Phase mixing vs. nonlinear advection in drift-kinetic plasma turbulence

A. A. Schekochihin,[1,2]† J. T. Parker,[3,4] E. G. Highcock,[1,4] P. J. Dellar,[3] W. Dorland[5,2,1] and G. W. Hammett[6,2,1]

[1]Rudolf Peierls Centre for Theoretical Physics, University of Oxford, 1 Keble Road, Oxford OX1 3NP, UK

[2]Merton College, Merton Street, Oxford OX1 4JD, UK

[3]OCIAM, Mathematical Institute, University of Oxford, Andrew Wiles Building, Radcliffe Observatory Quarter, Woodstock Road, Oxford OX2 6GG, UK

[4]Brasenose College, Radcliffe Square, Oxford OX1 4AJ, UK

[5]Department of Physics, University of Maryland, College Park, Maryland 20742, USA

[6]Plasma Physics Laboratory, Princeton University, P. O. Box 451, Princeton, New Jersey 08543, USA

(Received 28 March 2016)

A scaling theory of long-wavelength electrostatic turbulence in a magnetised, weakly collisional plasma (e.g., drift-wave turbulence driven by ion temperature gradients) is proposed, with account taken both of the nonlinear advection of the perturbed particle distribution by fluctuating $\mathbf{E} \times \mathbf{B}$ flows and of its phase mixing, which is caused by the streaming of the particles along the mean magnetic field and, in a linear problem, would lead to Landau damping. It is found that it is possible to construct a consistent theory in which very little free energy leaks into high velocity moments of the distribution function, rendering the turbulent cascade in the energetically relevant part of the wave-number space essentially fluid-like. The velocity-space spectra of free energy expressed in terms of Hermite-moment orders are steep power laws and so the free-energy content of the phase space does not diverge at infinitesimal collisionality (while it does for a linear problem); collisional heating due to long-wavelength perturbations vanishes in this limit (also in contrast with the linear problem, in which it occurs at the finite rate equal to the Landau-damping rate). The ability of the free energy to stay in the low velocity moments of the distribution function is facilitated by the "anti-phase-mixing" effect, whose presence in the nonlinear system is due to the stochastic version of the plasma echo (the advecting velocity couples the phase-mixing and anti-phase-mixing perturbations). The partitioning of the wave-number space between the (energetically dominant) region where this is the case and the region where linear phase mixing wins its competition with nonlinear advection is governed by the "critical balance" between linear and nonlinear timescales (which for high Hermite moments splits into two thresholds, one demarcating the wave-number region where phase mixing predominates, the other where plasma echo does).

1. Introduction

Turbulence is a process whereby energy injected into a system (via some mechanism usually associated with the system being out of equilibrium) is transferred nonlinearly—and therefore leading to chaotic and multiscale states—from the scale(s) at which it is

† Email: alex.schekochihin@physics.ox.ac.uk

injected to much smaller scales at which it is thermalised through microphysical dissipation channels available in the system. The system is forced to seek ways of transferring energy across a range of scales because the injection and dissipation physics are usually unrelated to each other and operate at disparate scales. It is the bridging of the gap between these scales that brings about turbulent cascades, broad-range power-law spectra, and so on. In fluid systems, however varied and multi-physics they are, most turbulence theories are basically extensions and generalisations of the ideas of Richardson (1922) and Kolmogorov (1941*b*) of a local-in-scale cascade maintaining a constant flux of energy away from the injection and towards the dissipation scales (e.g., Zakharov *et al.* 1992; Davidson 2013). This type of thinking has been tremendously successful in making sense of experimental and numerical evidence in both fluids and plasmas.

In plasmas, however, a straightforward application of such "fluid" thinking to any physical regime that is not collisionally dominated skirts over the obvious complication that the kinetic phase space includes the particle velocities as well as their positions, and the (free) energy is generally free to travel across this entire 6D space. Its ability—and propensity—to do so is, in fact, manifest in what is probably the most important phenomenon that makes plasmas conceptually different from fluids—the Landau (1946) damping of electromagnetic perturbations in a collisionless plasma. Viewed in energy terms, it involves the transfer of free energy from electromagnetic perturbations into perturbations of the particle distribution function, which develops ever finer structure in velocity space ("phase mixing") until this transfer (which looks like damping if one only tracks the electromagnetic fields) is made irreversible by coarse graining of the velocity-space structure. The physical agent of this coarse graining is collisions, even if they would appear to be infinitesimally small. Mathematically, the Landau (1936) collision operator is a diffusion operator in velocity space and so even small collision frequencies are enough to thermalise any amount of energy, provided sufficiently large velocity-space gradients develop.

In a *linear* plasma system, Landau damping, or, more generally, phase mixing, is the only available thermalisation route. It provides an adequate mechanism to process any injected free energy at any fixed wave number (since the process is linear, energy will stay in the wave number into which it is injected; there is no coupling), leading to a finite effective damping rate and filling up the phase space with free energy. If one uses a Hermite decomposition to quantify "scales" in velocity space, one finds that, in a steady-state system continuously pumped via low Hermite moments and dissipating free energy via high ones, the free energy will accumulate in phase space to a level that diverges if the collisionality is taken to zero; the collisional heating rate in this limit is finite and equal to the phase-mixing rate (Kanekar *et al.* 2015). How does this mechanism coexist and compete with the refinement of spatial scales caused by coupling between scales—a well-nigh inevitable consequence of nonlinearity?

In this paper, we address this question using a simple archetypal example of plasma turbulence—electrostatic turbulence in a drift-kinetic plasma. We will describe this example in section 2, along with all the relevant preliminaries: the concept of free energy, the Hermite decomposition, and the existing Kolmogorov-style "fluid" turbulence theory for this problem (Barnes *et al.* 2011). In section 3, we will introduce the phase-space formalism that explicitly separates the phase-mixing and the "anti-phase-mixing" perturbations (the latter activated by the plasma echo effect), both of which turn out to be inevitable in a nonlinear system, and provides a useful starting point for a substantive theoretical treatment of phase-space turbulence. In section 4, a phenomenological scaling theory of this turbulence will be proposed. While we will describe in detail how free energy and its fluxes are distributed in the inertial range—leading to some interesting

and testable scalings—the main conclusion will be that phase mixing is quite heavily suppressed in a turbulent system. Section 5 is devoted to summarising this and other findings and to discussing their implications, as well as future directions of travel. A reader only interested in a digest can skip to this section now.

2. Preliminaries

This section contains a rather extended tutorial on a number of topics constituting elementary but necessary background to what will follow. Readers who are sufficiently steeped in these matters can skim through this section and then dedicate themselves more seriously to sections 3 and 4 (where references to relevant parts of section 2 will be supplied).

2.1. *Prototypical kinetic problem*

We consider a plasma near Maxwellian equilibrium, in which case the distribution function for particles of species s can be expressed as

$$f_s = F_{\mathrm{M}s} + \delta f_s, \tag{2.1}$$

where $F_{\mathrm{M}s}$ is a Maxwellian distribution and δf_s a small perturbation.

We assume this plasma to be in a uniform strong magnetic field $\mathbf{B} = B\hat{\mathbf{z}}$ ($\hat{\mathbf{z}}$ is the unit vector in the direction of this field, designated the z axis). We consider low-frequency perturbations, which will be highly anisotropic with respect to the field:

$$\omega \ll \Omega_s, \quad k_\parallel \ll k_\perp, \tag{2.2}$$

where Ω_s is the Larmor frequency.

We assume these perturbations to be electrostatic, viz.,

$$\delta\mathbf{E} = -\boldsymbol{\nabla}\phi, \quad \delta\mathbf{B} = 0, \tag{2.3}$$

where ϕ is the scalar potential. We use Gaussian electromagnetic units.

We consider only long wavelengths,

$$k_\perp \rho_s \ll 1, \tag{2.4}$$

where ρ_s is the Larmor radius.

Finally, we assume a Boltzmann electron response (which arises via expansion of the electron drift-kinetic equation in the electron-to-ion mass ratio):†

$$\frac{e\phi}{T_e} = \frac{\delta n_e}{n_e} = \frac{\delta n_i}{n_i} = \frac{1}{n_i}\int \mathrm{d}^3 v\, \delta f_i, \tag{2.5}$$

where e is the electron charge, T_s and n_s are the equilibrium temperatures and number densities, respectively, and δn_s are density perturbations ($s = e$ for electrons, $s = i$ for ions). The second equality in equation (2.5) is a consequence of plasma quasineutrality. It is useful to denote

$$\varphi = \frac{Ze\phi}{T_i}, \tag{2.6}$$

where Z is the ratio of the ion to electron charge.

Under these assumptions, we may integrate out the dependence of the ion distribution

† The intricacies of the $k_\parallel = 0$ electron response (see section 2.1.1) do not affect the inertial-range theory to be presented here.

function on perpendicular velocities, so we introduce

$$g(t,\mathbf{r},v_\parallel)=\frac{1}{n_i}\int \mathrm{d}^2\mathbf{v}_\perp \delta f_i, \tag{2.7}$$

and write the drift-kinetic equation for g in a 4D phase space:

$$\frac{\partial g}{\partial t}+v_\parallel\nabla_\parallel(g+\varphi F_\mathrm{M})+\mathbf{u}_\perp\cdot\boldsymbol{\nabla}_\perp g=C[g]+\chi, \tag{2.8}$$

$$\varphi=\alpha\int \mathrm{d}v_\parallel g,\quad \alpha=\frac{ZT_e}{T_i}, \tag{2.9}$$

where F_M is a 1D Maxwellian with thermal speed v_th,

$$F_\mathrm{M}=\frac{1}{\sqrt{\pi}}\,e^{-v_\parallel^2/v_\mathrm{th}^2},\quad v_\mathrm{th}=\sqrt{\frac{2T_i}{m_i}}, \tag{2.10}$$

$\mathbf{u}_\perp$ is the $\mathbf{E}\times\mathbf{B}$ drift velocity,

$$\mathbf{u}_\perp=c\,\frac{\delta\mathbf{E}\times\mathbf{B}}{B^2}=\frac{\rho_i v_\mathrm{th}}{2}\,\hat{\mathbf{z}}\times\boldsymbol{\nabla}_\perp\varphi, \tag{2.11}$$

$C[g]$ is the collision operator and χ a source term—both of which need a little further discussion, which we will provide in section 2.1.2.

Note that while we will be referring to "slab" ion-temperature-gradient (ITG) turbulence (e.g., Cowley *et al.* 1991; Ottaviani *et al.* 1997; Horton 1999) as the main physical instantiation that we have in mind of the kinetic problem described above, there will be nothing in our theory that would make it inapplicable to the (inertial range of) electron-temperature-gradient (ETG) turbulence (Dorland *et al.* 2000; Jenko *et al.* 2000), or indeed to a generic case of electrostatic drift-kinetic turbulence with energy injection at long wavelengths.

2.1.1. *A nuance: Boltzmann closure and zonal flows*

In this context, we must come clean on an important detail. The Boltzmann closure (2.5) for the electron density is, in fact, only valid for perturbations with $k_\parallel\neq 0$ because it relies on electrons streaming quickly along the magnetic field lines to short out the parallel electric field. In tokamak plasmas, where magnetic shear imposes a link between $k_\parallel$ and k_y, the Boltzmann closure is normally amended (Dorland & Hammett 1993; Hammett *et al.* 1993) to remove from the electron density the response associated with perturbations that have $k_y=k_\parallel=0$ (the "zonal flows"), namely,

$$\frac{\delta n_e}{n_e}=\frac{e(\phi-\overline{\phi})}{T_e}, \tag{2.12}$$

where $\overline{\phi}$ is the flux-surface average, which in our context is an average over y and z (formally, one gets this by first deriving the density response in a toroidal, magnetically sheared system, as is done, e.g., in §J.2 of Abel & Cowley 2013, then taking the magnetic shear and curvature to be small and passing to the slab limit).

This implies that, in order to find the (y,z)-averaged (zonal) part of φ from the ion distribution function, at least the lowest-order finite-Larmor-radius correction has to be kept (physically representing the polarisation drift; see Krommes 1993, 2010), leading to α in equation (2.9) for the zonal part of φ having to be replaced with $\alpha=2/k_x^2\rho_i^2$, or

$$-\frac{1}{2}\,\rho_i^2\partial_x^2\overline{\varphi}=\int \mathrm{d}v_\parallel\overline{g}. \tag{2.13}$$

In ITG turbulence far from marginal stability, these changes affect important quantitative details of the interaction between zonal flows and drift waves at the outer scale (Rogers *et al.* 2000; see also discussion around equation (2.43)), but do not matter for the inertial-range physics that we will focus on in this paper.

A reader who is unconvinced may observe that equation (2.9) can be used without these modifications if, instead of considering ITG turbulence, we consider ETG turbulence (Dorland *et al.* 2000; Jenko *et al.* 2000). In this case, it is the ions that have a Boltzmann response (due to their large Larmor orbits, over which the density response from electron-scale fluctuations averages out),

$$\frac{\delta n_i}{n_i} = -\frac{Ze\phi}{T_i} \tag{2.14}$$

(which is $= \delta n_e/n_e$ by quasineutrality). The required modifications in equations (2.8) and (2.9) are

$$\varphi \to \frac{e\phi}{T_e}, \quad \alpha \to -\frac{T_i}{ZT_e}, \quad \rho_i \to \rho_e, \quad v_{\mathrm{th}} \to v_{\mathrm{the}} = \sqrt{\frac{2T_e}{m_e}}, \tag{2.15}$$

and $\varphi F_{\mathrm{M}} \to -\varphi F_{\mathrm{M}}$ in equation (2.8). None of this affects anything essential in the upcoming theoretical developments.

2.1.2. *Injection, phase mixing, advection, dissipation*

The precise nature of the source term χ in equation (2.8) will not matter in our theory, as long as it does not contain any sharp dependence on $v_\parallel$ (i.e., is confined to low velocity moments). A random forcing is often a convenient choice for analytical theory (e.g., Plunk 2013; Plunk & Parker 2014; Kanekar *et al.* 2015), but a more physical form in the context of electrostatic drift-kinetic turbulence in plasmas (e.g., Horton 1999) arises from accounting for the presence of equilibrium density and temperature gradients, taken, conventionally, to be in the negative x direction:†

$$\chi = -\frac{\mathbf{u}_\perp \cdot \boldsymbol{\nabla}(n_i F_{\mathrm{M}})}{n_i} = -\frac{\rho_i v_{\mathrm{th}}}{2}\frac{\partial \varphi}{\partial y}\left[\frac{1}{L_n} + \left(\frac{v_\parallel^2}{v_{\mathrm{th}}^2} - \frac{1}{2}\right)\frac{1}{L_T}\right] F_{\mathrm{M}}, \tag{2.16}$$

$$\frac{1}{L_n} = -\frac{1}{n_i}\frac{\mathrm{d}n_i}{\mathrm{d}x}, \quad \frac{1}{L_T} = -\frac{1}{T_i}\frac{\mathrm{d}T_i}{\mathrm{d}x}.$$

We shall see in section 2.3 that these terms render the system linearly unstable and thus extract energy from the equilibrium gradients and inject it into the perturbed distribution.

The resulting perturbations are subject to two influences, linear and nonlinear, encoded by the second ($v_\parallel \nabla_\parallel g$) and fourth ($\mathbf{u}_\perp \cdot \boldsymbol{\nabla}_\perp g$) terms on the left-hand side of equation (2.8), respectively. The nonlinear term represents advection of the distribution function by the mean perpendicular flow, itself determined by the former. This involves coupling between different wave numbers and thus usually leads to spatial mixing (generation of small spatial scales) of the perturbed distribution. The linear term represents phase mixing—generation of small velocity-space scales in the perturbed distribution function. The simplest way to understand this is to notice that the homogeneous solution to the linear kinetic equation in Fourier space, $\partial_t g + iv_\parallel k_\parallel g = \ldots$, is $g \sim e^{-iv_\parallel k_\parallel t}$ and the velocity gradient of that grows secularly with time, $\partial_{v_\parallel} g = -ik_\parallel t g$.

As fine structure in phase space is generated, there must be a means for removing it.

† The erudite reader given pause by 1/2 rather than 3/2 in the prefactor of $1/L_T$ in equation (2.16) will recall that we have integrated out the $v_\perp$ dependence.

This is why, even for a "collisionless" (meaning in fact weakly collisional) plasma, the collision operator $C[g]$ must be included in equation (2.8). We hasten to acknowledge that, in pretending that the collision operator operates purely on g, we are ignoring that the $\mathbf{v}_\perp$ dependence cannot in fact be integrated out of it: collisions will strive to isotropise the distribution and so the collision operator must necessarily couple $\mathbf{v}_\perp$ and $v_\parallel$. However, non-rigorously, when the collision frequency is small,

$$\nu \ll \omega,\ k_\parallel v_{\rm th},\ k_\perp u_\perp, \tag{2.17}$$

the collision operator's essential contribution will be simply to iron out fine structure in velocity space and, given an initial distribution and a source that are smooth in $\mathbf{v}$, only fine structure in $v_\parallel$ can arise. Thus, it should suffice to assume a simple model form for $C[g]$: for example, the Lenard & Bernstein (1958) operator,

$$C[g] = \nu \frac{\partial}{\partial \hat{v}_\parallel}\left(\frac{1}{2}\frac{\partial}{\partial \hat{v}_\parallel} + \hat{v}_\parallel\right) g, \quad \hat{v}_\parallel = \frac{v_\parallel}{v_{\rm th}}. \tag{2.18}$$

The fact that this operator does not conserve momentum or energy, while easily repaired if one strives for quantitatively precise energetics (Kirkwood 1946), will not cause embarrassment as collisions will only matter for high velocity moments (because large gradients with respect to $v_\parallel$ are necessary to offset the smallness of ν). It is not hard to estimate the velocity-space scales at which collisions can become important: balancing $C[g] \sim \omega g$, where $\omega \sim k_\parallel v_{\rm th}$ and/or $k_\perp u_\perp$ is the typical frequency scale of the collisionless dynamics, we find that the requisite velocity scale is

$$\frac{\delta v_\parallel}{v_{\rm th}} \sim \left(\frac{\nu}{\omega}\right)^{1/2}, \tag{2.19}$$

so the structure gets ever finer as $\nu \to +0$.

Finally, we are going to assume implicitly that equation (2.8) contains some regularising term to ensure a cutoff in $k_\perp$—a necessity because of the spatial mixing associated with the nonlinear advection. Physically, the advection term will drive the system out of the domain of validity of the drift-kinetic approximation, to $k_\perp \rho_i \sim 1$ and larger. The precise way in which the energy is thermalised at these Larmor and sub-Larmor scales is a rich and interesting topic in its own right, involving a kinetic cascade in a 5D phase space (with nonlinear phase mixing in $v_\perp$ now also occurring)—but these matters are outside the scope of this treatment (see Schekochihin *et al.* 2008, 2009; Tatsuno *et al.* 2009; Plunk *et al.* 2010; Bañón Navarro *et al.* 2011*b*).

It is the competition between the two ways—linear phase mixing vs. nonlinear advection—of generating small-scale structure in phase space and thus enabling the energy injected by the source to be thermalised that will be the subject of this paper.

2.2. *Free energy*

We have referred to injection and thermalisation of energy many times now, and so defining precisely what we mean by "energy" has become overdue.

Energy in δf kinetics (i.e., in near-equilibrium kinetics) is the free energy associated

with the perturbed distribution:†

$$\mathcal{F} = -\sum_s T_s \delta S_s = \sum_s T_s \, \delta \int \mathrm{d}^3\mathbf{v} \langle f_s \ln f_s \rangle = \sum_s \int \mathrm{d}^3\mathbf{v} \frac{T_s \langle \delta f_s^2 \rangle}{2F_{\mathrm{M}s}}, \tag{2.20}$$

where angle brackets denote spatial averaging and δS_s is the mean additional (negative!) entropy associated with the perturbed distribution of species s. The last expression in equation (2.20) was obtained by letting $f_s = F_{\mathrm{M}s} + \delta f_s$ and expanding $\langle f_s \ln f_s \rangle$ to second order in δf_s (see, e.g., Schekochihin *et al.* 2008; note that $\langle \delta f_s \rangle = 0$ because $\langle f_s \rangle = F_{\mathrm{M}s}$ by definition). It is now not hard to establish that $\mathcal{F} = n_i T_i W$, where

$$W = \int \mathrm{d}v_\parallel \frac{\langle g^2 \rangle}{2F_{\mathrm{M}}} + \frac{\langle \varphi^2 \rangle}{2\alpha} \tag{2.21}$$

is the quadratic quantity conserved by equation (2.8). This can be shown either by using the Boltzmann-electron closure in equation (2.20) (viz., $\delta f_e = (e\phi/T_e) F_{\mathrm{M}e}$, so the $s = e$ term in $\sum_s$ gives rise to the $\langle \varphi^2 \rangle$ term in W) or directly starting from equation (2.8), which gives us the following law of evolution of the free energy:

$$\frac{\mathrm{d}W}{\mathrm{d}t} = \int \mathrm{d}v_\parallel \left(\frac{\langle g\chi \rangle}{F_{\mathrm{M}}} + \langle \varphi\chi \rangle \right) + \int \mathrm{d}v_\parallel \frac{\langle gC[g] \rangle}{F_{\mathrm{M}}}. \tag{2.22}$$

The first term on the right-hand side is the energy-injection term, which turns into the usual flux term for ITG (or ETG) turbulence if we substitute χ from equation (2.16) (see equation (2.37) below), and the second, negative definite, term is the collisional thermalisation of this energy flux.

The Landau damping of the electrostatic perturbations is simply the transfer of free energy, via phase mixing, from the $\langle \varphi^2 \rangle$ part of W to $\langle g^2 \rangle$‡: since $\varphi = \alpha \int \mathrm{d}v_\parallel g$, small-scale velocity-space structure in g is washed out in φ but of course remains as free energy in $\langle g^2 \rangle$ (cf. Hammett & Perkins 1990; Hammett *et al.* 1992).

† The understanding that this is the case can be traced back through a sequence of papers, from early, somewhat forgotten, insights to a more recent surge in appreciation (Kruskal & Oberman 1958; Bernstein 1958; Fowler 1963, 1968; Krommes & Hu 1994; Krommes 1999; Sugama *et al.* 1996; Hallatschek 2004; Howes *et al.* 2006; Candy & Waltz 2006; Schekochihin *et al.* 2008, 2009; Scott 2010; Bañón Navarro *et al.* 2011*b*,*a*; Abel *et al.* 2013; Kunz *et al.* 2015; Parker & Dellar 2015). Note that we have not included in equation (2.20) the energy of the electric and magnetic field, $(\langle E^2 \rangle + \langle \delta B^2 \rangle)/8\pi$ (which is part of the general expression for the free energy; see, e.g., Schekochihin *et al.* 2008) because we are considering electrostatic perturbations ($\delta B = 0$) at scales much longer than the Debye length ($\langle E^2 \rangle$ is negligible).

‡ To be precise, from $\langle \varphi^2 \rangle$ and low-order ("fluid") velocity moments of g to higher-order ("kinetic") moments (see section 2.3.2).

