CrossMark

J. Plasma Phys. (2015), *vol.* 81, 905810501 Copyright © Cambridge University Press 2015
doi:10.1017/S0022377815000495

Generalized universal instability: transient linear amplification and subcritical turbulence

Matt Landreman[1,†], Gabriel G. Plunk[2] and William Dorland[1]

[1]Institute for Research in Electronics and Applied Physics, University of Maryland, College Park, MD 20742, USA

[2]Max Planck Institute for Plasma Physics, 17491 Greifswald, Germany

(Received 13 January 2015; revised 10 April 2015; accepted 10 April 2015)

In this work we numerically demonstrate both significant transient (i.e. non-modal) linear amplification and sustained nonlinear turbulence in a kinetic plasma system with no unstable eigenmodes. The particular system considered is an electrostatic slab with magnetic shear, kinetic electrons and ions, weak collisions and a density gradient, but with no temperature gradient. In contrast to hydrodynamic examples of non-modal growth and subcritical turbulence, here there is no sheared flow in the equilibrium. Significant transient linear amplification is found when the magnetic shear and collisionality are weak. It is also demonstrated that nonlinear turbulence can be sustained if initialized at sufficient amplitude. We prove that these two phenomena are related: when sustained turbulence occurs without unstable eigenmodes, states that are typical of the turbulence must yield transient linear amplification of the gyrokinetic free energy.

1. Introduction

Despite the usefulness of linear eigenmode analysis in many contexts, it has been established that eigenmode analysis can in practice give misleading predictions about the stability of various systems (Trefethen & Embree 2005). In particular, turbulence may be sustained in a number of fluid systems despite the absence of any unstable eigenmodes, a phenomenon that may be called 'subcritical turbulence'. The first systems known to display this behaviour were certain sheared flow patterns in neutral fluids, such as plane Poiseuille or plane Couette flow (Tillmark & Alfredsson 1992; Trefethen *et al.* 1993; Grossman 2000). Later, sustained turbulence was seen in fluid simulations of several plasma systems with no absolute linear instabilities (Waltz 1985; Scott 1990, 1992; Drake, Zeiler & Biskamp 1995). More recently, turbulence in the absence of linear instability has been seen in kinetic plasma simulations involving sheared equilibrium flow (Highcock *et al.* 2010; Barnes *et al.* 2011). These examples demonstrate that linear eigenmode analysis can give deceiving predictions for nonlinear behaviour, a conclusion supported by other studies of plasma turbulence in which linear instability exists, but the fluctuations in saturated nonlinear turbulence are different in character from the linear eigenmodes (Biskamp & Zeiler 1995; Friedman *et al.* 2014; Friedman & Carter 2014).

† Email address for correspondence: mattland@umd.edu

https://doi.org/10.1017/S0022377815000495 Published online by Cambridge University Press

For the cases of turbulence without absolute linear instability in the presence of sheared flows, for both neutral fluids and plasmas, the phenomenon of transient linear amplification appears to be important (Trefethen *et al.* 1993; Farrell & Ioannou 1996; Grossman 2000; DelSole 2004*a*; Trefethen & Embree 2005; Schmid 2007; Newton, Cowley & Loureiro 2010; Schekochihin, Highcock & Cowley 2012). Also known as non-modal amplification, transient amplification is a period of growth before the exponential decay sets in. A necessary (but not sufficient) condition for transient amplification is that the linear operator must not have a complete orthogonal set of eigenvectors. Analysis of transient amplification has also been applied to plasma systems in which linear instability exists but transient linear behaviour may be more important (Camargo, Tippett & Caldas 1998; Camporeale, Passot & Burgess 2010; Friedman & Carter 2014; Squire & Bhattacharjee 2014*a*,*b*).

In the present work, we consider the phenomena of transient amplification and turbulence without instability for the case of a weakly collisional plasma drift wave driven by a density gradient (Galeev, Oraevsky & Sagdeev 1963; Krall & Rosenbluth 1965), neglecting toroidal effects. This oscillation has been called the 'universal instability' as it was at first thought to be linearly unstable in any plasma with a density gradient. The oscillation is electrostatic in nature, and linear instability can indeed be found in slab geometry without magnetic shear. It was later established that when any amount of magnetic shear is included in the model, no matter how small, the eigenmodes with $k_y\rho_i \ll 1$ become linearly damped, where k_y is the perpendicular wavenumber and ρ_i is the ion gyroradius (Antonsen 1978; Ross & Mahajan 1978; Tsang *et al.* 1978). However, instability can persist at $k_y\rho_i \geqslant 1$ when the collisionality is very low (Landreman, Antonsen & Dorland 2015). When both magnetic shear and modest collisions are included, the mode becomes linearly stable at all $k_y\rho_i$. It is this latter situation that we will consider here. (As an aside, the small-but-finite-shear eigenmodes are different from zero-shear modes not only for the universal instability, but also for the slab ion-temperature-gradient (ITG) instability. Specifically, in Plunk *et al.* (2014), considering the slab mode by setting curvature drift effects ω_d to 0, the small-shear limit of (19) is different from the zero-shear result (12).)

Fluid versions of this drift wave system were investigated by Scott (1990, 1992) and Drake *et al.* (1995). In these models, all linear modes were damped, yet nonlinearly, turbulence could be sustained. In Drake *et al.* (1995) it was pointed out that an important secondary instability in this system is a drift wave rotated by $\pi/2$ in the plane perpendicular to the magnetic field, since the stabilizing effects of magnetic shear are reduced in this orientation. In this rotated drift instability, the role of the equilibrium density gradient is replaced in a sense by the perturbed density gradient, the latter represented in the nonlinear term. However, since the nonlinearity is conservative, the rotated drift wave does not inject energy into the system in the way that can be done by a perturbation feeding off the density gradient of the equilibrium. Several authors have also investigated transient linear effects in fluid drift wave models (Camargo *et al.* 1998; Friedman *et al.* 2014; Friedman & Carter 2014). However, these studies did not include magnetic shear, which we will find to have a very strong effect on some aspects of the transient behaviour.

In the following sections, we establish three results concerning the drift wave system in slab geometry with magnetic shear, extending previous work by considering a kinetic description and exploring the role of transient linear amplification. For these studies, the initial-value gyrokinetic code gs2 (Kotschenreuther, Rewoldt & Tang 1995) will be used for numerical computations. First, in § 3 we demonstrate that

https://doi.org/10.1017/S0022377815000495 Published online by Cambridge University Press

significant transient linear amplification is possible in this system. The transiently growing fluctuations resemble the zero-shear eigenmodes in the dependence of their instantaneous growth rate on wavenumber, although the transients are sensitive to magnetic shear in their global time dependence. Second, in § 4 we demonstrate that universal-mode turbulence without linear instability, seen previously in fluid models (Scott 1990, 1992; Drake *et al.* 1995), is also possible in the gyrokinetic model. Finally, we prove that there is a relationship between transient linear amplification and nonlinear turbulence in this system, using similar arguments to the proof for hydrodynamics in DelSole (2004*b*) and arriving at similar conclusions. Specifically, if the physical parameters are such that turbulence can be sustained, typical states of the turbulence must be amplified linearly for some amount of time, even if these states eventually decay exponentially. Two more precise statements of this result and their proof will be given in § 5.

2. Gyrokinetic model

We consider slab geometry with magnetic shear, so the magnetic field in Cartesian coordinates (x', y', z') is $\boldsymbol{B} = [\boldsymbol{e}_{z'} + (x'/L_s)\boldsymbol{e}_{y'}]B$ for some constant B and shear length L_s. A field-aligned coordinate system is introduced: $x = x'$, $y = y' - x'z'/L_s$, $z = z'$, so $\boldsymbol{B}\cdot\boldsymbol{\nabla}x = \boldsymbol{B}\cdot\boldsymbol{\nabla}y = 0$ and $\boldsymbol{B}\cdot\boldsymbol{\nabla}z = B$. In this geometry, we consider the nonlinear electrostatic gyrokinetic equation (Frieman & Chen 1982), dropping toroidal effects:

$$\frac{\partial h_s}{\partial t} + v_\parallel \frac{\partial h_s}{\partial z} + \frac{1}{B}\left[\frac{\partial \langle\Phi\rangle_{\boldsymbol{R}_s}}{\partial X_s}\frac{\partial h_s}{\partial Y_s} - \frac{\partial \langle\Phi\rangle_{\boldsymbol{R}_s}}{\partial Y_s}\frac{\partial h_s}{\partial X_s}\right] - \langle C_s\{h_s\}\rangle_{\boldsymbol{R}_s}$$
$$= \frac{q_s}{T_s} f_{Ms}\frac{\partial \langle\Phi\rangle_{\boldsymbol{R}_s}}{\partial t} - \frac{f_{Ms}}{BL_n}\frac{\partial \langle\Phi\rangle_{\boldsymbol{R}_s}}{\partial Y_s}\left[1 + \left(\mathscr{E}_s - \frac{3}{2}\right)\eta_s\right], \tag{2.1}$$

together with the quasineutrality condition

$$-\frac{ne\Phi}{T_i} + \int \mathrm{d}^3 v \langle h_i\rangle_{\boldsymbol{r}} = \frac{ne\Phi}{T_e} + \int \mathrm{d}^3 v \langle h_e\rangle_{\boldsymbol{r}}. \tag{2.2}$$

Here, $s \in \{i, e\}$ subscripts denote particle species, $\Phi(t, \boldsymbol{r})$ is the electrostatic potential, $h_s(t, \boldsymbol{R}_s, \mathscr{E}_s, \mu)$ is the non-adiabatic distribution function, $\mathscr{E}_s = m_s v^2/(2T_s)$, $\mu = v_\perp^2/(2B)$ and f_{Ms} is the leading-order Maxwellian. The scale length of the equilibrium density n is $L_n = -n/(\mathrm{d}n/\mathrm{d}x)$, $q_i = e = -q_e$ is the proton charge, T_s is the species temperature, $\eta_s = (\mathrm{d}\ln T_s/\mathrm{d}x)/(\mathrm{d}\ln n/\mathrm{d}x)$, C_s is the collision operator, and $\langle\rangle_{\boldsymbol{r}}$ and $\langle\rangle_{\boldsymbol{R}_s}$ denote gyroaverages at fixed particle position $\boldsymbol{r} = (x, y, z)$ and guiding-centre position $\boldsymbol{R}_s = (X_s, Y_s, Z_s)$. The distinction between z and Z_s may be neglected. We consider a flux tube domain, in which the fluctuating quantities can be taken to be periodic in x and y. Expanding the fluctuating quantities in spatial Fourier modes as $\Phi = \sum_{k_x,k_y}\Phi_{k_x,k_y}(t, z)\exp(\mathrm{i}k_x x + \mathrm{i}k_y y)$ and $h_s = \sum_{k_x,k_y} h_{s,k_x,k_y}(t, z, \mathscr{E}_s, \mu)\exp(\mathrm{i}k_x X_s + \mathrm{i}k_y Y_s)$, we may replace the gyroaverages with Bessel functions: $\langle\Phi_{k_x,k_y}\exp(\mathrm{i}k_x x + \mathrm{i}k_y y)\rangle_{\boldsymbol{R}_s} = J_{0s}\Phi_{k_x,k_y}\exp(\mathrm{i}k_x X_s + \mathrm{i}k_y Y_s)$ and $\langle h_{s,k_x,k_y}\exp(\mathrm{i}k_x X_s + \mathrm{i}k_y Y_s)\rangle_{\boldsymbol{r}} = J_{0s}h_{s,k_x,k_y}\exp(\mathrm{i}k_x x + \mathrm{i}k_y y)$, with $J_{0s} = J_0(k_\perp v_\perp/\Omega_s)$ and $\Omega_s = eB/m_s$. The perpendicular wavenumber is given by

$$k_\perp = \sqrt{k_y^2 + (k_x - k_y z/L_s)^2}. \tag{2.3}$$

https://doi.org/10.1017/S0022377815000495 Published online by Cambridge University Press

