $-i\Omega_D = \lambda_+(\boldsymbol{z}_0,\epsilon_0) \equiv \lambda_1^+(\boldsymbol{z}_0,\epsilon_0)$ and $-i\Omega_S \equiv \lambda_3^+(\boldsymbol{z}_0,\epsilon_0)$. Therefore

$$\boldsymbol{\mathcal{U}}^{(1)} = (\boldsymbol{u}_1^{(1)}, \boldsymbol{u}_{3,\Omega_S}^{(1)})^T = D^{(1)}\boldsymbol{\mathcal{Q}}_D + \text{c.c.}, \tag{D18}$$

where $\boldsymbol{\mathcal{Q}}_D$ is the $\boldsymbol{\mathcal{Q}}_0$ vector with elements multiplied by the corresponding harmonic dependences. The solution of Eq. (D17b) is given by $\boldsymbol{u}_2^{(1)} = \boldsymbol{z}^{(1)}\sin(2k_x x)$. Thus the $O(\delta)$ solution of the original system (5) is

$$\boldsymbol{u}^{(1)} = \boldsymbol{u}_1^{(1)} + \boldsymbol{u}_2^{(1)} + \boldsymbol{u}_{3,\Omega_S}^{(1)}. \tag{D19}$$

The quantities $D^{(1)}(T,\boldsymbol{R})$ and $\boldsymbol{z}^{(1)}(T,\boldsymbol{R})$ are amplitude modulations on the large time and space scales. For the DW frequencies and the eigenvectors $\boldsymbol{\mathcal{P}}_0$ and $\boldsymbol{\mathcal{Q}}_0$, refer to Appendix II B.

The other possible DW-like modes in the system are also solutions of Eq. (D17a); however, being linearly stable, they will eventually be driven to zero on the t_0 time scale even if they are initially excited. It is important to note that the other possible ZF-like modes [other than $(2k_x, 0)$] in the system also give nontrivial solutions to (D17b). We constrain these to vanish in our simplified model; our minimal model with only one ZF mode is sufficient to understand the basics of the collisionless dynamics.

$O(\delta^2)$: Upon collecting the $O(\delta^2)$ terms, we have

$$(\partial_{t_0}\mathsf{I} - \mathcal{M}^{(0)}) \cdot \boldsymbol{\mathcal{U}}^{(2)} = -\partial_{T_1}\boldsymbol{\mathcal{U}}^{(1)} + \mathcal{M}^{(1)} \cdot \boldsymbol{\mathcal{U}}^{(1)} + \widehat{\boldsymbol{\mathcal{N}}}^{(0)}(\boldsymbol{u}^{(1)}, \boldsymbol{u}^{(1)}), \tag{D20a}$$

$$\partial_{t_0}\boldsymbol{u}_2^{(2)} = -\partial_{T_1}\boldsymbol{u}_2^{(1)} + \widehat{\boldsymbol{N}}^{(0)}(\boldsymbol{u}^{(1)}, \boldsymbol{u}^{(1)}), \tag{D20b}$$

with $\widehat{\boldsymbol{\mathcal{N}}}^{(0)} \doteq (\widehat{\boldsymbol{N}}^{(0)}, \widehat{\boldsymbol{N}}^{(0)})^T$.

Upon acting on Eq. (D20a) by the projection operator

$$\widehat{\mathsf{P}}_0 = \int_0^T \frac{dt}{T} \int_0^{L_x} \frac{dx}{L_x} \int_0^{L_y} \frac{dy}{L_y}\, \boldsymbol{\mathcal{Q}}_0 \boldsymbol{\mathcal{P}}_D^\dagger, \tag{D21}$$

we get a constraint on the $D^{(1)}$ amplitude on the T_1 time scale:

$$\partial_{T_1} D^{(1)} = \boldsymbol{\mathcal{P}}_0^\dagger \cdot \mathcal{M}^{(1)} \cdot \boldsymbol{\mathcal{Q}}_0 D^{(1)} + \boldsymbol{\mathcal{P}}_0^\dagger \cdot \left(\boldsymbol{N}_1^{(0)}(\boldsymbol{u}_1^{(1)}, \boldsymbol{u}_2^{(1)}), \boldsymbol{0}\right)^T, \tag{D22}$$

where the first term

$$\boldsymbol{\mathcal{P}}_0^\dagger \cdot \mathcal{M}^{(1)} \cdot \boldsymbol{\mathcal{Q}}_0 = -\boldsymbol{V}_D \cdot \partial_{\boldsymbol{R}_1} D^{(1)} + \breve{\lambda}_+^{[2]} D^{(1)} \tag{D23}$$

describes propagation of the DW's at the group velocity $\boldsymbol{V}_D$ and linear growth at the rate $\breve{\lambda}_+^{[2]} \approx \breve{\lambda}_+^{[\infty]} = \lambda_+(\boldsymbol{z}_0, \epsilon) - \lambda_+(\boldsymbol{z}_0, \epsilon_0)$. In the nonlinear term in Eq. (D22) we used the fact that $\boldsymbol{N}_{3,\Omega_S}^{(0)}(\boldsymbol{u}_1^{(1)}, \boldsymbol{u}_2^{(1)}) = \boldsymbol{0}$; the subscript emphasizes that the averaging in Eq. (D9) is done over frequency Ω_S (if a frequency is omitted from a subscript, it is assumed to be Ω_D).