2.3. *Hermite decomposition*

A natural way to separate the "fluid" part of the problem from the "kinetic" one and represent phase mixing is to expand the perturbed distribution in Hermite polynomials:¶

$$g(v_\parallel) = \sum_{m=0}^{\infty} \frac{H_m(\hat{v}_\parallel) F_\mathrm{M}(v_\parallel)}{\sqrt{2^m m!}}\, g_m, \tag{2.23}$$

$$g_m = \int \mathrm{d}v_\parallel \frac{H_m(\hat{v}_\parallel)}{\sqrt{2^m m!}}\, g(v_\parallel), \tag{2.24}$$

where $\hat{v}_\parallel = v_\parallel/v_\mathrm{th}$ and the "physicist's" Hermite polynomials are

$$H_m(\hat{v}_\parallel) = (-1)^m e^{\hat{v}_\parallel^2} \frac{\mathrm{d}^m}{\mathrm{d}\hat{v}_\parallel^m}\, e^{-\hat{v}_\parallel^2}, \quad \int \mathrm{d}v_\parallel \frac{H_m(\hat{v}_\parallel) H_n(\hat{v}_\parallel)}{2^m m!}\, F_\mathrm{M}(v_\parallel) = \delta_{mn}. \tag{2.25}$$

The first three Hermite moments are the (ion) density (δn), mean-parallel-velocity ($u_\parallel$) and parallel-temperature ($\delta T_\parallel$) perturbations:

$$H_0(\hat{v}_\parallel) = 1 \quad \Rightarrow \quad g_0 = \frac{\delta n}{n} = \frac{\varphi}{\alpha}, \tag{2.26}$$

$$H_1(\hat{v}_\parallel) = 2\hat{v}_\parallel \quad \Rightarrow \quad g_1 = \sqrt{2}\,\frac{u_\parallel}{v_\mathrm{th}}, \tag{2.27}$$

$$H_2(\hat{v}_\parallel) = 4\left(\hat{v}_\parallel^2 - \frac{1}{2}\right) \quad \Rightarrow \quad g_2 = \frac{1}{\sqrt{2}}\frac{\delta T_\parallel}{T}. \tag{2.28}$$

Noting further that the source term, equation (2.16), is

$$\chi = -\frac{\rho_i v_\mathrm{th}}{2}\frac{\partial \varphi}{\partial y}\left[\frac{H_0(\hat{v}_\parallel)}{L_n} + \frac{H_2(\hat{v}_\parallel)}{4L_T}\right] F_\mathrm{M} \equiv \left[\chi_0 + \frac{H_2(\hat{v}_\parallel)}{2\sqrt{2}}\,\chi_2\right] F_\mathrm{M} \tag{2.29}$$

and that the streaming term in equation (2.8), $v_\parallel \nabla_\parallel g$, couples Hermite moments of adjacent orders via the formula

$$\hat{v}_\parallel H_m(\hat{v}_\parallel) = \frac{1}{2}\, H_{m+1}(\hat{v}_\parallel) + m H_{m-1}(\hat{v}_\parallel), \tag{2.30}$$

we arrive at the following Hermite representation of equation (2.8):

$$\frac{\partial}{\partial t}\frac{\varphi}{\alpha} + v_\mathrm{th}\nabla_\parallel \frac{u_\parallel}{v_\mathrm{th}} = \chi_0 = -\frac{v_\mathrm{th}}{2L_n}\rho_i \frac{\partial \varphi}{\partial y}, \tag{2.31}$$

$$\left(\frac{\partial}{\partial t} + \mathbf{u}_\perp \cdot \boldsymbol{\nabla}_\perp\right)\frac{u_\parallel}{v_\mathrm{th}} + v_\mathrm{th}\nabla_\parallel\left(\frac{1}{2}\frac{\delta T_\parallel}{T} + \frac{1+\alpha}{\alpha}\,\varphi\right) = 0, \tag{2.32}$$

$$\left(\frac{\partial}{\partial t} + \mathbf{u}_\perp \cdot \boldsymbol{\nabla}_\perp\right)\frac{\delta T_\parallel}{T} + v_\mathrm{th}\nabla_\parallel\left(\sqrt{3}\, g_3 + 2\frac{u_\parallel}{v_\mathrm{th}}\right) = \sqrt{2}\,\chi_2 = -\frac{v_\mathrm{th}}{2L_T}\rho_i\frac{\partial \varphi}{\partial y}, \tag{2.33}$$

and, for $m \geqslant 3$, a universal equation retaining no traces of the temperature-gradient drive or Boltzmann-electron physics:

$$\left(\frac{\partial}{\partial t} + \mathbf{u}_\perp \cdot \boldsymbol{\nabla}_\perp\right) g_m + v_\mathrm{th}\nabla_\parallel\left(\sqrt{\frac{m+1}{2}}\, g_{m+1} + \sqrt{\frac{m}{2}}\, g_{m-1}\right) = -\nu m g_m. \tag{2.34}$$

¶ This has attracted recurring bursts of attention over many years, especially recently (Grad 1949; Armstrong 1967; Grant & Feix 1967; Eltgroth 1974; Crownfield 1977; Hammett *et al.* 1993; Parker & Carati 1995; Smith 1997; Ng *et al.* 1999; Watanabe & Sugama 2004; Zocco & Schekochihin 2011; Black *et al.* 2013; Loureiro *et al.* 2013; Hatch *et al.* 2013, 2014; Plunk & Parker 2014; Kanekar *et al.* 2015; Kanekar 2015; Parker & Dellar 2015).

Note that we have taken advantage of the fact that Hermite polynomials are eigenfunctions of the Lenard–Bernstein operator (2.18), but ignored collisions in the $m = 1$ and $m = 2$ equations (this is allowed because we are assuming $\nu \to +0$ and so collisions will only be important at $m \gg 1$).

2.3.1. *Energy injection: slab ITG instability*

The first three equations are the standard three-field fluid system that describes an ITG-unstable plasma at long wavelengths in an unsheared slab (Cowley *et al.* 1991). The quickest way to obtain the slab ITG instability (Rudakov & Sagdeev 1961; Coppi *et al.* 1967; Cowley *et al.* 1991) is to balance the two terms on the left-hand side of equation (2.31), the first with the third term in equation (2.32), and the first term on the left-hand side with the temperature-gradient term on the right-hand side of equation (2.33). The resulting dispersion relation has three roots, of which one is unstable:

$$\omega^3 \approx \frac{\alpha}{2}(k_\parallel v_{\rm th})^2 \omega_{*T} \quad \Rightarrow \quad \omega \approx \left(-\frac{1}{2} + i\,\frac{\sqrt{3}}{2}\right)\left(\frac{\alpha}{2}\right)^{1/3}(k_\parallel v_{\rm th})^{2/3}\omega_{*T}^{1/3}, \tag{2.35}$$

where $\omega_{*T} = k_y \rho_i v_{\rm th}/2L_T$. This approximation is valid provided $L_n/L_T \gg 1$ and $\omega_{*T} \gg k_\parallel v_{\rm th}$, although, as the growth rate grows with $k_\parallel$, the fastest growth is in fact achieved for $k_\parallel v_{\rm th} \sim \omega_{*T}$, when the dispersion relation is a more complicated and somewhat unedifying equation. At $k_\parallel v_{\rm th} \gg \omega_{*T}$, the ITG mode is replaced by a sound wave, which, in a kinetic system, is heavily Landau damped.†

2.3.2. *Free-energy flows*

The temperature-gradient instability injects energy into the φ, $u_\parallel/v_{\rm th}$ and $\delta T_\parallel/T$ perturbations, all of which are comparable to each other in magnitude when $k_\parallel v_{\rm th} \sim \omega_{*T}$. Because the three-field system is not closed,‡ there is a transfer of energy from $\delta T_\parallel/T$ to higher Hermite moments: the g_3 term in equation (2.33) provides the energy sink from the unstable ("forced") moments and the g_2 term in equation (2.34) at $m = 3$ is the source for the higher moments; the energy thus received by them is eventually thermalised via collisions.

To be more precise about these statements, let us rewrite the free energy (2.21) in terms of Hermite moments:

$$W = \frac{1+\alpha}{2\alpha^2}\langle\varphi^2\rangle + \frac{\langle u_\parallel^2\rangle}{v_{\rm th}^2} + \frac{1}{4}\frac{\langle \delta T_\parallel^2\rangle}{T^2} + \frac{1}{2}\sum_{m=3}^{\infty}\langle g_m^2\rangle. \tag{2.36}$$

† An elementary analysis of the slab ITG dispersion relation can be found, e.g., in Appendix B.2 of Schekochihin *et al.* (2012). Note that, in section 2.4.3, we will argue that the inertial-range fluctuations in fact have $k_\parallel v_{\rm th} \gg \omega_{*T}$, and in section 4, we will show that their Landau damping is suppressed in the nonlinear regime.

‡ The only rigorous way to turn it into a closed system is to assume $\nu \gg \omega$, $k_\parallel v_{\rm th}$, $k_\perp u_\perp$ in equation (2.34), whence $g_m \gg g_{m+1}$ and so the heat flux is expressible in terms of the temperature gradient: $\sqrt{3}\,g_3 \approx (v_{\rm th}/\sqrt{2}\,\nu)\nabla_\parallel g_2 = (v_{\rm th}/2\nu)\nabla_\parallel \delta T_\parallel/T$. Putting this into equation (2.33) gives rise to a parallel heat conduction term. We are not, however, interested in this collisional limit.

Its "fluid" and "kinetic" parts satisfy:¶

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{1+\alpha}{\alpha^2}\langle\varphi^2\rangle+\frac{\langle u_\parallel^2\rangle}{v_{\mathrm{th}}^2}+\frac{1}{4}\frac{\langle\delta T_\parallel^2\rangle}{T^2}\right)=\frac{\langle\delta T_\parallel u_x\rangle}{2TL_T}-\frac{\sqrt{3}\,v_{\mathrm{th}}}{2T}\langle\delta T_\parallel\nabla_\parallel g_3\rangle, \tag{2.37}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{1}{2}\sum_{m=3}^{\infty}\langle g_m^2\rangle=\frac{\sqrt{3}\,v_{\mathrm{th}}}{2T}\langle\delta T_\parallel\nabla_\parallel g_3\rangle-\nu\sum_{m=3}^{\infty}m\langle g_m^2\rangle. \tag{2.38}$$

The first term on the right-hand side of equation (2.37) is the injected energy flux and the second term on the right-hand side of equation (2.38) is the dissipation of that flux by collisions. In steady state, $\mathrm{d}\langle\ldots\rangle/\mathrm{d}t=0$, we must have

$$\langle\delta T_\parallel\nabla_\parallel g_3\rangle\geqslant 0 \tag{2.39}$$

because the collision term is negative-definite in equation (2.38), and, therefore,

$$\langle\delta T_\parallel u_x\rangle\geqslant 0 \tag{2.40}$$

to achieve balance in equation (2.37). The inequality (2.39) implies a non-negative mean energy flux to higher Hermite moments (cf. Krommes & Hu 1994; Nakata *et al.* 2012).

How that flux is processed from being injected at $m=3$ to being dissipated at $m\gg 1$ (assuming $\nu\to+0$) is handled by equation (2.34). This equation contains in a beautifully explicit form the two effects to which this paper is devoted: the phase mixing is manifest in that g_m is coupled to g_{m+1} and g_{m-1}, providing a mechanism for pushing energy to higher m's; simultaneously, all Hermite moments g_m are advected (spatially mixed towards smaller scales) by the same fluctuating velocity $\mathbf{u}_\perp$, determined, via equation (2.11), by the zeroth Hermite moment, $\varphi=\alpha g_0$.

2.4. *"Fluid" turbulence theory*

Barnes *et al.* (2011) proposed a Kolmogorov-style theory of ITG turbulence, essentially ignoring the possibility of a leakage of free energy from the low Hermite moments to the high. While their theory is by no means uncontroversial or the only offering on the market (e.g., Gürcan *et al.* 2009; Plunk *et al.* 2015), it does appear to match the results of numerical experiments (in the strongly unstable regime) and so it is worth both reviewing how it is constructed and examining to what extent it contradicts the statement made in the previous subsection that free energy must leak to higher Hermite moments.

The scaling argument of Barnes *et al.* (2011) addresses two main questions (as would any such argument aspiring to be a complete theory):

(i) what is the effective outer (energy-containing) scale of the turbulence and the fluctuation level at that scale;

(ii) what is the spatial structure of the turbulence in the "inertial range" between that outer scale and the small-scale cutoff?

2.4.1. *Outer scale*

The first question would be trivial for turbulence forced externally at some fixed scale, but for temperature-gradient-driven turbulence, relying on a linear instability, the system must decide where to have its energy-containing scale. This is not simply the peak scale of the growth rate, because at long wave lengths the growth rate will generally grow with wave number less quickly than will the nonlinear cascade rate (as we shall see in

¶ Note that both equations (2.37) and (2.38) will also contain sinks accounting for energy losses at small spatial scales.

section 2.4.2) and so in fact it is the largest scale at which there is an instability that will end up being the energy-containing scale.

Barnes *et al.* (2011) conjectured that this infrared cutoff will be set by the largest *parallel* scale available to fluctuations:

$$k_{\parallel 0} \sim \frac{1}{L_\parallel}, \tag{2.41}$$

where $L_\parallel$ in our idealised homogeneous system is simply the parallel extent of the "box"—in a tokamak, it would be the magnetic connection length between unstable and stable parts of the plasma ($L_\parallel \sim qR$, where q is safety factor and R the major radius). The perpendicular energy-containing scale is then given by

$$\omega_{*T} = k_{y0}\rho_i \frac{v_{\rm th}}{L_T} \sim k_{\parallel 0} v_{\rm th} \quad \Rightarrow \quad k_{y0}\rho_i \sim \frac{L_T}{L_\parallel} \tag{2.42}$$

because the instability would be supplanted by stable (in fact, Landau-damped) sound waves at smaller k_y. Note that we require $L_T/L_\parallel \ll 1$ for the turbulence to occur in a scale range consistent with the drift-kinetic approximation $k_\perp \rho_i \ll 1$. Finally, it is further conjectured that the zonal flows generated by the turbulence will have a typical shearing rate $S_{\rm ZF}$ comparable to the nonlinear decorrelation rate $\tau_{\rm nl0}^{-1}$ at the outer scale (cf. Rogers *et al.* 2000) and, therefore, will isotropise the turbulence:†

$$k_{x0} \sim S_{\rm ZF} k_{y0} \tau_{\rm nl0} \sim k_{y0} \sim k_{\perp 0}. \tag{2.43}$$

This is the only place in the theory where the zonal flows make an appearance as it is assumed that they do not completely dominate the nonlinear dynamics, in contrast to their alleged behaviour in the near-threshold regime (e.g., Dimits *et al.* 2000; Diamond *et al.* 2005, 2011; Gürcan *et al.* 2009; Nakata *et al.* 2012; Ghim *et al.* 2013; Connaughton *et al.* 2014; Makwana *et al.* 2014).

The energy-containing scale given by equation (2.42), the amount of energy it contains is estimated by balancing the rate of injection by instability, ω_{*T}, against the rate $\tau_{\rm nl0}^{-1}$ of nonlinear removal of this energy to smaller scales via advection by the turbulent flow:

$$\omega_{*T} \sim k_{\perp 0}\rho_i \frac{v_{\rm th}}{L_T} \sim \tau_{\rm nl0}^{-1} \sim k_{\perp 0} u_{\perp 0} \sim \rho_i v_{\rm th} k_{\perp 0}^2 \varphi_0 \quad \Rightarrow \quad \varphi_0 \sim \frac{1}{k_{\perp 0} L_T} \sim \frac{\rho_i L_\parallel}{L_T^2}. \tag{2.44}$$

Finally, if one's overriding practical concern is the calculation of the effective heat transport caused by the turbulence, one concludes from the above that the turbulent thermal diffusivity and the heat flux are

$$D_{\rm turb} \sim u_{\perp 0}^2 \tau_{\rm nl0} \sim \frac{u_{\perp 0}}{k_{\perp 0}} \sim \rho_i v_{\rm th} \varphi_0, \quad \Rightarrow \quad Q \sim \frac{n D_{\rm turb} T}{L_T} \sim n \rho_i^2 v_{\rm th} \frac{L_\parallel}{L_T^3}. \tag{2.45}$$

All of this is not particularly sensitive to the fact that, in making the argument that led to equation (2.44), we completely ignored the possibility (in fact, the inevitability)

† It is possible to imagine (or conjecture) variants of drift-wave turbulence in which zonal flows are not strong enough to do this. In such systems, the saturated state at the outer scale is dominated by "streamers," anisotropic structures with $k_{x0} \ll k_{y0}$, whose radial extent is probably determined by the size of the system (Drake *et al.* 1988, 1991; Cowley *et al.* 1991; Rogers *et al.* 1998; Dorland *et al.* 2000; Jenko *et al.* 2000). In order for these structures to survive, they must be immune to the secondary instability that would otherwise give rise to zonal flows, which would in turn break up the streamers (Rogers *et al.* 2000; Quinn *et al.* 2013; Connaughton *et al.* 2014). How a streamer-dominated outer-scale state channels its energy into an inertial-range cascade is not entirely well understood. However, we do not expect that the physics of this inertial range to be much different from that described below.

that some of the energy injected at the outer scale might be removed not by the nonlinear advection, as if the system were purely fluid, but also by the phase mixing towards high m's. The presence of such a transfer (of which there is, in fact, numerical evidence; see, e.g., Watanabe & Sugama 2006; Hatch *et al.* 2011*a*,*b*; Nakata *et al.* 2012) would only break our argument if the rate $\sim k_{\parallel 0} v_{\rm th}$ of this transfer were substantially larger than the nonlinear advection rate and so if the dominant balance were $\omega_{*T} \sim k_{\parallel 0} v_{\rm th} \gg k_{\perp 0} u_{\perp 0}$. But this is obviously impossible as one cannot saturate a linear instability by a linear mechanism: there would not be anything in the theory to determine the saturated amplitude.‡ In view of equation (2.42), the phase mixing rate is, in fact, of the same order as both ω_{*T} and $k_{\perp 0} u_{\perp 0}$. Therefore, it cannot affect the basic scalings—although for the purposes of quantitative transport modelling, it is quite crucial to know by what fraction of order unity it might cut the nonlinear mixing rate, a key preoccupation in the development of "Landau-fluid" closures for plasma turbulence in fusion contexts (Hammett *et al.* 1992, 1993; Dorland & Hammett 1993; Beer & Hammett 1996; Snyder & Hammett 2001*b*).

A question that is much more sensitive to whether phase mixing is nonnegligible is the structure of the inertial range.

2.4.2. *Inertial range: perpendicular spectrum*

How is the energy injected at $(k_{\perp 0}, k_{\parallel 0})$ cascaded to smaller scales? Ignoring phase mixing, Barnes *et al.* (2011) proposed to calculate the dependence of the turbulent amplitudes on scale via the Kolmogorov assumption of constant energy flux: at each scale $k_\perp^{-1}$, energy φ^2 is transferred (locally) to the next smaller scale over the cascade time $\tau_{\rm nl}$:

$$\frac{\varphi^2}{\tau_{\rm nl}} \sim k_\perp u_\perp \varphi^2 \propto k_\perp^2 \varphi^3 = \text{const} \quad \Rightarrow \quad \varphi \propto k_\perp^{-2/3}, \tag{2.46}$$

where we used $u_\perp \propto k_\perp \varphi$ (see equation (2.11)). Note that, both here and in similar arguments that will follow, we do not make a distinction between the energy content of low-m moments, assuming†

$$\varphi \sim \frac{u_\parallel}{v_{\rm th}} \sim \frac{\delta T_\parallel}{T} \tag{2.47}$$

and possibly also $\sim$ a few more low-m moments of g, although we do assume that there is not a substantial energy leakage to asymptotically large m's. The 1D (perpendicular) spectrum is then

$$E_\varphi^\perp(k_\perp) = 2\pi k_\perp \int \mathrm{d}k_\parallel \langle |\varphi_{\mathbf{k}}|^2 \rangle \sim \frac{\varphi^2}{k_\perp} \propto k_\perp^{-7/3}, \tag{2.48}$$

where $\langle \cdots \rangle$ now denotes a time or ensemble average. This scaling is supported both by numerical simulations of Barnes *et al.* (2011) and, apparently, by those done by other groups (Hatch *et al.* 2013, 2014; Plunk *et al.* 2015, who confirm finding the same scaling, without, however, providing plots).

2.4.3. *Critical balance*

The structure of the turbulence in the parallel direction can now be inferred via a causality argument known in the astrophysical MHD literature as "critical balance" (Goldreich & Sridhar 1995, 1997; Boldyrev 2005) and emerging as a universal scaling principle

‡ Again, focusing on turbulence far above the threshold, we are going to ignore the possibility of a more sophisticated scheme involving zonal flows.

† This is because the typical rate for coupling these moments is $k_\parallel v_{\rm th}$, which will shortly be argued to be comparable to the nonlinear rate at which these moments change, equation (2.50).

for strong turbulence in wave-supporting systems (Cho & Lazarian 2004; Schekochihin *et al.* 2009; Nazarenko & Schekochihin 2011): fluctuations cannot stay correlated at parallel distances longer than those over which linear communication happens at the same rate as the nonlinear decorrelation: thus, fluctuations are uncorrelated for

$$k_\parallel v_{\rm th} \lesssim k_\perp u_\perp \propto k_\perp^{4/3} \quad \Rightarrow \quad k_\parallel L_\parallel \lesssim \left(\frac{k_\perp}{k_{\perp 0}}\right)^{4/3}. \tag{2.49}$$

Here and in what follows, we shall adopt a nondimensionalisation $L_\parallel = 1$ and $k_{\perp 0} = 1$, so the above condition will henceforth be written $k_\parallel \lesssim k_\perp^{4/3}$.

This argument implies that, at any given $k_\perp$, the "energy-containing" parallel scale will be given by the "critical-balance" wave number:

$$k_{\parallel {\rm c}} v_{\rm th} \sim k_\perp u_\perp \quad \Rightarrow \quad k_{\parallel {\rm c}} \sim k_\perp^{4/3}, \tag{2.50}$$

another scaling that was confirmed numerically by Barnes *et al.* (2011). The consequent scaling of the 1D parallel spectrum is (using equation (2.50) in equation (2.46))†

$$\varphi \propto k_\parallel^{-1/2} \quad \Rightarrow \quad E_\varphi^\parallel(k_\parallel) = 2\pi \int {\rm d}k_\perp k_\perp \langle |\varphi_{\mathbf{k}}|^2 \rangle \sim \frac{\varphi^2}{k_\parallel} \propto k_\parallel^{-2}. \tag{2.51}$$

Note that, under this scheme, the drift waves are slow in the inertial range because $\omega_{*T} \propto k_y$ whereas $k_\perp u_\perp \propto k_\perp^{4/3}$, so the relevant frequency in equation (2.49) is indeed $\sim k_\parallel v_{\rm th}$, not ω_{*T}. By the same token, energy injection by the temperature-gradient instability is slow compared to the nonlinear cascade rate, so, effectively, the instability only operates at the outer scale, while the fluctuations that carry the injected energy through the inertial range are more akin to ion sound waves than to drift waves.‡

2.4.4. *Constant flux is inconsistent with robust phase mixing*

In section 4.2, we will explain how to derive from these arguments the scaling of the 2D spectra for any $k_\perp$ and $k_\parallel$. However, we must first discuss the key point that the constant-flux assumption (2.46) cannot be consistent with both the idea that the energy resides along the "critical-balance curve" (2.50) *and* with phase mixing taking energy out to large m's at the rate $\sim k_\parallel v_{\rm th}$—simply because the latter would mean that the energy in the low-m moments is not conserved and so need not be fully transferred nonlinearly to smaller scales.

