Plasma Phys. Control. Fusion **50** (2008) 085014 (29pp) doi:10.1088/0741-3335/50/8/085014

Hamiltonian formulation and analysis of a collisionless fluid reconnection model

E Tassi[1], P J Morrison[2], F L Waelbroeck[2] and D Grasso[1]

[1] Burning Plasma Research Group, Dipartimento di Energetica and CNISM, Politecnico di Torino, Torino, 10129, Italy

[2] Department of Physics and Institute for Fusion Studies, University of Texas, Austin, Texas 78712, USA

E-mail: emanuele.tassi@polito.it

Received 27 February 2008, in final form 19 May 2008
Published 10 June 2008
Online at stacks.iop.org/PPCF/50/085014

Abstract

The Hamiltonian formulation of a plasma four-field fluid model that describes collisionless reconnection is presented. The formulation is noncanonical with a corresponding Lie–Poisson bracket. The bracket is used to obtain new independent families of invariants, so-called Casimir invariants, three of which are directly related to Lagrangian invariants of the system. The Casimirs are used to obtain a variational principle for equilibrium equations that generalize the Grad–Shafranov equation to include flow. Dipole and homogeneous equilibria are constructed. The linear dynamics of the latter is treated in detail in a Hamiltonian context: canonically conjugate variables are obtained; the dispersion relation is analyzed and exact thresholds for spectral stability are obtained; the canonical transformation to normal form is described; an unambiguous definition of negative energy modes is given; and thresholds sufficient for energy-Casimir stability are obtained. The Hamiltonian formulation is also used to obtain an expression for the collisionless conductivity and it is further used to describe the linear growth and nonlinear saturation of the collisionless tearing mode.

(Some figures in this article are in colour only in the electronic version)

1. Introduction

Due to the lower dimensionality of configuration space as compared with phase space, fluid models of the plasma have an intrinsic computational advantage over kinetic models. For this reason it is of great interest to develop and apply such models even in cases where the collisionality is too small to provide a firm justification for their use. In particular, fluid models have made important contributions to the understanding of magnetic reconnection, a phenomenon that plays a key role in events such as solar flares, magnetospheric substorms, and sawtooth oscillations in tokamaks [1–4]. They have also made key contributions to the understanding of plasma turbulence in the core [5–8] and edge [9–12] of magnetic confinement experiments in collisionless as well as collisional regimes. More generally, they offer the promise of being able to perform simulations of multiscale phenomena that are beyond the reach of kinetic models even after accounting for foreseeable advances in computation speed [13].

An important consideration when constructing new plasma fluid models is the existence of a Hamiltonian structure (see [14–16] for reviews). Because the fundamental laws governing charged particle dynamics are Hamiltonian, dissipative terms, which ultimately arise from simplifications, must be accompanied by phenomenological constants such as resistivity and viscosity. When such phenomenological quantities are neglected, it is desirable that the resulting model be Hamiltonian, as is the case for the most important kinetic and fluid models of plasma physics. The preservation of the Hamiltonian structure provides some confidence that the truncations that are used to derive the fluid model have not introduced unphysical sources of dissipation. The presence of the Hamiltonian structure has the additional benefit of providing important tools for calculations. For example, the magnetohydrodynamic (MHD) energy principle is a consequence of the Hamiltonian nature of MHD.

Since the discovery of the noncanonical Hamiltonian structure of MHD [17], many plasma fluid models have been shown to possess a Hamiltonian description in terms of noncanonical Poisson brackets, e.g. [4, 6–8, 18–23]. In some cases, the requirement that the dynamics has Hamiltonian form has been used to guide the construction [6,20] and has led to the identification of new and physically important terms [20]. In another case, the absence of a Hamiltonian structure for a given model was shown to lead to the violation of the solubility conditions for the equilibrium equations [24]. The Hamiltonian structure of fluid models has also been shown to be important for the consistent calculation of zonal flow dynamics [6–8].

Several fluid models have been proposed to study electromagnetic plasma dynamics (see [12, 25, 26] for reviews). Some of these models have been instrumental in advancing our understanding of magnetic reconnection [4, 20, 21, 25, 27]. In particular, the model of [20] led to the discovery of fast (compared with Sweet–Parker) magnetic reconnection by Aydemir [28, 29]. This model included the effects of finite ion temperature, but neglected electron inertia. Around the same time Hazeltine and Meiss (HM) [25] included electron inertia and made other improvements to the four-field model of Hazeltine, Kotschenreuther and Morrison [30, 31]. The HM model was originally used to provide a unified description of the formation of current channels in semi-collisional and collisionless regimes. Its Hamiltonian nature, however, has not been investigated until now.

Interest in the effects of electron inertia was recently revived by a controversy over its influence on the rate of collisionless magnetic reconnection. In collisionless magnetic reconnection, the ‘frozen-in’ condition of MHD is broken by the inclusion of electron inertia instead of resistivity [32,33]. This led Schep and collaborators to study two models that may be viewed as limiting cases of the HM model (aside from the fact that they avoid the Boussinesq approximation, retaining instead the density n in the form $\log n/n_0$) [4,21,27]. These authors constructed a Hamiltonian formulation for these models, which were subsequently used to demonstrate the role of phase mixing of the Lagrangian invariants during fast reconnection [22] and to investigate the role of instabilities of the nonlinearly developed current sheet [34]. More recently, Fitzpatrick and Porcelli (FP) have considered another limiting form of the HM model that, compared with the model derived by Schep *et al* is valid for a wider range of values of β, the latter indicating the ratio between the plasma pressure and the magnetic pressure based on the toroidal guide field. The FP model also extends the models of Schep *et al* by including the effects of parallel ion compressibility. It has subsequently been used to study two-fluid effects on the Taylor problem [35] and on the linear growth of tearing modes [36]. Rogers *et al* have recently shown that the predictions of the FP model for the linear growth rate of the tearing mode are in good agreement with those obtained with the gyrokinetic code GS2 when the ion temperature is not too large [37].

In this paper, we investigate the Hamiltonian structure of the FP version of the model of HM [25, 38]. The paper is organized as follows. In section 3 we present the noncanonical Poisson bracket and show that this bracket produces the equations of motion with the appropriate Hamiltonian. In section 4, we use the noncanonical Poisson bracket to obtain four infinite families of new Casimir invariants, three of which suggest that specific combinations of the field variables are Lagrangian invariants. In terms of these variables, the equations of motion and the Hamiltonian structure achieve a much simplified form. A preliminary version of these results was announced in [39].

The remaining sections describe a variety of applications that rely on the Hamiltonian structure. In section 5 we describe a variational principle for equilibria of the system and show that they are governed by a generalized Grad–Shafranov system of a pair of coupled elliptic equations. We treat two examples: dipole equilibria with Bessel function solutions and homogeneous equilibria that support wave motion. Section 6 explores the latter example in further detail by showing how to construct conventional canonical variables for the linear dynamics. We show that the system possesses Alfvén-like and drift-shear modes, obtain exact stability thresholds, and give a definition of negative energy modes. We note that the Hamiltonian form is indispensable for an unambiguous definition of negative energy modes. We also obtain energy-stability conditions, sufficient conditions for stability akin to the δW criterion of MHD. Section 7 contains a derivation of the collisionless conductivity that relies on the Jacobi identity of the Hamiltonian formulation, which we then use to obtain the tearing-layer parameter Δ' and the growth rate for the collisionless tearing mode. In section 8 we use the conservation of a Casimir invariant to obtain the nonlinearly saturated current profile and compare it with that obtained by Rutherford [40]. In section 9 we summarize and conclude.

2. Model equations

The model of [38] is given by the following equations:

$$\frac{\partial(\psi - d_e^2\nabla^2\psi)}{\partial t} + [\varphi, \psi - d_e^2\nabla^2\psi] - d_\beta[\psi, Z] = 0, \tag{1}$$

$$\frac{\partial Z}{\partial t} + [\varphi, Z] - c_\beta[v, \psi] - d_\beta[\nabla^2\psi, \psi] = 0, \tag{2}$$

$$\frac{\partial \nabla^2\varphi}{\partial t} + [\varphi, \nabla^2\varphi] + [\nabla^2\psi, \psi] = 0, \tag{3}$$

$$\frac{\partial v}{\partial t} + [\varphi, v] - c_\beta[Z, \psi] = 0. \tag{4}$$

Equation (1) is a reduced Ohm's law where the presence of finite electron inertia, which makes it possible for magnetic reconnection (MR) to take place, is indicated by the terms proportional to the electron skin depth d_e. Equations (2), (3) and (4) are obtained from the electron vorticity equation, the vorticity equation and the parallel momentum equation, respectively.

Considering a Cartesian coordinate system (x, y, z) and taking z as an ignorable coordinate, the fields ψ, Z, φ and v are related to the magnetic field $\boldsymbol{B}$ and to the velocity field $\boldsymbol{v}$ by the relations $\boldsymbol{B} = \nabla\psi \times \hat{z} + (B^{(0)} + c_\beta Z)\hat{z}$ and $\boldsymbol{v} = -\nabla\varphi \times \hat{z} + v\hat{z}$, respectively. Here $B^{(0)}$ is a constant guide field, whereas $c_\beta = \sqrt{\beta/(1+\beta)}$ and $d_\beta = d_i c_\beta$ with d_i indicating the ion skin depth. For small β, $d_\beta \approx \rho_s$, the sonic Larmor radius. The ions are assumed to be cold, but electron pressure perturbations are taken into account and are given by $p = P^{(0)} + B^{(0)} p_1$, with $P^{(0)}$ a constant background pressure and p_1 coupled to the magnetic field via the relation $p_1 \simeq -c_\beta Z$. Note that here the parameter β is defined as $\beta = (5/3)P^{(0)}/{B^{(0)}}^2$, and above all the quantities are expressed in a dimensionless form according to the following normalization: $\nabla = a\nabla$, $t = v_A t/a$, $\boldsymbol{B} = \boldsymbol{B}/B_p$, where a is a typical scale length of the problem, B_p is a reference value for the poloidal magnetic field and v_A is the Alfvén speed based on B_p and on the constant density. Finally, $[f, g] := \nabla f \times \nabla g \cdot \hat{\boldsymbol{z}}$, for generic fields f and g.

3. Hamiltonian formulation

A desirable property for fluid models of the plasma is that the non-dissipative part of their equation of motion should admit a noncanonical Hamiltonian formulation [14–16]. In short this means that it is possible to reformulate the ideal part of an n-field model as

$$\frac{\partial \xi_i}{\partial t} = \{\xi_i, H\}, \qquad i = 1, \ldots, n, \tag{5}$$

where ξ_i are suitable field variables, H is the Hamiltonian functional and $\{,\}$ is the Poisson bracket consisting of an antisymmetric bilinear form satisfying the Jacobi identity.

The first task in the derivation of a noncanonical Hamiltonian formulation is to identify a conserved functional, usually the energy, that can serve as the Hamiltonian of the model. If one considers a domain $\mathcal{D}$ in the x–y plane and fields such that boundary terms arising from integrations by parts vanish, the four-field model (1)–(4) admits the following constant of the motion:

$$H = \frac{1}{2}\int_{\mathcal{D}} d^2x\, (d_e^2 J^2 + |\nabla\psi|^2 + |\nabla\varphi|^2 + v^2 + Z^2) \tag{6}$$

with $J = -\nabla^2\psi$ indicating the parallel current density. The quantity H represents the total energy of the system. The first term refers to the kinetic energy due to the relative motion of the electrons with respect to ions along the z direction. The third and fourth terms account for the kinetic energy, whereas the second and last terms account for the magnetic energy.