The system is uniquely specified by the following dimensionless parameters: L_s/L_n, η_i, η_e, T_e/T_i and collisionality $\nu_{ee}L_n/v_i$, where $\nu_{ee} = \sqrt{2}\pi n e^4 \ln\Lambda/[(4\pi\epsilon_0)^2 m_e^{1/2} T_e^{3/2}]$ is the electron–electron collision frequency and $v_i = \sqrt{2T_i/m_i}$.

It should be noticed that the parallel coordinate z enters the system (2.1)–(2.3) only through (2.3). It should also be noticed that k_x enters the system (2.1)–(2.3) only through (2.3) and through the nonlinear term in (2.1). Therefore, a linear mode at non-zero k_x and non-zero k_y satisfies exactly the same equations as the corresponding $k_x = 0$ mode, but translated in z by $L_s k_x/k_y$. As a consequence, the linear growth or damping rates of $k_y \neq 0$ modes are independent of k_x, and the eigenmode structures are independent of k_x up to this translation in z.

One important property of the system (2.1)–(2.2) is the free-energy conservation equation (Sugama *et al.* 1996; Howes *et al.* 2006; Abel *et al.* 2013) obtained by applying the operation $\sum_s \int \mathrm{d}^3R_s \int \mathrm{d}^3v (T_s h_s/f_{Ms})(\cdots)$ to (2.1):

$$\frac{\mathrm{d}W}{\mathrm{d}t} = S(h_i, h_e), \tag{2.4}$$

where

$$W = \sum_s \int \mathrm{d}^3r \int \mathrm{d}^3v \frac{T_s \delta f_s^2}{2 f_{Ms}} \tag{2.5}$$

is the free energy, $\delta f_s = h_s - q_s \Phi f_{Ms}/T_s$ is the total departure of the distribution function from a Maxwellian and

$$\begin{aligned} S(h_i, h_e) \;=\; & \sum_s \int \mathrm{d}^3r \int \mathrm{d}^3v \frac{T_s h_s}{f_{Ms}} C_s\{h_s\} \\ & + \sum_s \int \mathrm{d}^3R_s \int \mathrm{d}^3v \frac{h_s}{B} \frac{\partial \langle \Phi \rangle_{\boldsymbol{R}_s}}{\partial Y_s} \frac{T_s}{n_s} \frac{\mathrm{d}n_s}{\mathrm{d}x} \left[1 + \left(\mathscr{E}_s - \frac{3}{2}\right)\eta_s\right] \end{aligned} \tag{2.6}$$

is the net energy source. (To derive (2.4)–(2.6), the equalities $\int \mathrm{d}^3R_s \int \mathrm{d}^3v \langle\cdots\rangle_{\boldsymbol{R}_s} = \int \mathrm{d}^3R_s \int \mathrm{d}^3v (\cdots) \approx \int \mathrm{d}^3r \int \mathrm{d}^3v (\cdots) = \int \mathrm{d}^3r \int \mathrm{d}^3v \langle\cdots\rangle_{\boldsymbol{r}}$ are used. The $v_\parallel \partial h_s/\partial z$ term in (2.1) has been integrated by parts; as detailed in Banón Navarro (2012, p. 66), the associated boundary term vanishes for a finite-z domain using a twist-and-shift parallel boundary condition (Beer, Cowley & Hammett 1995), taking the fluctuation amplitude to vanish at the ends of each set of linked domains.) The first term on the right-hand side of (2.6) is negative-definite (by the H-theorem). The second term has the form of the particle and heat fluxes multiplied by the density and temperature gradients, and this term increases W if the fluxes have the appropriate sign to relax the gradients. Thus, in a turbulent steady state where W is statistically constant, (2.4)–(2.6) indicate that energy is injected by the equilibrium gradient(s) and dissipated by collisions. It should be noted that the operation $\int \mathrm{d}^3R_s$ used in the derivation of (2.4) annihilates the nonlinear term in (2.1), and so (2.4)–(2.6) apply to both the linear and nonlinear dynamics.

The system (2.1)–(2.2) may or may not possess absolute linear instability, depending on the physical parameters. For the rest of this paper, we take $\eta_i = \eta_e = 0$, so it is certain that the ITG and electron-temperature-gradient (ETG) instabilities are suppressed. Even without temperature gradients, the system may still be unstable to the universal instability driven by the density gradient. For reference, the growth rates and frequencies of this instability in the absence of magnetic shear are given in appendix A. As detailed in appendix A, some insight into the instability can be

https://doi.org/10.1017/S0022377815000495 Published online by Cambridge University Press

obtained by expanding the plasma dispersion function Z for the electrons in the small-argument limit, $Z(\omega/(|k_\parallel|v_e)) \approx \mathrm{i}\sqrt{\pi}$ for $|\omega/(k_\parallel v_e)| \ll 1$, and by expanding the ion Z function in the large-argument limit, $Z(\omega/(|k_\parallel|v_i)) \approx -|k_\parallel|v_i/\omega$ for $|\omega/(k_\parallel v_i)| \gg 1$. One then finds a frequency with positive imaginary part $\propto \mathrm{i}\sqrt{\pi}/L_n$, indicating that the instability is due to resonant electrons in the presence of a density gradient. When an arbitrarily small amount of magnetic shear is introduced, the global eigenmodes are no longer recognizable as the unsheared slab modes, and indeed are absolutely stable for $k_y\rho_i \ll 1$ (Ross & Mahajan 1978; Tsang *et al.* 1978). One might reasonably expect weakly sheared systems with stable global eigenmodes to exhibit transient linear amplification, with the number of e-foldings of amplification tending to infinity as the magnetic shear tends to zero. In the following section, we will show this hypothesis to be true, and thereby demonstrate the apparently singular limit of zero shear to be a continuous limit of the general shear case.

3. Transient linear amplification

Turning now to linear initial-value simulations, we indeed find that significant transient linear amplification is possible for certain parameters. Figure 1(*a*) illustrates a linear computation for a single (k_x, k_y) in a case of very weak magnetic shear, $L_s/L_n = 300$. Other parameters used are $k_y\rho_i = 1$ (where $\rho_i = v_i/\Omega_i$), $\nu_{ee}L_n/v_i = 0.05$, $m_i = 3672m_e$, $T_e = T_i$ and $\eta_i = \eta_e = 0$. The results are independent of k_x up to a translation in z, as noted previously. Collisions are also included to eliminate absolute instability for $k_y\rho_i \geqslant 0.7$, as discussed in Landreman *et al.* (2015), using the collision operators detailed in Abel *et al.* (2008). As transient amplification generally depends on the choice of norm, here we compare three reasonable quadratic norms: W, $\int \mathrm{d}z\,|\Phi|^2$ and the maximum of $|\Phi|^2$ along z. The curve for each norm is scaled by a constant so that the initial amplitude is one. Amplification by several orders of magnitude is observed in all norms before exponential decay sets in.

For the calculation in figure 1, the initial condition used is the standard 'noise' initial condition of gs2: $g_s = h_s - \langle\Phi\rangle_{\boldsymbol{R}s} q_s f_{Ms}/T_s$ is set to a Maxwellian velocity distribution times a random complex number at each grid point in z. (The real and imaginary parts of these complex numbers are each taken from a uniform distribution on $[-1, 1]$.) While techniques exist to calculate the optimal perturbation that gives maximum amplification (Farrell & Ioannou 1996; Squire & Bhattacharjee 2014*a*,*b*), such techniques are numerically challenging for this kinetic model due to the high dimensionality of the problem. The noise initial condition is a convenient shortcut, ensuring that energy is likely to be given to a growing perturbation if one exists. It has been shown in Camargo *et al.* (1998) and Squire & Bhattacharjee (2014*b*) that for studies of transient linear amplification, random initial conditions represent a reasonable proxy for a directly calculated optimal initial condition. Figure 1(*b*) shows that when the calculation is repeated many times, the dynamics of the energy in figure 1(*a*) is typical.