The constraint on the $\boldsymbol{z}^{(1)}$ amplitude is easily obtained from Eq. (D20b):

$$\partial_{T_1}\boldsymbol{z}^{(1)} = \boldsymbol{N}_2^{(0)}(\boldsymbol{u}_1^{(1)}, \boldsymbol{u}_1^{(1)}). \tag{D24}$$

Upon introducing new variables according to $\widetilde{D} = \delta D^{(1)}$ and $\boldsymbol{z} = \delta \boldsymbol{z}^{(1)}$, then changing back to the T and $\boldsymbol{R}$ variables, we obtain from Eqs. (D22) and (D24)

$$\partial_{\xi_D}\widetilde{D} = \Gamma^{(1)}(\boldsymbol{z}, \boldsymbol{z}_0, \epsilon, \epsilon_0)\widetilde{D},\ \ \partial_T \boldsymbol{z} = \boldsymbol{g}(\boldsymbol{z}_0, \epsilon_0)|\widetilde{D}|^2. \tag{D25}$$

The functions $\Gamma^{(1)}(\boldsymbol{z},\boldsymbol{z}_0,\epsilon,\epsilon_0) \doteq \Gamma(\boldsymbol{z},\boldsymbol{z}_0,\epsilon,\epsilon_0) - \lambda_+(\boldsymbol{z}_0,\epsilon_0)$ and $\boldsymbol{g}(\boldsymbol{z}_0,\epsilon_0)$ result from the linear and nonlinear uniform terms in Eqs. (D22) and (D24). Refer to Appendix C 2 for these quantities.

With solvability satisfied by Eq. (D22), Eq. (D20a) transforms to an equation with the nontrivial solution

$$\boldsymbol{\mathcal{U}}^{(2)} = D^{(2)}\boldsymbol{\mathcal{Q}}_D + \boldsymbol{\mathcal{Y}}^{(2)}, \quad \boldsymbol{\mathcal{Y}}^{(2)} = \boldsymbol{\mathcal{Y}}_1^{(2)} + \boldsymbol{\mathcal{Y}}_3^{(2)}. \tag{D26}$$

For simplicity, we have dropped the second-order homogeneous solution; that can be treated the same way as in Sec. D 1. The modes described by $\boldsymbol{\mathcal{Y}}^{(2)}$ are linearly damped but have a nonzero amplitude due to nonlinear generation. $\boldsymbol{\mathcal{Y}}_1^{(2)}$ is a secondary DW:

$$\begin{aligned}\boldsymbol{\mathcal{Y}}_1^{(2)} &= (-i\Omega_D \mathsf{I} - \mathcal{M}^{(0)})^{\mathrm{INV}} \cdot \mathsf{H} \cdot \\ &\cdot \Big[\mathcal{M}^{(1)} \cdot D^{(1)}\boldsymbol{\mathcal{Q}}_D + \Big(\boldsymbol{N}_1^{(0)}(\boldsymbol{u}_1^{(1)},\boldsymbol{u}_2^{(1)})e^{-i\Omega_D t}, \mathbf{0}e^{-i\Omega_S t}\Big)^T \\ &\quad \times e^{ik_y y}\sin(k_x x)\Big] + \text{c.c.},\end{aligned} \tag{D27}$$

with $\mathsf{H} \doteq (\mathsf{I} - \boldsymbol{\mathcal{Q}}_0 \cdot \boldsymbol{\mathcal{P}}_0^{\dagger})$ and the notation INV referring to the generalized inverse calculated according to the Fredholm theorem. $\boldsymbol{\mathcal{Y}}_3^{(2)}$ is a SB mode:

$$\begin{aligned}\boldsymbol{\mathcal{Y}}_3^{(2)} \;=\; &(-i\Omega_D \mathsf{I} - \mathcal{M}^{(0)})^{\mathrm{INV}} \cdot \\ &\cdot \Big(\boldsymbol{N}_3^{(0)}(\boldsymbol{u}_1^{(1)},\boldsymbol{u}_2^{(1)})e^{-i\Omega_D t}, \mathbf{0}e^{-i\Omega_S t}\Big)^T \\ &\times e^{ik_y y}\sin(3k_x x) + \text{c.c.}\end{aligned} \tag{D28}$$

In Eqs. (D27) and (D28) we used $\boldsymbol{N}_{1,\Omega_S}^{(0)}(\boldsymbol{u}_1^{(1)},\boldsymbol{u}_2^{(1)}) = \mathbf{0}$ and $\boldsymbol{N}_{3,\Omega_S}^{(0)}(\boldsymbol{u}_1^{(1)},\boldsymbol{u}_2^{(1)}) = \mathbf{0}$.

The $O(\delta^2)$ solution to the original system (5) is

$$\boldsymbol{u}^{(2)} = \boldsymbol{u}_1^{(2)} + \boldsymbol{u}_{3,\Omega_S}^{(2)} + \boldsymbol{u}_2^{(2)} + \boldsymbol{y}_1^{(2)} + \boldsymbol{y}_3^{(2)}, \tag{D29}$$

where $D^{(2)}\boldsymbol{\mathcal{Q}}_D = (\boldsymbol{u}_1^{(2)}, \boldsymbol{u}_{3,\Omega_S}^{(2)})^T$ and $\boldsymbol{\mathcal{Y}}^{(2)} = (\boldsymbol{y}_1^{(2)} + \boldsymbol{y}_3^{(2)}, \mathbf{0})^T$.