Adopting $\psi_e = \psi - d_e^2\nabla^2\psi$, $U = \nabla^2\varphi$, Z, and v as field variables, i.e. $\xi = (\psi_e, U, Z, v)$, and (6) as Hamiltonian, it is possible to show that the model can indeed be cast in a noncanonical Hamiltonian form with the following Lie–Poisson bracket:

$$\{F, G\} = \int d^2x\, \left(U[F_\xi, G_\xi]_U + \psi_e[F_\xi, G_\xi]_{\psi_e} + Z[F_\xi, G_\xi]_Z + v[F_\xi, G_\xi]_v\right), \tag{7}$$

where

$$\begin{aligned}
&[F_\xi, G_\xi]_U = [F_U, G_U],\\
&[F_\xi, G_\xi]_Z = [F_Z, G_U] + [F_U, G_Z] - d_\beta d_e^2[F_{\psi_e}, G_{\psi_e}]\\
&\qquad + c_\beta d_e^2([F_v, G_{\psi_e}] + [F_{\psi_e}, G_v]) - \alpha[F_Z, G_Z] - c_\beta\gamma[F_v, G_v],\\
&[F_\xi, G_\xi]_{\psi_e} = [F_{\psi_e}, G_U] + [F_U, G_{\psi_e}] - d_\beta([F_Z, G_{\psi_e}] + [F_{\psi_e}, G_Z]) + c_\beta([F_v, G_Z] + [F_Z, G_v]),\\
&[F_\xi, G_\xi]_v = [F_v, G_U] + [F_U, G_v] + c_\beta d_e^2([F_Z, G_{\psi_e}] + [F_{\psi_e}, G_Z]) - c_\beta\gamma([F_v, G_Z] + [F_Z, G_v]),
\end{aligned} \tag{8}$$

with $\alpha = d_\beta + c_\beta d_e^2/d_i$, $\gamma = d_e^2/d_i$, and subscripts indicate functional differentiation. The Jacobi identity for the above Poisson bracket is readily established by applying the method described in [14, 41]. It is straightforward to show that equation (5) with the Hamiltonian given in equation (6) and the above bracket reproduces equations (1)–(4) of [38] and of this paper. In order to provide a few details about this derivation let us consider for instance equation (1). According to (5), (1) follows from $\partial\psi_e/\partial t = \{\psi_e, H\}$. One then needs to show that $\{\psi_e, H\} = -([\varphi, \psi_e] - d_\beta[\psi, Z])$. This can be done by making use of the explicit form of the bracket given in (7) and (8). For this purpose one needs to recall that functional differentiation yields

$$H_{\psi_e} = J, \qquad H_U = -\varphi, \qquad H_Z = Z, \qquad H_v = v,$$
$$\psi_e(x')_{\psi_e(x)} = \delta(x - x'), \qquad \psi_e(x')_{U(x)} = 0,$$
$$\psi_e(x')_{Z(x)} = 0, \qquad \psi_e(x')_{v(x)} = 0$$

and one also needs to make use of the relation

$$\int_{\mathcal{D}} d^2x f[g, h] = \int_{\mathcal{D}} d^2x h[f, g] = \int_{\mathcal{D}} d^2x g[h, f], \tag{9}$$

valid for generic fields f, g and h, provided the above conditions for vanishing of boundary terms hold.

Because the number of parameters is escalating, we record their definitions here for later referral:

$$d_\beta = c_\beta d_i, \qquad d = \sqrt{d_e^2 + d_i^2}, \qquad \gamma = d_e^2/d_i, \qquad \alpha = d_\beta + c_\beta d_e^2/d_i = c_\beta d^2/d_i. \tag{10}$$

The basic parameters of the model are c_β, d_e and d_i, while d_β, α, γ and d are useful shorthands.

4. Casimir invariants and bracket normal form

Lie–Poisson brackets for noncanonical Hamiltonian systems are accompanied by the presence of Casimir invariants. A Casimir invariant is a functional that annihilates the Lie–Poisson bracket when paired with any other functional. That is, a Casimir C must satisfy

$$\{F, C\} = 0, \tag{11}$$

for every functional F. Thus, Casimir invariants constrain the nonlinear dynamics generated by the Poisson bracket for any choice of Hamiltonian.

In order to identify the Casimirs of the four-field model we proceed in the following way. First, multiplying equation (4) times d_i and adding it to equation (1), we find

$$\frac{\partial D}{\partial t} + [\varphi, D] = 0, \tag{12}$$

where $D = \psi_e + d_i v$ is the ion canonical momentum. Equation (12) indicates that the field D is a Lagrangian invariant that is advected by the flow generated by the stream-function φ. The presence of this Lagrangian invariant also suggests that using D as one of the variables will simplify the Lie–Poisson bracket. Indeed, upon replacing ψ_e with D as field variable, equation (11) for the four-field model becomes

$$\begin{aligned}\{F, C\} = \int d^2x (&F_U[C_U, U] + F_D[C_U, D] + F_U[C_D, D]\\ &+ c_\beta F_v[C_Z, D] + c_\beta F_Z[C_v, D] + F_Z[C_U, Z] + F_U[C_Z, Z]\\ &- \alpha F_Z[C_Z, Z] - c_\beta \gamma F_v[C_v, Z] + F_v[C_U, v] + F_U[C_v, v]\\ &- \alpha F_v[C_Z, v] - \alpha F_Z[C_v, v]) = 0,\end{aligned} \tag{13}$$

a simpler bracket.

For equation (11) to be satisfied for any F, it is necessary for the coefficients of each of the functional derivatives of F in (13) to vanish separately. This leads to the following system of equations for C:

$$[C_U, D] = 0, \tag{14}$$

$$[C_U, U] + [C_D, D] + [C_Z, Z] + [C_v, v] = 0, \tag{15}$$

$$-c_\beta[C_v, D] - [C_U, Z] + \alpha([C_Z, Z] + [C_v, v]) = 0, \tag{16}$$

$$c_\beta[C_Z, D] - c_\beta\gamma[C_v, Z] + [C_U, v] - \alpha[C_Z, v] = 0. \tag{17}$$

The problem of finding the complete set of Casimir invariants is thus equivalent to solving the set of four coupled equations (14)–(17).

Beginning with (14), a functional integration shows that C is of the form

$$C(U, D, Z, v) = \int \mathrm{d}^2x(U\mathcal{F}(D) + g(D, Z, v)), \tag{18}$$

where $\mathcal{F}$ and g represent arbitrary functions of their arguments. The problem is now reduced to finding the functions $\mathcal{F}$ and g. Considering next equation (15), we see that this equation is automatically satisfied for any choice of C with an integrand that depends only upon the field variables and not their spatial derivatives, and therefore imposes no constraints. Substitution of (18) into (16) yields

$$\left(c_\beta g_{vv} + \alpha g_{vD}\right)[v, D] + \left(c_\beta g_{vZ} - \mathcal{F}'(D) + \alpha g_{DZ}\right)[Z, D] = 0\,, \tag{19}$$

where $'$ indicates derivative with respect to argument and subscripts on g denote partial derivatives. In (19) the coefficients multiplying the brackets '[,]' must vanish independently. These two relations lead to

$$c_\beta g_v + \alpha g_D = Z\mathcal{F}'(D) + K(D), \tag{20}$$

where K is an arbitrary function of D, and integration of (20) by the method of characteristics gives

$$g(Z, v, D) = \frac{Z}{\alpha}\mathcal{F}(D) + \mathcal{K}(D) + \mathcal{G}(Z, v - c_\beta D/\alpha)\,, \tag{21}$$

where $\mathcal{K}' = K/\alpha$. Similarly, insertion of (18) into (15) yields

$$\left(c_\beta g_{Zv} - \mathcal{F}'(D) + \alpha g_{ZD}\right)[D, v] + c_\beta\left(g_{ZZ} + \gamma g_{vD}\right)[D, Z] + \left(c_\beta\gamma g_{vv} - \alpha g_{ZZ}\right)[v, Z] = 0, \tag{22}$$

which gives three equations, but only $g_{ZZ} + \gamma g_{vD} = 0$ provides new information. Defining $V = v - c_\beta D/\alpha$, we see that $\mathcal{G}$ satisfies the wave equation

$$\mathcal{G}_{ZZ} - \frac{c_\beta\gamma}{\alpha}\mathcal{G}_{VV} = 0, \tag{23}$$

and thus has the general solution

$$\mathcal{G} = \sum_{\pm}\mathcal{G}_\pm\left(\pm\frac{1}{2\alpha}\sqrt{\frac{c_\beta}{\gamma\alpha}}\left(D - \frac{\alpha}{c_\beta}v\right) - \frac{Z}{2\alpha}\right). \tag{24}$$

Therefore, in light of (18), (21) and (24) we have the following four independent families of Casimir invariants:

$$C_1 = \int \mathrm{d}^2x\left(U + \frac{Z}{\alpha}\right)\mathcal{F}(D), \tag{25}$$

$$C_2 = \int \mathrm{d}^2x\,\mathcal{K}(D), \tag{26}$$

$$C_\pm = \int \mathrm{d}^2x\mathcal{G}_\pm\left(\pm\frac{d_\mathrm{i}^2}{2c_\beta d^3 d_\mathrm{e}}D \mp \frac{d_\mathrm{i}}{2c_\beta d d_\mathrm{e}}v - \frac{d_\mathrm{i}}{2c_\beta d^2}Z\right), \tag{27}$$

where we have scaled the arguments of $C_\pm$ for a reason that will become apparent soon.

Knowledge of the functional dependence of the Casimirs suggests a simplification of the Lie–Poisson bracket will occur if it is written in terms of the new variables

$$D = D, \tag{28}$$

$$\zeta = U + \frac{Z}{\alpha}, \tag{29}$$

$$T_{\pm} = \pm \frac{d_{\rm i}}{2c_{\beta} d^3 d_{\rm e}} \left(d_{\rm i} D - d^2 v \mp d d_{\rm e} Z \right), \tag{30}$$

which possess the inverse relations

$$D = D, \tag{31}$$

$$U = \zeta + T_+ + T_-, \tag{32}$$

$$Z = -\alpha (T_+ + T_-), \tag{33}$$

$$v = \frac{d_{\rm i}}{d^2} D - \frac{c_{\beta} d_{\rm e} d}{d_{\rm i}} (T_+ - T_-). \tag{34}$$

Indeed, in the new variables, the Lie–Poisson bracket reads

$$\{F, G\} = \int {\rm d}^2 x (\zeta [F_{\zeta}, G_{\zeta}] + D([F_D, G_{\zeta}] + [F_{\zeta}, G_D]) + T_-[F_{T_-}, G_{T_-}] + T_+[F_{T_+}, G_{T_+}]), \tag{35}$$

which is a bracket normal form that relies on the scaling used above. The bracket in terms of the new variables reveals its algebraic structure: it is identified as a sum of direct product and semi-direct product parts [15, 19, 41] and, consequently, the Jacobi identity follows from general theory. Making use of the variables suggested by the form of the Casimirs, the model equations can be rewritten in the compact form

$$\frac{\partial D}{\partial t} = -[\varphi, D], \tag{36}$$

$$\frac{\partial \zeta}{\partial t} = -[\varphi, \zeta] + d^{-2}[D, \psi], \tag{37}$$

$$\frac{\partial T_{\pm}}{\partial t} = -\left[\varphi_{\pm}, T_{\pm}\right], \tag{38}$$

where for convenience we have defined

$$\varphi_{\pm} := \varphi \pm \frac{c_{\beta} d}{d_{\rm e}} \psi. \tag{39}$$

Note that the variable ζ plays the role of a 'generalized' vorticity, and our development reveals the existence of the three Lagrangian invariants D, T_+ and T_- associated with the families of Casimirs C_2, C_+ and C_-, respectively. The existence of such invariants implies that the values of D, T_+ and T_- are constant on the contour lines of φ, φ_+, and φ_-, respectively. By choosing 'top-hat' functions for the free functions in the Casimirs C_2, C_+ and C_-, it follows that the area enclosed by the contour lines of the Lagrangian invariants remains constant. Notice also that T_+ and T_- in the limit $\beta \to 0$ and $d_{\rm i} \to \infty$ become proportional to the Lagrangian invariants $G_{\pm} = \psi - d_{\rm e}^2 \nabla^2 \psi \pm d_{\rm e} \rho_{\rm s} U$ of the two-field model derived in [4]. The family C_1 is of a different nature and one of the constraints imposed by it is that the total value of ζ within an area enclosed by a contour line of D remains constant.

5. Equilibria

Equations governing the equilibrium of the FP model are most easily obtained by setting the time derivatives equal to zero in equations (36)–(38) and solving for the fields D, ζ and $T_{\pm}$.