Results for the three norms shown in figure 1(*a*) are quite similar, although the curve for the W norm appears slightly higher than the other curves due to less decay in the initial period $tv_i/L_n < 20$. Indeed, we find that all results described or plotted for the remainder of this paper are quite similar for all of these norms, using the locally Maxwellian initial perturbations described above. However, there are certain non-Maxwellian initial perturbations that can cause arbitrarily fast growth in the electric field energy norms but not in the W norm (Hammett, G. W. & Abel, I. G. 2015, Personal Communication). Furthermore, it is W that is conserved by the

https://doi.org/10.1017/S0022377815000495 Published online by Cambridge University Press

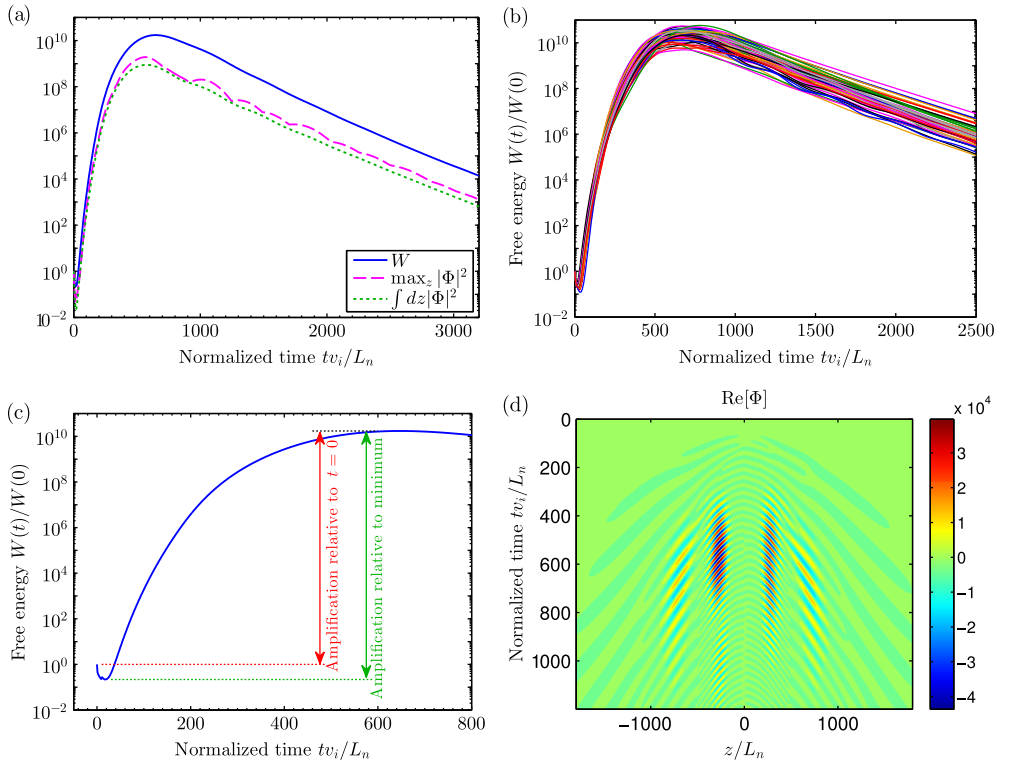


FIGURE 1. (*a*) Linear evolution for $L_s/L_n = 300$, $\eta_i = \eta_e = 0$, $\nu_{ee} = 0.05 v_i/L_n$ and $k_y\rho_i = 1$, illustrating transient amplification in three norms. The initial conditions are detailed in the main text. (*b*) On repeating the calculation 50 times, the results are always qualitatively similar. (*c*) The amount of transient amplification may be defined relative to $t = 0$ or relative to the minimum that occurs shortly thereafter. (*d*) The potential structure along a field line for this example.

nonlinearities (recall the conservation equation (2.4)), and thus the W norm that plays a role in determining bounds on the possibility of nonlinearly sustained turbulence, as we will show in § 5. For these reasons, we henceforth restrict our attention to the W norm.

Figure 1(*c*) shows the same computation, zooming in to the initial transient period. Before the amplification begins around $tv_i/L_n = 20$, there is an initial decay period due to strongly damped modes. An amplification factor may be defined either relative to this minimum amplitude or relative to the amplitude at $t = 0$. In both cases, we define the amplification factor to be 1 if there is no such transient amplification.

The transiently growing structures in the presence of magnetic shear are related to the zero-shear eigenmodes. To demonstrate this relationship, we consider a single (k_x, k_y), and define an effective $k_\perp$ at each time for the transient fluctuation by

$$\langle k_\perp \rangle(t) = \frac{\int \mathrm{d}z\, k_\perp(z) |\Phi_{k_x,k_y}(t,z)|^2}{\int \mathrm{d}z |\Phi_{k_x,k_y}(t,z)|^2}. \tag{3.1}$$

https://doi.org/10.1017/S0022377815000495 Published online by Cambridge University Press

Similarly, we define an effective $|k_\parallel|$ at each time for the transient structure by

$$\langle |k_\parallel| \rangle (t) = \frac{\int \mathrm{d}k_\parallel |k_\parallel| |\tilde{\Phi}_{k_x,k_y}(t, k_\parallel)|^2}{\int \mathrm{d}k_\parallel |\tilde{\Phi}_{k_x,k_y}(t, k_\parallel)|^2}, \tag{3.2}$$

where $\tilde{\Phi}_{k_x,k_y}$ is the Fourier transform of Φ_{k_x,k_y} in z. Using these effective values of $k_\perp$ and $k_\parallel$, the zero-shear dispersion relation may be evaluated. The resulting growth rate and real frequency can then be compared with the behaviour of the finite-shear transient structure. An example of such a comparison is given in figure 2, using the same simulation as in figure 1. For this case, we evaluate $\langle k_\perp \rangle$ and $\langle |k_\parallel| \rangle$ at $t = 300 L_n/v_i$. The figure shows that near this time, in the region $z \approx 300 L_n$ where the transient structure has greatest amplitude, the transient does resemble a wave packet with the real frequency and growth rate of the zero-shear eigenmode evaluated at $\langle k_\perp \rangle$ and $\langle |k_\parallel| \rangle$. The zero-shear eigenmodes cannot remain true eigenmodes when even a minute amount of magnetic shear is introduced, since introducing shear turns $k_\perp$ from a constant into a function of z. However, if this variation of $k_\perp(z)$ is on a longer scale length than $\sim 1/k_\parallel$ of the perturbation, then a localized wave packet can exist in a region of nearly constant $k_\perp$. We can then expect the zero-shear dispersion relation to give accurate predictions for the central frequency and growth rate of the packet. This packet will move in z, and as a result, the central values of $k_\perp$ and $k_\parallel$ will evolve, meaning that the equivalent zero-shear eigenmode at one time will differ from the equivalent zero-shear eigenmode at another time. Since the frequency and growth rate of the finite-shear structure thus evolve in time, this structure is not itself a true eigenmode; the true eigenmode is only established once all points along the field line have had time to communicate with each other. However, at each instant of the transient growth phase, the finite-shear structure grows at precisely the rate expected from the zero-shear instability at the average wavenumbers $\langle k_\perp \rangle$ and $\langle |k_\parallel| \rangle$.

To demonstrate this principle across a range of parameters, figure 3 shows that the correspondence between the finite-shear transients and zero-shear eigenmodes is maintained throughout the evolution of the transient, for a range of values of magetic shear. In this figure, we display 25 independent simulations of linear transients, varying L_s/L_n between 20 and 300, and plotting the evolution of $\gamma = (2W(t))^{-1}\,\mathrm{d}W/\mathrm{d}t$ from the time of peak growth rate to the time at which the energy begins to decay. For all simulations shown, $k_y \rho_i = 2$ and $\nu_{ee} = 0.05 v_i/L_n$. Both $\langle |k_\parallel| \rangle$ and $\langle k_\perp \rangle$ increase with time during the transient growth, corresponding to movement upward and rightward in figure 10(*a*). As a result, the effective zero-shear growth rate decreases over time. For all transients we examined, there is very close agreement between the instantaneous growth rate of the transient and the 'predicted' growth rate from the zero-shear dispersion relation. Thus, we can conclude that the transiently growing structures are closely related to the zero-shear eigenmodes.

For the range of $k_\parallel$ and $k_\perp$ values computed for the transient structures of figure 3, the zero-shear dispersion relation is well approximated by the approximate analytic dispersion relation (A 2). Thus, we can also conclude that the transient structures, like the zero-shear eigenmodes, are driven by the combination of resonant electrons and the density gradient.

In contrast to the agreement shown in figure 3 for the amplification phase following the maximum growth rate, we do not find such close agreement between the zero-shear dispersion relation and the finite-shear dynamics for very early times, before

https://doi.org/10.1017/S0022377815000495 Published online by Cambridge University Press

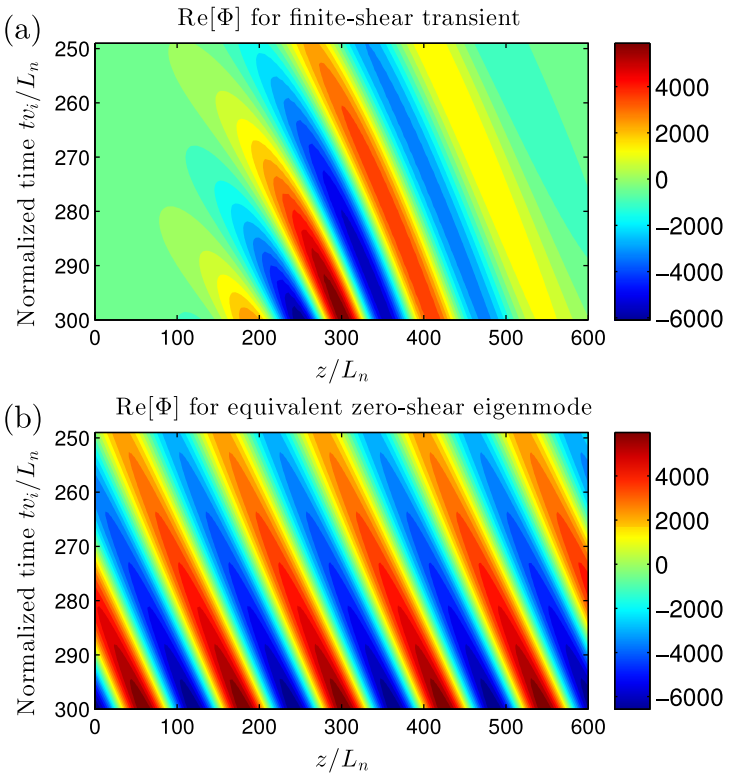


FIGURE 2. Comparison of the finite-shear transient from figure 1 with a zero-shear eigenmode, where $k_\perp$ and $k_\parallel$ for the latter are chosen to be $\langle k_\perp \rangle$ and $\langle |k_\parallel| \rangle$ from the transient at $t = 300 L_n/v_i$. It can be seen that near this time, the transient resembles a wave packet with roughly the same complex frequency as the zero-shear eigenmode. (The amplitude and phase of the eigenmode in (*b*) are arbitrary.)

the maximum growth rate occurs. In this initial period, a single dominant wave packet has not yet overtaken other fluctuations present in the initial random noise. We also do not see such close agreement between the zero-shear and finite-shear behaviour during the decay phase of the latter, i.e. at large times. The finite-shear decay rate as $t \to \infty$ corresponds to an eigenvalue, and it was established in previous work (Antonsen 1978; Ross & Mahajan 1978; Tsang *et al.* 1978) that the finite-shear eigenvalues are not continuously related to the zero-shear eigenvalues. In an eigenmode, points at all z have a coherent time dependence. Thus, this damping rate is determined by a coherent interaction between points of different $k_\perp(z) = k_y\sqrt{1 + (z/L_s)^2}$, and between all $k_\parallel$ Fourier components. This damping rate is thus not well approximated by the zero-shear dispersion relation which assumes a single $k_\perp$ and single $k_\parallel$.