3. Collisionless system ($\epsilon = \epsilon_*$)

Suppose that at some $\boldsymbol{z} = \boldsymbol{z}_{\mathcal{F}}^*$ and $\epsilon = \epsilon_*$ we have $\boldsymbol{N}_2^{(0)}(\boldsymbol{u}_1^{(1)},\boldsymbol{u}_1^{(1)}) = \mathbf{0}$; that is, the $\boldsymbol{z}$ dynamics on the T_1 scale (D24) are trivial:

$$\partial_{T_1}\boldsymbol{z}^{(1)} = \mathbf{0}. \tag{D30}$$

$O(\delta^3)$: Upon collecting the $O(\delta^3)$ terms, we have

$$\begin{aligned}(\partial_{t_0}\mathsf{I}-\mathcal{M}^{(0)})\cdot\boldsymbol{\mathcal{U}}^{(3)} &= -\partial_{T_1}\boldsymbol{\mathcal{U}}^{(2)}-\partial_{T_2}\boldsymbol{\mathcal{U}}^{(1)}\\ &+\mathcal{M}^{(1)}\cdot\boldsymbol{\mathcal{U}}^{(2)}+\mathcal{M}^{(2)}\cdot\boldsymbol{\mathcal{U}}^{(1)}\\ &+\widehat{\boldsymbol{\mathcal{N}}}^{(0)}(\boldsymbol{u}^{(1)},\boldsymbol{u}^{(2)})+\widehat{\boldsymbol{\mathcal{N}}}^{(1)}(\boldsymbol{u}^{(1)},\boldsymbol{u}^{(1)}).\end{aligned}\tag{D31}$$

Upon acting on Eq. (D31) by the $\widehat{\mathsf{P}}_0$ operator, we obtain the constraint on the $D^{(1)}$ amplitude involving the T_2 time scale. Similarly, upon averaging over the ZF harmonic dependence, we obtain a constraint on the $\boldsymbol{z}^{(1)}$ amplitude involving the T_2 time scale:

$$\partial_{T_2}\boldsymbol{z}^{(1)}=\boldsymbol{N}_2^{(0)}(\boldsymbol{u}_1^{(1)},\boldsymbol{y}_1^{(2)}+\boldsymbol{y}_3^{(2)}).\tag{D32}$$

The nontrivial solution of (D31) includes both corrections to already existing modes found above as well as various new SB and ZF various modes. We do not write down this solution since it is not important for the dynamics of the time- and space-scales in which we are interested.

Upon again introducing new variables according to

$$\widetilde{D}=\delta D^{(1)}+\delta^2 D^{(2)},\quad \boldsymbol{u}_{1,2}=\delta\boldsymbol{u}_{1,2}^{(1)}+\delta^2\boldsymbol{u}_{1,2}^{(2)},\quad \boldsymbol{z}=\delta\boldsymbol{z}^{(1)}+\delta^2\boldsymbol{z}^{(2)},\tag{D33a}$$

$$\boldsymbol{\mathcal{Y}}_{1,3}=(\boldsymbol{y}_{1,3},\boldsymbol{0})^T,\quad \boldsymbol{y}_{1,3}=\delta^2\boldsymbol{y}_{1,3}^{(2)}\tag{D33b}$$

and changing back to the T and $\boldsymbol{R}$ variables, we sum up the constraints on the $D^{(1)}$ amplitude arising from the various time scales to obtain

$$\begin{aligned}\partial_{\xi_D}\widetilde{D} &= \Gamma^{(1)}(\boldsymbol{z},\boldsymbol{z}_0,\epsilon,\epsilon_0)\widetilde{D}\\ &+\boldsymbol{\mathcal{P}}_0^\dagger\cdot\mathcal{M}^{(1)}\cdot\boldsymbol{\mathcal{Y}}_1+\boldsymbol{\mathcal{P}}_0^\dagger\cdot\mathcal{M}^{(2)}\cdot\boldsymbol{\mathcal{Q}}_0\widetilde{D}\\ &+\boldsymbol{\mathcal{P}}_0^\dagger\cdot\left(\boldsymbol{N}_1^{(0)}(\boldsymbol{u}_2,\boldsymbol{y}_1+\boldsymbol{y}_3),\boldsymbol{0}\right)^T\\ &+\boldsymbol{\mathcal{P}}_0^\dagger\cdot\left(\boldsymbol{N}_1^{(1)}(\boldsymbol{u}_1,\boldsymbol{u}_2),\boldsymbol{0}\right)^T,\end{aligned}\tag{D34a}$$

$$\partial_T\boldsymbol{z} = \boldsymbol{N}_2^{(0)}(\boldsymbol{u}_1,\boldsymbol{y}_1+\boldsymbol{y}_3).\tag{D34b}$$