Alternatively, it is possible to derive equilibrium equations from a variational principle, the existence of which is assured by the Hamiltonian nature of the equations [15]. The construction of the variational principle is more laborious than the direct approach but the extra work is richly rewarded by the well-known benefits of variational principles. In particular, the variational principle provides a basis for studying the stability as well as the equilibria of the system.

For a system with a collection of Casimirs, extrema of the free energy functional $F = H + \sum C$ are equilibria of the system. Such extrema can be derived by setting the first variation δF equal to zero. If ξ_i, $i = 1, 2, \ldots$, denotes the field variables of the system, this amounts to solving the equations $H_{\xi_i} + \sum C_{\xi_i} = 0$, where the subscript indicates functional derivative with respect to ξ_i. One advantage of this variational approach is that for equilibria obtained as extrema of F, the second variation of F provides criteria sufficient for stability.

5.1. General equilibria

For the model here, the free energy functional reads

$$F[\xi] = \int_{\mathcal{D}} \mathrm{d}^2x \left[\frac{d_\mathrm{e}^2 \left(\nabla^2 \mathcal{L}\psi_\mathrm{e}\right)^2}{2} + \frac{|\nabla \mathcal{L}\psi_\mathrm{e}|^2}{2} + \frac{|\nabla\nabla^{-2}U|^2}{2} + \frac{v^2}{2} + \frac{Z^2}{2} + \mathcal{K}(\psi_\mathrm{e} + d_\mathrm{i} v) \right.$$
$$\left. + \left(U + \frac{Z}{\alpha}\right) \mathcal{F}(\psi_\mathrm{e} + d_\mathrm{i} v) + \sum_{\pm} \mathcal{G}_\pm \left(\pm \frac{d_\mathrm{i}^2}{2c_\beta d^3 d_\mathrm{e}} D \mp \frac{d_\mathrm{i}}{2c_\beta d d_\mathrm{e}} v - \frac{d_\mathrm{i}}{2c_\beta d^2} Z \right) \right], \quad (40)$$

where the linear operator $\mathcal{L}$ is defined by $\mathcal{L}^{-1}\psi = \psi - d_\mathrm{e}^2 \nabla^2 \psi = \psi_\mathrm{e}$. Note that the arguments of the functions present in the Casimirs are of course much less compact when written in terms of the variables $\xi = (\psi_\mathrm{e}, U, Z, v)$. On the other hand, in terms of the variables $\Upsilon := (D, \zeta, T_+, T_-)$ the free energy functional reads

$$F[\Upsilon] = \int_{\mathcal{D}} \mathrm{d}^2x \left[\frac{c_\beta^2 d^4}{d_\mathrm{i}^2} (T_+^2 + T_-^2) + \frac{D^2}{2d^2} - \frac{1}{2} (\zeta + T_+ + T_-) \nabla^{-2} (\zeta + T_+ + T_-) \right.$$
$$- \frac{1}{2} \left(\frac{d_\mathrm{e}}{d^2} D + c_\beta d (T_+ - T_-) \right) \mathcal{L} \left(\frac{d_\mathrm{e}}{d^2} D + c_\beta d (T_+ - T_-) \right)$$
$$\left. + \mathcal{K}(D) + \zeta \mathcal{F}(D) + \mathcal{G}_+(T_+) + \mathcal{G}_-(T_-) \right]. \quad (41)$$

Equation (41) shows that in terms of the variables that are 'natural' for the Casimirs, the expression for the Hamiltonian becomes complicated. Unfortunately, there exists no preferred set of variables in terms of which both the Hamiltonian and the Casimirs take a simple form. In order to obtain equilibrium solutions by means of the variational principle, it is convenient to choose the Υ variables that are natural for the Casimirs, calculate the required functional derivatives of the Casimirs with respect to these variables and then use the functional chain rule to obtain the functional derivatives of H in terms of the variables ξ. More specifically, by setting $\delta F = 0$, the resulting equilibrium equations are given by

$$F_\zeta = H_\zeta + \sum_j C_{j\,\zeta} = H_\zeta + \mathcal{F}(D) = 0, \quad (42)$$

$$F_D = H_D + \sum_j C_{j\,D} = H_D + \zeta \mathcal{F}'(D) + \mathcal{K}'(D) = 0, \quad (43)$$

$$F_{T_\pm} = H_{T_\pm} + \sum_j C_{j\,T_\pm} = H_{T_\pm} + \mathcal{G}_\pm'(T_\pm) = 0, \quad (44)$$

where the index j ranges over the set $\{1, 2, +, -\}$. The functional chain rule can then be used to evaluate the functional derivatives of the Hamiltonian,

$$H_D = \frac{d_e^2}{d^2} H_{\psi_e} + \frac{d_i}{d^2} H_v,$$
$$H_\zeta = H_U,$$
$$H_{T_\pm} = \pm c_\beta d_e d\, H_{\psi_e} + H_U - \alpha\, H_Z \mp \frac{\alpha d_e}{d} H_v.$$

The functional derivatives H_{ψ_e}, H_U, H_Z and H_v can themselves easily be obtained from the Hamiltonian written in the form (6). The equilibrium equations (42)–(44) are then given by

$$-\varphi + \mathcal{F}(D) = 0, \tag{45}$$
$$-\frac{d_e^2}{d^2}\nabla^2\psi + \frac{d_i v}{d^2} + \zeta\mathcal{F}'(D) + \mathcal{K}'(D) = 0, \tag{46}$$
$$\mp c_\beta d_e d\,\nabla^2\psi - \varphi - \alpha Z \mp \frac{\alpha d_e v}{d} + \mathcal{G}_\pm'(T_\pm) = 0. \tag{47}$$

These equations are expressed in a mixture of the ξ and Υ variables. In order to simplify them we eliminate $\nabla^2\psi$ from the last two equations by using

$$d_e^2\nabla^2\psi = \psi - D + d_i v.$$

Using this in equations (46) and (47), we find

$$-d^{-2}\psi + \zeta\mathcal{F}'(D) + \hat{\mathcal{K}}'(D) = 0, \tag{48}$$
$$-\varphi_\pm \pm \frac{c_\beta d}{d_e d_i}(d_i D - d^2 v \mp d d_e Z) + \mathcal{G}_\pm'(T_\pm) = 0, \tag{49}$$

where $\hat{\mathcal{K}}(D) = \mathcal{K}(D) + D^2/2d^2$. Equation (49) may be simplified further by expressing Z and v in terms of the Υ variables using (33) and (34). This leads to the following complete system of equilibrium equations:

$$-\varphi + \mathcal{F}(D) = 0, \tag{50}$$
$$-\varphi_\pm + \hat{\mathcal{G}}_\pm'(T_\pm) = 0, \tag{51}$$
$$-d^{-2}\psi + \zeta\mathcal{F}'(D) + \hat{\mathcal{K}}'(D) = 0, \tag{52}$$

where $\hat{\mathcal{G}}_\pm(T_\pm) = \mathcal{G}_\pm(T_\pm) + \alpha^2 T_\pm^2$. One easily verifies that equations (50)–(52) describe equilibrium states by substituting them into equations (36) and (38).

Continuing our present approach of eliminating the ξ variables would now lead us to express the ψ, φ and $\varphi_\pm$ in terms of integral operators acting on the Υ variables. Clearly this is undesirable. Instead, we note that the above four equations express a dependence between the six quantities $\nabla^2\psi$, $\nabla^2\varphi$, φ, ψ, Z and v. It is thus possible in principle to use these equations to express four of these quantities in terms of the remaining two. If we choose φ and ψ as the independent fields we will obtain a closed system of equilibrium equations of the form

$$\nabla^2\psi = S(\psi, \varphi), \tag{53}$$
$$\nabla^2\varphi = P(\psi, \varphi). \tag{54}$$

Equation (53) is a generalized version of the Grad–Shafranov equation, whereas (54) is an analogous equation that determines the equilibrium polarization.

In order to calculate the form of the functions S and P we invert equations (50)–(52):

$$D = a(\varphi), \tag{55}$$
$$T_\pm = t_\pm(\varphi_\pm), \tag{56}$$
$$\zeta = d^{-2}\psi a'(\varphi) + b(\varphi), \tag{57}$$

where a and $t_\pm$ are the inverses of $\mathcal{F}$ and $\hat{\mathcal{G}}_\pm$, respectively, and $b(\varphi) = -\hat{\mathcal{K}}'(a(\varphi))a'(\varphi)$. Solving these equations for $\nabla^2\varphi$ and $\nabla^2\psi$ then yields (53) and (54) with

$$S(\psi,\varphi) := \frac{\psi}{d_e^2} - \frac{a(\varphi)}{d^2} - \frac{c_\beta d}{d_e}[t_+(\varphi_+) - t_-(\varphi_-)], \tag{58}$$

$$P(\psi,\varphi) := b(\varphi) + t_+(\varphi_+) + t_-(\varphi_-) + a'(\varphi)\frac{\psi}{d^2}. \tag{59}$$

We complete the system by expressing the two remaining unknowns Z and v in terms of φ and ψ. From equations (33) and (34), using (55), (56) there follows

$$v(\psi,\varphi) = d_i\, a(\varphi)/d^2 - c_\beta d_e d(t_+(\varphi_+) - t_-(\varphi_-))/d_i; \tag{60}$$

$$Z(\psi,\varphi) = -\alpha(t_+(\varphi_+) + t_-(\varphi_-)). \tag{61}$$

We have thus shown that solving the equilibrium equations amounts to solving the coupled system of (53) and (54) for the unknowns ψ and φ. This requires making choices for the free functions $t_\pm$, a and b. If one is only interested in solving the equilibrium problem, one may determine these free functions directly from physical considerations and the relationship between these functions and those appearing in the variational principle may be ignored. These relationships become important, however, if one wishes to use the variational principle either to solve the equilibrium or the stability problem (the variational principle has well-known advantages both for numerical and analytic applications). In this case the functions $\mathcal{G}_\pm$, $\mathcal{F}$, and $\mathcal{K}$, appearing in the variational form may be determined in terms of a, b, and $t_\pm$ as described above. Variational treatments of two-fluid equilibria have been given by [42–44], and applications of variational principles to stability are discussed by [42, 43, 45, 46].

We conclude this section by noting two difficulties with the equilibrium equations (53) and (54). The first is that these equations may become hyperbolic in the presence of strong flows. Recent analyses of this problem can be found in [44, 47]. The second difficulty is that the right-hand sides are singular in the limits $d_e \ll L$ and $d_i \ll L$, where L is a macroscopic scale length [42, 43, 48]. That is, for macroscopic equilibria the derivatives in equations (53)–(54) are multiplied by a small parameter, so that these equations form a stiff system. This has led to considerable grief, in particular for field-reversed configuration (FRC) devices where a similar set of equations is encountered. Steinhauer has proposed a method for dealing with this problem that he named the 'nearby fluid' approximation [48, 49]. In the following section we outline a similar approach to solving equations (53) and (54).

5.2. Perturbative solution for macroscopic equilibria

In order to deal with the singular nature of the equilibrium equations (53) and (54), we expand the fields in powers of the small parameters d_e and d_i and solve term by term. A byproduct of this procedure is the clarification of the physical meaning of the profile functions.

We begin by considering the limit $d_e \to 0$. For convenience we define

$$\hat{t}_\pm\left(\pm\psi + \frac{d_e\varphi}{c_\beta d}\right) := d_e c_\beta d\, t_\pm(\varphi_\pm)$$

and expand this and other profile functions in powers of d_e according to

$$\hat{t}_\pm(\phi) = \hat{t}_\pm^{(0)}(\phi) + d_e\hat{t}_\pm^{(1)}(\phi) + d_e^2\hat{t}_\pm^{(2)}(\phi) + \cdots.$$

We then consider the equilibrium equations order by order in d_e. From (54) and (59) we obtain at lowest order $\hat{t}_+^{(0)}(\psi) = -\hat{t}_-^{(0)}(-\psi)$; using this result in (53) and (58) we obtain to lowest order $\hat{t}_+^{(0)}(\phi) = \hat{t}_-^{(0)}(\phi) = \phi/2$. To next order, (53) and (58) yield $\hat{t}_-^{(1)}(-\psi) = \hat{t}_+^{(1)}(\psi)$;

substituting this result in (54) and (59), we find the following equation for the vorticity:

$$\nabla^2\varphi = \hat{b}(\varphi) + a'(\varphi)\psi/{d_{\mathrm{i}}}^2 - h(\psi)/d_\beta,$$

where $\hat{b} := b + \varphi/d_\beta^2$ is a profile function for the vorticity and $h(\psi) := -2\hat{t}_+^{(1)}(\psi)$.