Transient amplification is exponentially reduced by increasing magnetic shear, as illustrated in figure 4. For this figure, $k_y\rho_i$ is fixed at 2, and other parameters are the same as in figure 1. (We choose $k_y\rho_i = 2$ here since wavenumbers near this value will turn out to be the most amplified, as we will find shortly.) There is some scatter in figure 4, associated with the random initial condition, but the exponential trend with L_s/L_n is clear. It should be noticed that amplification only appears to be

https://doi.org/10.1017/S0022377815000495 Published online by Cambridge University Press

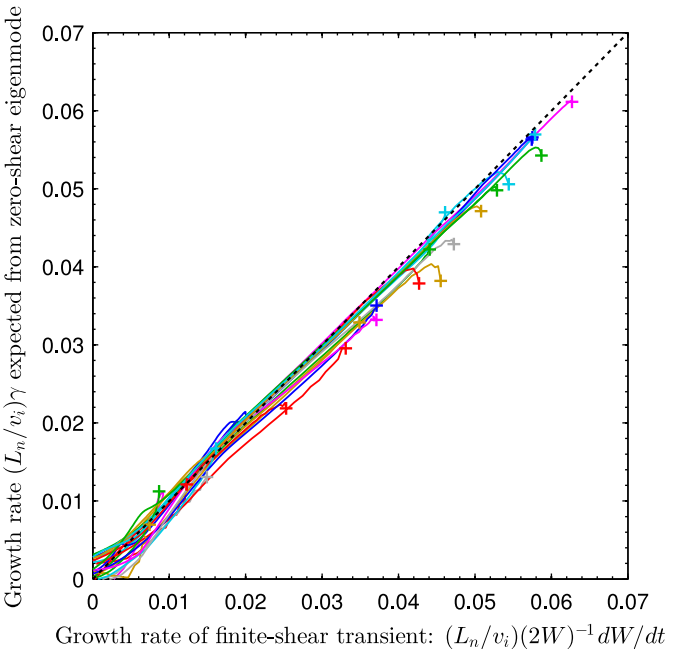


FIGURE 3. Although a change from zero to small finite magnetic shear causes a fundamental change in the eigenvalues (which change from unstable to decaying), the transient finite-shear structures retain a similarity to the zero-shear modes. As shown here, the instantaneous growth rate of the transient finite-shear structure is nearly identical to the growth rate of a zero-shear eigenmode at the dominant $k_\perp$ and $k_\parallel$ values (3.1)–(3.2). Each trajectory shown is an independent linear initial-value calculation, with L_s/L_n varying from 20 to 300 between simulations. The growth rate is plotted from the time of maximum growth rate (at the + symbols) until decay begins.

possible if $L_s/L_n \gtrsim 20$. Values of L_s/L_n far in excess of 20 are obtained in the edge of tokamak experiments. (For a high-aspect-ratio tokamak, $L_s \sim qR/\hat{s}$, where q is the safety factor, R is the major radius, $\hat{s} = (r/q)\,\mathrm{d}q/\mathrm{d}r$ and r is the minor radius.) Figure 4(*b*) demonstrates that the time over which amplification occurs varies approximately linearly with L_s/L_n. The figure compares two measures of the duration of amplification, measuring to the time of maximum either from $t = 0$ or from the time of minimum W, akin to figure 1(*b*) but in the time coordinate. These two measures of amplification duration display a similar linear trend.

While figures 2 and 3 made the case that the transients are insensitive to shear in one sense, it is not a contradiction that the transients are sensitive to shear in the sense of figure 4(*a,b*). Figures 2 and 3 show that the dominant complex frequency of the transient at each instant depends on time only through $\langle k_\perp\rangle(t)$, $\langle|k_\parallel|\rangle(t)$ and the zero-shear dispersion relation, with no significant additional dependence on shear. However, $\langle k_\perp\rangle(t)$ and $\langle|k_\parallel|\rangle(t)$ are sensitive to shear, yielding the sensitivity of the overall $W(t)$ curve for each transient seen in figure 4(*a,b*).

Figure 5 shows the dependence of the amplification factor on k_y at fixed $L_s/L_n = 200$. Significant amplification can be observed for a range of k_y surrounding $1/\rho_i$. Results are shown for two values of collisionality, $\nu_{ee}L_n/v_i = 0.1$ and 0.5. (When the collisionality is much smaller than these values, absolute linear instability is present

https://doi.org/10.1017/S0022377815000495 Published online by Cambridge University Press

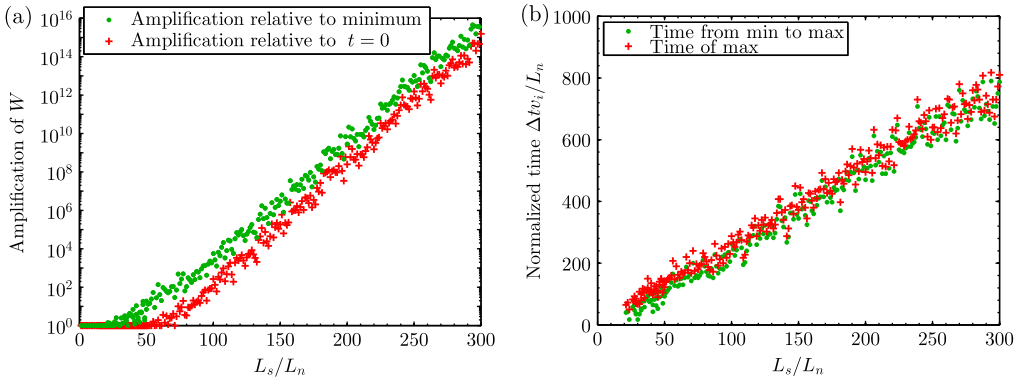


FIGURE 4. (*a*) The number of *e*-foldings of transient amplification and (*b*) the duration of the amplification both scale with the magnetic shear length scale L_s, consistent with the appearance of absolute instability in the limit of vanishing shear ($L_s \to \infty$). The parameters used are $k_y\rho_i = 2$ and $\nu_{ee} = 0.05 v_i/L_n$.

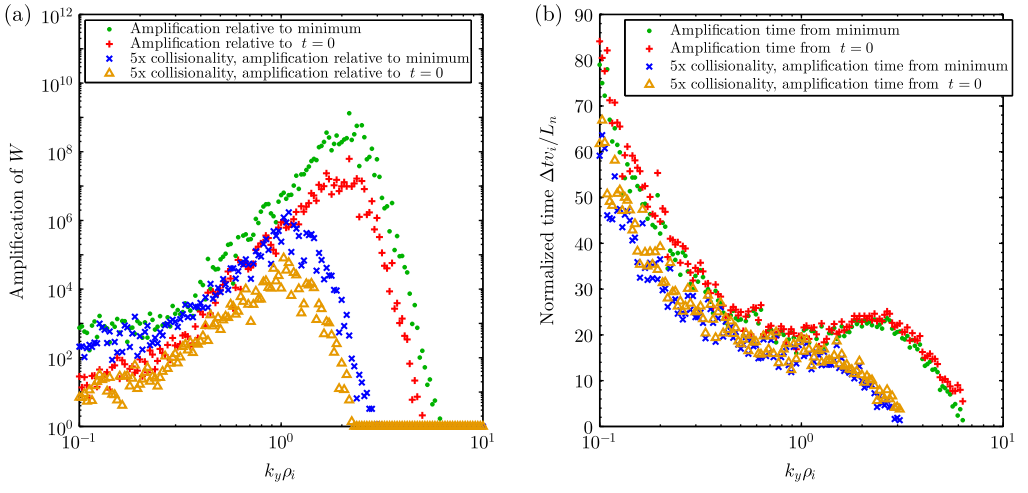


FIGURE 5. Transient amplification can be observed for a range of k_y surrounding $1/\rho_i$. For this figure, $L_s/L_n = 200$. The two collisionalities shown are $\nu_{ee}L_n/v_i = 0.1$ and 0.5.

(Landreman *et al.* 2015).) It can be seen that increasing collisionality tends to limit amplification at larger k_y, while having much less effect at smaller k_y. This kinetic problem is similar to conventional hydrodynamics in that for both cases non-modal amplification decreases with increasing collisional dissipation (associated with viscosity in the Navier–Stokes equation for the hydrodynamic case).

In the context of sheared flows in neutral fluids, a relationship between transient growth and eigenvalue sensitivity has been formally proved using the notion of a pseudospectrum (Reddy & Henningson 1993; Trefethen *et al.* 1993; Trefethen & Embree 2005). The ϵ-pseudospectrum of a linear operator can be defined as the union of the sets of eigenvalues obtained when the operator is perturbed by any other operator of norm $\leqslant\epsilon$, providing a precise notion of the sensitivity of eigenvalues. One can rigorously bound the amount of possible transient growth by the extent of the

https://doi.org/10.1017/S0022377815000495 Published online by Cambridge University Press

pseudospectrum into the unstable half of the complex plane (Reddy & Henningson 1993), summarized in equation (12) of Trefethen *et al.* (1993). However, for the universal instability, it is not clear that changing the magnetic shear from 0 to a small finite value has the form of the operator perturbation used to define the pseudospectrum. The reason is that the z domain is infinite, and for any non-zero shear, there is always some z for which the argument of J_0 is changed significantly compared with the zero-shear case. Indeed, all evidence suggests that when the shear changes from 0 to any small finite value, the eigenvalues (for a range of k_y) change from unstable to stable discontinuously rather than continuously (Ross & Mahajan 1978; Tsang *et al.* 1978). Thus, it may or may not be possible to understand the role of magnetic shear in the universal instability using the formal theory of pseudospectra, and we leave this careful mathematical analysis for future work.