4. Collisionless system (local analysis around $\boldsymbol{z_0=0}$ and $\boldsymbol{\epsilon=0}$)

In the special case of $\boldsymbol{z}_0=\boldsymbol{0}$ and $\epsilon=0$, Eqs. (D34) may be rewritten as

$$\partial_{\xi_D}\widetilde{D} = [\Gamma^{(1)}(\boldsymbol{z},\epsilon)+\widehat{\Lambda}_{\boldsymbol{R}}^{(2)}]\widetilde{D}+\widehat{\boldsymbol{\Lambda}}_{\boldsymbol{R}}^{(1)}(\boldsymbol{z},\widetilde{D}),\tag{D35a}$$

$$\partial_T\boldsymbol{z} = (-\epsilon\boldsymbol{a}+\mathsf{A}\cdot\boldsymbol{z})|\widetilde{D}|^2+\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{R}}^{(1)}(\widetilde{D},\overline{\widetilde{D}}).\tag{D35b}$$

Many terms in Eqs. (D34) yield unimportant corrections to the $\Gamma^{(1)}(\boldsymbol{z}, \epsilon)$ and $\boldsymbol{V}_D$ coefficients in Eqs. (D35). Refer to Appendix C 3 for the details of calculation of the uniform terms in Eqs. (D35).

---

* Electronic address: rkolesni@pppl.gov

[1] Z. Lin, T. S. Hahm, W. W. Lee, W. M. Tang, and R. B. White, Science **281**, 1835 (1998).

[2] S. J. Zweben, R. J. Maqueda, D. P. Stotler, A. Keesee, J. Boedo, C. E. Bush, S. M. Kaye, B. LeBlanc, J. L. Lowrance, V. J. Mastrocola, et al., Nucl. Fusion **44**, 134 (2004).

[3] F. L. Hinton and M. N. Rosenbluth, Plasma Phys. Control. Fusion **41**, A653 (1999).

[4] Z. Lin, T. S. Hahm, W. W. Lee, W. M. Tang, and P. H. Diamond, Phys. Rev. Lett. **83**, 3645 (1999).

[5] P. Diamond, S.-I. Itoh, K. Itoh, and T. Hahm, Plasma Phys. Control. Fusion **47**, R35 (2005).

[6] A. M. Dimits, G. Bateman, M. A. Beer, B. I. Cohen, W. Dorland, G. W. Hammett, C. Kim, J. E. Kinsey, M. Kotschenreuther, A. H. Kritz, et al., Phys. Plasmas **7**, 969 (2000).

[7] S. Dastgeer, S. Mahajan, and J. Weiland, Phys. Plasmas **9**, 4911 (2002).

[8] J. Weiland, S. Dastgeer, R. Moestam, I. Holod, and S. Gupta, J. Plasma Fusion Research **6**, 84 (2004).

[9] B. N. Rogers, W. Dorland, and M. Kotschenreuther, Phys. Rev. Lett. **85**, 5336 (2000).

[10] W. Horton, G. Hu, and G. Laval, Phys. Plasmas **3**, 2912 (1996).

[11] G. Hu and W. Horton, Phys. Plasmas **4**, 3262 (1997).

[12] J. A. Krommes and C.-B. Kim, Phys. Rev. E **62**, 8508 (2000).

[13] K. Nozaki, T. Taniuti, and K. Watanabe, J. Phys. Soc. Japan **46**, 983 (1979).

[14] J. Guckenheimer and P. Holmes, *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields* (Springer, New York, 1983).

[15] Y. A. Kuznetsov, *Elements of Applied Bifurcation Theory* (Springer, New York, 1998).

[16] P. Holmes, J. L. Lumley, and G. Berkooz, *Turbulence, Coherent Structures, Dynamical Systems and Symmetry* (Cambridge University Press, Cambridge, 1996).

[17] R. Kolesnikov and J. Krommes, Phys. Rev. Lett. **94**, 235002/1 (2005).

[18] J. A. Krommes, Phys. Rep. **360**, 1 (2002).

[19] Further discussion of this and related interchanges of limits can be found in J. A. Krommes and

G. Hu, Phys. Plasmas **1**, 3211 (1994).

[20] M. Ottaviani, M. Beer, S. Cowley, W. Horton, and J. Krommes, Phys. Rep. **283**, 121 (1997).

[21] M. A. Beer, Ph.D. thesis, Princeton University (1995).

[22] R. White, L. Chen, and F. Zonca, Phys. Rev. Lett. **92**, 75004 (2004).

[23] For considerable discussion of the difficulties of global analysis in general, see Refs. 14 and 15.

[24] M. A. Beer and G. W. Hammett, in *Theory of Fusion Plasmas: Proceedings of the Joint Varenna-Lausanne International Workshop, Aug. 31–Sept. 4, 1998* (Soc. Italiana di Fisica, Bologno, Italy, 1999), p. 19.

[25] A. C. Newell and J. A. Whitehead, J. Fluid Mech. **38**, 279 (1969).

[26] M. T. Landahl and E. Mollo-Christensen, *Turbulence and Random Processes in Fluid Mechanics* (Cambridge University Press, 1992), 2nd ed.

[27] J. T. Stuart, F. R. S., and R. C. DiPrima, Proc. R. Soc. Lond. A **362**, 27 (1978).