From the terms of order unity in (53) and (58) we next find

$$\nabla^2\psi = -a(\varphi)/d_{\mathrm{i}}^2 + h'(\psi)\varphi/d_\beta + I(\psi), \tag{62}$$

where $I(\psi) = \hat{t}_-^{(2)}(-\psi) - \hat{t}_+^{(2)}(\psi)$. Equation (62) is the Grad–Shafranov equation, where the term proportional to φ is the polarization current and $I(\psi)$ describes the inductive current.

The parallel velocity may be obtained from (55): to order d_{e}^0, $D = \psi - d_{\mathrm{e}}^2\nabla^2\psi + d_{\mathrm{i}}v = a(\varphi)$, yields

$$d_{\mathrm{i}}v + \psi = a(\varphi). \tag{63}$$

Expanding (61) to order d_{e}^0, gives

$$Z + \varphi/d_\beta = h(\psi). \tag{64}$$

The sum $Z + \varphi/d_\beta$ represents the electron stream-function. The fact that the electron stream-function is constant on the surfaces of constant flux, as expressed by equation (64), is a statement of the frozen-in property.

We may carry out the limit $d_{\mathrm{i}} \to 0$ in a similar way. From the ion momentum conservation equation (63), we obtain $\psi = a(\varphi)$ showing that the electrostatic potential must be a flux function to lowest order. It is convenient to introduce $\Phi(\psi) := a^{-1}(\psi) = \varphi$ to denote the inverse of a. We also define the Alfvénic Mach number $M(\psi) := \mathrm{d}\Phi/\mathrm{d}\psi$. Note that equation (64) specifies that to lowest order, $h(\psi) = \Phi(\psi)/d_\beta$. In terms of these quantities, the vorticity equation shows that to lowest order

$$\hat{b}(\varphi) = a(\varphi)a'(\varphi)/d_{\mathrm{i}}^2 - \varphi/d_\beta^2.$$

In order to eliminate φ from the Grad–Shafranov equation it is necessary to calculate the correction to the electrostatic potential. This is given by the vorticity equation,

$$M'(\nabla\psi)^2 + M\nabla^2\psi = (M^{-2} - c_\beta^{-2})\varphi^{(2)} - h^{(1)}(\psi).$$

Note that the potential exhibits a singularity for $M = c_\beta$ corresponding to the sound-wave resonance. Lastly, after eliminating the electrostatic potential from equation (62) we recover the MHD version of the Grad–Shafranov equation,

$$(1 - M^2)\nabla^2\psi - MM'(\nabla\psi)^2 = \hat{I}(\psi).$$

In the following sections we present some explicit solutions of the equilibrium equations for simple profile functions.

5.3. Quadratic Casimirs–dipole equilibria

The case of quadratic Casimir invariants is easily tractable. Choosing

$$\mathcal{K}(D) = \frac{A_D}{2}D^2, \qquad \mathcal{F}(D) = A_\zeta D, \qquad \mathcal{G}_\pm(T_\pm) = \frac{A_\pm}{2}T_\pm^2, \tag{65}$$

and following the steps of section 5.1 leads to

$$a(\varphi) = \frac{\varphi}{A_\zeta}, \qquad b(\varphi) = -\frac{\varphi}{A_\zeta^2}\left(A_D + \frac{1}{d^2}\right), \qquad t_\pm(\varphi_\pm) = \frac{\varphi_\pm}{A_\pm + 2\alpha^2}. \tag{66}$$

Upon inserting (66) into (53) and (54), we obtain

$$\nabla^2\psi = S_1\psi + S_2\varphi \qquad \text{and} \qquad \nabla^2\varphi = P_1\psi + P_2\varphi, \tag{67}$$

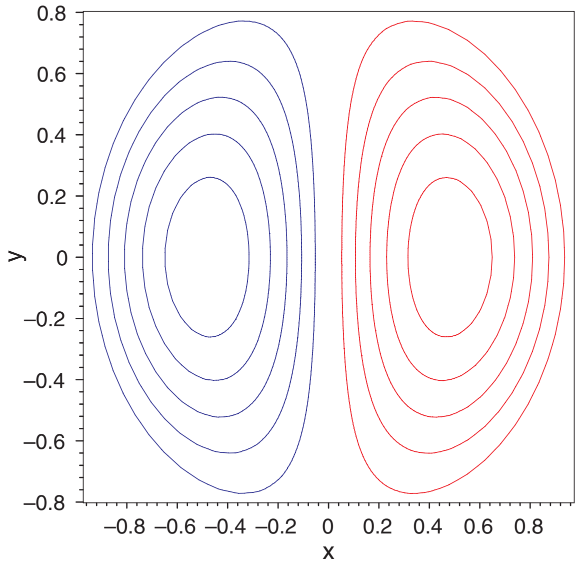

Figure 1. χ contours for dipole solution of (69).

where $S_{1,2}$ and $P_{1,2}$ are constants that depend on $A_\pm$, A_ζ, A_D, and the parameters of the system. Note that $S_{1,2}$ and $P_{1,2}$ are arbitrary except that $S_2 = -P_1$. Consequently, equations (67) have a variety of solutions that are closely related to the double-Beltrami flows investigated by Yoshida *et al* [42, 43]. Specifically, [42, 43] neglect electron inertial effects that are kept here, but they allow for more general geometry. In general, (67) can be diagonalized resulting in two decoupled equations of the form

$$\nabla^2 \chi_i = -\lambda_i \chi_i, \qquad i = 1, 2, \tag{68}$$

where $\lambda_{1,2} = -(S_1 + P_2 \pm \sqrt{(S_1 - P_2)^2 - 4P_1^2})/2$ and the χ_is are linear combinations of φ and ψ. If a solution of this system is found, then one obtains v and Z as particular linear combinations of φ and ψ as described in section 5.1.

Rather than describe the general solution, we give an example representative of the kinds of solutions that are possible. We assume a circular domain of unit radius, adopt polar coordinates (r, θ), and adjusting the parameters S_1, P_1, and P_2 so that $\sqrt{\lambda_{1,2}}$ are zeros (possibly distinct) of the first order Bessel function, i.e. $J_1(\sqrt{\lambda_{1,2}}) = 0$. With these assumptions, we obtain the solution

$$\chi_i(r, \theta) = A_i J_1(\sqrt{\lambda_i} r) \cos\theta, \tag{69}$$

where the A_is are constants and each of the χs has a dipolar structure such as that depicted in figure 1.

5.4. Homogeneous equilibria

The quadratic Casimirs of (65) also yield homogeneous equilibria, i.e. equilibria for which the linear dynamics has constant coefficients. For this choice, the free energy functional of (41) can be written as follows:

$$F = \frac{1}{2} \int_{\mathcal{D}} \mathrm{d}^2 x \left(\xi^{\mathrm{T}} \hat{H} \xi + \Upsilon^{\mathrm{T}} \hat{A} \Upsilon \right), \tag{70}$$

where

$$\hat{H} = \begin{pmatrix} -\mathcal{L}\nabla^2 & 0 & 0 & 0 \\ 0 & -\nabla^{-2} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad \text{and} \quad \hat{A} = \begin{pmatrix} A_D & A_\zeta & 0 & 0 \\ A_\zeta & 0 & 0 & 0 \\ 0 & 0 & A_+ & 0 \\ 0 & 0 & 0 & A_- \end{pmatrix}. \tag{71}$$

Recall $\xi = (\psi_e, U, Z, v)$ and $\Upsilon := (D, \zeta, T_+, T_-)$. Equations (28)–(30) amount to $\Upsilon = \mathcal{T}\xi$, where the matrix $\mathcal{T}$ is given by

$$\mathcal{T} = \begin{pmatrix} 1 & 0 & 0 & d_i \\ 0 & 1 & \dfrac{d_i}{c_\beta d^2} & 0 \\ \dfrac{d_i^2}{2c_\beta d_e d^3} & 0 & \dfrac{-d_i}{2c_\beta d^2} & \dfrac{-d_e d_i}{2c_\beta d^3} \\ \dfrac{-d_i^2}{2c_\beta d_e d^3} & 0 & \dfrac{-d_i}{2c_\beta d^2} & \dfrac{d_e d_i}{2c_\beta d^3} \end{pmatrix}. \tag{72}$$

The free energy functional can then be written as a quadratic form

$$F = \frac{1}{2}\int_{\mathcal{D}} d^2x \left[\xi^{\mathrm{T}}(\hat{H} + \mathcal{T}^{\mathrm{T}}\hat{A}\mathcal{T})\xi\right], \tag{73}$$

whence the equilibrium equations are obtained upon setting the functional derivatives F_ξ to zero,

$$(\hat{H} + \mathcal{T}^{\mathrm{T}}\hat{A}\mathcal{T})\xi = 0. \tag{74}$$

Here we have assumed the formal self-adjointness of the operators $\mathcal{L}\nabla^2$ and ∇^{-2}. Equation (74) is a linear homogeneous system of four equations with the four unknowns ψ_e, U, Z and v. The equilibria treated in section 5.3 are solutions of this system, but of interest here are the homogeneous equilibria

$$\psi_0 = \alpha_\psi x, \qquad \varphi_0 \equiv 0, \qquad Z_0 = \alpha_Z x, \qquad v_0 = \alpha_v x, \tag{75}$$

where the α_ψ is the Alfvén speed, and α_Z and α_v describe density and velocity shear, respectively. These are clearly equilibrium solutions and can be related to the chosen Casimirs. Evidently, we are seeking solutions with $\nabla^2\psi \equiv 0$ and $U \equiv 0$, so (74) reduces to

$$(\hat{I}_2^0 + \mathcal{T}^{\mathrm{T}}\hat{A}_3\mathcal{T})\xi_3 = 0\,, \tag{76}$$

where $\xi_3 := (\psi, Z, v)$,

$$\hat{I}_2^0 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{and} \quad \hat{A}_3 = \begin{pmatrix} A_D & 0 & 0 \\ 0 & A_+ & 0 \\ 0 & 0 & A_- \end{pmatrix}. \tag{77}$$

This is a linear homogeneous system of three equations with unknowns ψ, Z and v. The existence of non-trivial solutions requires $\det(\hat{I}_2^0 + \mathcal{T}^{\mathrm{T}}\hat{A}\mathcal{T}) = 0$, which fixes a condition on A_D, A_+ and A_- that assures ψ, Z and v are linearly dependent. Thus, they all can depend on x and be proportional, consistent with (75). These equilibria correspond to uniform poloidal magnetic fields, and to toroidal magnetic and velocity fields proportional to x.

6. Normal forms for homogeneous equilibria

Here we work out the linear canonical Hamiltonian form for the dynamics obtained by expansion about the equilibria of section 5.4. In terms of the Υ variables the equilibrium is

$$D_0 = \alpha_D x, \qquad \zeta_0 = \alpha_\zeta x, \qquad T_{\pm 0} = \alpha_\pm x, \tag{78}$$

whence e.g. $\alpha_D = \alpha_\psi + d_{\rm i}\alpha_v$ and (34) give $\alpha_\psi = c_\beta d d_{\rm e}(\alpha_+ - \alpha_-) + \alpha_D(1 - d_{\rm i}^2/d^2)$. We obtain the Poisson bracket and Hamiltonian for the linear dynamics, obtain the dispersion relation, then we find a set of canonical variables and discuss canonical transformations to normal forms. En route we define the meaning of negative energy modes.