4. Turbulence

We now demonstrate that when the nonlinear term in (2.1) is retained, sustained turbulence is possible in this system. For the demonstration here we consider the parameters $L_s/L_n = 20$, $T_e = T_i$, $m_i = 3672m_e$, $\eta_i = \eta_e = 0$ and $\nu_{ee}L_n/v_i = 0.05$. We find that this point in parameter space is approximately on the edge of where sustained turbulence is possible: if the magnetic shear or collisionality is increased much beyond these values, turbulence cannot be sustained indefinitely (though it may persist for a finite time). It should be noted that $L_s/L_n = 20$ is also approximately the maximum shear in figure 4 for which transient linear amplification is observed for random initial conditions. (While absolute linear instability was found in Landreman *et al.* (2015) for $L_s/L_n \geqslant 17$ in the limit of vanishing collisionality, here the finite collisionality is enough to suppress this linear instability.) Before presenting nonlinear simulations, we first illustrate the linear behaviour of the system using all the physical and numerical parameters that will be used for the nonlinear runs. Figure 6 shows the time evolution of W for all the dealiased k_y modes that will be used for nonlinear simulations, including twist-and-shift boundary conditions, but with the nonlinearity turned off, and using the random initial conditions described in § 3. More precisely, the quantities plotted are W_{k_x,k_y}, summed over the k_x that link to $k_x = 0$, where

$$W_{k_x,k_y} = L_x L_y \sum_s \int \mathrm{d}z \int \mathrm{d}^3 v \frac{T_s}{2f_{Ms}} |\delta f_{s,k_x,k_y}|^2 \tag{4.1}$$

are the 2D Fourier contributions to $W = \sum_{k_x,k_y} W_{k_x,k_y}$. Here, L_x and L_y are the sizes of the simulation domain in x and y, and $\delta f_s = \sum_{k_x,k_y} \delta f_{s,k_x,k_y} \exp(\mathrm{i}k_x x + \mathrm{i}k_y y)$. Curves for the other k_x in the simulation look similar to those in figure 6, as expected from the z-translation symmetry of the system (2.1)–(2.2); see the discussion following (2.3). The amplitude of each of these curves decreases for large t, confirming that there is no instability for these physical and numerical parameters. The decrease is mostly monotonic, but there is typically a small amount of transient growth by a factor of $\lesssim 2\times$ for $k_y\rho_i$ between 1 and 2. Thus, this point in parameter space is on the edge of the region in which transient linear amplification occurs, as anticipated from figure 4.

Throughout this section, the numerical parameters used are as follows: 12×2 grid points in pitch angle and $\mathrm{sgn}(v_\parallel)$, 12 grid points in energy, 24 grid points per cell in z (with cells at $k_y = j\Delta k_y$ linked if separated in k_x by $4j\Delta k_x$ for each integer $j \geqslant 1$), Courant–Friedrichs–Lewy (CFL) number 0.5, 256×64 modes in x and y before

https://doi.org/10.1017/S0022377815000495 Published online by Cambridge University Press

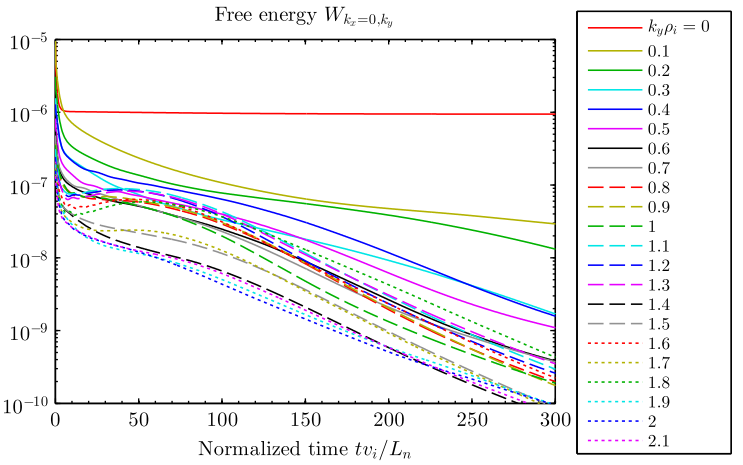


FIGURE 6. Linear dynamics for the parameters used in the nonlinear simulations of figure 7. The large-t behaviour demonstrates that there is no absolute linear instability. A small amount of transient amplification can be seen.

dealiasing and box size $(L_x, L_y, L_z) = (31.4\rho_i, 62.8\rho_i, 160L_n)$. It was verified that both the linear and nonlinear results reported in this section were not significantly altered under factor-of-2 changes in any of these numerical parameters.

One noteworthy linear property apparent in figure 6 is that damping of modes with $k_y = 0$ is very weak, nearly invisible on the scale of the figure. Linear modes with $k_y = 0$ and $\partial/\partial z = 0$ are damped only by collisions. (In the absence of any collisions, any h_e and h_i that depend only on $\mathscr{E}_s^o$ and μ are steady-state solutions of the linearized system.) Collisions introduce damping of these $k_y = 0$ modes, with a damping rate that increases with k_x^2 due to the classical transport terms in the gyroaveraged collision operator.

Finally, figure 7 shows two nonlinear simulations for the same parameters as figure 6, differing only in the initial amplitude of the $k_y = 0.1$ part of the perturbation. For sufficiently large initial amplitude, turbulence is sustained, demonstrating nonlinear instability.

A variety of strategies can be used to initiate the nonlinear instability in numerical simulations. Large-amplitude noise can be used, but this method is numerically inconvenient (at least in gs2) due to the rapid drop in amplitude of strongly damped modes over the first few time steps, which causes significant changes in the CFL-stable time step associated with the perpendicular $\boldsymbol{E} \times \boldsymbol{B}$ speed, thereby requiring recomputation of the implicit time-advance operator. Another method for initiating the nonlinear instability is to begin with a sufficient temperature gradient to bring about linear instability, and then to reduce the temperature gradient once the turbulence reaches a sufficient amplitude (Drake *et al.* 1995). A third effective method, the one used in figure 7, is to initialize with one of the least damped linear eigenmodes scaled up to appropriate amplitude, with noise in the other modes. We find this method to be particularly convenient since variation in the stable time step is minimized. (In both simulations shown in figure 7, the $k_y\rho_i = 0$ Fourier component is initialized to 0, but this component and the $k_y\rho_i = \pm 0.2$ components grow up quickly for the following reason. Since observable quantities are real, there is a large $k_y\rho_i = -0.1$ Fourier component present in the initial condition which is the complex

https://doi.org/10.1017/S0022377815000495 Published online by Cambridge University Press

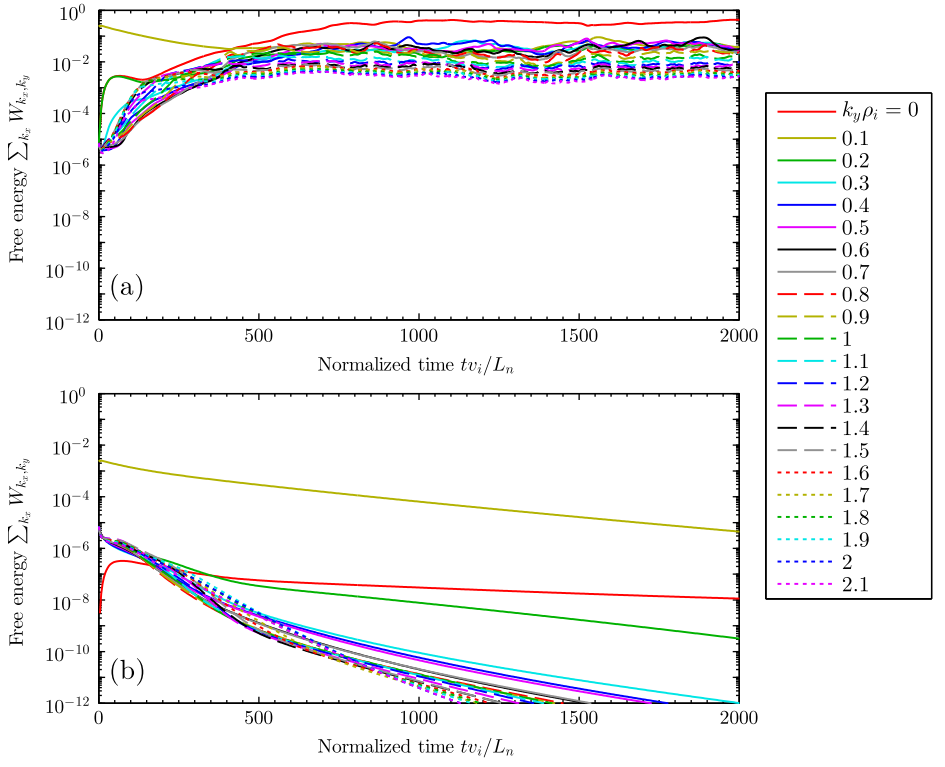


FIGURE 7. Two simulations for exactly the same physical and numerical parameters as the decaying linear simulation in figure (6), but now including nonlinearity and varying the initial amplitude of the $k_y = 0.1$ component. (*a*) When the initial amplitude is sufficient, turbulence grows and is sustained, whereas (*b*) for small initial amplitude, all modes decay as they do in a linear simulation.

conjugate of the large $k_y\rho_i = +0.1$ component shown, and these two components interact nonlinearly to drive components at the sum and difference wavenumbers.)

As another demonstration that the system displays nonlinear instability without linear instability, figure 8 shows the result of turning off the nonlinear term after a period of saturated turbulence. The nonlinear term is turned off at time $tv_i/L_n = 2064$, shown as a vertical dotted line. All modes eventually decay. However, some modes display a modest amount ($\lesssim 3\times$) of transient amplification before decaying.

The particle and heat fluxes for a long nonlinear simulation at these parameters are plotted in figure 9. As the time average of S in (2.4)–(2.6) must vanish, and since the collision term in (2.6) is negative-definite, there must be an average particle flux in the appropriate direction to relax the equilibrium density gradient, and indeed we find such a flux. Although there is no equilibrium temperature gradient, non-zero average fluxes of both ion and electron heat are found. Both heat fluxes have the same sign as the particle flux, meaning that heat flows in the $-\nabla n$ direction.

The fluxes in figure 9 are normalized by gyro-Bohm values defined using the length L_n and including the two in the thermal speed, e.g. the particle flux is normalized by $(\rho_i/L_n)^2 n v_i$. Converting to dimensional units for a deuterium plasma, we obtain the following estimate for the turbulent particle diffusivity: $D \approx (0.4\ \mathrm{m}^2\ \mathrm{s}^{-1})(T/500\ \mathrm{eV})^{3/2}(B/2\ \mathrm{T})^{-2}(L_n/1\ \mathrm{cm})^{-1}$. This magnitude of diffusion

https://doi.org/10.1017/S0022377815000495 Published online by Cambridge University Press

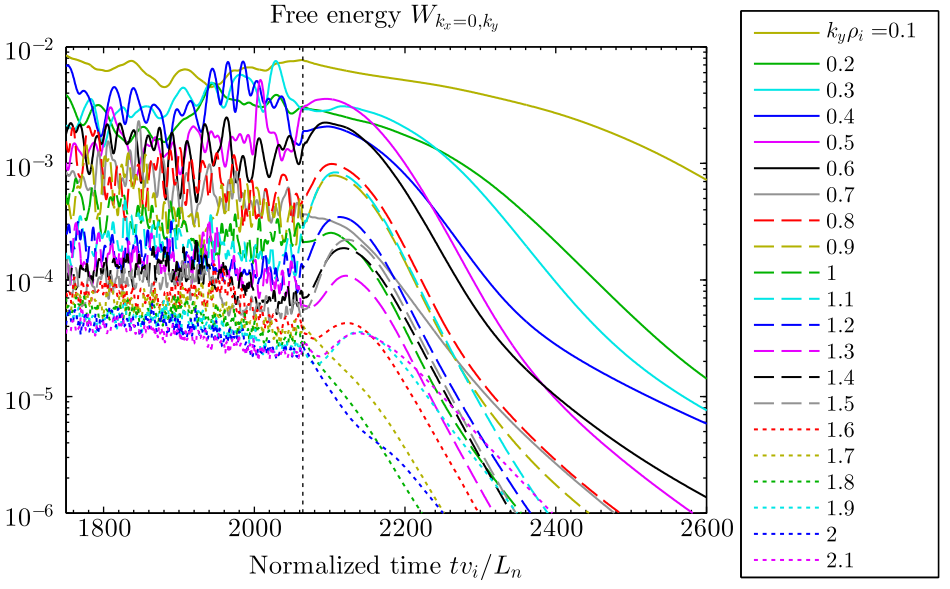


FIGURE 8. Demonstration of the concepts in § 5. The curves show the free energy in the 2D modes with $k_x = 0$ (including intervals linked by the twist-and-shift parallel boundary condition) for various k_y. After a period of saturated turbulence, the nonlinear term is turned off at normalized time 2064, denoted by the vertical dashed line. All modes eventually decay, but transient amplification is apparent first for some of the 2D modes. Such amplification must occur in at least one 2D mode in such a linear 'restart' from steady turbulence, as proved in § 5.

coefficient may be large enough to be of interest experimentally. A larger diffusion coefficient is obtained if L_s/L_n is increased.