The simplest way to explain the implications of this for the spectra is to replace equation (2.46) by a simple mock-up of an evolution equation for $E_\varphi^\perp(k_\perp)$ (cf. Batchelor

† Another way of arriving at this spectrum and at the critical balance (Beresnyak 2015) is to start with the constant-flux conjecture applied to the scaling of amplitudes with *frequencies*, rather than wavenumbers: $\varphi^2\omega \sim$ const $\Rightarrow \varphi \propto \omega^{-1/2}$ (Corrsin 1963). The frequencies of the perturbations will be $\omega \sim k_\parallel v_{\rm th}$, hence the parallel scaling (2.51).

‡ This also explains why the Barnes *et al.* (2011) cascade should asymptotically override the nonlinear transfer proposed by Gürcan *et al.* (2009): the latter authors argue, effectively, that the cascading of the energy to small scales is done by the nonlocal shearing of the drift waves by zonal flows, which they assume to occur at the rate $\sim k_\perp u_{\rm ZF}$, where $u_{\rm ZF}$ is a scale-independent zonal velocity; this, via a constant-flux argument analogous to (2.46), gives $E_\varphi^\perp \propto k_\perp^{-2}$. However, if the zonal shearing rate is comparable to the energy-injection rate at the outer scale (which we also assume; see equations (2.43) and (2.44): $k_{\perp 0} u_{\rm ZF} \sim S_{\rm ZF} \sim \tau_{\rm nl0}^{-1} \sim \omega_{*T}$), then it will be smaller than $k_\perp u_\perp$ for $k_\perp > k_{\perp 0}$. Note also that a nonlinear transfer rate $\propto k_\perp$ could not effectively dominate the injection rate, ω_{*T}, which is also $\propto k_\perp$.

1953; Howes *et al.* 2008):

$$\frac{\partial E_\varphi^\perp}{\partial t} = -\frac{\partial \varepsilon}{\partial k_\perp} - \gamma E_\varphi^\perp, \quad \varepsilon \sim \frac{k_\perp E_\varphi^\perp}{\tau_{\rm nl}}, \quad \gamma \sim k_\parallel v_{\rm th} \sim \tau_{\rm nl}^{-1} \sim k_\perp^2 \sqrt{k_\perp E_\varphi^\perp}, \tag{2.52}$$

where ε is the energy flux and γ is the effective rate of phase mixing (Landau damping), which, by the critical-balance conjecture (2.50), is of the same order as the cascade rate $\tau_{\rm nl}^{-1}$. Assuming steady state in equation (2.52) and letting $\gamma\tau_{\rm nl} = \xi = \text{const} \sim 1$ (independent of $k_\perp$, as per critical balance), we get

$$\frac{\partial \varepsilon}{\partial k_\perp} = -\frac{\xi}{k_\perp}\,\varepsilon \quad \Rightarrow \quad \varepsilon \propto k_\perp^{-\xi} \quad \Rightarrow \quad E_\varphi^\perp(k_\perp) \propto k_\perp^{-(7+2\xi)/3}. \tag{2.53}$$

Thus, the flux decreases with increasing wave number and so the spectrum is steeper than the constant-flux solution (2.48). The power laws that emerge in such dissipative systems are generally hard to predict and probably nonuniversal (cf. Bratanov *et al.* 2013; Passot & Sulem 2015)—in our case, because they depend on an order-unity prefactor (ξ) in the critical-balance relation (2.50), rather than on some dimensionally and physically inevitable scaling.† However, numerical—or, indeed, experimental—evidence does not appear to support spectra that are significantly steeper than $k_\perp^{-2}$ at long (above the Larmor scale) wavelengths (e.g., Hennequin *et al.* 2004; Görler & Jenko 2008; Casati *et al.* 2009; Vermare *et al.* 2011; Barnes *et al.* 2011; Kobayashi & Gürcan 2015). Furthermore, numerical investigations by Teaca *et al.* (2012, 2014) and Bañón Navarro *et al.* (2014) confirm local nonlinear energy transfer and possibly even constant fluxes, albeit with a number of caveats regarding non-asymptoticity of the simulations, consequent possible non-universality of their results, as well as distinctly measurable, if not dominant, amounts of dissipation (meaning, in their context, phase mixing) everywhere.

In what follows, we shall see that, in a sufficiently collisionless plasma, the constant-flux assumption is safer than it might appear.

2.5. *Hermite "cascade"*

As the last bit of essential background, let us consider what happens with free energy in phase space if we treat phase mixing as the dominant process and ignore nonlinearity—the opposite extreme to that pursued in section 2.4.

Returning to equation (2.34) and dropping the advection term $\mathbf{u}_\perp \cdot \boldsymbol{\nabla}_\perp$ for the time being, we perform a Fourier transform in the parallel direction and introduce the following very useful functions (Zocco & Schekochihin 2011):

$$\tilde{g}_m(k_\parallel) = (i \operatorname{sgn} k_\parallel)^m g_m(k_\parallel), \tag{2.54}$$

† It is easy to see that $\xi < 1$. Indeed, the nonlinear cascade rate that follows from equation (2.53) is $k_\perp u_\perp \propto k_\perp^{(4-\xi)/3}$, which can only overcome the injection rate associated with the temperature gradient if $\xi < 1$ (see discussion at the end of section 2.4.3). The extreme case $\xi = 1$ gives $E_\varphi^\perp \propto k_\perp^{-3}$. One can obtain such a spectrum if one assumes that the fluctuation energy present *at each scale*, not just at the outer scale, is determined by the balance between the instability growth rate, the nonlinear decorrelation rate—and also the phase mixing, which removes the energy to high m's, so there is no need for a constant flux. Then each scale behaves as the outer scale described in section 2.4.1 ($\varphi \propto k_\perp^{-1}$, as in equation (2.44)). We consider this scenario much too fanciful (it would require quite a complicated set of arrangements in the $(k_\perp, k_\parallel)$ space) and rather unlikely for a system far from the threshold. Note also that the restriction $\xi < 1$ would not apply in a system where the energy injection rate is not proportional to $k_\perp$, e.g., one where χ in equation (2.8) is just a large-scale force and so the injection occurs only at the scale of the force. Then the non-universal spectrum (2.53) can be steeper than $k_\perp^{-3}$, although we must have $\xi < 4$ in order for $k_\perp u_\perp$ to increase with $k_\perp$ and so for the nonlinear transfer to stay local. The steepest possible spectrum in this case is, therefore, $E_\varphi^\perp \propto k_\perp^{-5}$.

where $g_m(k_\parallel)$ are the Fourier–Hermite harmonics. The (linearised) equation (2.34) then becomes

$$\frac{\partial \tilde{g}_m}{\partial t} + \frac{|k_\parallel| v_{\rm th}}{\sqrt{2}}(\sqrt{m+1}\,\tilde{g}_{m+1} - \sqrt{m}\,\tilde{g}_{m-1}) = -\nu m \tilde{g}_m. \tag{2.55}$$

The point of these manipulations is that they have made the phase-mixing term on the left-hand side of equation (2.55) look like a derivative with respect to m. Indeed, assuming, in the limit of $m \gg 1$, that we can treat $\tilde{g}_m$ as though it were continuous and differentiable in m (an assumption that will come under close scrutiny in section 3.1), i.e., $\tilde{g}_{m\pm1} \approx \tilde{g}_m \pm \partial_m \tilde{g}_m$, we have

$$\begin{aligned}\sqrt{m+1}\,\tilde{g}_{m+1} - \sqrt{m}\,\tilde{g}_{m-1} &= \sqrt{m}\left(\sqrt{1+\frac{1}{m}}\,\tilde{g}_{m+1} - \tilde{g}_{m-1}\right) \\ &\approx \sqrt{m}\left(\frac{\tilde{g}_m}{2m} + 2\frac{\partial \tilde{g}_m}{\partial m}\right) = 2m^{1/4}\frac{\partial}{\partial m} m^{1/4}\tilde{g}_m. \end{aligned} \tag{2.56}$$

Thus, equation (2.55) becomes

$$\frac{\partial \tilde{g}_m}{\partial t} + \sqrt{2}\,|k_\parallel| v_{\rm th}\, m^{1/4}\frac{\partial}{\partial m} m^{1/4}\tilde{g}_m = -\nu m \tilde{g}_m. \tag{2.57}$$

Introducing the Fourier-Hermite free-energy spectrum $C_m(k_\parallel) = \langle |\tilde{g}_m(k_\parallel)|^2\rangle = \langle |g_m(k_\parallel)|^2\rangle$, we find

$$\frac{\partial C_m}{\partial t} + \frac{\partial}{\partial m}|k_\parallel| v_{\rm th}\sqrt{2m}\, C_m = -2\nu m C_m. \tag{2.58}$$

In steady state, the solution is (Zocco & Schekochihin 2011; cf. Watanabe & Sugama 2004)

$$C_m = \frac{A(k_\parallel)}{\sqrt{m}}\, e^{-(m/m_{\rm c})^{3/2}}, \quad m_{\rm c} = \left(\frac{3|k_\parallel| v_{\rm th}}{2\sqrt{2}\,\nu}\right)^{2/3}, \tag{2.59}$$

where $A(k_\parallel)$ is the constant of integration. Below the collisional cutoff, $m \ll m_{\rm c}$, the power-law scaling $C_m \propto m^{-1/2}$ is the solution corresponding to constant free-energy flux in Hermite space (the Hermite flux is the expression under ∂_m in equation (2.58)). It is possible to show quite rigorously (by direct Hermite transformation of the Landau response function) that this is indeed the Hermite-space solution that arises in a linear system with external forcing at low m and Landau damping (Kanekar *et al.* 2015).

The solution (2.59) has two important properties. Firstly, the free-energy dissipation associated with it (the last term in equation (2.38)) is dominated by Hermite moments with $m \sim m_{\rm c}$ and does not explicitly depend on the collision frequency (assuming $A(k_\parallel)$ does not),

$$D = \nu \sum_m m\langle g_m^2\rangle = \nu \sum_{k_\parallel}\sum_m m C_m(k_\parallel) \approx \nu \sum_{k_\parallel}\int_{\sim 1}^{\infty} {\rm d}m\, m C_m(k_\parallel) = \sum_{k_\parallel}\frac{|k_\parallel| v_{\rm th} A(k_\parallel)}{\sqrt{2}}. \tag{2.60}$$

Secondly, the total amount of free energy stored in the phase space in order to achieve this finite dissipation (corresponding to finite amount of injected power) diverges as $\nu \to 0$:

$$W \approx \frac{1}{2}\sum_{k_\parallel}\int_{\sim 1}^{\infty} {\rm d}m\, C_m(k_\parallel) = \sum_{k_\parallel}\frac{\Gamma(1/3)}{3^{2/3}\sqrt{2}}\frac{A(k_\parallel)}{(|k_\parallel| v_{\rm th})^{2/3}\nu^{1/3}} \tag{2.61}$$

(Kanekar *et al.* 2015).

Thus, if we thought that Landau damping in a turbulent system works in the same

way as it does in a linear one, we might have to conclude that, rather than staying in low m's and being nonlinearly cascaded to small spatial scales, as in a fluid problem, the free energy fills up phase space and dissipates on collisions. A dedicated study of the Hermite spectra of slab ITG turbulence by Hatch *et al.* (2013, 2014) showed that this does *not* happen, with Hermite spectrum of the free energy following a much steeper power law than equation (2.59) and the wavenumber spectrum consistent with equation (2.48). In what follows, we will show how such a solution can emerge (section 4.4.2 has the answer and appendix C the physical basis for it; see section 4.7 for the nonlinear versions of equations (2.61) and (2.60)).

3. Formalism

3.1. *Phase mixing and anti-phase-mixing*

Our first order of business in constructing an appropriate mathematical description for phase-space turbulence is to reexamine our rather blithe assumption in section 2.5 that $\tilde{g}_m(k_\parallel)$, defined by equation (2.54) and satisfying equation (2.55) (to which the nonlinearity will be restored in section 3.2), can be treated as continuous in m.

Consider

$$1 \ll m \ll \left(\frac{|k_\parallel| v_{\rm th}}{\nu}\right)^2. \tag{3.1}$$

If we assume that the rate of change of $\tilde{g}_m$ is small compared to $\sqrt{m}\,|k_\parallel| v_{\rm th}$, equation (2.55) tells us that, to lowest approximation,

$$\sqrt{m+1}\,\tilde{g}_{m+1} - \sqrt{m}\,\tilde{g}_{m-1} = 0 \quad \Rightarrow \quad \tilde{g}_{m+1} \approx \tilde{g}_{m-1}. \tag{3.2}$$

This has two solutions:

$$\tilde{g}_{m+1} \approx \pm \tilde{g}_m, \tag{3.3}$$

so, in fact, either $\tilde{g}_m$ or $(-1)^m \tilde{g}_m$ can be treated as continuous in m. We therefore introduce the following decomposition (which we already used in Kanekar *et al.* 2015 and Parker & Dellar 2015)

$$\tilde{g}_m = \tilde{g}^+_m + (-1)^m \tilde{g}^-_m, \tag{3.4}$$

where

$$\tilde{g}^+_m = \frac{\tilde{g}_m + \tilde{g}_{m+1}}{2}, \quad \tilde{g}^-_m = (-1)^m \frac{\tilde{g}_m - \tilde{g}_{m+1}}{2} \tag{3.5}$$

can both be assumed continuous in m. Evolution equations for these two types of modes can be derived by adding or subtracting evolution equations (2.55) for $\tilde{g}_m$ and $\tilde{g}_{m+1}$ and then expanding in large m in the same fashion as we did in section 2.5. The result is

$$\frac{\partial \tilde{g}^\pm_m}{\partial t} \pm \sqrt{2}\,|k_\parallel| v_{\rm th}\, m^{1/4} \frac{\partial}{\partial m} m^{1/4} \tilde{g}^\pm_m = -\nu m \tilde{g}^\pm_m. \tag{3.6}$$

Manifestly, the "+" modes are the phase-mixing modes, propagating from small to large m, whereas the "−" modes propagate from large to small m and thus represent "anti-phase-mixing": free energy coming back from phase space, a possibility earlier mooted, in somewhat different terms, by Hammett *et al.* (1993) and Smith (1997). We shall discuss the energetics of this process more quantitatively in section 3.4

In a linear problem, in the absence of free-energy sources at high m, the only solution that satisfies the boundary condition $\tilde{g}_{m\to\infty} \to 0$ is $\tilde{g}^-_m = 0$, so there will be no anti-

phase-mixing and the treatment in section 2.5 is correct.† As we are about to see, the situation changes once nonlinearity is accounted for.

3.2. *Nonlinear coupling and plasma echo*

Let us now restore the nonlinear advection (the second term on the left-hand side of equation (2.34)) and Fourier transform it in the parallel direction:

$$\left(\frac{\partial g_m}{\partial t}\right)_{\rm nl} = -\left[\mathbf{u}_\perp \cdot \boldsymbol{\nabla}_\perp g_m\right](k_\parallel) = -\sum_{p_\parallel + q_\parallel = k_\parallel} \mathbf{u}_\perp(p_\parallel) \cdot \boldsymbol{\nabla}_\perp g_m(q_\parallel). \tag{3.7}$$

Then the nonlinear term that must be added to the right-hand side of equation (2.55) is

$$\left(\frac{\partial \tilde{g}_m}{\partial t}\right)_{\rm nl} = -(i\,\mathrm{sgn}\,k_\parallel)^m \left[\mathbf{u}_\perp \cdot \boldsymbol{\nabla}_\perp g_m\right](k_\parallel) = -\sum_{p_\parallel + q_\parallel = k_\parallel} \mathbf{u}_\perp(p_\parallel) \cdot \boldsymbol{\nabla}_\perp \left(\frac{\mathrm{sgn}\,k_\parallel}{\mathrm{sgn}\,q_\parallel}\right)^m \tilde{g}_m(q_\parallel). \tag{3.8}$$

Finally, adding or subtracting the above for the m-th and $(m+1)$-st Hermite moments, and using the decomposition (3.4), we find the nonlinear term for equation (3.6):

$$\frac{\partial \tilde{g}_m^\pm}{\partial t} \pm \sqrt{2}\,|k_\parallel| v_{\rm th}\, m^{1/4} \frac{\partial}{\partial m} m^{1/4} \tilde{g}_m^\pm + \nu m \tilde{g}_m^\pm = \\ -\sum_{p_\parallel + q_\parallel = k_\parallel} \mathbf{u}_\perp(p_\parallel) \cdot \boldsymbol{\nabla}_\perp \left[\delta^+_{k_\parallel, q_\parallel} \tilde{g}_m^\pm(q_\parallel) + \delta^-_{k_\parallel, q_\parallel} \tilde{g}_m^\mp(q_\parallel)\right], \tag{3.9}$$

where $\delta^\pm_{k_\parallel, q_\parallel} = \left[1 \pm \mathrm{sgn}(k_\parallel q_\parallel)\right]/2$, i.e., δ^+ is non-zero (and equals unity) only if $k_\parallel$ and $q_\parallel$ have the same sign, and δ^- is non-zero (and equals unity) only if they have the opposite sign.

The key development manifest in equation (3.9) is that the advecting velocity field can couple parallel wave numbers of opposite signs and thus produce anti-phase-mixing "$-$" modes out of phase-mixing "$+$" ones and vice versa; $\tilde{g}_m^- = 0$ is no longer a valid solution. This is a manifestation of the textbook plasma-physics phenomenon known as plasma echo (Gould *et al.* 1967; Malmberg *et al.* 1968). The importance of it in our discussion is that once the free-energy flux through phase space is not compelled to be unidirectional towards high m's (as it was in the naive treatment of section 2.5), all bets are off as to the effectiveness of Landau damping/phase mixing as a dissipation mechanism in a nonlinear system.

3.3. *Dual kinetic equation in phase space*

Equation (3.9) can be recast in a remarkably simple form if we introduce a change of variables and a rescaling of $\tilde{g}_m^\pm$:

$$s = \sqrt{m}, \quad \tilde{f}(s, k_\parallel) = m^{1/4} \cdot \begin{cases} \tilde{g}_m^+(k_\parallel) & \text{if} \quad k_\parallel \geqslant 0, \\ \tilde{g}_m^-(k_\parallel) & \text{if} \quad k_\parallel < 0. \end{cases} \tag{3.10}$$

† Kanekar *et al.* (2015) showed that in a (forced) linear problem, the spectrum of the "$-$" modes is $\propto m^{-3/2}$ and so subdominant to the spectrum (2.59) of the "$+$" modes. This does not mean that there is some small subdominant amount of anti-phase-mixing in a linear system, but is rather due to the interpretation of $\tilde{g}_m^+$ and $\tilde{g}_m^-$ as being forward and backward propagating modes in m space being correct only to lowest order in $1/m$. Note that this interpretation breaks down also at such large m that the inequality (3.1) is no longer satisfied. When $m \gg (|k_\parallel v_{\rm th}|/\nu)^2$, the collisional term in the right-hand side of equation (2.55) is dominant and the solution is $\tilde{g}_m \approx (|k_\parallel| v_{\rm th}/\nu\sqrt{2m})\tilde{g}_{m-1} \ll \tilde{g}_{m-1}$. Therefore, in this approximation, the two modes are $\tilde{g}_m^+ \approx \tilde{g}_m/2 \approx (-1)^m \tilde{g}_m^-$, and so they formally have the same energy.

For any given $k_\parallel \geqslant 0$, the original distribution function is reconstructed in the following way, via equations (2.54) and (3.4):

$$g_m(k_\parallel) = (-i)^m \left[\tilde{f}(\sqrt{m}, k_\parallel) + (-1)^m \tilde{f}(\sqrt{m}, -k_\parallel)\right], \quad g_m(-k_\parallel) = g_m^*(k_\parallel). \tag{3.11}$$

The new function $\tilde{f}$ satisfies

$$\frac{\partial \tilde{f}}{\partial t} + \frac{k_\parallel v_{\rm th}}{\sqrt{2}} \frac{\partial \tilde{f}}{\partial s} + \nu s^2 \tilde{f} = -\sum_{p_\parallel} \mathbf{u}_\perp(p_\parallel) \cdot \boldsymbol{\nabla}_\perp \tilde{f}(k_\parallel - p_\parallel). \tag{3.12}$$

The echo effect in this equation looks explicitly like mode coupling from positive to negative parallel wave numbers, or vice versa, whereas the phase mixing and anti-phase-mixing are simply propagation in s with velocity $k_\parallel v_{\rm th}/\sqrt{2}$. We will make repeated references to this equation in the scaling arguments of section 4.

Equation (3.12) is a kinetic equation in phase space dual to the original kinetic equation (2.8), with the variable s (or $\sqrt{2}\, s/v_{\rm th}$) effectively acting as a Fourier dual to $v_\parallel$—this is not a huge surprise because for $m \gg 1$, Hermite polynomials are well approximated by trigonometric functions in $v_\parallel$, with "frequency" $\sqrt{2m}/v_{\rm th}$:

$$H_m(\hat{v}_\parallel) e^{-\hat{v}_\parallel^2/2} \approx \sqrt{2} \left(\frac{2m}{e}\right)^{m/2} \cos\left(\frac{\sqrt{2m}}{v_{\rm th}} v_\parallel - \frac{\pi m}{2}\right). \tag{3.13}$$

It is worth stressing that, while the functions $\tilde{g}_m^\pm(k_\parallel)$ are subject to reality conditions, inherited from g_m via $\tilde{g}_m$ (see definitions (2.54) and (3.5)),

$$g_m(-k_\parallel) = g_m^*(k_\parallel) \quad \Rightarrow \quad \tilde{g}_m(-k_\parallel) = \tilde{g}_m^*(k_\parallel) \quad \Rightarrow \quad \tilde{g}_m^\pm(-k_\parallel) = \left[\tilde{g}_m^\pm(k_\parallel)\right]^*, \tag{3.14}$$

the function $\tilde{f}(k_\parallel)$ has no such property because it has been spliced together from the positive-$k_\parallel$ values of $\tilde{g}_m^+$ and the negative-$k_\parallel$ values of $\tilde{g}_m^-$ and there is, *a priori*, no symmetry between the "+" and "−" modes.