6.1. Linear Hamiltonian form

Denoting the linear variables with a tilde, i.e. $\zeta = \zeta_0 + \tilde{\zeta}$, etc, and expanding the bracket (35) gives the bracket for the linear dynamics (see [15])

$$\{F, G\}_L = \int {\rm d}^2x \left(\zeta_0 \left[\frac{\delta F}{\delta \tilde{\zeta}}, \frac{\delta G}{\delta \tilde{\zeta}}\right] + D_0 \left(\left[\frac{\delta F}{\delta \tilde{D}}, \frac{\delta G}{\delta \tilde{\zeta}}\right] + \left[\frac{\delta F}{\delta \tilde{\zeta}}, \frac{\delta G}{\delta \tilde{D}}\right]\right) + T_{\pm 0}\left[\frac{\delta F}{\delta \tilde{T}_\pm}, \frac{\delta G}{\delta \tilde{T}_\pm}\right]\right),$$

where a sum over the $\pm$ terms is implied. Upon integration by parts this bracket becomes

$$\{F, G\}_L = -\int {\rm d}^2x \left[\alpha_\zeta \frac{\delta F}{\delta \tilde{\zeta}} \frac{\partial}{\partial y} \frac{\delta G}{\delta \tilde{\zeta}} + \alpha_D \left(\frac{\delta F}{\delta \tilde{D}} \frac{\partial}{\partial y} \frac{\delta G}{\delta \tilde{\zeta}} + \frac{\delta F}{\delta \tilde{\zeta}} \frac{\partial}{\partial y} \frac{\delta G}{\delta \tilde{D}}\right) + \alpha_\pm \frac{\delta F}{\delta \tilde{T}_\pm} \frac{\partial}{\partial y} \frac{\delta G}{\delta \tilde{T}_\pm}\right]. \tag{79}$$

The above bracket together with the following quadratic Hamiltonian:

$$H_L = \tfrac{1}{2} \int {\rm d}^2x \left(\tilde{\psi}_{\rm e} \tilde{J} + |\nabla \tilde{\phi}|^2 + \tilde{v}^2 + \tilde{Z}^2 + A_D \tilde{D}^2 + A_\pm \tilde{T}_\pm^2\right), \tag{80}$$

where

$$A_D = \frac{1}{d^2}\left(\frac{\alpha_\psi}{\alpha_D} - 1\right) \qquad \text{and} \qquad A_\pm = c_\beta \left(\pm \frac{d}{d_{\rm e}} \frac{\alpha_\psi}{\alpha_\pm} - 2\frac{c_\beta d^4}{d_{\rm i}^2}\right), \tag{81}$$

when written entirely in terms of the variables $(\tilde{\zeta}, \tilde{D}, \tilde{T}_+, \tilde{T}_-)$, yield the linearized equations of motion in Poisson bracket form.

Unlike the nonlinear semi-direct product bracket of (35), the linear bracket of (79) can be brought into direct product form by the transformation

$$\bar{D} = -\alpha_\zeta \tilde{D} + \alpha_D \tilde{\zeta}, \tag{82}$$

which yields the convenient form

$$\{F, G\}_L = -\int {\rm d}^2x \left[\alpha_\zeta \frac{\delta F}{\delta \tilde{\zeta}} \frac{\partial}{\partial y} \frac{\delta G}{\delta \tilde{\zeta}} + \alpha_{\bar{D}} \frac{\delta F}{\delta \bar{D}} \frac{\partial}{\partial y} \frac{\delta G}{\delta \bar{D}} + \alpha_\pm \frac{\delta F}{\delta \tilde{T}_\pm} \frac{\partial}{\partial y} \frac{\delta G}{\delta \tilde{T}_\pm}\right], \tag{83}$$

where $\alpha_{\bar{D}} := -\alpha_D^2 \alpha_\zeta$.

6.2. Canonical coordinates

Because the equilibrium equations do not depend explicitly on x and y, we Fourier expand $\tilde{\zeta}$ as

$$\tilde{\zeta}(x, y, t) = \sum_{k_x, k_y = -\infty}^{\infty} \zeta_{k_x, k_y}(t)\, {\rm e}^{-{\rm i}(k_x x + k_y y)} \tag{84}$$

and similarly for $\bar{D}$, $\tilde{T}_+$ and $\tilde{T}_-$. For convenience we suppress the sum over k_x and set $k_y = k$. The variable k_x will only appear in the combination $k_\perp^2 := k_x^2 + k_y^2$. It is not difficult to prove (see, e.g. [50]) the following general functional derivative relationship:

$$\frac{\delta F}{\delta \tilde{\zeta}} = \sum_{k=-\infty}^{\infty} \left(\frac{\delta F}{\delta \tilde{\zeta}}\right)_k \mathrm{e}^{-\mathrm{i}ky} = \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} \frac{\partial \bar{F}}{\partial \zeta_{-k}} \mathrm{e}^{-\mathrm{i}ky}, \tag{85}$$

where $F[\zeta] = \bar{F}[\zeta_k]$. Note that in the above expression there is an extra factor of 2π that occurs in the denominator of the suppressed sum on k_x, but this factor is compensated by a factor of 2π that accompanies a suppressed sum in the Hamiltonian.

Inserting (85) and similar relations for the other variables into the bracket of (79) gives

$$\begin{aligned}\{\bar{F}, \bar{G}\} = \sum_{k=1}^{\infty} \frac{\mathrm{i}k}{2\pi} \Bigg[&\alpha_\zeta \left(\frac{\partial \bar{F}}{\partial \zeta_k} \frac{\partial \bar{G}}{\partial \zeta_{-k}} - \frac{\partial \bar{F}}{\partial \zeta_{-k}} \frac{\partial \bar{G}}{\partial \zeta_k} \right) + \alpha_{\bar{D}} \left(\frac{\delta \bar{F}}{\delta \bar{D}_k} \frac{\delta \bar{G}}{\delta \bar{D}_{-k}} - \frac{\delta \bar{F}}{\delta \bar{D}_{-k}} \frac{\delta \bar{G}}{\delta \bar{D}_k} \right) \\ &+ \alpha_+ \left(\frac{\partial \bar{F}}{\partial T_{+k}} \frac{\partial \bar{G}}{\partial T_{+-k}} - \frac{\partial \bar{F}}{\partial T_{+-k}} \frac{\partial \bar{G}}{\partial T_{+k}} \right) + \alpha_- \left(\frac{\partial \bar{F}}{\partial T_{-k}} \frac{\partial \bar{G}}{\partial T_{--k}} - \frac{\partial \bar{F}}{\partial T_{--k}} \frac{\partial \bar{G}}{\partial T_{-k}} \right) \Bigg],\end{aligned} \tag{86}$$

where in order to facilitate the transformation to canonical variables the positive and negative values of k are assumed to be independent and the sum is written over only positive values of k.

A set of real valued canonical variables is given by

$$\begin{aligned} q_k^{(1)} &= -\sqrt{\frac{\pi}{k|\alpha_{\bar{D}}|}} \left(\bar{D}_k + \bar{D}_{-k}\right), \qquad & p_k^{(1)} &= \mathrm{i}\sqrt{\frac{\pi}{k|\alpha_{\bar{D}}|}} \left(\bar{D}_k - \bar{D}_{-k}\right), \\ q_k^{(2)} &= \sqrt{\frac{\pi}{k\alpha_\zeta}} \left(\zeta_k + \zeta_{-k}\right), & p_k^{(2)} &= \mathrm{i}\sqrt{\frac{\pi}{k\alpha_\zeta}} \left(\zeta_k - \zeta_{-k}\right), \\ q_k^{(3)} &= \sqrt{\frac{\pi}{k\alpha_+}} \left(T_{+k} + T_{+-k}\right), & p_k^{(3)} &= \mathrm{i}\sqrt{\frac{\pi}{k\alpha_+}} \left(T_{+k} - T_{+-k}\right), \\ q_k^{(4)} &= \sqrt{\frac{\pi}{k\alpha_-}} \left(T_{-k} + T_{--k}\right), & p_k^{(4)} &= \mathrm{i}\sqrt{\frac{\pi}{k\alpha_-}} \left(T_{-k} - T_{--k}\right), \end{aligned} \tag{87}$$

with the inverse transformation

$$\begin{aligned} \bar{D}_k &= -\frac{1}{2}\sqrt{\frac{k|\alpha_{\bar{D}}|}{\pi}} \left(q_k^{(1)} + \mathrm{i}p_k^{(1)}\right), \qquad & \bar{D}_{-k} &= -\frac{1}{2}\sqrt{\frac{k|\alpha_{\bar{D}}|}{\pi}} \left(q_k^{(1)} - \mathrm{i}p_k^{(1)}\right), \\ \zeta_k &= \frac{1}{2}\sqrt{\frac{k\alpha_\zeta}{\pi}} \left(q_k^{(2)} - \mathrm{i}p_k^{(2)}\right), & \zeta_{-k} &= \frac{1}{2}\sqrt{\frac{k\alpha_\zeta}{\pi}} \left(q_k^{(2)} + \mathrm{i}p_k^{(2)}\right), \\ T_{+k} &= \frac{1}{2}\sqrt{\frac{k\alpha_+}{\pi}} \left(q_k^{(3)} - \mathrm{i}p_k^{(3)}\right), & T_{+-k} &= \frac{1}{2}\sqrt{\frac{k\alpha_+}{\pi}} \left(q_k^{(3)} + \mathrm{i}p_k^{(3)}\right), \\ T_{-k} &= \frac{1}{2}\sqrt{\frac{k\alpha_-}{\pi}} \left(q_k^{(4)} - \mathrm{i}p_k^{(4)}\right), & T_{--k} &= \frac{1}{2}\sqrt{\frac{k\alpha_-}{\pi}} \left(q_k^{(4)} + \mathrm{i}p_k^{(4)}\right). \end{aligned} \tag{88}$$

Above we have assumed that α_ζ, α_+ and α_- are positive; in light of its definition, $\alpha_{\bar{D}}$ is negative and there is an intrinsic parity difference between the $\bar{D}_{\pm k}$ and the $\zeta_{\pm k}$ degrees of freedom. This is a fundamental property of linearized semi-direct product brackets. If $\alpha_{\bar{D}}$ is positive, then α_ζ is negative, and one must alter the $\zeta_{\pm k}$ transformations. In general, if any of the α's are negative, then a suitable canonizing transformation is obtained by inserting an absolute value inside the square root and replacing the corresponding q_k by minus q_k.

The Hamiltonian corresponding to (86) is

$$H_L = 2\pi \sum_{k=1}^{\infty} \frac{a_1}{\alpha_\zeta^2}|\bar{D}_k|^2 - \frac{a_1}{\alpha_\zeta}\sqrt{|\alpha_{\bar{D}}|/\alpha_\zeta}\left(\bar{D}_k\zeta_{-k} + \bar{D}_{-k}\zeta_k\right) + \left(a_2 + a_1|\alpha_{\bar{D}}|/\alpha_\zeta\right)|\zeta_k|^2$$
$$+a_3|T_{+k}|^2 + a_4|T_{-k}|^2 + a_5\left(T_{+k}T_{--k} + T_{+-k}T_{-k}\right) + \frac{a_6}{\alpha_\zeta}\left(\bar{D}_kT_{+-k} + \bar{D}_{-k}T_{+k}\right)$$
$$-\frac{a_6}{\alpha_\zeta}\left(\bar{D}_kT_{--k} + \bar{D}_{-k}T_{-k}\right) + \left(a_2 - \frac{a_6}{\alpha_\zeta}\sqrt{|\alpha_{\bar{D}}|/\alpha_\zeta}\right)\left(\zeta_kT_{+-k} + \zeta_{-k}T_{+k}\right)$$
$$+\left(a_2 + \frac{a_6}{\alpha_\zeta}\sqrt{|\alpha_{\bar{D}}|/\alpha_\zeta}\right)\left(\zeta_kT_{--k} + \zeta_{-k}T_{-k}\right). \tag{89}$$