As mentioned earlier, for certain values of shear, collisionality and initial perturbation amplitude, a third type of behaviour was sometimes observed besides the two outcomes shown in figure 7: turbulence resembling figure 7(*a*) can persist for a finite time but eventually decay linearly as in figure 7(*b*). This finite-lifetime plasma turbulence is reminiscent of turbulent puffs in neutral (Navier–Stokes) pipe flow. In neutral pipe flow it has been found (Hof *et al.* 2006; Tél & Lai 2008) that in fact the lifetime of turbulence puffs does not become infinite at a critical value of the relevant physical parameter (Reynolds number), but rather the lifetime varies with the Reynolds number exponentially, so the turbulence lifetime can become longer than any reasonable observation time. An interesting open question (which we leave for future research) is whether a similar scaling of the turbulence lifetime may be present in the plasma system we consider here.

5. Turbulence implies transient amplification of typical states

It seems plausible that the previous two sections are related, i.e. that a connection exists between transient linear amplification and sustained nonlinear turbulence. Let us now prove this relationship rigorously. Here, we present two such arguments, both of which yield the conclusion that non-modal amplification is a necessary condition for sustained turbulence. One argument for the case of hydrodynamics was given in DelSole (2004*b*) and § 4 of DelSole (2004*a*), and below we adapt this reasoning

https://doi.org/10.1017/S0022377815000495 Published online by Cambridge University Press

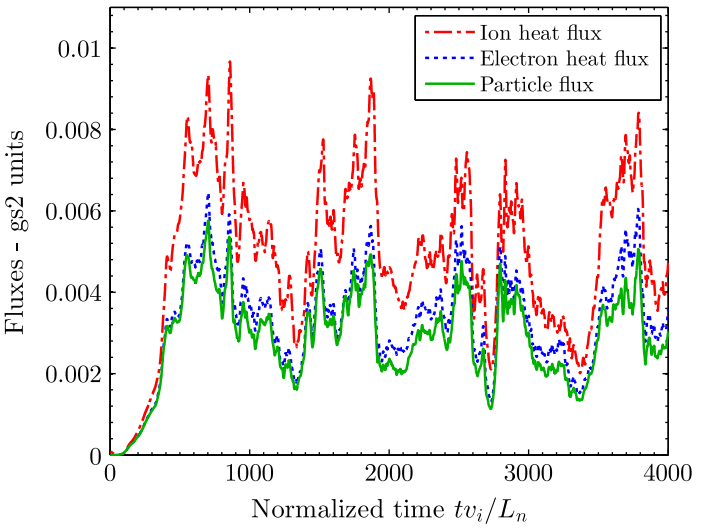


FIGURE 9. Fluxes for the simulation of figure 7(*a*) with $\nu_{ee}L_n/v_i = 0.05$ and $L_s/L_n = 20$, approximately the maximum shear at which turbulence can be sustained at this collisionality. Even in the absence of equilibrium temperature gradients, non-zero average ion and electron heat fluxes are obtained.

to gyrokinetics. Before we do so, we give an alternative argument that relies less on linear algebra. For this discussion, it is useful to define a '2D mode' to be any structure with given k_x and k_y, but with no restriction on the z dependence (so that the structure need not be an eigenmode).

Statement 1. Consider a turbulent solution of the nonlinear gyrokinetic system (2.1)–(2.2), i.e. a solution for which W does not decay to zero at large t. Choose any time t_0 at which the free energy satisfies $\mathrm{d}W/\mathrm{d}t > 0$; such times must occur in steady turbulence as W fluctuates about its mean value. Suppose that at t_0 we turn off the nonlinear term and re-evolve the system in time, as in figure 8. Then, at least one 2D mode must grow in the W norm for some time following t_0, i.e. there must be transient linear amplification.

Proof. Recall that (2.4) is a moment of both the nonlinear and linear dynamical equations. When the nonlinear term is turned off at t_0, $S(h_i, h_e)$ does not immediately change, because the system state does not immediately change. Therefore, $\mathrm{d}W/\mathrm{d}t$ must vary continuously when the nonlinear term is turned off, and so $\mathrm{d}W/\mathrm{d}t$ is positive for some period following t_0 under the action of the linear dynamics. Recall that W is a sum of positive contributions (4.1) from each 2D mode: $W = \sum_{k_x,k_y} W_{k_x,k_y}$. Therefore $\mathrm{d}W_{k_x,k_y}/\mathrm{d}t$ must be positive in the linear dynamics for at least one 2D mode, so this mode must be growing transiently in the W norm. □

Roughly speaking, the upshot of this argument is that we should not be surprised that there is transient amplification after the nonlinear term is turned off in figure 8. Such amplification must occur. Put another way, a necessary condition for turbulence (in the absence of linear instability) is that typical states of the turbulence can be transiently amplified. Furthermore, the energy growth in the 2D modes that are amplified must be sufficient to balance the energy decay from the damped 2D modes.

https://doi.org/10.1017/S0022377815000495 Published online by Cambridge University Press

We note that while $\mathrm{d}W/\mathrm{d}t$ is continuous when the nonlinear term is turned off, each $\mathrm{d}W_{k_x,k_y}/\mathrm{d}t$ is generally not. The reason is that in order to annihilate the nonlinear term in the derivation of (2.4)–(2.6), it is necessary to integrate over x and y, i.e. sum over k_x and k_y. Thus, while (2.4) is identical for the nonlinear and linear dynamics, the analogous equation for $\mathrm{d}W_{k_x,k_y}/\mathrm{d}t$ is different for the nonlinear and linear dynamics. Removal of the nonlinear term therefore causes a discontinuous change in $\mathrm{d}W_{k_x,k_y}/\mathrm{d}t$.

Next, we present a second argument for the necessity of transient amplification. This argument is based on linear algebra and closely follows DelSole (2004*b*). For this second approach, we let $\boldsymbol{h}$ denote a column vector that specifies the state of the system $\{h_i, h_e\}$ in some convenient basis. For example, $\boldsymbol{h}$ could represent the array of values specifying the distribution function in a code.

Statement 2. Given a set of physical parameters for which the nonlinear gyrokinetic system (2.1)–(2.2) permits statistically steady turbulence, there exists at least one state $\boldsymbol{h}_0$ such that (1) $\mathrm{d}W/\mathrm{d}t > 0$ if $\boldsymbol{h} = \boldsymbol{h}_0$ and (2) the squared projection of the turbulent state $\boldsymbol{h}(t)$ onto $\boldsymbol{h}_0$, defined below, has a non-zero average.

Proof. Let $\boldsymbol{L}$ denote the time-independent linear operator associated with (2.1)–(2.2) in the absence of the nonlinear term, so the linearized (2.1)–(2.2) can be written

$$\mathrm{d}\boldsymbol{h}/\mathrm{d}t = \boldsymbol{L}\boldsymbol{h}. \tag{5.1}$$

(Here we will take $\boldsymbol{L}$ to be real since (2.1)–(2.2) are real, although the following argument may be generalized to complex $\boldsymbol{L}$ by replacing the transpose operation with Hermitian conjugation in the steps that follow.) Moreover, let $\boldsymbol{E}$ be the time-independent operator that defines the free-energy norm:

$$W = \boldsymbol{h}^{\mathrm{T}}\boldsymbol{E}\boldsymbol{h}, \tag{5.2}$$

where $^{\mathrm{T}}$ denotes the transpose. It should be noted that $\boldsymbol{E}$ is real and symmetric ($\boldsymbol{E} = \boldsymbol{E}^{\mathrm{T}}$). Differentiating (5.2),

$$\mathrm{d}W/\mathrm{d}t = \boldsymbol{h}^{\mathrm{T}}[(\boldsymbol{E}\boldsymbol{L})^{\mathrm{T}} + \boldsymbol{E}\boldsymbol{L}]\boldsymbol{h}, \tag{5.3}$$

which is equivalent to (2.4). As discussed previously, the presence of the nonlinear term does not alter (2.4) (since the nonlinear term is annihilated in the integration over the perpendicular directions X_s and Y_s), so (5.3) holds for both the linear and nonlinear gyrokinetic equations. Introducing a time average $\langle\cdots\rangle$ over statistically steady turbulence, then

$$\langle \boldsymbol{h}^{\mathrm{T}}[(\boldsymbol{E}\boldsymbol{L})^{\mathrm{T}} + \boldsymbol{E}\boldsymbol{L}]\boldsymbol{h}\rangle = 0. \tag{5.4}$$