[28] In some cases, perturbation expansion is not possible because the shape of the CM is beyond all orders in the expansion parameter, e.g. it has the form $e^{-1/\epsilon}$.

[29] P. Manneville, *Dissipative Structures and Weak Turbulence* (Academic Press, Boston, 1990).

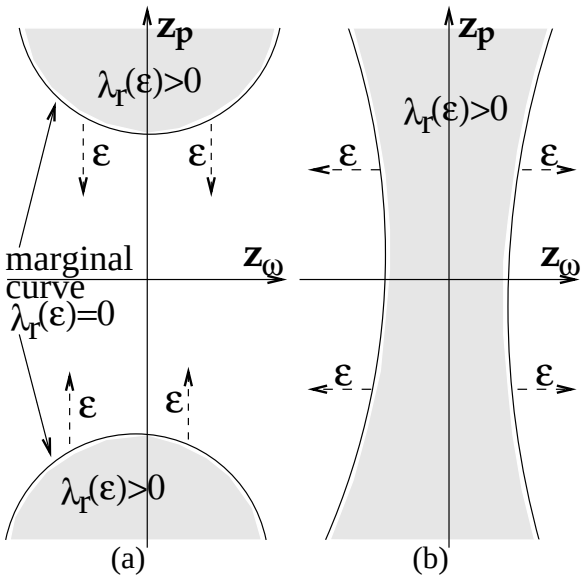

FIG. 1: The $\boldsymbol{z}$ plane is divided by the marginal curve into regions with positive and negative $\mathrm{Re}[\lambda_+(\boldsymbol{z}, \epsilon_0)]$. As ϵ increases the two parts of the curve move towards each other (a), coalesce, then move away from each other (b).

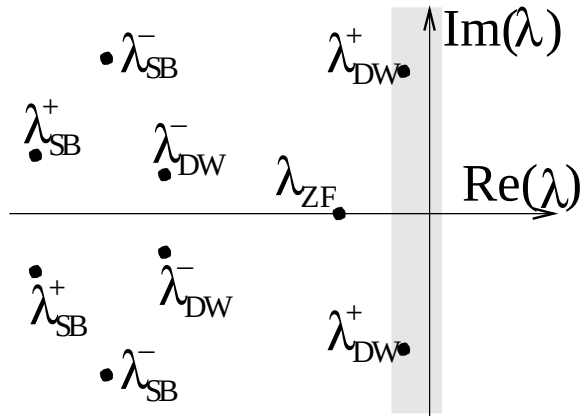

FIG. 2: For a collisional system, a schematic representation of the collection of eigenvalues in the complex λ plane. There are in total four drift wave modes $\lambda^{\pm}_{\rm DW} = \lambda^{\pm}_1(\mathbf{0},\epsilon)$, four side-band modes $\lambda^{\pm}_{\rm SB} = \lambda^{\pm}_3(\mathbf{0},\epsilon)$, and two zonal modes $\lambda_{\rm ZF} = -\eta_2$. Only the two DW modes $\lambda^{\pm}_{\rm DW}$ are included in the center eigenspace; the system undergoes a Hopf bifurcation when those modes cross the imaginary axis as a parameter ϵ goes through ϵ_c.

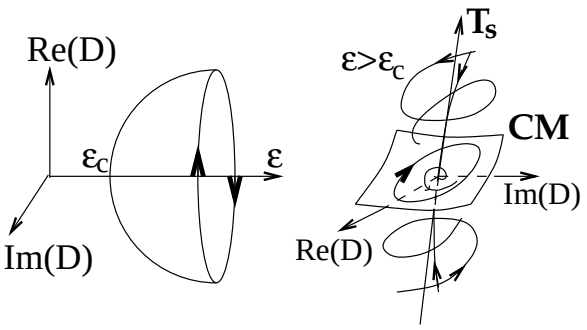

FIG. 3: The collisional system undergoes a supercritical Hopf bifurcation at the linear threshold $\epsilon = \epsilon_c$. The model has a 2D stable eigenspace T^c (and consequently a 2D CM) and an 8D stable manifold T^s (shown in 1D). See also similar figures in Refs. 14 and 15.

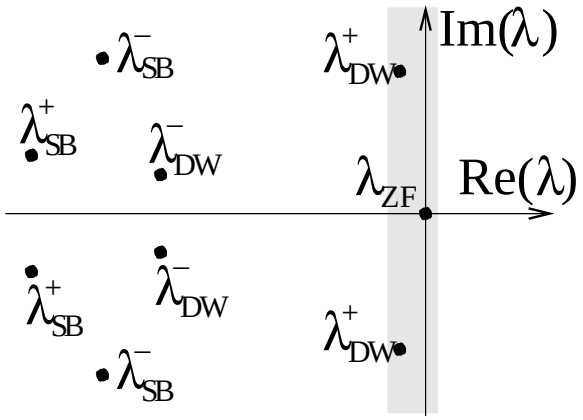

FIG. 4: The collection of eigenmodes in a collisionless system in the complex λ eigenvalue plane; $\lambda^{\pm}_{\rm DW} = \lambda^{\pm}_1(\boldsymbol{z}_0, \epsilon)$, $\lambda^{\pm}_{\rm SB} = \lambda^{\pm}_3(\boldsymbol{z}_0, \epsilon)$ for arbitrary $\boldsymbol{z}_0$. In addition to two DW modes $\lambda^{\pm}_{\rm DW}$, two undamped ZF modes $\lambda_{\rm ZF}$ are also included in the center eigenspace.