In terms of the canonical variables this Hamiltonian becomes

$$H_L = \frac{1}{2}\sum_{k=1}^{\infty}\sum_{i,j=1}^{4} k\left(M_{ij}p_k^{(i)}p_k^{(j)} + K_{ij}q_k^{(i)}q_k^{(j)}\right), \tag{90}$$

where the symmetric matrices M and K are given by

$$M_{11} = K_{11} = \frac{|\alpha_{\bar{D}}|}{\alpha_\zeta^2}a_1, \qquad M_{22} = K_{22} = \frac{|\alpha_{\bar{D}}|}{\alpha_\zeta^2}a_1 + \frac{a_2}{\alpha_\zeta}, \qquad M_{33} = K_{33} = \alpha_+a_3,$$
$$M_{44} = K_{44} = \alpha_-a_4, \qquad M_{12} = -K_{12} = -\frac{|\alpha_{\bar{D}}|}{\alpha_\zeta^2}a_1, \qquad M_{13} = -K_{13} = \frac{\sqrt{|\alpha_{\bar{D}}|\alpha_+}}{\alpha_\zeta}a_6,$$
$$M_{14} = -K_{14} = -\frac{\sqrt{|\alpha_{\bar{D}}|\alpha_-}}{\alpha_\zeta}a_6, \qquad M_{34} = K_{34} = \sqrt{\alpha_+\alpha_-}a_5, \tag{91}$$
$$M_{23} = K_{23} = \sqrt{\alpha_+\alpha_\zeta}a_2 - \frac{\sqrt{|\alpha_{\bar{D}}|\alpha_+}}{\alpha_\zeta}a_6, \qquad M_{24} = K_{24} = \sqrt{\alpha_\zeta\alpha_-}a_2 + \frac{\sqrt{|\alpha_{\bar{D}}|\alpha_-}}{\alpha_\zeta}a_6,$$

with

$$a_1 = \frac{1}{d^2} + A_D - \frac{d_e^2}{d^4(1+d_e^2k_\perp^2)} = \frac{\alpha_\psi}{\alpha_Dd^2} - \frac{d_e^2}{d^4(1+d_e^2k_\perp^2)},$$
$$a_2 = \frac{1}{k_\perp^2},$$
$$a_{3,4} = \frac{1}{k_\perp^2} + A_\pm + \frac{2c_\beta^2d^4}{d_i^2} - \frac{c_\beta^2d^2}{1+d_e^2k_\perp^2} = \frac{1}{k_\perp^2} \pm \frac{c_\beta d}{d_e}\frac{\alpha_\psi}{\alpha_\pm} - \frac{c_\beta^2d^2}{1+d_e^2k_\perp^2},$$
$$a_5 = \frac{1}{k_\perp^2} + \frac{c_\beta^2d^2}{1+d_e^2k_\perp^2},$$
$$a_6 = \frac{c_\beta d_e}{(1+d_e^2k_\perp^2)d}. \tag{92}$$

Note that the matrices M and K have commuting and anti-commuting parts, which can be traced to the intrinsic parity difference mentioned above.

6.3. Stability, signature, and normal forms

One could proceed directly by linearizing the system of equations (1)–(4) about the equilibrium (75) and obtain a system for the linear dynamics. Because we are using variables indexed by k and $-k$ as independent variables, we obtain the combined linear system

$$\dot{\tilde{\xi}}_k = \mathcal{R}_k\cdot\tilde{\xi}_k \qquad \text{and} \qquad \dot{\tilde{\xi}}_{-k} = \mathcal{R}_{-k}\cdot\tilde{\xi}_{-k}, \tag{93}$$

where $\tilde{\xi}_{\pm k}$ is a four dimensional vector and $\mathcal{R}_{\pm k}$ is a 4×4 matrix, for each value of $k = 1, 2, \ldots$. This means we have an eight-dimensional system for each value of k, and upon assuming $\tilde{\xi} \sim \exp(\mathrm{i}\omega t)$ we obtain the dispersion relation from

$$\det\begin{pmatrix} \mathrm{i}\omega I_4 - \mathcal{R}_k & O_4 \\ O_4 & \mathrm{i}\omega I_4 - \mathcal{R}_{-k} \end{pmatrix} = \det(\mathrm{i}\omega I_4 - \mathcal{R}_k)\cdot\det(\mathrm{i}\omega I_4 - \mathcal{R}_{-k}) = \det(-\omega^2 I_4 + \mathcal{R}_k^2) = 0, \tag{94}$$

where O_4 is a 4×4 matrix of zeros, I_4 is the 4×4 identity matrix and use has been made of $\mathcal{R}_{-k} = -\mathcal{R}_k$ which is easily verified for the case at hand. Alternatively, one can assume $(q_k^{(i)}, p_k^{(i)}) = (\hat{q}_k^{(i)}\exp(\mathrm{i}\omega t), \hat{p}_k^{(i)}\exp(\mathrm{i}\omega t))$, and obtain the dispersion relation from Hamilton's equations in the form

$$\mathrm{i}\omega\hat{q}_k^{(i)} = \sum_j kM_{ij}\,\hat{p}_k^{(j)} \qquad \text{and} \qquad \mathrm{i}\omega\hat{p}_k^{(i)} = -\sum_j kK_{ij}\hat{q}_k^{(j)}, \tag{95}$$

whence one obtains the dispersion relation as $\det(\omega^2\mathcal{I}_4 - k^2MK) = 0$, a relation equivalent to (94) with $MK = \mathcal{R}_k^2$. Because of this special form, the dispersion relation can be factored, and reduced to a 4×4 determinant that provides the frequencies for both positive and negative k. Thus the symmetry $\mathcal{R}_{-k} = -\mathcal{R}_k$ allows for a simplification that is manifested by the special form of the Hamiltonian of (90). We note that this special form occurs for all the basic fluid and plasma models, because they have real variable Hamiltonian form.

For convenience, we introduce the dimensionless variables

$$v_{\mathrm{A}} = \alpha_\psi, \qquad \kappa_\perp = k_\perp d_{\mathrm{e}}, \qquad r = \frac{d_{\mathrm{e}}}{d_\beta}, \qquad N = \frac{\omega r}{kv_{\mathrm{A}}},$$
$$\delta = c_\beta r = \frac{d_{\mathrm{e}}}{d_{\mathrm{i}}} = \sqrt{\frac{m_{\mathrm{e}}}{m_{\mathrm{i}}}}, \qquad s = \frac{\alpha_v d_\beta r}{v_{\mathrm{A}}}, \qquad \nu = \frac{\alpha_Z d_\beta r}{v_A}, \tag{96}$$

in terms of which we derive the following dispersion relation:

$$(1+\kappa_\perp^2)N^4 - \nu N^3 - \left(\delta^2 + \delta^2\kappa_\perp^2 + \delta s + r^2 + \kappa_\perp^2\right)N^2 + \nu r^2 N + r^2\delta(s+\delta) = 0, \tag{97}$$

From the form of (97) it is clear that N is a function of $\kappa_\perp^2$ alone; moreover, it can be rewritten as

$$\kappa_\perp^2 = -\frac{(N^2 - N_r^2)(N - N_+)(N - N_-)}{N^2(N^2 - N_\delta^2)}, \tag{98}$$

where

$$N_\pm = \frac{\nu \pm \sqrt{\nu^2 + 4\delta(\delta + s)}}{2}, \qquad N_r = r \qquad \text{and} \qquad N_\delta = \sqrt{1+\delta^2}. \tag{99}$$

Analysis of the dispersion relation reveals that the four roots correspond to two that are Alfvén-like, two that are a combination of a drift-like wave that arises from density shear ν and a Kelvin–Helmholtz-like wave that arises from the parallel velocity shear s.

For example, if one sets $\delta = 0$ and $\nu = 0$, then (98) yields in dimensional variables $\omega = \pm kv_{\mathrm{A}}\sqrt{(1 + k_\perp^2 d_\beta^2)/(1 + k_\perp^2 d_{\mathrm{e}}^2)}$. The square root in this dispersion relation displays the slippage of flux, which comes from two sources, a numerator that is dependent on the electron temperature through d_β and a denominator that is dependent on electron inertia through d_{e}. Both of the terms in the square root break the MHD frozen-in condition and for $d_{\mathrm{e}} = d_\beta = 0$ one recovers the Alfvén wave dispersion relation. For $d_{\mathrm{e}} = 0$ and $d_\beta \neq 0$ this dispersion relation reduces to that for a version of the 'kinetic' Alfvén wave, while for large k one obtains the phase velocity $\omega/k \sim v_{\mathrm{A}} d_\beta/d_{\mathrm{e}} = \sqrt{T_{\mathrm{e}}/m_{\mathrm{e}}}$, the electron thermal speed. For $d_\beta = 0$ and large k, the lower hybrid frequency $\omega \sim v_{\mathrm{A}}/d_{\mathrm{e}} = v_{\mathrm{A}}\omega_{\mathrm{pe}}/c = eB/c\sqrt{m_{\mathrm{e}}m_{\mathrm{i}}}$ is obtained.

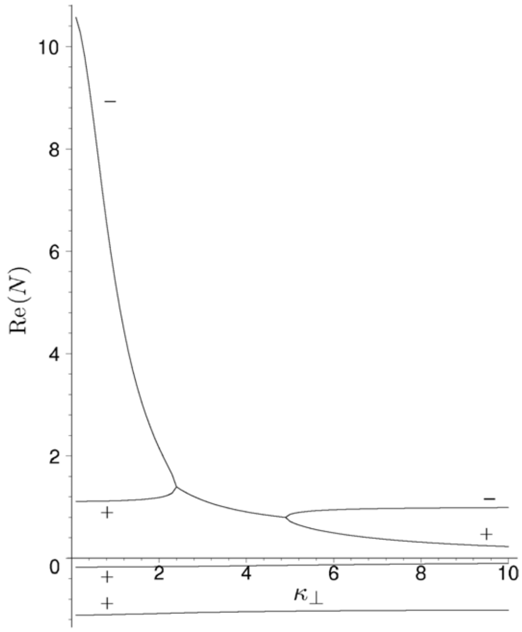

Figure 2. Solution of the dispersion relation of (97). The real part of N is plotted versus $\kappa_\perp$ for $r = 1.11$, $\nu = 10.5$, $\delta = 0.10$, $s = 18.5$, which corresponds to positive parallel velocity shear with $\delta + s > 0$. The energy signatures of the modes are indicated.

Similar limits reveal the presence of drift-waves associated with ν and Kelvin–Helmholtz modes associated with s. For example, if we set $\delta = 0$ and suppose N is small for large wave numbers, then (98) gives the drift-wave dispersion relation, $\omega = kv_*/(1 + k_\perp^2 d_\beta^2)$, where $v_* = d_\beta \alpha_Z$.

Equation (98) is convenient for obtaining stability criteria, by examination of the zeros and divergences of its right-hand side, and by noting that it asymptotes to unity for large $|N|$. Because δ and r are always positive, the sign of the divergence at $N = 0$ is governed by $\delta + s$, and this sign distinguishes two cases. The *first case*, $\delta + s > 0$, corresponds to positive or weakly negative parallel velocity shear. It is convenient to define the central band by $\mathcal{C}_\delta := \{N||N| < N_\delta\}$, which is bordered by the divergences, and the set of frequencies at which zeros occur, $\mathcal{N} := \{N_+, N_-, N_r, -N_r\}$. If any two elements of $\mathcal{N}$ are contained in $\mathcal{C}_\delta$, then the system is stable. If three elements of $\mathcal{N}$ are contained in $\mathcal{C}_\delta$, then the system is stable, and if all four of the elements of $\mathcal{N}$ are contained in $\mathcal{C}_\delta$, then the system is stable. The *second case*, $\delta + s < 0$, corresponds to strong negative velocity shear. If $\nu^2 + 4\delta(\delta + s) > 0$, then the system is unstable for large enough $\kappa_\perp$, but always possesses two stable modes. In the case of very strong negative velocity shear $\nu^2 + \delta(\delta + s) < 0$, the set of zeros becomes $\mathcal{N} := \{N_r, -N_r\}$, and the system is unstable for all $\kappa_\perp$, but again always possesses two stable modes.

In figures 2 and 3 solutions of (98), with the real part of N versus $\kappa_\perp$, are plotted. In figure 2 the parameters $r = 1.11$, $\nu = 10.5$, $\delta = 0.10$ and $s = 18.5$ are used, while in figure 3 the same values except $s = -13.4$ are used. For small $\kappa_\perp$, the uppermost curve corresponds to the drift-shear wave that has $N = N_+$ at $\kappa_\perp = 0$. The second to upper is the Alfven wave, which has $N = N_r$ at $\kappa_\perp = 0$ corresponding to $\omega/k = v_A$, and the lowermost curve is its negative counterpart. The remaining wave is a drift-shear mode with $N = N_-$ at $\kappa_\perp = 0$. For larger $\kappa_\perp$ there exist regions of instability.