Next, observe that the operator $(\boldsymbol{E}\boldsymbol{L})^{\mathrm{T}} + \boldsymbol{E}\boldsymbol{L}$ in (5.3)–(5.4) is real and symmetric, so it possesses a complete set of orthogonal eigenvectors with real eigenvalues. Let $\boldsymbol{u}_j$ and λ_j denote these eigenvectors (normalized so that $\boldsymbol{u}_j^{\mathrm{T}}\boldsymbol{u}_k = \delta_{j,k}$) and their associated eigenvalues. In the hydrodynamics literature, the eigenvectors $\boldsymbol{u}_k$ are sometimes called instantaneous optimals; the $\boldsymbol{u}_j$ with maximum λ_j is the fluctuation that maximizes the instantaneous $\mathrm{d}W/\mathrm{d}t$. We may decompose the state vector in the $\boldsymbol{u}_j$ basis: $\boldsymbol{h}(t) = \sum_j a_j(t)\boldsymbol{u}_j$ for some amplitudes $a_j(t)$. Use of this decomposition in (5.4) gives

$$\sum_j \lambda_j \langle |a_j(t)|^2\rangle = 0. \tag{5.5}$$

Neglecting the uninteresting case in which all terms in the sum (5.5) are zero, there must be at least one positive eigenvalue λ_{j_0} with an associated non-zero average weight $\langle |a_{j_0}(t)|^2\rangle$. Considering the associated $\boldsymbol{u}_{j_0}$ in (5.3), this state $\boldsymbol{u}_{j_0}$ must give positive $\mathrm{d}W/\mathrm{d}t$, proving statement 2. □

https://doi.org/10.1017/S0022377815000495 Published online by Cambridge University Press

Thus, we have shown in two ways that transient linear amplification is a necessary condition for turbulence in the associated nonlinear system. We have also shown that the transiently growing state(s) must have a significant amplitude in the turbulent state, in the following senses. In the first line of argument, not only must there be at least a single growing 2D mode, but there must be enough growing 2D modes of sufficient amplitude to balance the negative $\mathrm{d}W_{k_x,k_y}/\mathrm{d}t$ of all the decaying 2D modes. In the second argument, in (5.5) there must not only be at least one state j with an instantaneous growth rate $\lambda_j > 0$, but the amplitudes of states with positive and negative λ_j must be in balance so that $\langle \mathrm{d}W/\mathrm{d}t \rangle = 0$. In a statistically steady state, there must be a balance between energy input from the instantaneously growing linear states and energy extraction from the instantaneously decaying linear states.

Both of the arguments above rely on the fact that the norm W satisfies the same equation (2.4) in both the linear and nonlinear dynamics. Thus, the 'nonlinear invariant' W is a special norm, and the arguments do not generally apply to other norm-like quantities such as $\int \mathrm{d}^3 r |\Phi|^2$ in which one might choose to measure transient amplification. (In fact, for any operator with damped eigenmodes, even if transient amplification exists in one norm, another norm always exists in which transient amplification is impossible. The sum of squared eigenmode coefficients is such a norm (Farrell & Ioannou 1993).)

While the arguments here rely on the gyrokinetic conservation law (2.4), the same reasoning may be applied to other models and systems in which an energy conservation equation holds. For example, equation (6) in Drake *et al.* (1995) gives the appropriate conservation law for the fluid model in that work, and the first line of (6) gives the associated norm to use in place of W.

6. Discussion and conclusions

The 'universal mode' system (slab geometry with a density gradient, electrostatic, no sheared equilibrium flow and weak temperature gradients) is known to have a dramatic change in linear stability as the magnetic shear is raised from zero to small finite values (Antonsen 1978; Ross & Mahajan 1978; Tsang *et al.* 1978; Landreman *et al.* 2015). As the eigenvalues are so sensitive to magnetic shear, we might expect traditional modal analysis to give a misleading picture of typical behaviour, as with other systems that have sensitive eigenvalues (Trefethen & Embree 2005). Indeed, here we have demonstrated that significant transient linear amplification is possible in the drift wave system even when all eigenmodes are decaying. As shown in figure 3, the transient structures grow at the zero-shear growth rate, when evaluated at the dominant $k_\perp$ and $k_\parallel$ at that instant. We have demonstrated that the amount and duration of amplification scale with the magnetic shear length L_s, so these quantities naturally become unbounded in the limit of vanishing shear. The non-modal amplification also decreases with increasing collisionality, particularly at the shortest perpendicular wavelengths, and the amplification tends to be largest at wavelengths near $k_y \rho_i \sim 1$–2. In contrast to more famous examples of non-modal amplification from hydrodynamics (Trefethen *et al.* 1993), and to some recent examples from plasma physics (Highcock *et al.* 2010; Newton *et al.* 2010; Barnes *et al.* 2011; Schekochihin *et al.* 2012), the amplification in the plasma drift wave system here does not depend on sheared equilibrium flow.

We have also demonstrated using nonlinear gyrokinetic simulations that sustained turbulence is possible in this system even when all linear modes are damped. This

https://doi.org/10.1017/S0022377815000495 Published online by Cambridge University Press

phenomenon had been seen previously in fluid models (Scott 1990, 1992; Drake *et al.* 1995). This phenomenon, clearly visible in figure 7, is opposite to the 'Dimits shift' (Dimits *et al.* 2000) seen at different physical parameters, in which nonlinear fluctuations are suppressed despite the presence of linear instability.

In this system without unstable eigenvalues, we find that sustained nonlinear turbulence and transient linear amplification both become possible around $L_s/L_n \gtrsim 20$ (for the collisionality $\nu_{ee}L_n/v_i = 0.05$ and other parameters considered). As discussed in § 5, the fact that these two phenomena appear together is not a coincidence. Transient linear amplification in the free-energy norm W is a necessary condition for statistically steady turbulence, and the turbulent state must have a significant projection onto states that are linearly growing in the W norm. When the nonlinear term is turned off in a simulation of turbulence, as in figure 8, amplification in the W norm must occur in at least one of the 2D modes. The amplitude of the instantaneously growing states (instantaneous optimals) must be significant, enough for the associated input of free energy to balance the states that are decaying.

Considering the nonlinear and linear results together, the existence of transient amplification and turbulence evidently depends on whether the magnetic shear is sufficiently weak that unsheared slab mode dynamics can persist. This geometric property can be controlled independently of the energy source of the turbulence (i.e. the density gradient), in contrast to turbulence driven by sheared flows where the non-uniformity of the flow plays two roles at once: shaping the global structure of the growing structures and providing energy to the turbulence. This feature makes the present system an interesting case study of subcritical turbulence, and we are unaware of other such examples.

It should be noted that the transient linear process emphasized here and secondary instability are not mutually exclusive. Both are required to support turbulence. Some form of secondary instability (as discussed in Drake *et al.* 1995) must be present to transfer energy out of the transiently growing linear structures before these structures begin to decay substantially. However, nonlinear processes cannot directly contribute to the system's free energy W, due to the annihilation of the nonlinear term in the derivation of (2.4).

Several effects that may be important for realistic plasma turbulence were not included in the present work, including toroidicity and trapped particles, and electromagnetic fluctuations. Therefore, further work is required to determine how relevant the nonlinear turbulence considered here is for realistic experimental parameters. However, we at least conclude that linear stability need not preclude the possibility of significant turbulent particle and energy fluxes.

Acknowledgements

This material is based upon work supported by the US Department of Energy, Office of Fusion Energy Science, under Award Numbers DE-FG02-93ER54197 and DE-FC02-08ER54964, and the SciDAC Center for the Study of Plasma Microturbulence. Computations were performed on the Edison system at the National Energy Research Scientific Computing Center, a DOE Office of Science User Facility supported by the Office of Science of the US Department of Energy under Contract No. DE-AC02-05CH11231. Conversations with T. Antonsen, M. Barnes, G. Colyer, J. Drake, E. Highcock, G. Hammett and A. Hassam are gratefully acknowledged.

https://doi.org/10.1017/S0022377815000495 Published online by Cambridge University Press

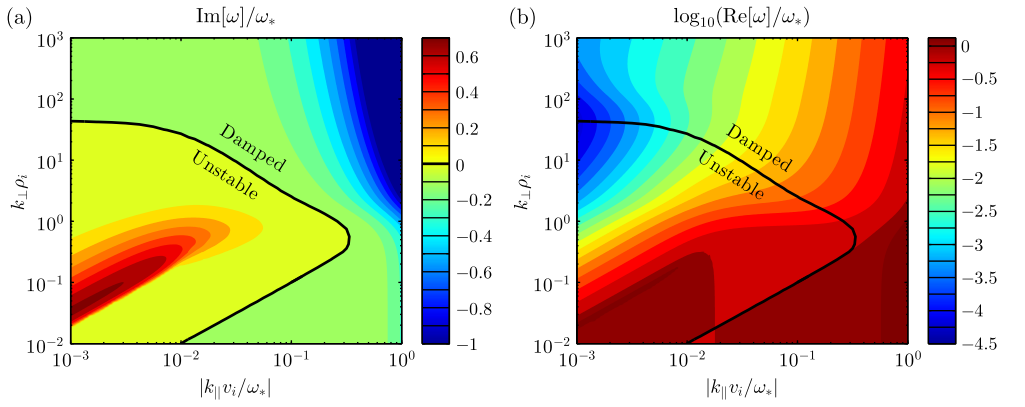


FIGURE 10. Solution of the linear dispersion relation in the zero-shear (local) limit, i.e. the root of (A 1) with largest imaginary part, for $\eta_i = \eta_e = 0$, $T_e = T_i$ and $m_i = 3672 m_e$. The imaginary (*a*) and real (*b*) parts of the frequency are shown.

Appendix A. Linear instability in the absence of magnetic shear

For reference, here we present the growth/damping rates and real frequencies of the universal instability in the limit of vanishing magnetic shear and vanishing collisions. In the absence of shear, we can replace $\partial/\partial z$ in (2.1) with $\mathrm{i}k_\parallel$ for some constant $k_\parallel$. Using (2.2), and after some algebra, the local dispersion relation may then be written

$$\frac{T_e}{T_i} - \zeta_i \frac{\omega_*}{\omega}\left\{-\left[\frac{\omega T_e}{\omega_* T_i}+1\right] Z_i \Gamma_{0i} + \eta_i Z_i \left[\left(\frac{1}{2}+b_i\right)\Gamma_{0i} - b_i \Gamma_{1i}\right] - \eta_i \zeta_i [1+\zeta_i Z_i]\Gamma_{0i}\right\}$$
$$+1+\zeta_e \frac{\omega_*}{\omega}\left\{\left[\frac{\omega}{\omega_*}-1\right] Z_e \Gamma_{0e} + \eta_e Z_e \left[\left(\frac{1}{2}+b_e\right)\Gamma_{0e} - b_e \Gamma_{1e}\right] - \eta_e \zeta_e [1+\zeta_e Z_e]\Gamma_{0e}\right\}$$
$$= 0. \qquad \text{(A 1)}$$

Here, $\zeta_s = \omega/(|k_\parallel| v_s)$, $v_s = \sqrt{2T_s/m_s}$ is the thermal speed, $Z_s = Z(\zeta_s)$, Z is the plasma dispersion function, $\Gamma_{js} = I_j(b_s)e^{-b_s}$, I_j is a modified Bessel function, $b_s = k_\perp^2 v_s^2/(2\Omega_s^2)$ and $\omega_* = k_y T_e/(eBL_n)$. Figure 10 shows the solution of (A 1) with largest imaginary part for the case $T_e = T_i$ and the deuterium–electron mass ratio. We take $\eta_i = \eta_e = 0$, so it is certain that the ITG and ETG modes are suppressed. Figure 10 was generated by applying nonlinear root-finding to (A 1), verifying that the solution agreed with initial-value gs2 simulations for various values of $k_\parallel$ and $k_\perp$. It can be seen that absolute linear instability exists for sufficiently small values of $|k_\parallel|$ when $k_\perp \rho_i < 40$.