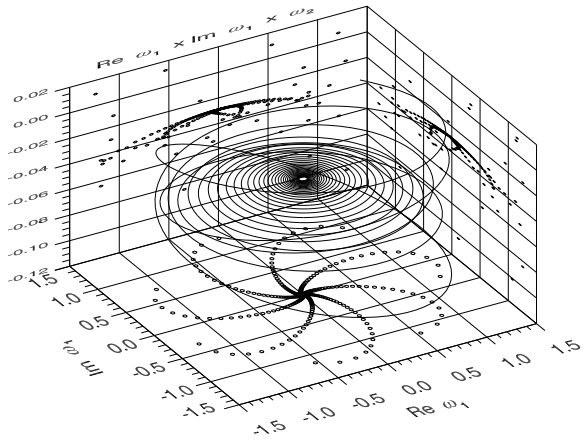

FIG. 5: Projection of the 10D phase space into the two DW directions ϖ_1 and P_1 and the ZF direction ϖ_2. For ease of visualization, the trajectory is divided into ten differently colored segments (online only); points equally spaced in time are also projected onto the coordinate planes. Following an initial transient, the trajectory is attracted to a hypersurface in the phase space. Ultimately, it evolves to a fixed point for which the drift-wave and sideband amplitudes vanish but the zonal amplitudes are nonzero. Color online.

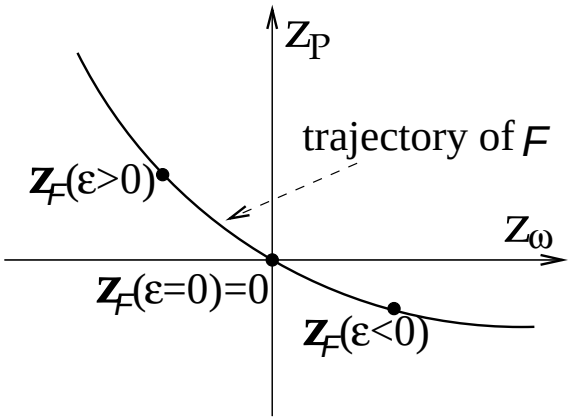

FIG. 6: The trajectory of the nontrivial fixed point $\mathcal{F}$ on the z plane. Note that $\mathcal{F}$ exists even for $\epsilon < 0$.

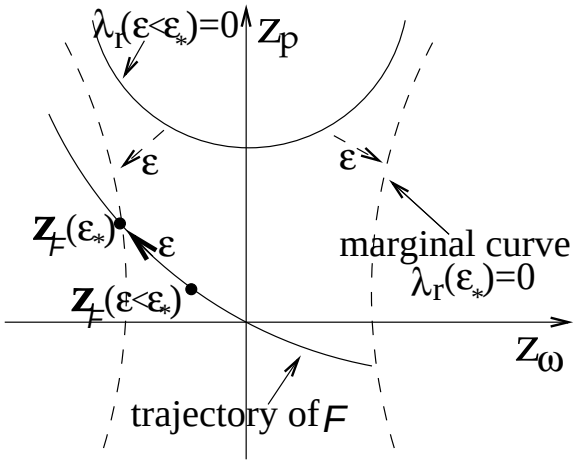

FIG. 7: The fixed point $\mathcal{F}$ crosses the marginal curve at $\epsilon = \epsilon_*$ making the system change topology (and thus undergo a bifurcation).

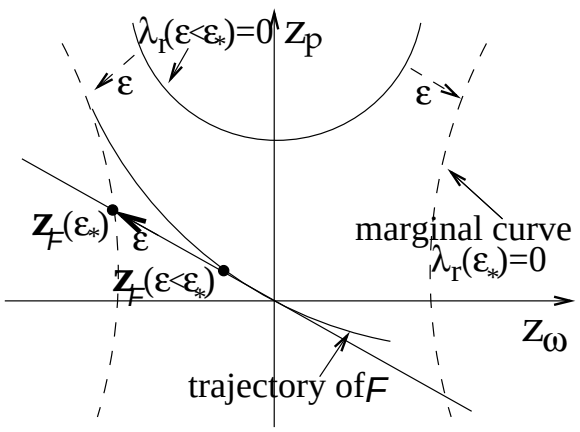

FIG. 8: The bifurcation point ϵ_* is found from the intersection of the line tangent to the trajectory of $\mathcal{F}$ at $\boldsymbol{z} = \mathbf{0}$ with the marginal curve.