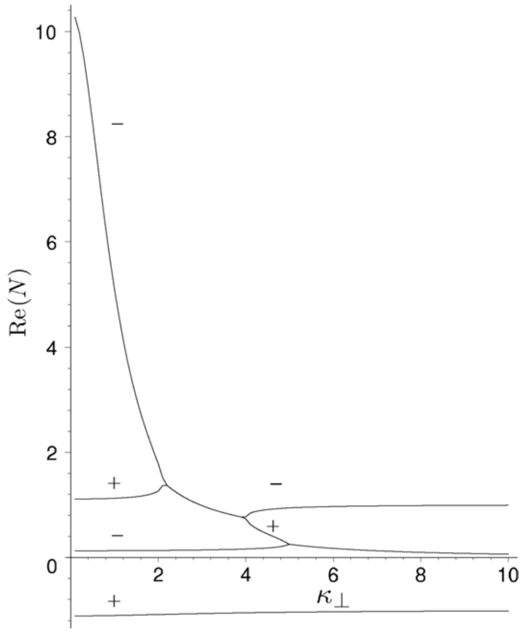

Figure 3. Solution of the dispersion relation of (97). The real part of N is plotted versus $\kappa_\perp$ for $r = 1.11$, $\nu = 10.5$, $\delta = 0.10$, $s = -13.4$, which corresponds to strong negative parallel velocity shear with $\delta + s < 0$. The energy signatures of the modes are indicated.

For the stable degrees of freedom of Hamiltonian systems there exists a special form, a so-called normal form, to which all such systems can be mapped by a canonical transformation, $(q, p) \longleftrightarrow (Q, P)$. The algorithm for this, which uses the real and imaginary parts of the linear eigenvectors, was first proven in total generality in [51]. (See, e.g. [52] for a more recent source for a version of the algorithm and [53] where a plasma example is worked out.) For the stable modes this normal form Hamiltonian is given by

$$H_L' = \frac{1}{2}\sum_k{}' \sum_{i=1}^{4} \sigma_k^{(i)}\, \omega_k^{(i)} \left(P_k^{(i)2} + Q_k^{(i)2} \right) , \tag{100}$$

where the prime on the sum means the k_y and k_x values for the unstable modes are absent (recall the suppressed sum on k_x) and the frequencies $\omega_k^{(i)} > 0$. This Hamiltonian is merely that for a collection of simple harmonic oscillators, except for the presence of the signature $\sigma_k^{(i)} \in \{1, -1\}$. Modes with signature $\sigma_k^{(i)} = -1$ oscillate, but are negative energy modes.

The signatures of the modes can be obtained by inserting the eigenvectors into the Hamiltonian (89). If $\Upsilon_k^{(i)}$ is the eigenvector associated with the mode indexed by i and k, then its signature is given by the sign of $H_L = \Upsilon_k^{(i)*} \hat{H}_L \Upsilon_k^{(i)}$, where $\hat{H}_L$ is the matrix of the bilinear form (89). To determine the signature of all modes, it is only necessary to do this for $\kappa_\perp \to \infty$ and $\kappa_\perp \to 0$.

From (98) it follows for $\kappa_\perp \to \infty$ that $\omega\kappa_\perp/k \sim \pm v_{\mathrm{A}}\sqrt{\delta(\delta+s)/(1+\delta^2)}$. For these two modes in this limit, (80) is dominated by contributions from the terms $|J|^2$, $|\nabla\varphi|^2$, v^2 and D^2, giving

$$H_L \sim \frac{\kappa_\perp^2 |\varphi_k|^2}{d_{\mathrm{e}}^2}, \qquad N \to 0, \qquad \kappa_\perp \to \infty, \tag{101}$$

and thus they both have positive signature. Note that, because $s + \delta < 0$, these modes are unstable for figure 3 and (101) does not apply. For the remaining two modes the energy in this limit behaves as

$$H_L \sim \frac{\kappa_\perp^4 |\varphi_k|^2 (1+\delta^2)^2}{r^2 d_e^2 (N_\delta \mp N_+)(N_\delta \mp N_-)}, \qquad N \to \pm N_\delta, \qquad \kappa_\perp \to \infty. \tag{102}$$

From (102) if follows that for the examples of figures 2 and 3, the mode approaching N_δ is negative while that approaching $-N_\delta$ is positive.

In the limit $\kappa_\perp \to 0$ the modes have the values $\pm N_r$ and $N_\pm$, plus corrections to these values of order $\kappa_\perp^2$. For the Alfven waves that emerge from $\pm N_r$ the energy vanishes to leading order, reflecting the vanishing of line-bending in this long wavelength limit, but proceeding to order $\kappa_\perp^2$ gives positive energies for all the stable Alfven waves of figures 2 and 3. Note that, for figure 2, the signature of the negative Alfven wave $-N_r$ is consistent with the result of the $\kappa_\perp \to \infty$ calculation: because this mode does not traverse zero or suffer a bifurcation to instability, it cannot change signature. Similarly we obtain that the drift-shear mode emerging from N_+ has negative energy for the values of both figures 2 and 3, while the N_- mode has positive energy for figure 2 and negative energy for figure 3. As s goes from 18.5 to -13.4, its frequency goes through zero and the mode changes signature.

In figure 2 there is an unstable gap that lies between $2.39 \leqslant \kappa_\perp \leqslant 4.91$. At the value $\kappa_\perp \approx 2.39$, we have a 'collision' where the frequencies of the Alfven wave and the upper drift-shear wave match and a transition to instability occurs. This transition is one of two types that occur in Hamiltonian systems; the other occurs at zero frequency as is the case for ideal MHD instabilities of static equilibria. A necessary condition for the transition to instability at nonzero frequency is that one of the modes must be a negative energy mode. This result is known as the Krein–Moser theorem [54]. In figure 2 the transition that occurs at $\kappa_\perp \approx 2.39$ and its inverse that occurs at $\kappa_\perp \approx 4.91$ are both Krein–Moser transitions. Consequently one of the modes involved *must* be a negative energy mode, and we have shown this to be the case. It is important to note that the Hamiltonian formalism is the only reliable way to determine the existence of negative energy modes of a system. If one chooses an initial condition that only excites the drift-shear wave, the energy of the system will be negative. This is essentially what is shown when the sign of $H_L = \Upsilon_k^{(i)*} \hat{H}_L \Upsilon_k^{(i)}$ is obtained.

The existence of negative energy modes is a necessary condition for the transition to instability, but it is not a sufficient condition. Upon collision, modes may merely pass through each other and remain on the stable axis. The Hamiltonian is not a frame dependent quantity, and it is possible to change the signature of a mode with a frame change. For systems that possess a momentum invariant, the energy in the new frame is the same as the old, but with doppler shifted frequencies. Sometimes it is possible to remove a negative energy mode by this procedure and thus obtain a Liapunov type stability argument that precludes the transition to instability [55,56]. In any event, when a transition at finite frequency occurs, one is certain of the existence of a negative energy mode.

After the transition, the energy drops to zero. Unstable modes always have zero energy, and thus they have no signature. After the transition, the unstable gap modes have eigenvalues that occur in quartet form, $\pm\omega_R \pm \omega_I$, and the unstable normal form is

$$H_L'' = \sum_k{}'' \Bigg[\omega_{Rk} \left(P_k^{(1)} Q_k^{(2)} - P_k^{(2)} Q_k^{(1)} \right) - \omega_{Ik} \left(P_k^{(1)} Q_k^{(1)} + P_k^{(2)} Q_k^{(2)} \right) + \frac{1}{2} \sum_{i=3,4} \sigma_k^{(i)} \omega_k^{(i)} \left(P_k^{(i)2} + Q_k^{(i)2} \right) \Bigg], \tag{103}$$

where the double prime indicates a sum over k-values of the gap, the indices 1 and 2 denote the upper Alfven and drift modes, and the indices 3 and 4 denote the remaining two stable modes. In the new coordinates, the total Hamiltonian is given by $H_L = H_L' + H_L''$. Recall, in all of these sums, the sum over the x-components of the wave numbers have been suppressed.

In figure 3 the unstable gap arising from the collision of the Alfven and drift modes occurs for values $2.12 \leqslant \kappa_\perp \leqslant 3.99$, and then another collision occurs when a stable wave with positive signature collides with the lower drift-shear mode at $\kappa_\perp \approx 4.95$. This lower mode changed from its positive energy value in figure 2 to a negative energy mode upon traversing $N = 0$. For $\kappa_\perp > 4.95$, the normal form is like (103), except now the stable modes are the lower Alfven wave and the mode that limits to N_δ, and the other two modes are unstable for arbitrarily large $\kappa_\perp$.

From equations (101) and (102) we see that if all modes are stable they will have positive energy if $(N_\delta \mp N_+)(N_\delta \mp N_-) > 0$ for both signs. When this is the case, the system is energy stable, which can be traced back to the positive definiteness of $\delta^2 F$. This kind of stability is sometimes called energy-Casimir stability (see [15] for references to the early plasma literature where this idea was introduced). Being energy stable means that our Hamiltonian H_L can be brought into the form (100) for all k, with $\sigma_k^{(i)} > 0$ for all i and k.

7. Collisionless conductivity and tearing modes

In this section we consider another linear application. We show that the Jacobi identity of the bracket '[,]', an identity essential for the Hamiltonian description of section 3 and its Jacobi identity for the bracket '{ , }', can be used to show that the azimuthal current density responds *locally* to the electric field on each flux surface. That is, the current is proportional to the parallel electric field and is independent of its gradient across flux surfaces. This would not be true in the presence of particle diffusivity or in the presence of a finite electron gyroradius. The locality of the conduction is important as it allows the linearized system of equations to be reduced to a system that differs from that describing resistive tearing modes in a sheared slab by merely replacing the collisional conductivity σ by a spatially dependent, ac collisionless conductivity.

7.1. Collisionless conductivity

We consider perturbations of a plasma slab in which all equilibrium fields vary only in the x direction, i.e. the field variables have the following form:

$$\psi = \psi_0(x) + \tilde{\psi}_k(x)\mathrm{e}^{\mathrm{i}\omega t - \mathrm{i}ky} + \mathrm{c.c.}, \tag{104}$$

$$\varphi = -\omega x/k + \tilde{\varphi}_k(x)\mathrm{e}^{\mathrm{i}\omega t - \mathrm{i}ky} + \mathrm{c.c.}, \tag{105}$$

$$Z = \alpha_Z x + \tilde{Z}_k(x)\mathrm{e}^{\mathrm{i}\omega t - \mathrm{i}ky} + \mathrm{c.c.}, \tag{106}$$

$$v = \alpha_v x + \tilde{v}_k(x)\mathrm{e}^{\mathrm{i}\omega t - \mathrm{i}ky} + \mathrm{c.c.}, \tag{107}$$

where, to avoid clutter, below we drop the subscript k on the tilde variables. The above equations are the same as those of section 6.2, except here we assume that $\psi_0(x)$ of (104) is a polynomial of at most quadratic order in x. This assumption makes it possible to apply the final result to two cases, namely, the homogeneous equilibrium of section 5.4 and the case of a magnetic equilibrium with a resonant surface at $x = 0$. The presence of an equilibrium flow in equation (105) corresponds to choosing a reference frame moving with the phase velocity of the perturbation. This choice is convenient because it allows us to ignore the terms with the time derivatives in our calculations below.

The Jacobi identity for [,] is

$$[\varphi, [\psi, \xi]] + [\xi, [\varphi, \psi]] + [\psi, [\xi, \varphi]] = 0, \tag{108}$$

where ξ represents any of the field variables of (104)–(107) and consequently it has the form $\xi = \xi_0(x) + \tilde{\xi}(x)\exp(\mathrm{i}\omega t - \mathrm{i}ky) + \mathrm{c.c.}$ Linearizing (108) and retaining terms of first order give

$$-\frac{\omega}{k}\frac{\partial}{\partial y}[\psi, \xi]_L + \xi_0'\frac{\partial}{\partial y}[\varphi, \psi]_L + \psi_0'\frac{\partial}{\partial y}[\xi, \varphi]_L = 0, \tag{109}$$

where $[f, g]_L$ is a short-hand notation for the linearized Poisson bracket between two fields, e.g. $[\psi, \xi]_L = -\mathrm{i}k(\psi_0'\tilde{\xi} - \xi_0'\tilde{\psi})\exp(\mathrm{i}\omega t - \mathrm{i}ky)$.