Let us reinforce this point by showing that a solution of equation (3.12) can only have the property

$$\tilde{f}(-k_\parallel) = \tilde{f}^*(k_\parallel) \tag{3.15}$$

if the phase mixing is ignorable (this is worth noting because if equation (3.15) does hold, then the free-energy flux in Hermite space vanishes, as per equation (3.27); we will make good use of this argument in section 4.3). Taking the complex conjugate of equation (3.12) and subtracting from it the same equation written for $\tilde{f}(-k_\parallel)$, we get

$$\left(\frac{\partial}{\partial t} + \nu s^2\right)\left[\tilde{f}^*(k_\parallel) - \tilde{f}(-k_\parallel)\right] + \frac{k_\parallel v_{\rm th}}{\sqrt{2}} \frac{\partial}{\partial s}\left[\tilde{f}^*(k_\parallel) + \tilde{f}(-k_\parallel)\right] = \\ -\sum_{p_\parallel} \mathbf{u}_\perp(p_\parallel) \cdot \boldsymbol{\nabla}_\perp \left[\tilde{f}^*(k_\parallel + p_\parallel) - \tilde{f}(-k_\parallel - p_\parallel)\right], \tag{3.16}$$

where have used $\mathbf{u}_\perp^*(p_\parallel) = \mathbf{u}_\perp(-p_\parallel)$ and then changed the summation variable $p_\parallel \to -p_\parallel$ in the sum involving $\tilde{f}^*$. Equation (3.16) is compatible with the condition (3.15) only if the phase mixing term can be ignored—which might happen because $k_\parallel$ is small and/or because $\tilde{f}$ depends on s in such a way that the phase-mixing term is subdominant at, say, high s.

3.4. *Free-energy spectrum and free-energy flux*

Since we are going to discuss free-energy spectra and free-energy fluxes in phase space, let us provide the formal definitions and evolution equations for them.

We define, in the same way as we did in section 2.5,

$$C_m(k_\parallel) = \langle |\tilde{g}_m(k_\parallel)|^2 \rangle = \langle |g_m(k_\parallel)|^2 \rangle. \tag{3.17}$$

Then, using equation (2.55) with the nonlinear term given by equation (3.8), we have

$$\frac{\partial C_m}{\partial t} + \Gamma_m - \Gamma_{m-1} + 2\nu m C_m = 2\mathrm{Re}\left\langle \left(\frac{\partial \tilde{g}_m}{\partial t}\right)_{\mathrm{nl}} \tilde{g}_m^* \right\rangle \equiv \left(\frac{\partial C_m}{\partial t}\right)_{\mathrm{nl}}, \tag{3.18}$$

where the free-energy flux in Hermite space is (cf. Watanabe & Sugama 2004)

$$\Gamma_m(k_\parallel) = \sqrt{2(m+1)}\,|k_\parallel| v_{\mathrm{th}} \mathrm{Re}\langle \tilde{g}_{m+1}\tilde{g}_m^* \rangle = \sqrt{2(m+1)}\,k_\parallel v_{\mathrm{th}} \mathrm{Im}\langle g_{m+1} g_m^* \rangle. \tag{3.19}$$

These expressions are exact. In the limit of $m \gg 1$, both C_m and Γ_m are continuous in m (even if $\tilde{g}_m$ alternates sign, equation (3.3)), so we may rewrite equation (3.18) as follows

$$\frac{\partial C_m}{\partial t} + \frac{\partial \Gamma_m}{\partial m} + 2\nu m C_m = \left(\frac{\partial C_m}{\partial t}\right)_{\mathrm{nl}}. \tag{3.20}$$

Using the definition (3.5) of the "±" modes and defining their spectra

$$C_m^\pm(k_\parallel) = \langle |\tilde{g}_m^\pm(k_\parallel)|^2 \rangle, \tag{3.21}$$

we notice that, still exactly, for any m,

$$\Gamma_m = \sqrt{2(m+1)}\,|k_\parallel| v_{\mathrm{th}} \left(C_m^+ - C_m^-\right). \tag{3.22}$$

Thus, the Hermite flux is *exactly* proportional to the difference between the spectra of the "+" and "−" modes. The sum of these spectra is the free energy, but only approximately, for $m \gg 1$:

$$C_m^+ + C_m^- = \frac{C_m + C_{m+1}}{2} \approx C_m. \tag{3.23}$$

The evolution equations for $C_m^\pm$ at large m are best obtained from equation (3.9):

$$\begin{aligned}\frac{\partial C_m^\pm}{\partial t} \pm |k_\parallel| v_{\mathrm{th}} \frac{\partial}{\partial m}\sqrt{2m}\, C_m^\pm + 2\nu m C_m^\pm = \\ - 2\mathrm{Re}\sum_{p_\parallel + q_\parallel = k_\parallel} \left\langle \left[\tilde{g}_m^\pm(k_\parallel)\right]^* \mathbf{u}_\perp(p_\parallel)\cdot \boldsymbol{\nabla}_\perp \left[\delta^+_{k_\parallel,q_\parallel}\tilde{g}_m^\pm(q_\parallel) + \delta^-_{k_\parallel,q_\parallel}\tilde{g}_m^\mp(q_\parallel)\right]\right\rangle.\end{aligned} \tag{3.24}$$

The sum of these two equations gives us back equation (3.20) with Γ_m given by equation (3.22) (with $m \gg 1$). Another, more compact, way to write equation (3.24) is in terms of the spectrum of the function $\tilde{f}$ introduced in section 3.3. Defining

$$F(s, k_\parallel) = \langle |\tilde{f}(s, k_\parallel)|^2 \rangle, \tag{3.25}$$

we infer from equation (3.12):

$$\frac{\partial F}{\partial t} + \frac{k_\parallel v_{\mathrm{th}}}{\sqrt{2}}\frac{\partial F}{\partial s} + 2\nu s^2 F = -2\mathrm{Re}\sum_{p_\parallel}\left\langle \tilde{f}^*(k_\parallel)\mathbf{u}_\perp(p_\parallel)\cdot\boldsymbol{\nabla}_\perp \tilde{f}(k_\parallel - p_\parallel)\right\rangle. \tag{3.26}$$

Note that, whereas $C_m(k_\parallel)$, $C_m^\pm(k_\parallel)$ and $\Gamma_m(k_\parallel)$ must all be even in $k_\parallel$ because of the reality conditions (3.14), there is no such constraint on $F(k_\parallel)$ and, in fact, it is the odd part of $F(k_\parallel)$ that sets the Hermite flux: in view of equation (3.22),

$$\Gamma_m(k_\parallel) \approx \sqrt{2}\,|k_\parallel| v_{\mathrm{th}} \left[F(s, |k_\parallel|) - F(s, -|k_\parallel|)\right]. \tag{3.27}$$

The next step in the formal solution of the problem is to solve equation (3.26) for $F(s, k_\parallel)$. However, even in principle, this is only possible if a suitable closure is found

for the triple correlator in the right-hand side. A particular solvable model will be discussed in Schekochihin *et al.* (2016), but it will come at the price of decoupling the advecting velocity from the advected distribution function (i.e., considering a "kinetic passive scalar", rather than the fully self-consistent turbulence problem). In general, as always with turbulence problems, we are reduced to (or blessed with) having to resort to phenomenological scaling theories, which we will pursue in the next section.

4. Scaling theory

In constructing the scaling theory for our turbulence in phase space, we shall continue to consider as sensible and valid the arguments in section 2.4.1 that led to estimates of the outer scale (equation (2.42)) and the amplitude of φ at that scale (equation (2.44)). We will, therefore, focus on what happens in the inertial range.

The argument that is presented in this section is quite long because, even within the inertial range, the phase space splits into several regions, where different physics are at work—and building the full picture involves investigating each of these regions and matching free-energy spectra at their boundaries. A road map to what is done where is provided by the subsection headings and by the overall summary in section 5.1, which an impatient reader might find it useful to read first.

In what follows, wherever our expressions appear to be dimensionally incorrect, this is because the wave numbers are normalised to the outer scale:

$$\frac{k_\parallel}{k_{\parallel 0}} = k_\parallel L_\parallel \to k_\parallel, \quad \frac{k_\perp}{k_{\perp 0}} \to k_\perp; \tag{4.1}$$

we will also omit, wherever this makes exposition more rather than less transparent, such dimensional factors as $v_{\rm th}$, ρ_i, etc. We remind the reader that at the outer scale, the parallel-propagation/phase-mixing and the nonlinear-advection time scales are assumed comparable, $k_{\parallel 0} v_{\rm th} \sim k_{\perp 0} u_{\perp 0}$.

4.1. *Spectra in the phase-mixing-dominated region*

The first two terms on the left-hand side of equation (3.12) describe propagation of a perturbation in s with time, along the characteristic

$$s = \frac{k_\parallel v_{\rm th}}{\sqrt{2}}\, t. \tag{4.2}$$

If we consider $k_\parallel > 0$, perturbations will phase-mix in an unfettered way for at least a time comparable to the time it takes the nonlinearity to couple these perturbations to different wave numbers:

$$t \lesssim \tau_{\rm nl} \sim (k_\perp u_\perp)^{-1} \propto k_\perp^{-r}, \tag{4.3}$$

where r is the scaling exponent of the nonlinear decorrelation rate. This means that whatever spectrum, denoted $E_\varphi(k_\parallel, k_\perp)$, prevails at low s (and so low m), it will simply propagate to higher s as long as

$$s \lesssim k_\parallel v_{\rm th} \tau_{\rm nl} \sim \frac{k_\parallel v_{\rm th}}{k_\perp u_\perp} \sim \frac{k_\parallel}{k_\perp^r}. \tag{4.4}$$

We can rearrange this statement to mean that, for any given $m = s^2$, the part of the wave-number space satisfying

$$k_\parallel v_{\rm th} \gtrsim \sqrt{m}\, k_\perp u_\perp \quad \Leftrightarrow \quad k_\parallel \gtrsim \sqrt{m}\, k_\perp^r \tag{4.5}$$

will contain an exact replica of the low-m spectrum:†

$$E_{\tilde{f}}(s, |k_\parallel|, k_\perp) = \sqrt{m}\, E_m^+(k_\parallel, k_\perp) \sim E_\varphi(k_\perp, k_\parallel) \sim k_\perp^b k_\parallel^{-a}. \tag{4.6}$$

Here we have defined 2D spectra

$$\begin{aligned} E_{\tilde{f}}(s, k_\parallel, k_\perp) &= 2\pi k_\perp \langle |\tilde{f}(s, k_\parallel, \mathbf{k}_\perp)|^2 \rangle, \\ E_m^\pm(k_\parallel, k_\perp) &= 2\pi k_\perp \langle |\tilde{g}_m^\pm(k_\parallel, \mathbf{k}_\perp)|^2 \rangle, \\ E_\varphi(k_\parallel, k_\perp) &= 2\pi k_\perp \langle |\varphi(k_\parallel, \mathbf{k}_\perp)|^2 \rangle, \end{aligned} \tag{4.7}$$

where $\varphi(k_\parallel, \mathbf{k}_\perp)$ etc. are Fourier transforms of the original fields in all three spatial directions. We shall refer to the lower bound on $k_\parallel$ (or upper bound on $k_\perp$) defined by the condition (4.5), $k_\parallel \sim \sqrt{m}\, k_\perp^r$, as the *phase-mixing threshold*.

The scaling exponents a and b in equation (4.6) are as yet unknown. One of them, b, can be determined in a purely "kinematic" way: since it describes the low-$k_\perp$ (see equation (4.5)) asymptotic behaviour of the spectrum, it must, in a homogeneous isotropic system, be $b = 3$ (the derivation of this result, which is quite standard, is given in appendix A—it describes the spectrum at perpendicular wavelengths that are longer than the perpendicular correlation scale of perturbations with a given $k_\parallel$).

Thus, we have found a *phase-mixing-dominated region* (as we shall henceforth call it) of the phase space, with spectra

$$E_m^+(k_\parallel, k_\perp) \sim \frac{k_\perp^3 k_\parallel^{-a}}{\sqrt{m}}, \quad E_m^-(k_\parallel, k_\perp) \ll E_m^+(k_\parallel, k_\perp), \quad k_\parallel \gtrsim \sqrt{m}\, k_\perp^r. \tag{4.8}$$

These and all subsequent spectra that will emerge are sketched in figure 2, which the reader is invited to consult for illustration (and preview) of the upcoming results, as they emerge.

Unsurprisingly, in equation (4.8) we have a $1/\sqrt{m}$ Hermite spectrum—the standard linear result already derived in section 2.5. The anti-phase-mixing component of the free energy (E_m^-) must be small compared to the phase-mixing one here because this is the part of phase space where the nonlinearity has no time to exert any influence and so there will not be any echo effect.

While we do not yet know the exponent a (it will be deduced, in two different ways, in sections 4.2 and 4.4), it is clear that $E_m^+(k_\parallel, k_\perp)$ must decay sufficiently fast with $k_\parallel$ in order for the total free energy not to diverge at short parallel wave lengths. This tendency for the free-energy spectrum to decay sharply at parallel wave numbers bounded from below (or, equivalently, at perpendicular wave numbers bounded from above) by the phase-mixing threshold $k_\parallel v_{\rm th} \sim k_\perp u_\perp$ (one might also call this threshold the "phase-space critical balance") was recently reported by Hatch *et al.* (2013, 2014) (cf. Watanabe & Sugama 2004) in their simulations of slab ITG turbulence (they, however, had a different explanation for it).

4.2. *Spectra of low moments*

As we explained in section 4.1, the spectrum (4.8) is inherited (propagated by phase mixing) from low m's, so we must have

$$E_\varphi(k_\parallel, k_\perp) \sim k_\perp^3 k_\parallel^{-a}, \quad k_\parallel \gtrsim k_\perp^r. \tag{4.9}$$

† We assume that there is no discontinuity in the Hermite spectrum at low m, i.e., that the low-s limit of the solution to equation (3.12) (which is technically only valid for $s = \sqrt{m} \gg 1$) will smoothly connect onto the spectra of low-m "fluid" moments φ, $u_\parallel$, $\delta T_\parallel$, etc. and also that the spectra of these quantities all have the same scaling with $k_\parallel$ and $k_\perp$.

Thus, this is the 2D spectrum of the electrostatic turbulence on the short-parallel-wavelength side of the critical-balance condition (2.50).

As we argued in section 2.4.3, the critical balance is essentially a causality condition and so the spectrum at the long-parallel-wavelength side of the critical balance, $k_\parallel \lesssim k_\perp^r$, must reflect the fact that the perturbations at these parallel scales are essentially uncorrelated. The spectrum of such uncorrelated perturbations is the spectrum of white noise, so

$$E_\varphi(k_\parallel, k_\perp) \sim k_\parallel^0 k_\perp^{-c}, \quad k_\parallel \lesssim k_\perp^r. \tag{4.10}$$

Matching this with equation (4.9) along the curve $k_\parallel \sim k_\perp^r$ gives

$$a = \frac{3+c}{r}. \tag{4.11}$$

If $a > 1$, then $k_\parallel \sim k_\perp^r$ is the energy-containing parallel scale for any given $k_\perp$. The 1D perpendicular spectrum is, therefore,

$$E_\varphi^\perp(k_\perp) = \int \mathrm{d}k_\parallel\, E_\varphi(k_\parallel, k_\perp) \sim \int_0^{k_\perp^r} \mathrm{d}k_\parallel\, k_\parallel^0 k_\perp^{-c} \sim k_\perp^{-(c-r)}. \tag{4.12}$$

This immediately implies a consistency relation between c and r:†

$$k_\perp^r \sim k_\perp u_\perp \sim k_\perp^2 \varphi \sim k_\perp^2 (k_\perp E_\varphi^\perp)^{1/2} \quad \Rightarrow \quad r = 5 - c. \tag{4.13}$$

Finally, the 1D parallel spectrum for any given $k_\parallel$ is dominated by $k_\perp \sim k_\parallel^{1/r}$:

$$E_\varphi^\parallel(k_\parallel) = \int \mathrm{d}k_\perp\, E_\varphi(k_\parallel, k_\perp) \sim \int_{k_\parallel^{1/r}}^\infty \mathrm{d}k_\perp\, k_\parallel^0 k_\perp^{-c} \sim k_\parallel^{-(c-1)/r}. \tag{4.14}$$

4.2.1. *Scaling exponents under constant-flux conjecture*

Note that so far, we have invoked no cascade physics, but, in order to determine the exponent c, we do now need to make an assumption as to how energy is passed from scale to scale. Energetically, only the wave-number region $k_\parallel \lesssim k_\perp^r$ matters because at larger $k_\parallel$, the spectrum is assumed (and will be confirmed) to have a steep decay with $k_\parallel$ (equation (4.9)). We shall call it the *advection-dominated region* and anticipate that phase mixing there will not be a significant energy sink, i.e., the anti-phase-mixing energy flux due to the echo effect will on average cancel the phase-mixing flux, leading to effective conservation of $\langle\varphi^2\rangle$. Then we can return to the constant-flux argument of section 2.4.2:

$$\varphi^2 k_\perp u_\perp \sim \text{const} \quad \Rightarrow \quad k_\perp E_\varphi^\perp(k_\perp) \sim \varphi^2 \sim k_\perp^{-r} \quad \Rightarrow \quad c = 1 + 2r, \tag{4.15}$$

where equation (4.12) was used to obtain the last relation. Combining equations (4.15), (4.13) and (4.11), we find

$$r = \frac{4}{3}, \quad c = \frac{11}{3}, \quad a = 5. \tag{4.16}$$

† We remind the reader that φ here and in all similar calculations in this paper is *not* the Fourier transform of the potential, but rather its amplitude corresponding to the scale $k_\perp^{-1}$ (this can be thought of, for example, as the typical magnitude of the potential's increment across a distance $k_\perp^{-1}$). Its relationship to the Fourier transform $\varphi_{\mathbf{k}}$ and to the 1D spectrum $E_\varphi^\perp(k_\perp)$ was given in equation (2.48). This can be understood dimensionally or by noticing that the energy associated with a given scale $k_\perp^{-1}$ is the integral over the energies contained in the wave number $k_\perp$ and larger, $\varphi^2 \sim \int_{k_\perp}^\infty \mathrm{d}k'_\perp E_\varphi^\perp(k'_\perp) \sim k_\perp E_\varphi^\perp(k_\perp)$ (the latter relation holds as long as the 1D spectrum is steeper than $k_\perp^{-1}$).

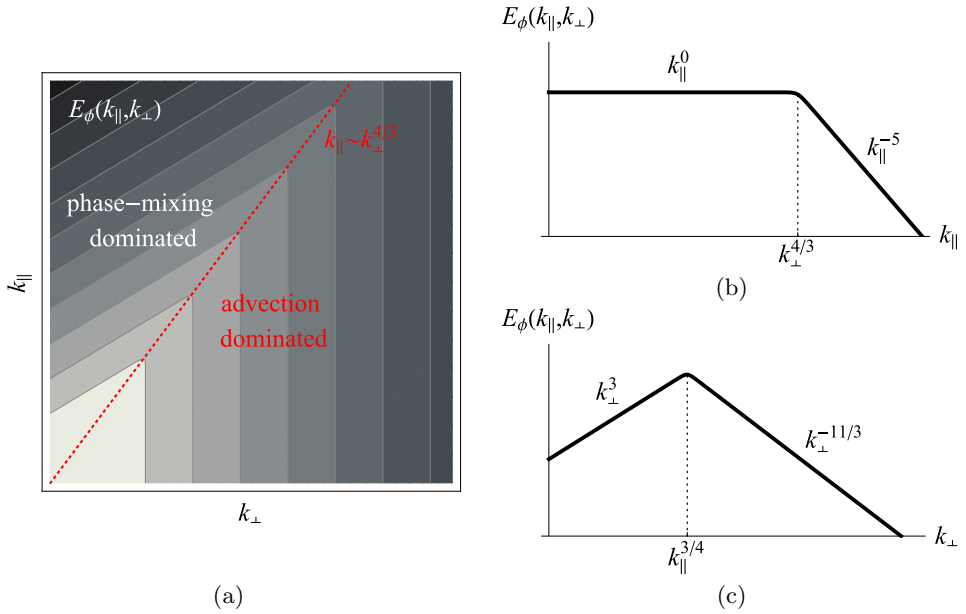

FIGURE 1. Spectra of low Hermite moments, $E_\varphi(k_\parallel, k_\perp)$: (a) in the $(k_\perp, k_\parallel)$ plane, (c) vs. $k_\parallel$ at constant $k_\perp$, (b) vs. $k_\perp$ at constant $k_\parallel$. All plots are logarithmic.

This gives us back the Barnes *et al.* (2011) 1D spectra:

$$E_\varphi^\perp(k_\perp) \sim k_\perp^{-7/3}, \quad E_\varphi^\parallel(k_\parallel) \sim k_\parallel^{-2} \tag{4.17}$$

via equations (4.12) and (4.14), respectively. We have now also learned what the full 2D spectrum behind these 1D ones is: combining equations (4.9) and (4.10) with the scaling exponents (4.16), we have

$$E_\varphi(k_\parallel, k_\perp) \sim \begin{cases} k_\parallel^0 k_\perp^{-11/3} & \text{if} \quad k_\parallel \lesssim k_\perp^{4/3}, \\ k_\perp^3 k_\parallel^{-5} & \text{if} \quad k_\parallel \gtrsim k_\perp^{4/3}. \end{cases} \tag{4.18}$$

These spectra are sketched in figure 1.

Since we have fixed the value of the spectral exponent a and since the spectra of φ in the phase-mixing-dominated region, where this exponent applies, propagate to higher m's, we now also have determined $E_m(k_\parallel, k_\perp)$ for $k_\parallel \gtrsim \sqrt{m}\, k_\perp^r$: see equation (4.8).

Physically, the validity of the argument that led to the last set of results (equation (4.15) onwards) hinges on our ability to produce phase-space spectra that are consistent with substantial cancellation of the phase-mixing flux at $k_\parallel \lesssim k_\perp^r$ and thus with the majority of the free energy residing in the low Hermite moments. Note that it is actually not controversial that the phase mixing should be negligible for $k_\parallel \ll k_\perp^r$ because this means the phase-mixing rate is low compared to the nonlinear advection rate, $k_\parallel v_{\rm th} \ll k_\perp u_\perp$, but, as we saw above, the energy is substantially dominated by the critical-balance curve $k_\parallel \sim k_\perp^r$, where the two rates are comparable (recall our critique of the "fluid" theory in section 2.4.4). In what follows, we shall build a case for the spectra that we have just derived—and so we will carefully avoid using the constant-flux argument (4.15) and keep all the scaling exponents general.

4.3. *Spectra of higher moments in the advection-dominated region*

Let us now consider the higher Hermite moments, $m \gg 1$. The condition for the phase-mixing rate to be negligible compared to the nonlinear advection rate is the opposite of the condition (4.5):

$$k_\parallel v_{\rm th} \ll \sqrt{m}\, k_\perp u_\perp \quad \Leftrightarrow \quad k_\parallel \ll \sqrt{m}\, k_\perp^r \tag{4.19}$$

(i.e., the part of the phase space on the side of the phase-mixing threshold opposite to the phase-mixing-dominated region). In section 4.1, the phase-mixing threshold was derived by arguing that it represented the value of s up to which the "+" perturbations at some given $k_\parallel$ and $k_\perp$ could propagate before being diverted to different wave numbers by the nonlinear coupling. More generally and more formally, we may simply argue that, if, say, the s dependence of $\tilde{f}$ is a power law (which it must be; see appendix C), the size of the phase-mixing term in equation (3.12) can be estimated as $\sim (k_\parallel v_{\rm th}/s)\tilde{f}$, which is negligible compared to the nonlinear term if the condition (4.19) is satisfied.