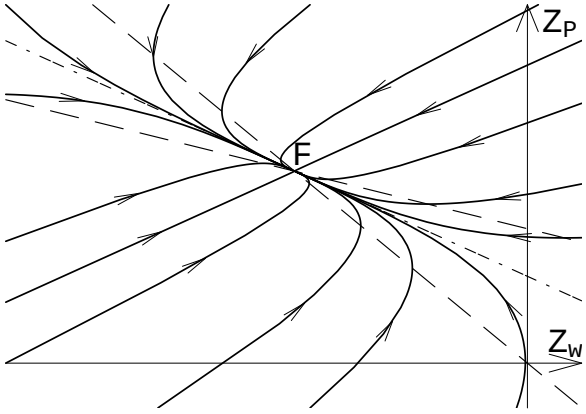

FIG. 9: Representative phase trajectories in the $\boldsymbol{z}$ plane, showing the nontrivial fixed point $\mathcal{F}$ (stable in the $\boldsymbol{z}$ plane) the overall stability of which determines the Dimits shift. Horizontal solid line: z_ϖ axis; vertical solid line: z_P axis; upper dashed line: $v(\boldsymbol{z}) = 0$ (slope $dz_P/dz_\varpi = 0$); lower dashed line: $u(\boldsymbol{z}) = 0$ (slope $= \infty$); dash-dotted lines: eigenvector directions (one such line is obscured by the trajectory at approximately 45°). Figure reprinted with permission from R. Kolesnikov and J. Krommes, Phys. Rev. Lett. 94, 235002/1 (2005). Copyright (2005) by the American Physical Society.

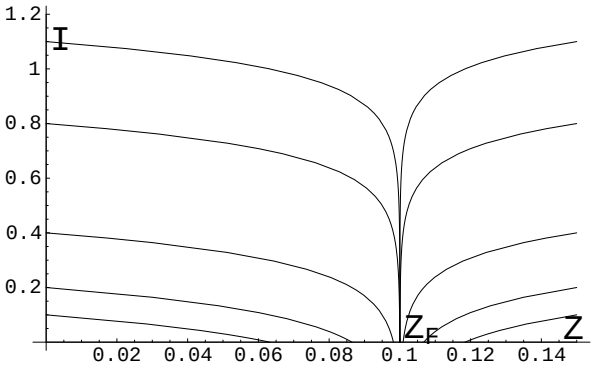

FIG. 10: Level curves for the dynamics of Eqs. (24) with $z_{\mathcal{F}} = 0.1$ and $\lambda = 0.1$. In spite of the presence of a fixed point at $z = z_{\mathcal{F}}$, only the single trajectory with $I_0 = \infty$ actually ends up at the fixed point at $t = \infty$.

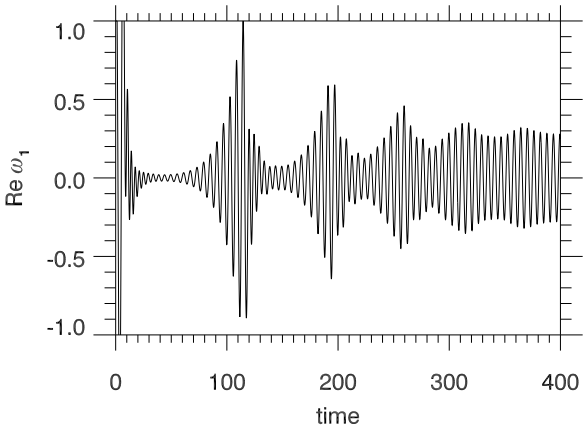

FIG. 11: Time dependence of DW vorticity in an 82-mode, energy-conserving truncation of a weakly collisional ITG system. The bursting behavior relaxes to a quasiregular nonturbulent state. The parameters used are $k_\perp \rho_s = 1$, $\tau = 0.1$, $\mu = 0.01$ (but $\eta_2 = 0.01$), $\nu = 1$, and $R/L_T = 2.2$.