Some insight into the density-gradient-driven instability can be obtained by examining the limit $|\zeta_e| \ll 1$ so $Z_e \approx \mathrm{i}\sqrt{\pi}$, and $|\zeta_i| \gg 1$ so $Z_i \approx -1/\zeta_i$. Considering solutions with $|\omega/\omega_*| \ll 1$, one finds

$$\omega = \frac{\omega_* \Gamma_{0i}}{\left(\dfrac{T_e}{T_i}+1\right)^2 + \left(\sqrt{\pi}\dfrac{\omega_*}{|k_\parallel| v_e}\right)^2}\left(\frac{T_e}{T_i}+1+\mathrm{i}\sqrt{\pi}\frac{\omega_*}{|k_\parallel| v_e}\right). \qquad \text{(A 2)}$$

The imaginary part of ω, proportional to $Z_e \approx \mathrm{i}\sqrt{\pi}$, evidently arises due to the electron resonance.

https://doi.org/10.1017/S0022377815000495 Published online by Cambridge University Press

REFERENCES

ABEL, I. G., BARNES, M., COWLEY, S. C., DORLAND, W. & SCHEKOCHIHIN, A. A. 2008 Linearized model Fokker–Planck collision operators for gyrokinetic simulations. I. Theory. *Phys. Plasmas* **15**, 122509, 1–11.

ABEL, I. G., PLUNK, G. G., WANG, E., BARNES, M., COWLEY, S. C., DORLAND, W. & SCHEKOCHIHIN, A. A. 2013 Multiscale gyrokinetics for rotating tokamak plasmas: fluctuations, transport and energy flows. *Rep. Prog. Phys.* **76**, 116201, 1–69.

ANTONSEN, T. M. 1978 Stability of bound eigenmode solutions for the collisionless universal instability. *Phys. Rev. Lett.* **41**, 33–36.

BANÓN NAVARRO, A. 2012 Gyrokinetic large eddy simulations. PhD thesis, Université Libre de Bruxelles.

BARNES, M., PARRA, F. I., HIGHCOCK, E. G., SCHEKOCHIHIN, A. A., COWLEY, S. C. & ROACH, C. M. 2011 Turbulent transport in tokamak plasmas with rotational shear. *Phys. Rev. Lett.* **106**, 175004, 1–4.

BEER, M. A., COWLEY, S. C. & HAMMETT, G. W. 1995 Field-aligned coordinates for nonlinear simulations of tokamak turbulence. *Phys. Plasmas* **2**, 2687–2700.

BISKAMP, D. & ZEILER, A. 1995 Nonlinear instability mechanism in 3D collisional drift-wave turbulence. *Phys. Rev. Lett.* **74**, 706–709.

CAMARGO, S. J., TIPPETT, M. K. & CALDAS, I. L. 1998 Nonmodal energetics of resistive drift waves. *Phys. Rev.* E **58**, 3693–3704.

CAMPOREALE, E., PASSOT, T. & BURGESS, D. 2010 Implications of a non-modal linear theory for the marginal stability state and the dissipation of fluctuations in the solar wind. *Astrophys. J.* **715**, 260–270.

DELSOLE, T. 2004*a* Stochastic models of quasigeostrophic turbulence. *Surv. Geophys.* **25**, 107–149.

DELSOLE, T. 2004*b* The necessity of instantaneous optimals in stationary turbulence. *J. Atmos. Sci.* **61**, 1086–1091.

DIMITS, A. M., BATEMAN, G., BEER, M. A., COHEN, B. I., DORLAND, W., HAMMETT, G. W., KIM, C., KINSEY, J. E., KOTSCHENREUTHER, M., KRITZ, A. H., LAO, L. L., MANDREKAS, J., NEVINS, W. M., PARKER, S. E., REDD, A. J., SHUMAKER, D. E., SYDORA, R. & WEILAND, J. 2000 Comparisons and physics basis of tokamak transport models and turbulence simulations. *Phys. Plasmas* **7**, 969–983.

DRAKE, J. F., ZEILER, A. & BISKAMP, D. 1995 Nonlinear self-sustained drift-wave turbulence. *Phys. Rev. Lett.* **75**, 4222–4225.

FARRELL, B. F. & IOANNOU, P. J. 1993 Stochastic forcing of the linearized Navier–Stokes equations. *Phys. Fluids* A **5**, 2600–2609.

FARRELL, B. F. & IOANNOU, P. J. 1996 Generalized stability theory. Part I: autonomous operators. *J. Atmos. Sci.* **53**, 2025–2040.

FRIEDMAN, B. & CARTER, T. A. 2014 Linear technique to understand non-normal turbulence applied to a magnetized plasma. *Phys. Rev. Lett.* **113**, 205003, 1–5.

FRIEDMAN, B., CARTER, T. A., UMANSKY, M. V., SCHAFFNER, D. & DUDSON, B. 2014 Energy dynamics in a simulation of LAPD turbulence. *Phys. Plasmas* **19**, 102307, 1–12.

FRIEMAN, E. A. & CHEN, L. 1982 Nonlinear gyrokinetic equations for low-frequency electromagnetic waves in general plasma equilibria. *Phys. Fluids* **25**, 502–508.

GALEEV, A. A., ORAEVSKY, V. N. & SAGDEEV, R. Z. 1963 Universal instability of an inhomogeneous plasma in a magnetic field. *Sov. Phys. JETP* **17**, 615–620.

GROSSMAN, S. 2000 The onset of shear flow turbulence. *Rev. Mod. Phys.* **72**, 603–618.

HIGHCOCK, E. G., BARNES, M., SCHEKOCHIHIN, A. A., PARRA, F. I., ROACH, C. M. & COWLEY, S. C. 2010 Transport bifurcation in a rotating tokamak plasma. *Phys. Rev. Lett.* **105**, 215003, 1–4.

HOF, B., WESTERWEEL, J., SCHNEIDER, T. M. & ECKHARDT, B. 2006 Finite lifetime of turbulence in shear flows. *Nature* **443**, 59–62.

HOWES, G. G., COWLEY, S. C., DORLAND, W., HAMMETT, G. W., QUATAERT, E. & SCHEKOCHIHIN, A. A. 2006 Astrophysical gyrokinetics: basic equations and linear theory. *Astrophys. J.* **651**, 590–614.

https://doi.org/10.1017/S0022377815000495 Published online by Cambridge University Press

KOTSCHENREUTHER, M., REWOLDT, G. & TANG, W. M. 1995 Comparison of initial value and eigenvalue codes for kinetic toroidal plasma instabilities. *Comput. Phys. Commun.* **88**, 128–140.

KRALL, N. A. & ROSENBLUTH, M. N. 1965 Universal instability in complex field geometries. *Phys. Fluids* **8**, 1488–1503.

LANDREMAN, M., ANTONSEN, T. M. JR & DORLAND, W. 2015 Universal instability for wavelengths below the ion larmor scale. *Phys. Rev. Lett.* **114**, 095003, 1–5.

NEWTON, S. L., COWLEY, S. C. & LOUREIRO, N. F. 2010 Understanding the effect of sheared flow on microinstabilities. *Plasma Phys. Control. Fusion* **52**, 125001, 1–24.

PLUNK, G. G., HELANDER, P., XANTHOPOULOS, P. & CONNOR, J. W. 2014 Collisionless microinstabilities in stellarators. III. The ion-temperature-gradient mode. *Phys. Plasmas* **21**, 032112, 1–14.

REDDY, S. C. & HENNINGSON, D. S. 1993 Energy growth in viscous channel flows. *J. Fluid Mech.* **252**, 209–238.

ROSS, D. W. & MAHAJAN, S. M. 1978 Are drift-wave eigenmodes unstable? *Phys. Rev. Lett.* **40**, 324–327.

SCHEKOCHIHIN, A. A., HIGHCOCK, E. G. & COWLEY, S. C. 2012 Subcritical fluctuations and suppression of turbulence in differentially rotating gyrokinetic plasmas. *Plasma Phys. Control. Fusion* **54**, 055011, 1–19.

SCHMID, P. J. 2007 Nonmodal stability theory. *Ann. Rev. Fluid Mech.* **39**, 129–162.

SCOTT, B. D. 1990 Self-sustained collisional drift-wave turbulence in a sheared magnetic field. *Phys. Rev. Lett.* **65**, 3289–3292.

SCOTT, B. D. 1992 The mechanism of self-sustainment in collisional drift wave turbulence. *Phys. Fluids* B **4**, 2468–2494.

SQUIRE, J. & BHATTACHARJEE, A. 2014*a* Nonmodal growth of the magnetorotational instability. *Phys. Rev. Lett.* **113**, 025006, 1–5.

SQUIRE, J. & BHATTACHARJEE, A. 2014*b* Magnetorotational instability: nonmodal growth and the relationship of global modes to the shearing box. *Astrophys. J.* **797**, 67, 1–15.

SUGAMA, H., OKAMOTO, M., HORTON, W. & WAKATANI, M. 1996 Transport processes and entropy production in toroidal plasmas with gyrokinetic electromagnetic turbulence. *Phys. Plasmas* **3**, 2379–2394.

TÉL, T. & LAI, Y.-C. 2008 Chaotic transients in spatially extended systems. *Phys. Rep.* **460**, 245–275.

TILLMARK, N. & ALFREDSSON, P. H. 1992 Experiments on transition in plane Couette flow. *J. Fluid Mech.* **235**, 89–102.

TREFETHEN, L. & EMBREE, M. 2005 *Spectra and Pseudospectra: The Behavior of Nonnormal Matrices and Operators*. Princeton University Press.

TREFETHEN, L. N., TREFETHEN, A. E., REDDY, S. C. & DRISCOLL, T. A. 1993 Hydrodynamic stability without eigenvalues. *Science* **261**, 578–584.

TSANG, K. T., CATTO, P. J., WHITSON, J. C. & SMITH, J. 1978 Absolute universal instability is not universal. *Phys. Rev. Lett.* **40**, 327–331.

WALTZ, R. E. 1985 Subcritical magnetohydrodynamic turbulence. *Phys. Rev. Lett.* **55**, 1098–1101.

https://doi.org/10.1017/S0022377815000495 Published online by Cambridge University Press