The advection-dominated region

$$k_\parallel v_{\rm th} \lesssim k_\perp u_\perp \quad \Leftrightarrow \quad k_\parallel \lesssim k_\perp^r, \tag{4.20}$$

which, as we argued in section 4.2, contains most of the energy content of the low-m Hermite moments (collectively represented by φ) and, therefore, of the advecting velocity $\mathbf{u}_\perp$, is well within the domain of validity of the condition (4.19), provided $m \gg 1$. Therefore, if we restrict our attention to the wave numbers (4.20), we may neglect the phase-mixing term (second on the left-hand side) in equation (3.12) and thus deal with what is a purely "fluid" equation for $\tilde{f}(s)$ at each s. As we argued at the end of section 3.3, we can then have solutions satisfying $\tilde{f}^*(k_\parallel) = \tilde{f}(-k_\parallel)$, for which the "+" and "−" spectra are the same and the Hermite flux is zero (see equation (3.27)):

$$E_m^+(k_\parallel, k_\perp) \approx E_m^-(k_\parallel, k_\perp). \tag{4.21}$$

Physically, this is because, in the advection-dominated region, the nonlinear coupling between positive and negative $k_\parallel$ mediated by the velocity field $\mathbf{u}_\perp$ will be vigorous and fast—assuming, importantly, that interactions between $\mathbf{u}_\perp$ and $\tilde{f}$ are local in $k_\parallel$ and so, in the right-hand side of equation (3.12), the sum $\sum_{p_\parallel}$ is dominated by wave-number triads with $p_\parallel \sim k_\parallel \sim k_\parallel - p_\parallel$.

With the phase mixing neglected, the variance of $\tilde{f}$ is (approximately) conserved at each s. The field $\tilde{f}(s)$ is nonlinearly cascaded to smaller scales (larger $k_\perp$) by the advecting velocity $\mathbf{u}_\perp$, so the standard constant-flux argument gives us

$$\tilde{f}^2(s) k_\perp u_\perp \sim \text{const}(s) \quad \Rightarrow \quad \tilde{f}^2(s) \propto k_\perp^{-r} \quad \Rightarrow \quad E_m^\perp(k_\perp) \propto k_\perp^{-r-1}, \tag{4.22}$$

where $E_m^\perp(k_\perp)$ is the 1D perpendicular spectrum of g_m. Note that the spectrum has an m dependence, which cannot be determined via this argument.

Since, in the advection-dominated region (4.20), the parallel-communication times are long compared to the nonlinear-decorrelation times, the perturbations can be expected to have a white-noise spectrum in $k_\parallel$. Therefore, we can write their 2D spectrum as follows

$$E_m(k_\parallel, k_\perp) \sim E_m^\pm(k_\parallel, k_\perp) \sim \frac{k_\parallel^0 k_\perp^{-d}}{m^\sigma}, \quad k_\parallel \lesssim k_\perp^r, \tag{4.23}$$

where we have allowed for an as yet unknown scaling with m. The $k_\perp$-scaling exponent d can be determined via the requirement that equation (4.23) be consistent with equation (4.22): assuming that, for any given $k_\perp$, the region $k_\parallel \lesssim k_\perp^r$ contains most of the

energy,†

$$E_m^{\perp}(k_\perp) = \int \mathrm{d}k_\parallel\, E_m(k_\parallel, k_\perp) \sim \int_0^{k_\perp^r} \mathrm{d}k_\parallel\, \frac{k_\parallel^0 k_\perp^{-d}}{m^\sigma} \propto k_\perp^{-(d-r)} \quad \Rightarrow \quad d = 1 + 2r. \tag{4.24}$$

Since, as we have argued here, the Hermite flux is (approximately) zero in this region of wave-number space for all higher m's, there can be very little net free-energy flow out of low m's—and so, in retrospect, we were justified in assuming in section 4.2 that the energy of the low m's was conserved and so a constant-flux argument (4.15) could be used to deduce the scaling of φ. This allows us to adopt the scaling exponents (4.16) (which we have thus far avoided using), in particular, $r = 4/3$, and so, using equation (4.24), the free-energy spectrum is (see figure 2 for illustration)

$$E_m^{\pm}(k_\parallel, k_\perp) \sim \frac{k_\parallel^0 k_\perp^{-11/3}}{m^\sigma}, \quad k_\parallel \lesssim k_\perp^{4/3}. \tag{4.25}$$

Unsurprisingly, there is continuity between the scalings of $E_m^{\pm}$ and the scaling of E_φ. The scaling exponent σ will be found in section 4.4.1.

Before moving on to complete our scaling theory, we note that jumping to the result (4.25) already in this section was borne of pure impatience: we will discover in section 4.4 that, in fact, it is possible to determine the scaling exponent d without relying on the as yet perhaps somewhat unconvincing claim that a constant-flux argument is legitimate for φ despite the phase mixing being notionally not small along the critical-balance curve $k_\parallel \sim k_\perp^r$.

4.4. *Intermediate region and matching conditions*

We now have the form of the free-energy spectra in two regions, $k_\parallel \gtrsim \sqrt{m}\, k_\perp^r$ (phase-mixing dominated, very little free energy, equation (4.8)) and $k_\parallel \lesssim k_\perp^r$ (advection-dominated cascade, contains most of the free energy, equation (4.25)). It remains to determine the free-energy spectrum in the *intermediate region* between these two:

$$E_m^{+}(k_\parallel, k_\perp) \sim \frac{k_\parallel^{-a'} k_\perp^{-d'}}{m^{\sigma'}}, \quad k_\perp^r \lesssim k_\parallel \lesssim \sqrt{m}\, k_\perp^r \tag{4.26}$$

(the following argument will only apply to the "+" modes; both the reasons for this and the way to determine the spectrum of the "−" modes will be explained in section 4.5).

We have three new scaling exponents, but we also have the requirement to match equation (4.26) with equations (4.8) and (4.23) along the boundaries of the intermediate region. This gives us four relations

$$\sigma' = \sigma, \quad d' + a'r = d, \quad d' + a'r = ar - 3, \quad a' + 2\sigma' = 1 + a, \tag{4.27}$$

which we rearrange so:

$$a' = \frac{d - d'}{r}, \quad d = ar - 3, \quad \sigma = \sigma' = \frac{3 + r + d'}{2r}. \tag{4.28}$$

Let us combine these with two equally uncontroversial (i.e., requiring no leaps of physical intuition) matching and consistency relations from section 4.2: equations (4.11) and (4.13), which we can rewrite as

$$a = \frac{8 - r}{r}, \quad c = 5 - r. \tag{4.29}$$

† Technically speaking, we do not yet know this. We will justify this assumption *a posteriori* in section 4.4.1.

The second of equations (4.28) then gives

$$d = 5 - r = c. \tag{4.30}$$

Thus, the perpendicular scalings of $E_m^{\pm}$ and E_φ must be the same in the advection-dominated region $k_\parallel \lesssim k_\perp^r$. If we bring in equation (4.24), i.e., the constant-flux argument (4.22), we get immediately

$$r = \frac{4}{3}, \quad d = c = \frac{11}{3}, \quad a = 5. \tag{4.31}$$

These are the same as the exponents (4.16), except the need for the "fluid" constant-flux argument (4.15) for φ has now been obviated by the combination of a more solid phase-space argument (4.22) and a number of inevitable consistency relations. Equations (4.28) will give us a', σ' and σ if we know d'. In order to determine the latter, we must now consider why physics in the intermediate phase-space region $k_\perp^r \lesssim k_\parallel \lesssim \sqrt{m}\, k_\perp^r$ should be at all different from what happens in the advection-dominated region $k_\parallel \lesssim k_\perp^r$ considered in section 4.3 (and so why $d' \neq d$).

4.4.1. *Spectra in the intermediate region*

Except at the phase-mixing threshold $k_\parallel \sim \sqrt{m}\, k_\perp^r$ (which is considered with more care in appendix C), phase mixing in the intermediate region is dominated by nonlinear advection. However, interactions between $\mathbf{u}_\perp$ and $\tilde{f}$ cannot, unlike in the advection-dominated region discussed in section 4.3, be local in $k_\parallel$. Indeed, we know from section 4.2 that there is very little energy left in $\mathbf{u}_\perp$ at $k_\parallel \gg k_\perp^r$. Assuming interactions to be local in $k_\perp$, the wave-number sum in the right-hand side of equation (3.12) will be dominated by $p_\parallel \lesssim k_\perp^r$ (see appendix B for a careful analysis of the possible nonlocal interactions in the intermediate region). If $k_\parallel \gg k_\perp^r$, then $p_\parallel \ll k_\parallel$, $\tilde{f}(k_\parallel - p_\parallel) \approx \tilde{f}(k_\parallel)$, and so equation (3.12) now describes the advection of $\tilde{f}(s, k_\parallel)$ by an essentially two-dimensional velocity field (its parallel scale is much longer than that of $\tilde{f}$), with both s and $k_\parallel$ appearing as implicit parameters. This means that the variance of $\tilde{f}$ will be conserved for each individual s and $k_\parallel$†—and this in turn, by yet another constant-flux-in-$k_\perp$ argument, implies

$$\tilde{f}^2(s, k_\parallel) k_\perp u_\perp \sim \mathrm{const}(s, k_\parallel) \quad \Rightarrow \quad \tilde{f}^2(s, k_\parallel) \propto k_\perp^{-r} \quad \Rightarrow \quad E_m(k_\parallel, k_\perp) \propto k_\perp^{-r-1}. \tag{4.32}$$

This scaling is of the 2D spectrum, not of the 1D perpendicular one, because $k_\parallel$ is a fixed parameter, rather than a variable over which there can be any nonlinear coupling. Comparing equations (4.32) and (4.26), we read off d' and hence, with the aid of the first and third equations (4.28), complete the determination of all scaling exponents:

$$d' = r + 1 = \frac{7}{3}, \quad a' = 1, \quad \sigma' = \sigma = \frac{5}{2}. \tag{4.33}$$

Thus, the free-energy spectrum in the intermediate region is

$$E_m^+(k_\parallel, k_\perp) \sim \frac{k_\parallel^{-1} k_\perp^{-7/3}}{m^{5/2}}, \quad k_\perp^{4/3} \lesssim k_\parallel \lesssim \sqrt{m}\, k_\perp^{4/3}, \tag{4.34}$$

sketched in figure 2. The spectrum of the "$-$" modes will be discussed in section 4.5.

† This, incidentally, addresses a possible objection to the arguments in section 4.3 that might have been troubling a perceptive reader: were we really justified in assuming that the conserved variance of $\tilde{f}(s)$ could all be accounted for within the region $k_\parallel \lesssim k_\perp^r$, leading to the constant-flux argument (4.22)? The answer is that the free energy outside that region is either subdominant (section 4.1) or conserved separately (section 4.4.1).

4.4.2. *1D spectra*

In the run up to equation (4.24), we assumed that integrating the 2D spectrum $E_m(k_\parallel, k_\perp)$ with respect to $k_\parallel$ over the advection-dominated region $k_\parallel \lesssim k_\perp^r$ captures most of the free energy contained in any fixed $k_\perp$. Now in possession of equation (4.34), we see that this is not entirely correct: in fact, using now both equations (4.25) and (4.34), we find the 1D perpendicular spectrum to be

$$E_m^{\perp+}(k_\perp) = \int \mathrm{d}k_\parallel\, E_m^+(k_\parallel, k_\perp) \sim \int_0^{k_\perp^{4/3}} \mathrm{d}k_\parallel \frac{k_\parallel^0 k_\perp^{-11/3}}{m^{5/2}} + \int_{k_\perp^{4/3}}^{\sqrt{m}\,k_\perp^{4/3}} \mathrm{d}k_\parallel \frac{k_\parallel^{-1} k_\perp^{-7/3}}{m^{5/2}} \sim \frac{k_\perp^{-7/3}}{m^{5/2}} \left(1 + \ln\sqrt{m}\right). \tag{4.35}$$

So there is logarithmically more free energy in the intermediate region, but this does not affect the $k_\perp$ scaling, which is what we were after in equation (4.24), so the derivation in section 4.3 survives.

For completeness, let us also calculate the 1D parallel spectrum. Integration over $k_\perp$ is dominated by the wave numbers around the phase-mixing threshold $k_\perp \sim (k_\parallel/\sqrt{m})^{3/4}$, so

$$E_m^{\parallel+}(k_\parallel) = \int \mathrm{d}k_\perp\, E_m^+(k_\parallel, k_\perp) \sim \int_0^{(k_\parallel/\sqrt{m})^{3/4}} \mathrm{d}k_\perp \frac{k_\perp^3 k_\parallel^{-5}}{\sqrt{m}} + \int_{(k_\parallel/\sqrt{m})^{3/4}}^{k_\parallel^{3/4}} \mathrm{d}k_\perp \frac{k_\parallel^{-1} k_\perp^{-7/3}}{m^{5/2}} \sim \frac{k_\parallel^{-2}}{m^2}. \tag{4.36}$$

Note that this m^{-2} scaling appears to be in decent agreement with the Hermite-space spectra reported by Hatch *et al.* (2013, 2014).

A perceptive reader might be feeling a growing resentment over our use of the spectrum (4.34) at wave numbers around the phase-mixing threshold $k_\parallel \sim \sqrt{m}\,k_\perp^{4/3}$, even though, technically speaking, we have only justified equation (4.34) in the region $k_\perp^{4/3} \lesssim k_\parallel \ll \sqrt{m}\,k_\perp^{4/3}$, seeing that at $k_\parallel \sim \sqrt{m}\,k_\perp^{4/3}$, phase mixing cannot be neglected compared to the nonlinear advection. In appendix C, we show that it nevertheless makes sense simply to match the spectrum (4.26) to the phase-mixing-dominated spectrum (4.8) along the phase-mixing threshold.

4.5. *Anti-phase-mixing spectra*

The arguments about the intermediate-region spectra presented in section 4.4 only apply to the spectrum of the "+" modes. Since the advection velocity is effectively two-dimensional in the intermediate region (see section 4.4.1), there is no coupling between different parallel wave numbers and so no echo effect. Thus, if, as we argued in section 4.3, $E_m^- \approx E_m^+ \propto k_\perp^{-d}$ in the advection-dominated region, $k_\perp \gtrsim k_\parallel^{1/r}$, due to vigorous coupling between parallel wave numbers, and $E_m^- \ll E_m^+$ in the intermediate and phase-mixing-dominated region, $k_\perp \ll k_\parallel^{1/r}$, due to absence of any such coupling, we must expect that the energy-containing wave numbers for the "−" modes are ones along the critical-balance curve $k_\perp \sim k_\parallel^{1/r}$.

The $k_\perp \to 0$ asymptotic behaviour of the "−" spectrum must be the same as for any other field, $E_m^- \propto k_\perp^3$, because the reasons for it are purely kinematic (appendix A).

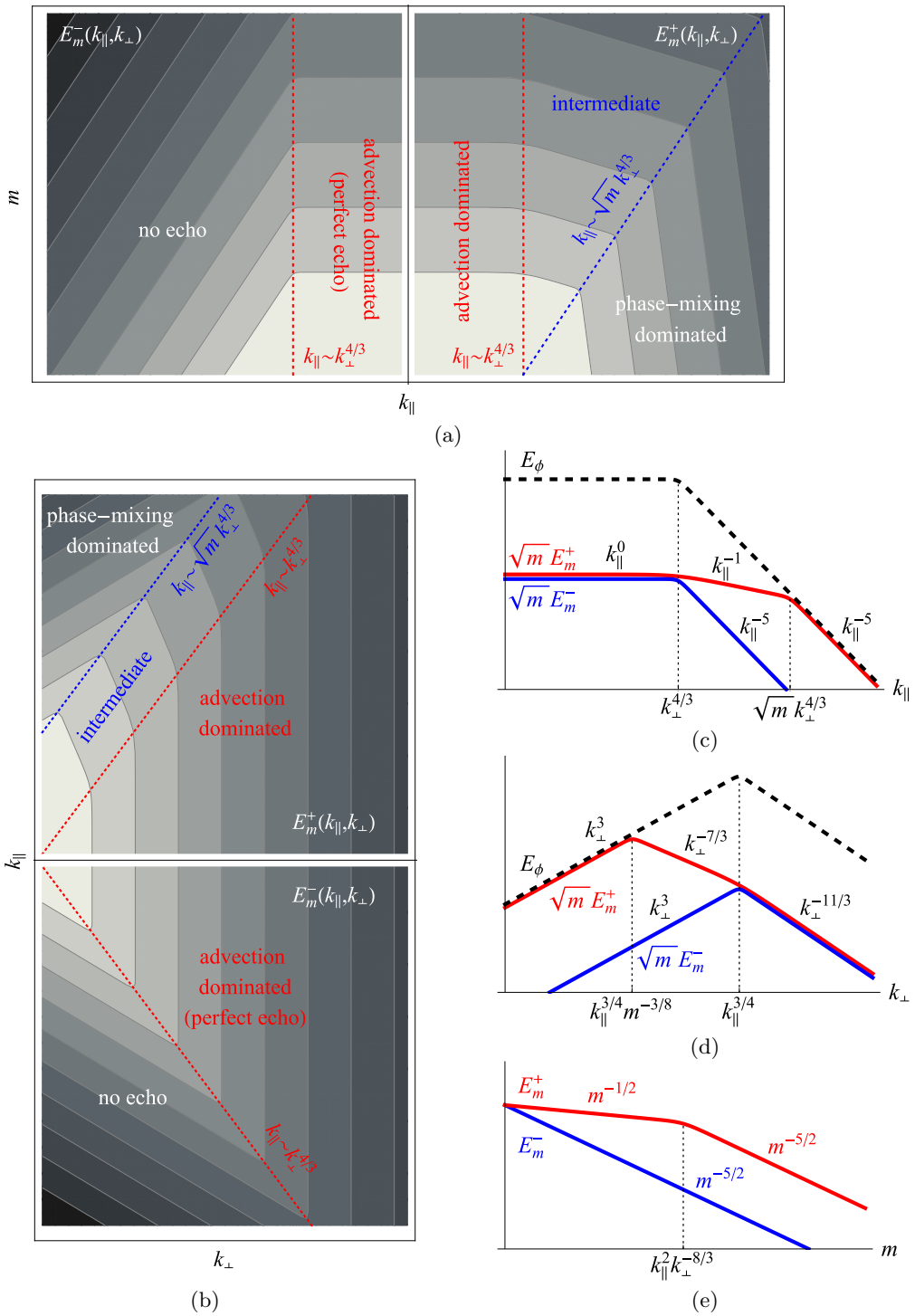

FIGURE 2. Spectra of higher Hermite moments, $E_m^\pm(k_\parallel,k_\perp)$: (a) in the $(k_\parallel,m)$ plane at constant $k_\perp$ (E_m^+ and E_m^- are shown in the right and left panels, respectively; in the left panel, $k_\parallel$ increases leftwards), (b) in the $(k_\perp,k_\parallel)$ plane at constant m (E_m^+ and E_m^- are shown in the upper and lower panels, respectively; in the lower panel, $k_\parallel$ increases downwards), (c) vs. $k_\parallel$ at constant $k_\perp$ and m, (d) vs. $k_\perp$ at constant $k_\parallel$ and m, (e) vs. m at constant $k_\parallel$ and $k_\perp$ (such that $k_\parallel > k_\perp^{4/3}$, otherwise the spectra are $E_m^\pm \sim m^{-5/2}$ at all m). All plots are logarithmic. The spectrum $E_\varphi(k_\parallel,k_\perp)$ (see figure 1) is given in (c) and (d) as a dashed line, for reference. Free-energy flow through phase space as represented in (a) and (b) is described in section 5.2.

Thus, we posit

$$E_m^-(k_\parallel, k_\perp) \sim \frac{k_\perp^3 k_\parallel^{-a''}}{m^{\sigma''}}, \quad k_\parallel \gtrsim k_\perp^r \tag{4.37}$$

and impose matching conditions between this spectrum and equation (4.23) at $k_\parallel \sim k_\perp^r$:

$$a'' = \frac{d+3}{r} = 5, \quad \sigma'' = \sigma = \frac{5}{2}. \tag{4.38}$$

This completes the determination of the phase-space spectra of the anti-phase-mixing component of the free energy:

$$E_m^-(k_\parallel, k_\perp) \sim \frac{1}{m^{5/2}} \cdot \begin{cases} k_\perp^3 k_\parallel^{-5}, & \text{if} \quad k_\parallel \gtrsim k_\perp^{4/3}, \\ k_\parallel^0 k_\perp^{-11/3}, & \text{if} \quad k_\parallel \lesssim k_\perp^{4/3}. \end{cases} \tag{4.39}$$

Figure 2 shows these and illustrates their relationship to other spectra derived above.

The 1D spectra that follow from equation (4.39) are

$$E_m^{\perp -}(k_\perp) = \int \mathrm{d}k_\parallel \, E_m^-(k_\parallel, k_\perp) \sim \frac{k_\perp^{-7/3}}{m^{5/2}}, \tag{4.40}$$

$$E_m^{\parallel -}(k_\parallel) = \int \mathrm{d}k_\perp \, E_m^-(k_\parallel, k_\perp) \sim \frac{k_\parallel^{-2}}{m^{5/2}}. \tag{4.41}$$

Note that these are both subdominant, in m, to the "+"-mode spectra (4.35) and (4.36).

4.6. *Effect of collisions*

4.6.1. *Collisional cutoff for phase-mixing modes*

In the phase-mixing-dominated regime (section 4.1), the collisional cutoff is set, in the same way as in the linear theory (section 2.5), by the competition between the phase-mixing rate $\sim k_\parallel v_{\mathrm{th}}/\sqrt{m}$ and the collision rate $\sim \nu m$. The perturbations are collisionally damped if

$$\nu m \gtrsim \frac{k_\parallel v_{\mathrm{th}}}{\sqrt{m}} \gtrsim k_\perp u_\perp \quad \Leftrightarrow \quad m \gtrsim \left(\frac{k_\parallel}{\nu}\right)^{2/3}, \quad k_\parallel \gtrsim \frac{k_\perp^2}{\sqrt{\nu}}, \tag{4.42}$$

giving a cutoff in Hermite space (cf. equation (2.59)).† In both the intermediate (section 4.4) and advection-dominated (section 4.3) regimes, the relevant comparison is between the collision rate and the nonlinear-advection rate:

$$\nu m \gtrsim k_\perp u_\perp \gtrsim \frac{k_\parallel v_{\mathrm{th}}}{\sqrt{m}} \quad \Leftrightarrow \quad m \gtrsim \frac{k_\perp^{4/3}}{\nu}, \qquad k_\parallel \lesssim \frac{k_\perp^2}{\sqrt{\nu}}. \tag{4.43}$$

These cutoffs are sketched in figure 3.