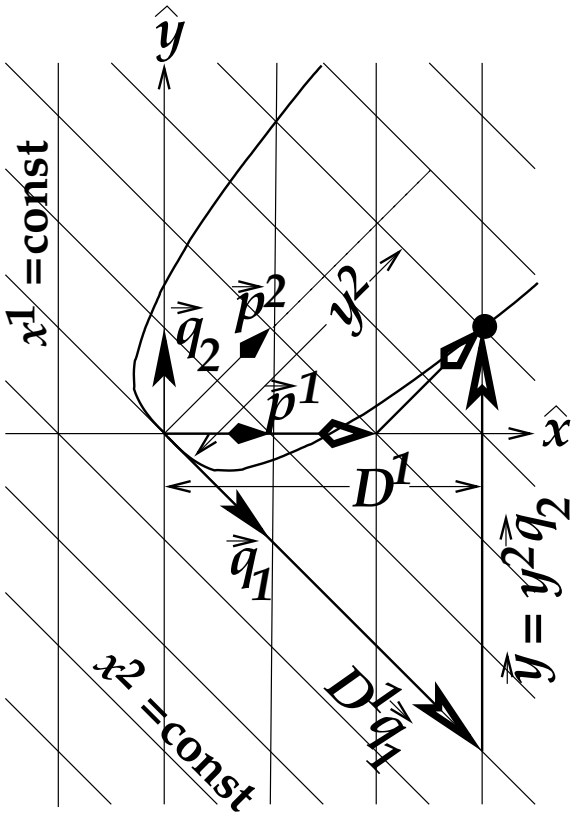

FIG. 12: An example of the contravariant representation of the center manifold (heaviest solid line). In a local Cartesian x–y coordinate system, the contravariant coordinates $x^1 = x$ and $x^2 = x + y$ are used. The basis vectors are $\boldsymbol{q}_1 = \partial\boldsymbol{x}/\partial x^1 = \widehat{\boldsymbol{x}} - \widehat{\boldsymbol{y}}$, $\boldsymbol{q}_2 = \partial\boldsymbol{x}/\partial x^2 = \widehat{\boldsymbol{y}}$, $\boldsymbol{p}^1 = \nabla x^1 = \widehat{\boldsymbol{x}}$, and $\boldsymbol{p}^2 = \nabla x^2 = \widehat{\boldsymbol{x}} + \widehat{\boldsymbol{y}}$. The center eigenspace is the $\boldsymbol{q}_1 \equiv \boldsymbol{q}$ axis; the stable eigenspace is the $\boldsymbol{y} \equiv \boldsymbol{q}_2$ axis. The CM is tangent to the center eigenspace at the origin. A point on the CM is represented by $\boldsymbol{u} = D\boldsymbol{q} + \boldsymbol{y}$ (vectors with concave open arrowheads), where in this example $D = D^1$, $\boldsymbol{y} = y^2\boldsymbol{q}_2$. The $\boldsymbol{y}$ direction is perpendicular to the direction reciprocal to $\boldsymbol{q}$: $\boldsymbol{p}^1 \cdot \boldsymbol{y} = 0$. The covariant representation of $\boldsymbol{u}$ is also indicated by the vectors with convex open arrowheads.

External Distribution

Plasma Research Laboratory, Australian National University, Australia
Professor I.R. Jones, Flinders University, Australia
Professor João Canalle, Instituto de Fisica DEQ/IF - UERJ, Brazil
Mr. Gerson O. Ludwig, Instituto Nacional de Pesquisas, Brazil
Dr. P.H. Sakanaka, Instituto Fisica, Brazil
The Librarian, Culham Science Center, England
Mrs. S.A. Hutchinson, JET Library, England
Professor M.N. Bussac, Ecole Polytechnique, France
Librarian, Max-Planck-Institut für Plasmaphysik, Germany
Jolan Moldvai, Reports Library, Hungarian Academy of Sciences, Central Research Institute for Physics, Hungary
Dr. P. Kaw, Institute for Plasma Research, India
Ms. P.J. Pathak, Librarian, Institute for Plasma Research, India
Dr. Pandji Triadyaksa, Fakultas MIPA Universitas Diponegoro, Indonesia
Professor Sami Cuperman, Plasma Physics Group, Tel Aviv University, Israel
Ms. Clelia De Palo, Associazione EURATOM-ENEA, Italy
Dr. G. Grosso, Instituto di Fisica del Plasma, Italy
Librarian, Naka Fusion Research Establishment, JAERI, Japan
Library, Laboratory for Complex Energy Processes, Institute for Advanced Study, Kyoto University, Japan
Research Information Center, National Institute for Fusion Science, Japan
Professor Toshitaka Idehara, Director, Research Center for Development of Far-Infrared Region, Fukui University, Japan
Dr. O. Mitarai, Kyushu Tokai University, Japan
Mr. Adefila Olumide, Ilorin, Kwara State, Nigeria
Dr. Jiangang Li, Institute of Plasma Physics, Chinese Academy of Sciences, People's Republic of China
Professor Yuping Huo, School of Physical Science and Technology, People's Republic of China
Library, Academia Sinica, Institute of Plasma Physics, People's Republic of China
Librarian, Institute of Physics, Chinese Academy of Sciences, People's Republic of China
Dr. S. Mirnov, TRINITI, Troitsk, Russian Federation, Russia
Dr. V.S. Strelkov, Kurchatov Institute, Russian Federation, Russia
Kazi Firoz, UPJS, Kosice, Slovakia
Professor Peter Lukac, Katedra Fyziky Plazmy MFF UK, Mlynska dolina F-2, Komenskeho Univerzita, SK-842 15 Bratislava, Slovakia
Dr. G.S. Lee, Korea Basic Science Institute, South Korea
Dr. Rasulkhozha S. Sharafiddinov, Theoretical Physics Division, Insitute of Nuclear Physics, Uzbekistan
Institute for Plasma Research, University of Maryland, USA
Librarian, Fusion Energy Division, Oak Ridge National Laboratory, USA
Librarian, Institute of Fusion Studies, University of Texas, USA
Librarian, Magnetic Fusion Program, Lawrence Livermore National Laboratory, USA
Library, General Atomics, USA
Plasma Physics Group, Fusion Energy Research Program, University of California at San Diego, USA
Plasma Physics Library, Columbia University, USA
Alkesh Punjabi, Center for Fusion Research and Training, Hampton University, USA
Dr. W.M. Stacey, Fusion Research Center, Georgia Institute of Technology, USA
Director, Research Division, OFES, Washington, D.C. 20585-1290

The Princeton Plasma Physics Laboratory is operated
by Princeton University under contract
with the U.S. Department of Energy.

Information Services
Princeton Plasma Physics Laboratory
P.O. Box 451
Princeton, NJ 08543

Phone: 609-243-2750
Fax: 609-243-2751
e-mail: pppl_info@pppl.gov
Internet Address: http://www.pppl.gov