At a fixed m, the above relations imply that there is an infrared collisional cutoff in the $(k_\parallel, k_\perp)$ space: perturbations are damped if

$$k_\parallel \lesssim \nu m^{3/2}, \quad k_\perp \lesssim (\nu m)^{3/4}. \tag{4.44}$$

These cutoffs will not, of course, be relevant in comparison with the outer scales ($k_{\parallel 0}$ and $k_{\perp 0}$; see section 2.4.1) except at high enough m or if the collision frequency approaches the characteristic phase-mixing and nonlinear-advection rates at the outer scale. In the

† In the last formula in equation (4.42), we implicitly nondimensionalised the collision frequency: $\nu L_\parallel / v_{\mathrm{th}} = L_\parallel / \lambda_{\mathrm{mfp}} \to \nu$, so the Hermite cutoff is $m_{\mathrm{c}} \sim (k_\parallel \lambda_{\mathrm{mfp}})^{2/3}$, where $\lambda_{\mathrm{mfp}} = v_{\mathrm{th}}/\nu$ is the mean free path. In other words, in rescaled units, one can replace $\nu \Leftrightarrow 1/\lambda_{\mathrm{mfp}}$ wherever this makes things more transparent.

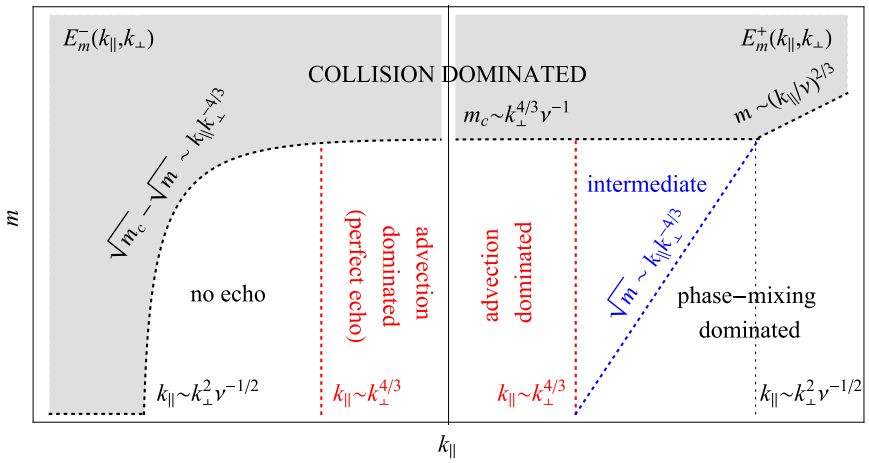

FIGURE 3. Partition of phase space, viz., $(k_\parallel, m)$ plane at fixed $k_\perp$, showing the collision-dominated region. Cf. figure 2(a). All axes are logarithmic (in the left panel, $k_\parallel$ increases leftwards).

latter case, one expects some amount of free energy to drain via collisional dissipation around the outer scale, petering out at larger $(k_\parallel, k_\perp)$, where the collisionless physics asserts itself (cf. Hatch *et al.* 2011*a*,*b*, 2013, 2014).

4.6.2. *Collisional cutoff for anti-phase-mixing modes*

The collisional cutoff (4.43) on the free-energy spectrum of the "+" modes in the advection-dominated regime must extend to the "−" (anti-phase-mixing) modes because nonlinear coupling is the only source of the latter (see section 4.3). But the anti-phase-mixing modes propagate from high to low m and so a zero "boundary condition" at high m will be imprinted onto a region of phase space at lower m's. To wit, arguing analogously to section 4.1 and considering now $k_\parallel < 0$ in equation (3.12), we note that anti-phase-mixing modes propagate along the characteristics

$$s = -\frac{|k_\parallel| v_{\rm th}}{\sqrt{2}}\, t + s_0, \tag{4.45}$$

where s_0 is a constant. Whatever anti-phase-mixing spectrum exists at $s = s_0$, it will be replicated over all s satisfying equation (4.45) for times shorter than the nonlinear time, $t \lesssim \tau_{\rm nl} \sim (k_\perp u_\perp)^{-1} \sim k_\perp^{4/3}$. Assuming that E_m^- is cut off for $s \gtrsim s_{\rm c} = k_\perp^{2/3}/\sqrt{\nu}$ (equation (4.43)) and letting $s_0 = s_{\rm c}$, we conclude that E_m^- must also be cut off for

$$s \gtrsim s_{\rm c} - \frac{|k_\parallel| v_{\rm th}}{k_\perp u_\perp} \quad \Leftrightarrow \quad \sqrt{m} \gtrsim \sqrt{m_{\rm c}} - \frac{|k_\parallel|}{k_\perp^{4/3}}, \quad m_{\rm c} = \frac{k_\perp^{4/3}}{\nu}. \tag{4.46}$$

This implies, in particular, that there is no anti-phase-mixing energy at any m's for wave numbers satisfying

$$|k_\parallel| \gtrsim \frac{k_\perp^2}{\sqrt{\nu}}. \tag{4.47}$$

The collision-dominated region of the phase space is sketched in figure 3.

Whereas the collisional cutoff is safely removed to infinite m in the limit $\nu \to 0$, in systems with only moderately low collision frequency, one should expect to see a finite reduction in the anti-phase-mixing flux at higher m's, as per equation (4.46).

4.7. *Total free energy and dissipation*

In the linear problem, where all energy injected into the system had to be removed by Landau damping (meaning phase mixing followed by collisional dissipation at high m), the free energy stored in phase space in a steady state had to diverge with vanishing collisionality (see equation (2.61)) in order for the dissipation to remain finite (equation (2.60)). In the nonlinear situation with which we are now preoccupied, the Hermite spectra are steep power laws and so the free energy will be finite and collisional dissipation vanish, with all of the injected energy having to be removed via dissipation at small spatial scales (sub-Larmor and so outside the regime of validity of this theory).

To demonstrate this a little more quantitatively, let us repeat the calculation of the 1D parallel spectrum (equation (4.36)), but now, in integrating the 2D spectrum over $k_\perp$, we assume that there is no free energy at perpendicular wave numbers below the collisional cutoff $k_\perp \sim (\nu m)^{3/4}$ (equation (4.44)). The following three cases correspond to the collisional cutoff falling into the phase-mixing-dominated, intermediate and advection-dominated regions, respectively:

$$E_m^{\parallel+}(k_\parallel) \sim \begin{cases} \dfrac{k_\parallel^{-2}}{m^2} & \text{if} \quad m \lesssim \left(\dfrac{k_\parallel}{\nu}\right)^{2/3} \sim m_\mathrm{c}, \\ \dfrac{k_\parallel^{-1}}{\nu m^{7/2}} & \text{if} \quad \left(\dfrac{k_\parallel}{\nu}\right)^{2/3} \lesssim m \lesssim \dfrac{k_\parallel}{\nu}, \\ \dfrac{k_\parallel^{0}}{\nu^2 m^{9/2}} & \text{if} \quad m \gtrsim \dfrac{k_\parallel}{\nu}. \end{cases} \tag{4.48}$$

Note that the last two cases are only relevant at very high m (because $k_\parallel \geqslant k_{\parallel 0} \sim 1$, the outer scale in our units). Now integrating these spectra over m, we find that the total free energy in a given $k_\parallel$ is completely dominated by low m's:

$$W(k_\parallel) \sim \int_{\sim 1}^{\infty} \mathrm{d}m\, E_m^{\parallel+}(k_\parallel) \sim k_\parallel^{-2}. \tag{4.49}$$

The total collisional-dissipation rate vanishes with ν:

$$D(k_\parallel) \sim \nu \int_{\sim 1}^{\infty} \mathrm{d}m\, m E_m^{\parallel+}(k_\parallel) \sim k_\parallel^{-2}\nu \int_{\sim 1}^{m_\mathrm{c}} \frac{\mathrm{d}m}{m} \sim k_\parallel^{-2}\nu \ln\left(\frac{k_\parallel}{\nu}\right)^{2/3} \to 0 \quad \text{as} \quad \nu \to 0. \tag{4.50}$$

Equations (4.49) and (4.50) are the nonlinear versions of equations (2.61) and (2.60), respectively—we see that, unlike in the linear problem, the free energy remains finite and collisional dissipation vanishes as $\nu \to +0$.

5. Conclusion

5.1. *Summary of free-energy spectra and of the method of deriving them*

Considering the full phase space $(k_\parallel, k_\perp, m)$, we posited a set of power-law relationships for the free-energy spectra and then determined the scaling exponents from a combination of matching conditions between different regions of the phase space and physical arguments about the free-energy flows, constrained by conservation laws. The spectra

are, for the phase-mixing modes (i.e., perturbations that propagate from low to high m),

$$E^{+}_{m}(k_\parallel,k_\perp)\sim\begin{cases}\dfrac{k_\parallel^{0}k_\perp^{-11/3}}{m^{5/2}} & \text{if}\quad k_\parallel\lesssim k_\perp^{4/3}\\ & \text{(advection dominated)},\\ \dfrac{k_\parallel^{-1}k_\perp^{-7/3}}{m^{5/2}} & \text{if}\quad k_\perp^{4/3}\lesssim k_\parallel\lesssim\sqrt{m}\,k_\perp^{4/3}\\ & \text{(intermediate)},\\ \dfrac{k_\parallel^{-5}k_\perp^{3}}{\sqrt{m}} & \text{if}\quad k_\parallel\gtrsim\sqrt{m}\,k_\perp^{4/3}\\ & \text{(phase-mixing dominated)},\end{cases}\tag{5.1}$$

for the anti-phase-mixing modes (propagating from high to low m),

$$E^{-}_{m}(k_\parallel,k_\perp)\sim\begin{cases}\dfrac{k_\parallel^{0}k_\perp^{-11/3}}{m^{5/2}} & \text{if}\quad k_\parallel\lesssim k_\perp^{4/3}\\ & \text{(advection dominated)},\\ \dfrac{k_\parallel^{-5}k_\perp^{3}}{m^{5/2}} & \text{if}\quad k_\parallel\gtrsim k_\perp^{4/3}\\ & \text{(no echo)},\end{cases}\tag{5.2}$$

and for the "fluid" (low-m) moments,

$$E_{\varphi}(k_\parallel,k_\perp)\sim\begin{cases}k_\parallel^{0}k_\perp^{-11/3} & \text{if}\quad k_\parallel\lesssim k_\perp^{4/3}\\ & \text{(advection dominated)},\\ k_\parallel^{-5}k_\perp^{3} & \text{if}\quad k_\parallel\gtrsim k_\perp^{4/3}\\ & \text{(phase-mixing dominated)}.\end{cases}\tag{5.3}$$

A graphical summary of these spectra is presented in figure 2.

As is manifest in the above formulae, the phase space is partitioned into several regions, where different physics controls the distribution of the free energy.

• In the *phase-mixing-dominated region* (section 4.1), the phase-mixing rate is greater than the rate of nonlinear advection, $k_\parallel v_{\rm th}/\sqrt{m}\gg k_\perp u_\perp$, and so whatever distribution of free energy exists at these wave numbers at low m's will simply be propagated to larger m's—this is the part of the wave-number space where modes are "Landau-damped" in the usual linear sense. The perpendicular spectrum in this region ($\propto k_\perp^3$; see equation (5.1)) is fixed on purely kinematic grounds (appendix A), the m scaling ($\propto m^{-1/2}$) is the same as in the linear problem, corresponding to constant Hermite flux (Zocco & Schekochihin 2011, Kanekar *et al.* 2015; see section 2.5), whereas the scaling exponent of the parallel spectrum ($\propto k_\parallel^{-a}$, $a=5$) is fixed by matching with the nonlinear dynamics (section 4.4; see equation (4.31)).

• The continual flow of free energy into high m's as described above sets the matching condition at the *phase-mixing threshold*, where the nonlinear advection rate becomes comparable to the phase-mixing rate, $k_\perp u_\perp\sim k_\parallel v_{\rm th}/\sqrt{m}$. The role of the nonlinear advection is to divert the free energy from flowing straight to higher m's to flowing to higher $k_\perp$'s. The competition between these two processes sets the prevailing dependence of the free energy on m, giving rise to the $m^{-5/2}$ scaling of the 2D spectra (equation (5.1); derived in section 4.4.1 and appendix C) and the m^{-2} overall scaling (equation (4.36); in reasonable agreement with recent numerical studies by Hatch *et al.* 2013, 2014). The situation at the phase-mixing threshold is so crucial because the free-energy spectra rise as $k_\perp$ increases and $k_\parallel$ decreases from the phase-mixing-dominated region towards the

phase-mixing threshold and then fall beyond it, at higher $k_\perp$ and lower $k_\parallel$, so it is along the phase-mixing threshold that the energy-containing scales in phase space lie.

- The *intermediate region* comprises the wave numbers at which the nonlinear-advection rate is already dominant compared to the phase-mixing rate of the high-m moments of the distribution function, but not of the low-m moments and, in particular, of the zeroth moment, φ, which is what sets the $\mathbf{E}\times\mathbf{B}$ flow velocity $\mathbf{u}_\perp$ that is doing the advection: $k_\parallel v_{\rm th}/\sqrt{m} \ll k_\perp u_\perp \ll k_\parallel v_{\rm th}$. The energy-containing wave numbers for the flow lie along the *critical-balance* curve $k_\perp u_\perp \sim k_\parallel v_{\rm th}$ (section 4.2)—and so the nonlinear interactions in the intermediate region are nonlocal in $k_\parallel$, with short-parallel-scale perturbations of the distribution function advected by a longer-scale flow, i.e., an effectively 2D velocity field (see appendix B). A constant-flux argument for the free-energy cascade in $k_\perp$ then fixes the $k_\perp^{-7/3}$ scaling of the 2D free-energy spectrum in this region, while its $k_\parallel^{-1}$ scaling follows from matching to the spectra at the phase-mixing threshold and at the critical-balance curve (the second scaling in equation (5.1); see section 4.4).
- Beyond the critical-balance curve, $k_\perp u_\perp \gg k_\parallel v_{\rm th}$, the nonlinear advection is completely dominant over phase mixing, giving rise to the *advection-dominated region*. The advecting flow is now 3D and another constant-flux argument gives the $k_\perp^{-11/3}$ scaling of the free-energy spectrum, whereas its $k_\parallel^0$ scaling is a white-noise spectrum deduced via a simple causality argument implying that perturbations with a certain perpendicular scale are decorrelated at parallel distances long enough that information cannot traverse them at the speed $\sim v_{\rm th}$ over one cascade time corresponding to that perpendicular scale (the first scaling in equation (5.1); see section 4.3).

The above arguments have all focused on the free energy contained in the perturbations that propagate from low to high m, i.e., ones prone to phase mixing (whether it is fast or slow compared to nonlinear advection). In a nonlinear system, an advecting flow that has a parallel spatial dependence, i.e., $k_\parallel \neq 0$, can couple these perturbations to others that have parallel wave numbers of opposite sign and so will propagate from high to low m, a phenomenon known as plasma echo (section 3.2). Separating all perturbations into these "+" and "−" components (section 3.1) allows us to express the free energy as the sum of their spectra and its flux in Hermite space as proportional to the difference between these spectra (section 3.4). In the advection-dominated region, vigorous nonlinear coupling implies that the "+" and "−" spectra are the same and so, statistically, there is no free-energy flux between different m's—i.e., the phase-mixing and the anti-phase-mixing energy fluxes cancel each other (see section 4.3 and the first scaling in equation (5.2)). In contrast, there is no echo effect and, therefore, no significant "−" energy either in the intermediate region (because the flow velocity there is effectively 2D and so cannot couple different $k_\parallel$'s) or in the phase-mixing-dominated region (because anti-phase-mixing modes do not propagate to higher m's). The "−" spectrum outside the advection-dominated region (the second scaling in equation (5.2), derived in section 4.5) is, therefore, determined by the kinematic constraint giving the $k_\perp^3$ scaling at long wavelengths and by the matching conditions along the boundary of that region—the critical-balance curve.

Finally, the spectra (5.3) of the low-m, "fluid" moments are basically a continuation of the high-m spectra (5.1) and (5.2) down to low m's. Physically, since the Hermite flux between different m's is on average shut down in the advection-dominated region, these scalings can be determined by assuming constant flux of the "fluid" part of the free-energy, i.e., effectively, by pretending that the turbulence is fluid-like (Barnes *et al.* 2011; see sections 2.4.2 and 4.2). Such a shortcut has always been tempting (e.g., Weiland 1992), but was not *a priori* justified for a kinetic system (section 2.4.4).

5.2. *Free-energy flows*

Although this was implicit in our discussion of the partition of phase space (section 5.1), it is worth spelling out what path the free energy takes through it. Let us start from some $(m, k_\parallel, k_\perp)$ in the *phase-mixing-dominated region* (lower right corner in the right panel of figure 2(a)). At first, the free energy will move (phase-mix) from there to higher m (vertically towards the blue line in figure 2(a)) until it reaches the phase-mixing threshold $m \sim k_\parallel^2/k_\perp^{8/3}$ (equation (4.5) with $r = 4/3$; the blue line in figures 2(a,b)). There it enters the *intermediate region*, where it is advected by an effectively 2D velocity field (see section 4.4.1 and appendix B) to higher $k_\perp$ while staying at fixed $k_\parallel$ (in the top panel of figure 2(b), horizontally from the blue towards the red line) until it reaches the critical-balance threshold $k_\perp \sim k_\parallel^{3/4}$ (equation (4.20); the red line in figure 2(b)). At that point it enters the *advection-dominated region*, where the advection is 3D and the energy flows along the critical-balance curve (diagonally upwards along the red line in the top panel of figure 2(b); see section 4.3). Since a 3D velocity is effective at coupling positive and negative $k_\parallel$'s, this flow of energy involves both "+" and "−" modes (the latter shown in the bottom panel of figure 2(b), where the critically balanced energy flow is also along the red line). There is not much flow of the "−" energy beyond the critical-balance threshold (to the left of the red line in the bottom panel of figure 2(b), or, equivalently, to the lower left of the red line in the left panel of figure 2(a)) because it nonlinearly couples back to "+" modes faster than it can anti-phase-mix to lower m's.†

5.3. *Implications and outlook*

The free-energy distribution in phase space summarised above has several important properties and implications.

The free-energy flux out of the "fluid" moments is heavily suppressed in the wave-number region bounded by the critical-balance curve, $k_\parallel \lesssim k_\perp^{4/3}$, which is also the region that contains most of the free energy flowing through the inertial range. Thus, at the energetically relevant wave numbers of the inertial range, Landau damping is effectively absent. The resulting Hermite spectra have steep power laws ($\propto m^{-2}$ for the total energy; see equation (4.36)) and so the total free energy contained in the phase space is finite, dominated by low m's (the energy in the "fluid" moments) and does not diverge at vanishing collisionality (equation (4.49))—in sharp contrast to its behaviour in the linear problem (see equation (2.61)). Furthermore, the total collisional dissipation vanishes in the nonlinear problem (equation (4.50)), again in contrast to the linear case, where the dissipation rate is finite and absorbs all of the energy that is injected into the system (Kanekar *et al.* 2015; see equation (2.60)). This means that most of the dissipation occurs at small *spatial* scales (i.e., beyond the Larmor scale, a region that we have left outside our detailed focus). This is indeed what was recently found numerically by Hatch *et al.* (2013, 2014): decreasing share of the collisional dissipation with decreasing collisionality. Note that finite collisionality imposes a cutoff on the free-energy spectra at high enough m's, or, equivalently, at low enough $k_\parallel$ and $k_\perp$ (equation (4.44)); when the collision frequency approaches the rates of phase-mixing and nonlinear-advection rates, a certain amount of collisional dissipation will occur at low wave numbers (cf. Watanabe & Sugama 2006; Hatch *et al.* 2011*a*,*b*).

It is inevitable that one must ask about the implications our results might have for

† It is possible for mode-coupling in $k_\parallel$ to combine with anti-phase-mixing to push some "−" energy towards larger $k_\parallel$, but that process is diffusive in $k_\parallel$ and will be slower than direct nonlinear coupling back into "+" modes. It becomes important when the advecting flow is scale-separated from the distribution function that is advected by it (Schekochihin *et al.* 2016).

the Landau-fluid closures as a viable modelling technique—a subject that has long been discussed and refined in the context of fusion plasmas (Hammett & Perkins 1990; Hammett *et al.* 1992, 1993; Weiland 1992; Mattor 1992; Hedrick & Leboeuf 1992; Dorland & Hammett 1993; Beer & Hammett 1996; Snyder *et al.* 1997; Snyder & Hammett 2001*a*,*b*; Ramos 2005) as well as, more recently, space and astrophysical ones (Passot & Sulem 2004, 2006, 2007; Goswami *et al.* 2005; Passot *et al.* 2012). While the basic idea of the Landau-fluid approach is to include into fluid equations damping terms ($\sim |k_\parallel| v_{\rm th}$) fit to capture correctly the linear Landau damping, it has long been known in this field that quantitatively these models work better when more Hermite moments are retained and this inclusion happens at the level of the highest of them (Smith 1997). Considering that the free energy scales steeply with m, as shown above, it stands to reason that, at low collisionality, Landau-fluid closures that retain a certain finite (independent of the collision frequency) number of moments may be sufficient for a full characterisation of kinetic turbulence—in that already just this finite number of moments will be enough to capture most of the echo flux from phase space back to "fluid" moments. The Landau closure terms affecting the highest of the retained moments will then serve to regularise the problem in the energetically subdominant part of the wave-number space—the phase-mixing region—where the free energy has a shallow scaling $\sim m^{-1/2}$ (section 4.1).

There is clearly space for further development of this line of reasoning, leading to more quantitative prescriptions for capturing the echo effect within the Landau-fluid framework. If one thinks of these closures in the same modelling spirit as one does about Large-Eddy-Simulation techniques in fluid dynamics (Smagorinsky 1963)—and, more recently, in gyrokinetics (Morel *et al.* 2011, 2012; Bañón Navarro *et al.* 2014),—the $m^{-5/2}$ spectrum we have derived can serve the useful role of providing the signature of a well-developed nonlinear phase-space "cascade," which, once formed, can be promptly and safely cut off by model dissipation terms.

Ranging somewhat further afield, we note that spacecraft measurements of compressive (density and magnetic-field strength) fluctuations in the inertial range† of the solar-wind turbulence (Celnikier *et al.* 1983, 1987; Marsch & Tu 1990; Bershadskii & Sreenivasan 2004; Hnat *et al.* 2005; Kellogg & Horbury 2005; Chen *et al.* 2011, 2014) show healthy Kolmogorov-like power-law spectra—in what is generally a $\beta \sim 1$ plasma, where the Landau damping of such fluctuations (Barnes 1966) ought to be of the same order as their nonlinear cascade rates (Schekochihin *et al.* 2009). Similarly robust power-law spectra at sub-ion-Larmor scales have also been measured (Sahraoui *et al.* 2009, 2010, 2013; Alexandrova *et al.* 2009, 2012; Chen *et al.* 2010, 2013) and found in kinetic simulations (Howes *et al.* 2011; Chang *et al.* 2011), even though Landau damping of kinetic Alfvén waves (Howes *et al.* 2006; Gary & Borovsky 2008) should be quantitatively noticeable at these scales (Howes *et al.* 2008; Podesta *et al.* 2010). Whereas attempts have been made to argue that in some of these situations the linear damping might be weak (Lithwick & Goldreich 2001; Howes *et al.* 2008; Schekochihin *et al.* 2009), it is a tempting—and more interesting—thought that the general mechanism for (statistical) suppression of phase mixing in a turbulent system proposed here is responsible for making collisionless plasma turbulence in the solar wind behave in a seemingly more "fluid-like" fashion

† These can be shown to be drift-kinetic fields passively advected by the turbulent velocity field $\mathbf{u}_\perp$ associated with Alfvénic perturbations. They satisfy equations that are quite similar to equation (2.8), although with an additional complication that particles stream along magnetic field that is also perturbed by the Alfvénic turbulence and so linear and nonlinear mixing are somewhat intertwined (Schekochihin *et al.* 2007, 2009; Kunz *et al.* 2015).

than theoreticians might have thought it had a right to do.‡ A numerical and theoretical investigation of this possibility is a subject of our current efforts.

To conclude, the considerations presented above appear to point to a number of promising directions for numerical experiment and further thought. We hope to explore some of those in the not so distant future.†

We are grateful to I. Abel, M. Barnes, S. Cowley, A. Kanekar, N. Loureiro, F. Parra, C. Staines, and L. Stipani for many important discussions on this and related topics. I. Abel, N. Loureiro and L. Stipani have read this paper in manuscript and made useful comments. Constructive critique from two anonymous but diligent referees have helped improve our exposition, for which we are thankful. A.A.S. is indebted to R. Jeffrey for his collaboration on an unpublished early precursor to this project. J.T.P. was supported by the UK Engineering and Physical Sciences Research Council through a Doctoral Training Grant award. E.G.H.'s work has been carried out within the framework of the EUROfusion Consortium and was supported by a EUROfusion Fusion Researcher Fellowship [WP14-FRF-CCFE/Highcock]. The views and opinions expressed herein do not necessarily reflect those of the European Commission. W.D. was supported by the US DoE grants DE-FG02-93ER54197 and DE-FC02-08ER54964. All authors are grateful to the Wolfgang Pauli Institute, University of Vienna, for its hospitality on several occasions.

Appendix A. Long-wavelength scaling of spectra

Here we review the standard argument that the spectrum of a 2D-isotropic homogeneous field $\varphi(k_\parallel, \mathbf{r}_\perp)$ has the low-wavelength asymptotic form

$$E_\varphi(k_\parallel, k_\perp) = 2\pi k_\perp \langle |\varphi(k_\parallel, \mathbf{k}_\perp)|^2 \rangle \propto k_\perp^3 \quad \text{as} \quad k_\perp \to 0. \tag{A 1}$$

Setting the perpendicular-Fourier-transform conventions to be

$$\varphi(k_\parallel, \mathbf{r}_\perp) = \left(\frac{L_\perp}{2\pi}\right)^2 \int \mathrm{d}^2\mathbf{k}_\perp e^{i\mathbf{k}_\perp\cdot\mathbf{r}_\perp} \varphi(k_\parallel, \mathbf{k}_\perp), \tag{A 2}$$

$$\varphi(k_\parallel, \mathbf{k}_\perp) = \int \frac{\mathrm{d}^2\mathbf{r}_\perp}{L_\perp^2}\, e^{-i\mathbf{k}_\perp\cdot\mathbf{r}_\perp} \varphi(k_\parallel, \mathbf{r}_\perp), \tag{A 3}$$

where $L_\perp$ is the box size, we find

$$\begin{aligned}\langle |\varphi(k_\parallel, \mathbf{k}_\perp)|^2 \rangle &= \int \frac{\mathrm{d}^2\mathbf{r}_{\perp 1}}{L_\perp^2} \int \frac{\mathrm{d}^2\mathbf{r}_{\perp 2}}{L_\perp^2}\, e^{-i\mathbf{k}_\perp\cdot(\mathbf{r}_{\perp 1}-\mathbf{r}_{\perp 2})} \langle \varphi(k_\parallel, \mathbf{r}_{\perp 1}) \varphi^*(k_\parallel, \mathbf{r}_{\perp 2}) \rangle \\ &= \int \frac{\mathrm{d}^2\mathbf{r}_\perp}{L_\perp^2}\, e^{-i\mathbf{k}_\perp\cdot\mathbf{r}_\perp} C(k_\parallel, r_\perp) = \frac{2\pi}{L_\perp^2} \int_0^\infty \mathrm{d}r_\perp r_\perp J_0(k_\perp r_\perp) C(k_\parallel, r_\perp), \end{aligned} \tag{A 4}$$

where $\mathbf{r}_\perp = \mathbf{r}_{\perp 1} - \mathbf{r}_{\perp 2}$, $C(k_\parallel, r_\perp)$ is the two-point correlation function of φ (which only depends on $r_\perp = |\mathbf{r}_\perp|$ because the field is statistically homogeneous and isotropic), and J_0 is the Bessel function of order zero.

‡ An immediate physically interesting conclusion from such an outcome, apart from power-law compressive spectra being theoretically legitimised, would be that one should not expect any ion heating associated with the inertial-range turbulence (see equation (4.50)), with the thermal fate of all turbulent energy determined at the ion Larmor scale, where the 4D drift-kinetic phase-space cascade morphs into a more complicated 5D gyrokinetic one (Schekochihin *et al.* 2009; Howes *et al.* 2011; Told *et al.* 2015).

† As this paper is going into press, the first dedicated numerical tests of our theory have been undertaken by Kanekar (2015), Parker (2016) and Parker *et al.* (2016), so far broadly supporting our conclusions.

If the correlation function $C(k_\parallel, r_\perp)$ decays sufficiently quickly with $r_\perp$, it will restrict the integral in equation (A 4) to values of $r_\perp$ that are smaller than or comparable to the perpendicular correlation length $r_{\perp \mathrm{c}}$ of the field $\varphi(k_\parallel, \mathbf{r}_\perp)$. Note that $r_{\perp \mathrm{c}}$ will be a function of $k_\parallel$, so it is not necessarily the outer scale—in section 4.2, we argue that it is the critical-balance scale, $r_{\perp \mathrm{c}} \sim k_{\perp c}^{-1} \sim k_\parallel^{-1/r}$. If we now consider $k_\perp r_{\perp \mathrm{c}} \ll 1$, we may expand the Bessel function $J_0(k_\perp r_\perp) = 1 - k_\perp^2 r_\perp^2/4 + \dots$ in equation (A 4), which then gives us

$$E_\varphi(k_\parallel, k_\perp) = 2\pi k_\perp \langle |\varphi(k_\parallel, \mathbf{k}_\perp)|^2 \rangle = k_\perp \left[\left(\frac{2\pi}{L_\perp} \right)^2 \int_0^\infty \mathrm{d}r_\perp r_\perp C(k_\parallel, r_\perp) \right] + k_\perp^3 \left[-\frac{1}{4} \left(\frac{2\pi}{L_\perp} \right)^2 \int_0^\infty \mathrm{d}r_\perp r_\perp^3 C(k_\parallel, r_\perp) \right] + \dots \quad \text{(A 5)}$$

The first term is proportional to $\int \mathrm{d}^2 \mathbf{r}_\perp \langle \varphi(k_\parallel, \mathbf{r}_\perp) \varphi^*(k_\parallel, 0) \rangle$ and so it vanishes if we assume that $\int \mathrm{d}^2 \mathbf{r}_\perp \varphi(k_\parallel, \mathbf{r}_\perp) = \varphi(k_\parallel, k_\perp = 0) = 0$, i.e., that there are no purely 1D parallel modes.† Hence we obtain the desired result (A 1), with the proviso that $C(k_\parallel, r_\perp)$ decays faster than $1/r_\perp^4$ as $r_\perp \to \infty$ and so the integral prefactor of $k_\perp^3$ in equation (A 5) converges.

Appendix B. Nonlocal interactions in the intermediate region

Consider the nonlinear coupling expressed by the right-hand side of equation (3.12), which we now rewrite as a wave-number convolution in both parallel and perpendicular directions:

$$\left(\frac{\partial \tilde{f}}{\partial t} \right)_{\mathrm{nl}} (k_\parallel, \mathbf{k}_\perp) = -i \mathbf{k}_\perp \cdot \sum_{p_\parallel, \mathbf{p}_\perp} \mathbf{u}_\perp(p_\parallel, \mathbf{p}_\perp) \tilde{f}(k_\parallel - p_\parallel, \mathbf{k}_\perp - \mathbf{p}_\perp). \quad \text{(B 1)}$$

In the intermediate wave-number region between the phase-mixing threshold and the critical balance,

$$\sqrt{m}\, k_\perp u_\perp \gtrsim k_\parallel v_{\mathrm{th}} \gtrsim k_\perp u_\perp \quad \Leftrightarrow \quad \sqrt{m}\, k_\perp^r \gtrsim k_\parallel \gtrsim k_\perp^r, \quad \text{(B 2)}$$

the coupling in equation (B 1) must be predominantly between disparate wave numbers (i.e., the coupling is nonlocal) because the energy-containing wave numbers for $\mathbf{u}_\perp$ are $p_\parallel \lesssim p_\perp^r$, which lie outside the region (B 2). There are two basic possibilities: coupling that is local in $k_\perp$ but nonlocal in $k_\parallel$ and coupling that is local in $k_\parallel$ but nonlocal in $k_\perp$. In analysing the rates of such interactions, we will consider $\tilde{f}$ to be at the phase-mixing threshold, $k_\parallel \sim \sqrt{m}\, k_\perp^r$, and $\mathbf{u}_\perp$ in critical balance, $p_\parallel \sim p_\perp^r$.

Suppose the perpendicular coupling is local, $p_\perp \sim |\mathbf{k}_\perp - \mathbf{p}_\perp| \sim k_\perp$. Then

$$p_\parallel \sim k_\perp^r \sim \frac{k_\parallel}{\sqrt{m}} \ll k_\parallel, \quad \text{(B 3)}$$

so the distribution function $\tilde{f}(k_\parallel - p_\parallel) \approx \tilde{f}(k_\parallel)$ is advected by an effectively two-

† In the theory of a passive scalar, the quantity $\int_0^\infty \mathrm{d}^2 \mathbf{r}_\perp C(k_\parallel, r_\perp)$ is known as the Corrsin (1951) invariant—the decay laws for a passive scalar can depend on whether this invariant is zero or finite because that sets the long-wavelength asymptotic behaviour of the scalar's spectrum (e.g., Eyink & Xin 2000; Schekochihin *et al.* 2004). The fact that this asymptotic behaviour is $\sim k_\perp^3$ in our theory, will have implications for the determination of the Hermite spectrum; see appendix C.

dimensional velocity field: back in real space, equation (B 1) becomes

$$\left(\frac{\partial \tilde{f}}{\partial t}\right)_{\rm nl} \approx -\mathbf{u}_\perp(z=0,\mathbf{r}_\perp)\cdot\boldsymbol{\nabla}_\perp \tilde{f}(k_\parallel,\mathbf{r}_\perp). \tag{B 4}$$

The rate of nonlinear advection of $\tilde{f}(k_\parallel)$ is, as usual,

$$k_\perp u_\perp \sim k_\perp^r. \tag{B 5}$$

Note that as there is no coupling in $k_\parallel$, there can be no echo.

Now suppose instead that it is the parallel coupling that is local, $p_\parallel \sim |k_\parallel - p_\parallel| \sim k_\parallel$. Then

$$p_\perp \sim k_\parallel^{1/r} \sim m^{1/2r} k_\perp \gg k_\perp, \tag{B 6}$$

so the distribution function $\tilde{f}$ is advected by a much-smaller-scale (in the perpendicular direction) velocity field. The net effect of such an advection will be turbulent diffusion of $\tilde{f}$ with the effective mixing length $\sim 1/p_\perp$ and the effective diffusion coefficient

$$D_{\rm turb} \sim \frac{u_\perp}{p_\perp} \sim p_\perp^{r-2} \sim k_\perp^{r-2} m^{(r-2)/2r}. \tag{B 7}$$

The rate of nonlinear advection associated with this process is then

$$D_{\rm turb} k_\perp^2 \sim \frac{k_\perp^r}{m^{(2-r)/2r}} \ll k_\perp^r, \tag{B 8}$$

provided $r < 2$ (which it is, considering it will turn out to be $r = 4/3$). This is much smaller than the local-in-$k_\perp$, nonlocal-in-$k_\parallel$ advection rate (B 5). Thus, the latter type of interactions will be the dominant ones—the claim we make in section 4.4.1, which this appendix is meant to back up.

Note that other kinds of interaction—of various degree of non-locality in both $k_\parallel$ and $k_\perp$—cannot prove faster because non-locality in $k_\perp$ will always slow down coupling (diffusion is slower than advection) while more or less non-locality in $k_\parallel$ simply makes the velocity $\mathbf{u}_\perp$ more or less two-dimensional compared to $\tilde{f}$, without changing the rate of advection.

Appendix C. Spectra near phase-mixing threshold and the free-energy decay in Hermite space

In a statistical steady state, the free-energy spectrum is independent of time and so described by equation (3.26):

$$\frac{k_\parallel v_{\rm th}}{\sqrt{2}}\frac{\partial F}{\partial s} = -2{\rm Re}\sum_{p_\parallel}\left\langle \tilde{f}^*(k_\parallel)\mathbf{u}_\perp(p_\parallel)\cdot\boldsymbol{\nabla}_\perp \tilde{f}(k_\parallel - p_\parallel)\right\rangle - 2\nu s^2 F. \tag{C 1}$$

Let us consider the wave numbers around the phase-mixing threshold, for which $k_\parallel v_{\rm th} \sim k_\perp u_\perp$, so the phase-mixing term is comparable to the nonlinear term. Ignoring collisions, assuming locality in $k_\perp$ (see appendix B) and expanding in $p_\parallel \sim k_\perp u_\perp / v_{\rm th} \ll \sqrt{m}\, k_\perp u_\perp / v_{\rm th} \sim k_\parallel$, we have (cf. equation (B 4))

$$\frac{k_\parallel v_{\rm th}}{\sqrt{2}}\frac{\partial}{\partial s}\left\langle |\tilde{f}(k_\parallel)|^2\right\rangle \approx -\left\langle \mathbf{u}_\perp(z=0,\mathbf{r}_\perp)\cdot\boldsymbol{\nabla}_\perp |\tilde{f}(k_\parallel)|^2\right\rangle. \tag{C 2}$$

Formally, this looks like an equation for the spectrum of a passive 2D field $\tilde{f}(k_\parallel,\mathbf{r}_\perp)$, parametrised by $k_\parallel$, advected by a 2D velocity field $\mathbf{u}_\perp(z=0,\mathbf{r}_\perp)$ and decaying with s,

which plays the role of time. While devising a specific quantitative closure for the triple correlator in the right-hand side of equation (C 2) is outside the scope of this paper, it is plausible that the solutions around the phase-mixing threshold $k_\parallel v_{\rm th} \sim s\, k_\perp u_\perp$ will satisfy, roughly,

$$\frac{\partial \tilde{f}^2}{\partial s} \sim \frac{k_\perp u_\perp}{k_\parallel v_{\rm th}}\, \tilde{f}^2 \sim \frac{\tilde{f}^2}{s} \quad \Rightarrow \quad \tilde{f}^2(s, k_\parallel) \propto \frac{1}{s^\mu}. \tag{C 3}$$

Thus, the decay must be a power law, as we indeed assumed in equation (4.26). This decay law is set at the "outer scale", which is the phase-mixing threshold: $k_\perp \sim (k_\parallel/s)^{1/r}$ ($k_\parallel$ is fixed). Below this scale, i.e., at $k_\perp \gg (k_\parallel/s)^{1/r}$, the phase-mixing term is small and the $\tilde{f}(s, k_\parallel)$ is simply cascaded subject to the constant-flux argument proposed in section 4.4.1 (i.e., the right-hand side of equation (C 2) must vanish to lowest order in $1/s$). This gives a spectrum of the form (4.26), inheriting its decay law $\sim 1/m^{\sigma'}$ from the "outer scale". The decay law of the spectrum in the advection-dominated region $k_\perp \gtrsim k_\parallel^{1/r}$, equation (4.25), is then the same, $\sigma = \sigma'$, via matching at the critical-balance curve $k_\perp \sim k_\parallel^{1/r}$ (see equation (4.28)).

What is the relationship between μ and σ' and how is this scaling exponent determined? Equation (C 3) effectively sets the 1D parallel-wavenumber spectrum, i.e., as explained above, the free-energy content of all wavenumbers $k_\perp \gtrsim (k_\parallel/s)^{1/r}$: using equation (4.26), we get

$$\frac{\tilde{f}^2(s, k_\parallel)}{s} \sim E_m^\parallel(k_\parallel) \sim \int_{(k_\parallel/\sqrt{m})^{1/r}}^\infty {\rm d}k_\perp E_m^+(k_\parallel, k_\perp) \sim \frac{k_\parallel^{-a'-(d'-1)/r}}{m^{\sigma'-(d'-1)/2r}}, \tag{C 4}$$

so $\mu = 2\sigma' - 1 - (d'-1)/r$. This decay exponent, or, equivalently, σ', is deduced (along with a') by matching the decay law (C 3) with the decay law of the total variance of $\tilde{f}^2(k_\parallel)$ contained at long wavelengths $k_\perp \lesssim (k_\parallel/s)^{1/r}$: using the asymptotic form (4.8), we get

$$\frac{\tilde{f}^2(k_\parallel)}{s} \sim E_m^\parallel(k_\parallel) \sim \int_0^{(k_\parallel/\sqrt{m})^{1/r}} {\rm d}k_\perp E_m^+(k_\parallel, k_\perp) \sim \frac{k_\parallel^{-a+4/r}}{m^{1/2+2/r}}, \tag{C 5}$$

so $\mu = 4/r$. Matching equations (C 4) and (C 5), we get two relations constraining σ' and a', which, combined with matching conditions at the critical-balance curve $k_\perp \sim k_\parallel^{1/r}$, are the same as equations (4.28). Note that using the set of exponents (4.31) and (4.33), we happily recover the 1D parallel spectrum (4.36) from either of equations (C 4) and (C 5). We also find that $\mu = 3$.

Note that deducing the decay law of a turbulent field by fixing the long-wavelength asymptotic behaviour of its spectrum ($E_m \propto k_\perp^3$ in our case) is a standard trick of the trade in turbulence theory (e.g., Kolmogorov 1941*a*; Corrsin 1951; Saffman 1967; Eyink & Xin 2000; Schekochihin *et al.* 2004; Davidson 2010, 2013).

REFERENCES

Abel, I. G. & Cowley, S. C. 2013 Multiscale gyrokinetics for rotating tokamak plasmas: II. Reduced models for electron dynamics. *New J. Phys.* **15**, 023041.

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.

Alexandrova, O., Lacombe, C., Mangeney, A., Grappin, R. & Maksimovic, M. 2012 Solar wind turbulent spectrum at plasma kinetic scales. *Astrophys. J.* **760**, 121.

Alexandrova, O., Saur, J., Lacombe, C., Mangeney, A., Mitchell, J., Schwartz, S. J. & Robert, P. 2009 Universality of solar-wind turbulent spectrum from MHD to electron scales. *Phys. Rev. Lett.* **103**, 165003.

Armstrong, T. P. 1967 Numerical studies of the nonlinear Vlasov equation. *Phys. Fluids* **10**, 1269.

Bañón Navarro, A., Morel, P., Albrecht-Marc, M., Carati, D., Merz, F., Görler, T. & Jenko, F. 2011*a* Free energy balance in gyrokinetic turbulence. *Phys. Plasmas* **18**, 092303.

Bañón Navarro, A., Morel, P., Albrecht-Marc, M., Carati, D., Merz, F., Görler, T. & Jenko, F. 2011*b* Free energy cascade in gyrokinetic turbulence. *Phys. Rev. Lett.* **106**, 055001.

Bañón Navarro, A., Teaca, B., Jenko, F., Hammett, G. W. & Happel, T. 2014 Applications of large eddy simulation methods to gyrokinetic turbulence. *Phys. Plasmas* **21**, 032304.

Barnes, A. 1966 Collisionless damping of hydromagnetic waves. *Phys. Fluids* **9**, 1483.

Barnes, M., Parra, F. I. & Schekochihin, A. A. 2011 Critically balanced ion temperature gradient turbulence in fusion plasmas. *Phys. Rev. Lett.* **107**, 115003.

Batchelor, G. K. 1953 *The Theory of Homogeneous Turbulence*. Cambridge: Cambridge University Press.

Beer, M. A. & Hammett, G. W. 1996 Toroidal gyrofluid equations for simulations of tokamak turbulence. *Phys. Plasmas* **3**, 4046.

Beresnyak, A. 2015 On the parallel spectrum in magnetohydrodynamic turbulence. *Astrophys. J.* **801**, L9.

Bernstein, I. B. 1958 Waves in a plasma in a magnetic field. *Phys. Rev.* **109**, 10.

Bershadskii, A. & Sreenivasan, K. R. 2004 Intermittency and the passive nature of the magnitude of the magnetic field. *Phys. Rev. Lett.* **93**, 064501.

Black, C., Germaschewski, K., Bhattacharjee, A. & Ng, C. S. 2013 Discrete kinetic eigenmode spectra of electron plasma oscillations in weakly collisional plasma: A numerical study. *Phys. Plasmas* **20**, 012125.

Boldyrev, S. 2005 On the spectrum of magnetohydrodynamic turbulence. *Astrophys. J.* **626**, L37.

Bratanov, V., Jenko, F., Hatch, D. R. & Wilczek, M. 2013 Nonuniversal Power-Law Spectra in Turbulent Systems. *Phys. Rev. Lett.* **111**, 075001.

Candy, J. & Waltz, R. E. 2006 Velocity-space resolution, entropy production, and upwind dissipation in Eulerian gyrokinetic simulations. *Phys. Plasmas* **13**, 032310.

Casati, A., Gerbaud, T., Hennequin, P., Bourdelle, C., Candy, J., Clairet, F., Garbet, X., Grandgirard, V., Gürcan, Ö. D., Heuraux, S., Hoang, G. T., Honoré, C., Imbeaux, F., Sabot, R., Sarazin, Y., Vermare, L. & Waltz, R. E. 2009 Turbulence in the TORE SUPRA tokamak: measurements and validation of nonlinear simulations. *Phys. Rev. Lett.* **102**, 165005.

Celnikier, L. M., Harvey, C. C., Jegou, R., Moricet, P. & Kemp, M. 1983 A determination of the electron density fluctuation spectrum in the solar wind, using the ISEE propagation experiment. *Astron. Astrophys.* **126**, 293.

Celnikier, L. M., Muschietti, L. & Goldman, M. V. 1987 Aspects of interplanetary plasma turbulence. *Astron. Astrophys.* **181**, 138.

Chang, O., Peter Gary, S. & Wang, J. 2011 Whistler turbulence forward cascade: Three-dimensional particle-in-cell simulations. *Geophys. Res. Lett.* **38**, L22102.

Chen, C. H. K., Bale, S. D., Salem, C. & Mozer, F. S. 2011 Frame dependence of the electric field spectrum of solar wind turbulence. *Astrophys. J.* **737**, L41.

Chen, C. H. K., Boldyrev, S., Xia, Q. & Perez, J. C. 2013 Nature of subproton scale turbulence in the solar wind. *Phys. Rev. Lett.* **110**, 225002.

Chen, C. H. K., Horbury, T. S., Schekochihin, A. A., Wicks, R. T., Alexandrova, O. & Mitchell, J. 2010 Anisotropy of solar wind turbulence between ion and electron scales. *Phys. Rev. Lett.* **104**, 255002.

Chen, C. H. K., Sorriso-Valvo, L., Šafránková, J. & Němeček, Z. 2014 Intermittency of solar wind density fluctuations from ion to electron scales. *Astrophys. J.* **789**, L8.

Cho, J. & Lazarian, A. 2004 The anisotropy of electron magnetohydrodynamic turbulence. *Astrophys. J.* **615**, L41.

Connaughton, C., Nazarenko, S. & Quinn, B. 2014 Rossby and drift wave turbulence and zonal flows: the Charney-Hasegawa-Mima model and its extensions. *arXiv:1407.1896* .

Coppi, B., Rosenbluth, M. N. & Sagdeev, R. Z. 1967 Instabilities due to temperature gradients in complex magnetic field configurations. *Phys. Fluids* **10**, 582.

Corrsin, S. 1951 On the spectrum of isotropic temperature fluctuations in an isotropic turbulence. *J. Aeronaut. Sci.* **18**, 417.

Corrsin, S. 1963 Estimates of the relations between Eulerian and Lagrangian scales in large Reynolds number turbulence. *J. Atmos. Sci.* **20**, 115.

Cowley, S. C., Kulsrud, R. M. & Sudan, R. 1991 Considerations of ion-temperature-gradient-driven turbulence. *Phys. Fluids B* **3**, 2767.

Crownfield, Jr., F. R. 1977 Plasma oscillations and Landau damping. *Phys. Fluids* **20**, 1483.

Davidson, P. A. 2010 On the decay of Saffman turbulence subject to rotation, stratification or an imposed magnetic field. *J. Fluid Mech.* **663**, 268.

Davidson, P. A. 2013 *Turbulence in Rotating, Stratified and Electrically Conducting Fluids*. Cambridge: Cambridge University Press.

Diamond, P. H., Hasegawa, A. & Mima, K. 2011 Vorticity dynamics, drift wave turbulence, and zonal flows: a look back and a look ahead. *Plasma Phys. Control. Fusion* **53**, 124001.

Diamond, P. H., Itoh, S.-I., Itoh, K. & Hahm, T. S. 2005 Zonal flows in plasma—a review. *Plasma Phys. Control. Fusion* **47**, 35.

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.

Dorland, W. & Hammett, G. W. 1993 Gyrofluid turbulence models with kinetic effects. *Phys. Fluids B* **5**, 812.

Dorland, W., Jenko, F., Kotschenreuther, M. & Rogers, B. N. 2000 Electron temperature gradient turbulence. *Phys. Rev. Lett.* **85**, 5579.

Drake, J. F., Guzdar, P. N. & Dimits, A. 1991 Three-dimensional simulation of ∇T_i-driven turbulence and transport. *Phys. Fluids B* **3**, 1937.

Drake, J. F., Guzdar, P. N. & Hassam, A. B. 1988 Streamer formation in plasma with a temperature gradient. *Phys. Rev. Lett.* **61**, 2205.

Eltgroth, P. G. 1974 Plasma heating calculations using a transform method. *Phys. Fluids* **17**, 1602.

Eyink, G. L. & Xin, J. 2000 Self-similar decay in the Kraichnan model of a passive scalar. *J. Stat. Phys.* **100**, 679.

Fowler, T. K. 1963 Lyapunov's stability criteria for plasmas. *J. Math. Phys.* **4**, 559.

Fowler, T. K. 1968 Thermodynamics of unstable plasmas. *Adv. Plasma Phys.* **1**, 201.

Gary, S. P. & Borovsky, J. E. 2008 Damping of long-wavelength kinetic Alfvén fluctuations: Linear theory. *J. Geophys. Res.* **113**, A12104.

Ghim, Y.-c., Schekochihin, A. A., Field, A. R., Abel, I. G., Barnes, M., Colyer, G., Cowley, S. C., Parra, F. I., Dunai, D. & Zoletnik, S. 2013 Experimental signatures of critically balanced turbulence in MAST. *Phys. Rev. Lett.* **110**, 145002.

Goldreich, P. & Sridhar, S. 1995 Toward a theory of interstellar turbulence. 2: Strong Alfvénic turbulence. *Astrophys. J.* **438**, 763.

Goldreich, P. & Sridhar, S. 1997 Magnetohydrodynamic turbulence revisited. *Astrophys. J.* **485**, 680.

Görler, T. & Jenko, F. 2008 Multiscale features of density and frequency spectra from nonlinear gyrokinetics. *Phys. Plasmas* **15**, 102508.

Goswami, P., Passot, T. & Sulem, P. L. 2005 A Landau fluid model for warm collisionless plasmas. *Phys. Plasmas* **12**, 102109.

Gould, R. W., O'Neil, T. M. & Malmberg, J. H. 1967 Plasma wave echo. *Phys. Rev. Lett.* **19**, 219.

Grad, H. 1949 On the kinetic theory of rarefied gases. *Comm. Pure Appl. Math.* **2**, 331.

Grant, F. C. & Feix, M. R. 1967 Fourier-Hermite solutions of the Vlasov equations in the linearized limit. *Phys. Fluids* **10**, 696.

Gürcan, Ö. D., Garbet, X., Hennequin, P., Diamond, P. H., Casati, A. & Falchetto, G. L. 2009 Wave-number spectrum of drift-wave turbulence. *Phys. Rev. Lett.* **102**, 255002.

Hallatschek, K. 2004 Thermodynamic potential in local turbulence simulations. *Phys. Rev. Lett.* **93**, 125001.

Hammett, G. W., Beer, M. A., Dorland, W., Cowley, S. C. & Smith, S. A. 1993 Developments in the gyrofluid approach to tokamak turbulence simulations. *Plasma Phys. Control. Fusion* **35**, 973.

Hammett, G. W., Dorland, W. & Perkins, F. W. 1992 Fluid models of phase mixing, Landau damping, and nonlinear gyrokinetic dynamics. *Phys. Fluids B* **4**, 2052.

Hammett, G. W. & Perkins, F. W. 1990 Fluid moment models for Landau damping with application to the ion-temperature-gradient instability. *Phys. Rev. Lett.* **64**, 3019.

Hatch, D. R., Jenko, F., Bañón Navarro, A. & Bratanov, V. 2013 Transition between saturation regimes of gyrokinetic turbulence. *Phys. Rev. Lett.* **111**, 175001.

Hatch, D. R., Jenko, F., Bratanov, V., Navarro, A. B. & Navarro 2014 Phase space scales of free energy dissipation in gradient-driven gyrokinetic turbulence. *J. Plasma Phys.* **80**, 531.

Hatch, D. R., Terry, P. W., Jenko, F., Merz, F. & Nevins, W. M. 2011*a* Saturation of Gyrokinetic Turbulence through Damped Eigenmodes. *Phys. Rev. Lett.* **106**, 115003.

Hatch, D. R., Terry, P. W., Jenko, F., Merz, F., Pueschel, M. J., Nevins, W. M. & Wang, E. 2011*b* Role of subdominant stable modes in plasma microturbulence. *Phys. Plasmas* **18**, 055706.

Hedrick, C. L. & Leboeuf, J.-N. 1992 Landau fluid equations for electromagnetic and electrostatic fluctuations. *Phys. Fluids B* **4**, 3915.

Hennequin, P., Sabot, R., Honoré, C., Hoang, G. T., Garbet, X., Truc, A., Fenzi, C. & Quéméneur, A. 2004 Scaling laws of density fluctuations at high-k on Tore Supra. *Plasma Phys. Control. Fusion* **46**, B121.

Hnat, B., Chapman, S. C. & Rowlands, G. 2005 Compressibility in solar wind plasma turbulence. *Phys. Rev. Lett.* **94**, 204502.

Horton, W. 1999 Drift waves and transport. *Rev. Mod. Phys.* **71**, 735.

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.

Howes, G. G., Cowley, S. C., Dorland, W., Hammett, G. W., Quataert, E. & Schekochihin, A. A. 2008 A model of turbulence in magnetized plasmas: Implications for the dissipation range in the solar wind. *J. Geophys. Res.* **113**, A05103.

Howes, G. G., Tenbarge, J. M., Dorland, W., Quataert, E., Schekochihin, A. A., Numata, R. & Tatsuno, T. 2011 Gyrokinetic simulations of solar wind turbulence from ion to electron scales. *Phys. Rev. Lett.* **107**, 035004.

Jenko, F., Dorland, W., Kotschenreuther, M. & Rogers, B. N. 2000 Electron temperature gradient driven turbulence. *Phys. Plasmas* **7**, 1904.

Kanekar, A., Schekochihin, A. A., Dorland, W. & Loureiro, N. F. 2015 Fluctuation-dissipation relations for a plasma-kinetic Langevin equation. *J. Plasma Phys.* **81**, 305810104.

Kanekar, A. V. 2015 Phase mixing in turbulent magnetized plasmas. *Ph. D. Thesis.* University of Maryland, College Park (http://drum.lib.umd.edu/handle/1903/16418).

Kellogg, P. J. & Horbury, T. S. 2005 Rapid density fluctuations in the solar wind. *Ann. Geophys.* **23**, 3765.

Kirkwood, J. G. 1946 The statistical mechanical theory of transport processes. I. General theory. *J. Chem. Phys.* **14**, 180.

Kobayashi, S. & Gürcan, Ö. D. 2015 Gyrokinetic turbulence cascade via predator-prey interactions between different scales. *Phys. Plasmas* **22**, 050702.

Kolmogorov, A. N. 1941*a* On the degeneration of isotropic turbulence in an incompressible viscous fluid. *Dokl. Acad. Nauk SSSR* **31**, 538.

Kolmogorov, A. N. 1941*b* The local structure of turbulence in incompressible viscous fluid at very large Reynolds numbers. *Dokl. Acad. Nauk SSSR* **30**, 299.

Krommes, J. A. 1993 Dielectric response and thermal fluctuations in gyrokinetic plasma. *Phys. Fluids B* **5**, 1066.

Krommes, J. A. 1999 Thermostatted δf. *Phys. Plasmas* **6**, 1477.

Krommes, J. A. 2010 Nonlinear gyrokinetics: a powerful tool for the description of microturbulence in magnetized plasmas. *Physica Scripta* **T142**, 014035.

Krommes, J. A. & Hu, G. 1994 The role of dissipation in the theory and simulations of homogeneous plasma turbulence, and resolution of the entropy paradox. *Phys. Plasmas* **1**, 3211.

Kruskal, M. D. & Oberman, C. R. 1958 On the stability of plasma in static equilibrium. *Phys. Fluids* **1**, 275.

Kunz, M. W., Schekochihin, A. A., Chen, C. H. K., Abel, I. G. & Cowley, S. C. 2015 Inertial-range kinetic turbulence in pressure-anisotropic astrophysical plasmas. *J. Plasma Phys.* **81**, 325810501.

Landau, L. 1936 Transport equation in the case of Coulomb interaction. *Zh. Eksp. Teor. Fiz.* **7**, 203.

Landau, L. 1946 On the vibration of the electronic plasma. *Zh. Eksp. Teor. Fiz.* **16**, 574.

Lenard, A. & Bernstein, I. B. 1958 Plasma oscillations with diffusion in velocity space. *Phys. Rev.* **112**, 1456.

Lithwick, Y. & Goldreich, P. 2001 Compressible magnetohydrodynamic turbulence in interstellar plasmas. *Astrophys. J.* **562**, 279.

Loureiro, N. F., Schekochihin, A. A. & Zocco, A. 2013 Fast collisionless reconnection and electron heating in strongly magnetized plasmas. *Phys. Rev. Lett.* **111**, 025002.

Makwana, K. D., Terry, P. W., Pueschel, M. J. & Hatch, D. R. 2014 Subdominant modes in zonal-flow-regulated turbulence. *Phys. Rev. Lett.* **112**, 095002.

Malmberg, J. H., Wharton, C. B., Gould, R. W. & O'Neil, T. M. 1968 Plasma wave echo experiment. *Phys. Rev. Lett.* **20**, 95.

Marsch, E. & Tu, C.-Y. 1990 Spectral and spatial evolution of compressible turbulence in the inner solar wind. *J. Geophys. Res.* **95**, 11945.

Mattor, N. 1992 Can Landau-fluid models describe nonlinear Landau damping? *Phys. Fluids B* **4**, 3952.

Morel, P., Bañón Navarro, A., Albrecht-Marc, M., Carati, D., Merz, F., Görler, T. & Jenko, F. 2012 Dynamic procedure for filtered gyrokinetic simulations. *Phys. Plasmas* **19**, 012311.

Morel, P., Navarro, A. B., Albrecht-Marc, M., Carati, D., Merz, F., Görler, T. & Jenko, F. 2011 Gyrokinetic large eddy simulations. *Phys. Plasmas* **18**, 072301.

Nakata, M., Watanabe, T.-H. & Sugama, H. 2012 Nonlinear entropy transfer via zonal flows in gyrokinetic plasma turbulence. *Phys. Plasmas* **19**, 022303.

Nazarenko, S. V. & Schekochihin, A. A. 2011 Critical balance in magnetohydrodynamic, rotating and stratified turbulence: towards a universal scaling conjecture. *J. Fluid Mech.* **677**, 134.

Ng, C. S., Bhattacharjee, A. & Skiff, F. 1999 Kinetic eigenmodes and discrete spectrum of plasma oscillations in a weakly collisional plasma. *Phys. Rev. Lett.* **83**, 1974.

Ottaviani, M., Beer, M. A., Cowley, S. C., Horton, W. & Krommes, J. A. 1997 Unanswered questions in ion-temperature-gradient-driven turbulence. *Phys. Rep.* **283**, 121.

Parker, J. T. 2016 Gyrokinetic simulations of fusion plasmas using a spectral velocity space representation. *D. Phil. Thesis.* University of Oxford (arXiv:1603.04727).

Parker, J. T. & Dellar, P. J. 2015 Fourier-Hermite spectral representation for the Vlasov-Poisson system in the weakly collisional limit. *J. Plasma Phys.* **81**, 305810203.

Parker, J. T., Highcock, E. G., Schekochihin, A. A. & Dellar, P. J. 2016 Suppression of phase mixing in drift-kinetic plasma turbulence. *arXiv:1603.06968* .

Parker, S. E. & Carati, D. 1995 Renormalized dissipation in plasmas with finite collisionality. *Phys. Rev. Lett.* **75**, 441.

Passot, T. & Sulem, P. L. 2004 A Landau fluid model for dispersive magnetohydrodynamics. *Phys. Plasmas* **11**, 5173.

Passot, T. & Sulem, P. L. 2006 A fluid model with finite Larmor radius effects for mirror mode dynamics. *J. Geophys. Res.* **111**, A04203.

Passot, T. & Sulem, P. L. 2007 Collisionless magnetohydrodynamics with gyrokinetic effects. *Phys. Plasmas* **14**, 082502.

Passot, T. & Sulem, P. L. 2015 A model for the non-universal power law of the solar wind sub-ion-scale magnetic spectrum. *Astrophys. J.* **812**, L37.

Passot, T., Sulem, P. L. & Hunana, P. 2012 Extending magnetohydrodynamics to the slow dynamics of collisionless plasmas. *Phys. Plasmas* **19**, 082113.

Plunk, G. G. 2013 Landau damping in a turbulent setting. *Phys. Plasmas* **20**, 032304.

Plunk, G. G., Bañón Navarro, A. & Jenko, F. 2015 Understanding nonlinear saturation in zonal-flow-dominated ion temperature gradient turbulence. *Plasma Phys. Control. Fusion* **57**, 045005.

Plunk, G. G., Cowley, S. C., Schekochihin, A. A. & Tatsuno, T. 2010 Two-dimensional gyrokinetic turbulence. *J. Fluid Mech.* **664**, 407.

Plunk, G. G. & Parker, J. T. 2014 Irreversible energy flow in forced Vlasov dynamics. *Eur. Phys. J. D* **68**, 296.

Podesta, J. J., Borovsky, J. E. & Gary, S. P. 2010 A kinetic Alfvén wave cascade subject to collisionless damping cannot reach electron scales in the solar wind at 1 AU. *Astrophys. J.* **712**, 685.

Quinn, B., Nazarenko, S., Connaughton, C., Gallagher, S. & Hnat, B. 2013 Modulational instability in basic plasma and geophysical models. *arXiv:1312.4256* .

Ramos, J. J. 2005 Fluid formalism for collisionless magnetized plasmas. *Phys. Plasmas* **12**, 052102.

Richardson, L. F. 1922 *Weather Prediction by Numerical Process*. Cambridge: Cambridge University Press.

Rogers, B. N., Dorland, W. & Kotschenreuther, M. 2000 Generation and stability of zonal flows in ion-temperature-gradient mode turbulence. *Phys. Rev. Lett.* **85**, 5336.

Rogers, B. N., Drake, J. F. & Zeiler, A. 1998 Phase space of tokamak edge turbulence, the L-H transition, and the formation of the edge pedestal. *Phys. Rev. Lett.* **81**, 4396.

Rudakov, L. I. & Sagdeev, R. Z. 1961 On the instability of inhomogeneous rarefied plasma in a strong magnetic field. *Dokl. Acad. Nauk SSSR* **138**, 581.

Saffman, P. G. 1967 The large-scale structure of homogeneous turbulence. *J. Fluid Mech.* **27**, 581.

Sahraoui, F., Goldstein, M. L., Belmont, G., Canu, P. & Rezeau, L. 2010 Three dimensional anisotropic k spectra of turbulence at subproton scales in the solar wind. *Phys. Rev. Lett.* **105**, 131101.

Sahraoui, F., Goldstein, M. L., Robert, P. & Khotyaintsev, Y. V. 2009 Evidence of a cascade and dissipation of solar-wind turbulence at the electron gyroscale. *Phys. Rev. Lett.* **102**, 231102.

Sahraoui, F., Huang, S. Y., Belmont, G., Goldstein, M. L., Rétino, A., Robert, P. & De Patoul, J. 2013 Scaling of the electron dissipation range of solar wind turbulence. *Astrophys. J.* **777**, 15.

Schekochihin, A. A., Cowley, S. C. & Dorland, W. 2007 Interplanetary and interstellar plasma turbulence. *Plasma Phys. Control. Fusion* **49**, 195.

Schekochihin, A. A., Cowley, S. C., Dorland, W., Hammett, G. W., Howes, G. G., Plunk, G. G., Quataert, E. & Tatsuno, T. 2008 Gyrokinetic turbulence: a nonlinear route to dissipation through phase space. *Plasma Phys. Control. Fusion* **50**, 124024.

Schekochihin, A. A., Cowley, S. C., Dorland, W., Hammett, G. W., Howes, G. G., Quataert, E. & Tatsuno, T. 2009 Astrophysical gyrokinetics: kinetic and fluid turbulent cascades in magnetized weakly collisional plasmas. *Astrophys. J. Suppl.* **182**, 310.

Schekochihin, A. A., Haynes, P. H. & Cowley, S. C. 2004 Diffusion of passive scalar in a finite-scale random flow. *Phys. Rev. E* **70**, 046304.

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.

Schekochihin, A. A., Stipani, L., Califano, F., Staines, C., Kanekar, A., Dorland, W. & Hammett, G. W. 2016 Phase mixing vs. nonlinear advection in drift-kinetic plasma turbulence. II. Kinetic passive scalar. *J. Plasma Phys.* In preparation.

Scott, B. 2010 Derivation via free energy conservation constraints of gyrofluid equations with finite-gyroradius electromagnetic nonlinearities. *Phys. Plasmas* **17**, 102306.

Smagorinsky, J. 1963 General circulation experiments with the primitive equations. *Mon. Weather Rev.* **91**, 99.

Smith, S. A. 1997 Dissipative closures for statistical moments, fluid moments, and subgrid scales in plasma turbulence. *Ph. D. Thesis.* Princeton University (http://w3.pppl.gov/∼hammett/sasmith/thesis.html).

Snyder, P. B. & Hammett, G. W. 2001*a* A Landau fluid model for electromagnetic plasma microturbulence. *Phys. Plasmas* **8**, 3199.

Snyder, P. B. & Hammett, G. W. 2001*b* Electromagnetic effects on plasma microturbulence and transport. *Phys. Plasmas* **8**, 744.

Snyder, P. B., Hammett, G. W. & Dorland, W. 1997 Landau fluid models of collisionless magnetohydrodynamics. *Phys. Plasmas* **4**, 3974.

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.

Tatsuno, T., Dorland, W., Schekochihin, A. A., Plunk, G. G., Barnes, M., Cowley, S. C. & Howes, G. G. 2009 Nonlinear phase mixing and phase-space cascade of entropy in gyrokinetic plasma turbulence. *Phys. Rev. Lett.* **103**, 015003.

Teaca, B., Navarro, A. B. & Jenko, F. 2014 The energetic coupling of scales in gyrokinetic plasma turbulence. *Phys. Plasmas* **21**, 072308.

Teaca, B., Navarro, A. B., Jenko, F., Brunner, S. & Villard, L. 2012 Locality and universality in gyrokinetic turbulence. *Phys. Rev. Lett.* **109**, 235003.

Told, D., Jenko, F., TenBarge, J. M., Howes, G. G. & Hammett, G. W. 2015 Multiscale nature of the dissipation range in gyrokinetic simulations of Alfvénic turbulence. *Phys. Rev. Lett.* **115**, 025003.

Vermare, L., Hennequin, P., Gürcan, Ö. D., Bourdelle, C., Clairet, F., Garbet, X., Sabot, R. & Tore Supra Team 2011 Impact of collisionality on fluctuation characteristics of micro-turbulence. *Phys. Plasmas* **18**, 012306.

Watanabe, T.-H. & Sugama, H. 2004 Kinetic simulation of steady states of ion temperature gradient driven turbulence with weak collisionality. *Phys. Plasmas* **11**, 1476.

Watanabe, T.-H. & Sugama, H. 2006 Velocity space structures of distribution function in toroidal ion temperature gradient turbulence. *Nucl. Fusion* **46**, 24.

Weiland, J. 1992 Nonlinear effects in velocity space and drift wave transport in tokamaks. *Phys. Fluids B* **4**, 1388.

Zakharov, V. E., L'vov, V. S. & Falkovich, G. 1992 *Kolmogorov Spectra of Turbulence I: Wave Turbulence*. Berlin: Springer.

Zocco, A. & Schekochihin, A. A. 2011 Reduced fluid-kinetic equations for low-frequency dynamics, magnetic reconnection, and electron heating in low-beta plasmas. *Phys. Plasmas* **18**, 102309.