By replacing ξ with Z in (109) and using (2) one obtains

$$-\frac{\omega}{k}\frac{\partial}{\partial y}[\psi, Z]_L + \alpha_Z\frac{\partial}{\partial y}[\varphi, \psi]_L + \psi_0'\frac{\partial}{\partial y}(d_\beta[J, \psi]_L - c_\beta[v, \psi]_L) = 0, \tag{110}$$

where $J = -\nabla^2\psi$. On the other hand by replacing ξ with v in (109), one gets

$$\frac{\partial}{\partial y}[v, \psi]_L = -\frac{k}{\omega}\left(\alpha_v\frac{\partial}{\partial y}[\varphi, \psi]_L + \psi_0'\frac{\partial}{\partial y}[v, \varphi]\right) = 0. \tag{111}$$

The latter expression can be used to replace $\partial/\partial y[v, \psi]_L$ in (110) and obtain

$$\begin{aligned} &-\frac{\omega}{k}\frac{\partial}{\partial y}[\psi, Z]_L + \alpha_Z\frac{\partial}{\partial y}[\varphi, \psi]_L + \psi_0' d_\beta\frac{\partial}{\partial y}[J, \psi]_L + \psi_0' c_\beta \\ &\qquad \times\left(\frac{k}{\omega}\alpha_v\frac{\partial}{\partial y}[\varphi, \psi]_L + \frac{k}{\omega}\psi_0'\frac{\partial}{\partial y}[v, \varphi]\right) = 0. \end{aligned} \tag{112}$$

Making use of (4), (112) can be reformulated as

$$-d_\beta\frac{\partial}{\partial y}[\psi, Z]_L = d_\beta\frac{(\omega\alpha_Z + \psi_0'\alpha_v c_\beta k)\frac{\partial}{\partial y}[\varphi, \psi]_L + \omega\psi_0' d_\beta\frac{\partial}{\partial y}[J, \psi]_L}{-\omega^2 + {\psi_0'}^2 c_\beta^2 k^2}. \tag{113}$$

Ohm's law (1), on the other hand, yields

$$-d_\beta\frac{\partial}{\partial y}[\psi, Z]_L = -\frac{\partial}{\partial y}[\varphi, \psi]_L - d_\mathrm{e}^2\frac{\partial}{\partial y}[\varphi, J]. \tag{114}$$

Given that in the linearized Poisson bracket the dependence on y enters only through the exponential, the derivative with respect to y amounts to a multiplication times $-\mathrm{i}k$. Bearing this in mind and combining (113) with (114) yields

$$\frac{d_\beta k\omega\alpha_Z + d_\beta\psi_0'\alpha_v c_\beta k^2 - \omega^2 + {\psi_0'}^2 c_\beta^2 k^2}{-\omega^2 + {\psi_0'}^2 c_\beta^2 k^2}[\varphi, \psi]_L = \left[\frac{\omega k\psi_0' d_\beta^2}{-\omega^2 + {\psi_0'}^2 c_\beta^2 k^2}\psi - d_\mathrm{e}^2\varphi, J\right]_L. \tag{115}$$

Now if one introduces in the linearized Poisson brackets the explicit expressions (104)–(107) for the fields, one obtains

$$\mathrm{i}\sigma(x)(\omega\tilde{\psi} + \psi_0' k\tilde{\varphi}) = -k^2\tilde{\psi} + \tilde{\psi}'', \tag{116}$$

where

$$\sigma(x) := \frac{1 - \omega_*/\omega - k_\parallel c_\beta d_\beta\omega_{\mathrm{KH}}/\omega^2 - {k_\parallel}^2 c_\beta^2/\omega^2}{\mathrm{i}\omega(d_\mathrm{e}^2 - {k_\parallel}^2 d^2 c_\beta^2/\omega^2)}, \tag{117}$$

with $\omega_* = d_\beta k\alpha_Z$, $k_\parallel = k\psi_0'$ and $\omega_{\mathrm{KH}} = k\alpha_v$. This shows that Ohm's law can be written as a proportionality relation between the amplitudes of the projection of the current density along z and of the electric field along the poloidal magnetic field. The quantity $\sigma(x)$ then plays the role of a spatially dependent conductivity. The various terms in σ have the following interpretation. The numerator represents the forces acting on the electrons: the electric field itself, the pressure force (represented by the ω_*/ω term) and the screening effect of parallel ion motion ($k_\parallel c_\beta d_\beta \omega_{\rm KH}/\omega^2 + {k_\parallel}^2 c_\beta^2/\omega^2$ where the term proportional to $\omega_{\rm KH}$ represents the effect of the Doppler shift caused by the background flow). The denominator represents the response to those forces, namely, the electron inertia ($d_{\rm e}^2$) and a term representing electron thermal diffusion along the field line (${k_\parallel}^2 d^2 c_\beta^2/\omega^2$).

The case $\psi_0(x) = x^2/2L_{\rm s}$ corresponds to an equilibrium with scale length $L_{\rm s}$ and with a resonant surface at $x = 0$. If for this case we restrict to a thin layer around the resonant surface, the system comprised of (116) and the vorticity equation (3) can be approximated by

$$\mathrm{i}\sigma(x)(\omega\tilde{\psi} + k_\parallel\tilde{\varphi}) = \tilde{\psi}'' \qquad \text{and} \qquad \omega\tilde{\phi}''(x) = -k_\parallel\tilde{\psi}'', \tag{118}$$

where $k_\parallel = kx/L_{\rm s}$ and y-derivatives, being negligible in the layer, have been dropped. Thus the layer equations for the present model take the same form as those of MHD,

$$\omega\tilde{\varphi}''(x) = x\tilde{\psi}''(x) \qquad \text{and} \qquad \sigma(x)E_\parallel = \tilde{\psi}''(x), \tag{119}$$

except for the replacement of the conductivity by the spatially varying ac conductivity $\sigma(x)$ of (117). Here $E_\parallel = \mathrm{i}(\omega\tilde{\psi} + x\tilde{\varphi})$.

The case $\psi_0(x) = \alpha_\psi x$ corresponds to the homogeneous equilibria of section 6. Here σ is constant and equations (118) become

$$\mathrm{i}\sigma(\omega\tilde{\psi} + \alpha_\psi k\tilde{\varphi}) = -k^2\tilde{\psi} + \tilde{\psi}'' \qquad \text{and} \qquad \omega(-k^2\tilde{\varphi} + \tilde{\varphi}'') = -\alpha_\psi k(-k^2\tilde{\psi} + \tilde{\psi}''). \tag{120}$$

The solvability condition for this system, for solutions with dependence on x of the form $\exp(-\mathrm{i}k_x x)$, gives again the dispersion relation (97).

7.2. *Collisionless tearing mode*

Now the dispersion relation derived above is used to obtain the growth rate for collisionless tearing modes. We restrict attention to the case of moderate Δ' where the constant-$\tilde{\psi}$ approximation applies. Mirnov *et al* [57] have recently described the opposite case of large Δ' using a two-fluid model analogous to that of [38].

In the constant-$\tilde{\psi}$ approximation, the dispersion relation for the tearing mode follows from the matching condition for the magnetic perturbation,

$$\Delta' = \frac{1}{\tilde{\psi}}\int_{-\infty}^{\infty} \mathrm{d}x\, \tilde{J}. \tag{121}$$

At the resonant surface, $k_\parallel = 0$, the conductivity is very high due to the high electron mobility, $\mathrm{i}\omega\sigma(0) = (1 - \omega_*/\omega)/d_{\rm e}^2$. Away from the resonant surface, however, the conductivity decreases rapidly due to the shielding of the electric field by the electron motion along the magnetic field. The shielding is described by the $k_\parallel d_\beta$ term in the denominator. The region of high conductivity is called the current channel and for moderate tearing parameter Δ', it contains most of the current in the reconnection layer. In the current channel, the conductivity may be approximated by

$$\sigma(x) \simeq -d_e^{-2}\frac{1-\omega_*/\omega}{\mathrm{i}\omega\left[\frac{k_\parallel^2 d_\beta^2}{\omega^2 d_e^2}-1\right]},$$

Substituting this in the matching integral and evaluating the integral gives

$$\Delta' = -\mathrm{i}\pi(\omega-\omega_*)\frac{L_s}{kd_\beta d_e},$$

whence we obtain the dispersion relation

$$\omega = \omega_* + \mathrm{i}\frac{\Delta' k d_\beta}{\pi L_s}d_e. \tag{122}$$

In the limit $\beta \ll 1$ (122) agrees with the kinetic result of [58], aside from a factor of $2/\sqrt{\pi} \simeq 1.13$ in the growth rate, and it agrees with the fluid result of [59] in the low β limit.

8. Saturation of the collisionless tearing mode

For our last application of the Hamiltonian formalism we use a Casimir invariant to find the nonlinear saturated state of the collisionless tearing mode. Indeed, although we are treating a Hamiltonian model, in reality viscosity will always be present to dissipate the kinetic energy so that the system could evolve toward a saturated equilibrium. In particular, the procedure adopted in this section, relies on the assumption that the viscous dissipation time is shorter than the resistive diffusion time, so that the conservation of electron momentum used in the calculation can still be considered valid on the time scale of viscous dissipation. This assumption is reasonable for many plasmas of interest.

In this analysis we follow an approach similar to that in [60], except that here we make use of the constant-ψ approximation to simplify the analysis. For simplicity we consider the cold plasma limit where $d_\beta = c_\beta = 0$. Inspection of (1) reveals that the Casimir C_2 of (26) survives but with D replaced by ψ_e, equation (3) is unaltered, and equations (2) and (4) ensure that if initially $Z, v \equiv 0$, then they will remain so.

We use the invariance of C_2 to describe what becomes of an unstable un-reconnected state, $\psi^{(0)} = B_0x^2/2L_s$, such as that described in section 7.2, as it evolves into a final approximate equilibrium state $\psi^{(\infty)} = B_0x^2/2L_s + \bar{\psi}\cos y$ that represents a magnetic island of half-width $w = (4L_s\bar{\psi}/B_0)^{1/2}$. For convenience, throughout this section we use dimensionless units where lengths are scaled with w and the flux with B_0/L_s. In these units, for example, $\psi^{(\infty)}$ becomes $\psi^{(\infty)} = x^2/2 + \cos y/4$.

Choosing $\mathcal{K}(\psi_e) = \delta(\psi_e - \hat{\psi}_e)/2\pi$ singles out the surface of constant $\hat{\psi}_e$, yielding

$$C_2(\hat{\psi}_e) = \frac{1}{2\pi}\oint\frac{\mathrm{d}y}{\partial_x\psi_e}, \tag{123}$$

where $\partial_x\psi_e(x, y, t)$ is to be evaluated at $x = \psi_e^{-1}(\hat{\psi}_e, y, t)$, and this can be done at any time. From $\psi^{(0)}$ we obtain $\psi_e^{(0)} = x^2/2 - \eta^2$, where $\eta^2 := d_e^2/w^2$; whence for the initial state $C_2(\hat{\psi}_e) = 1/\sqrt{2(\hat{\psi}_e + \eta^2)}$.

Because the final state is an equilibrium state, $\psi_e^{(\infty)}$ is a function of $\psi^{(\infty)}$ according to $\psi_e^{(\infty)} = \psi^{(\infty)} + \eta^2 I(\psi^{(\infty)})$, where I is the final current profile of the saturated island. Thus (123) becomes

$$C_2(\hat{\psi}_e) = \left(\frac{\mathrm{d}\psi_e^{(\infty)}}{\mathrm{d}\psi^{(\infty)}}\right)^{-1}\frac{1}{2\pi}\oint\frac{\mathrm{d}y}{\partial_x\psi^{(\infty)}}, \tag{124}$$

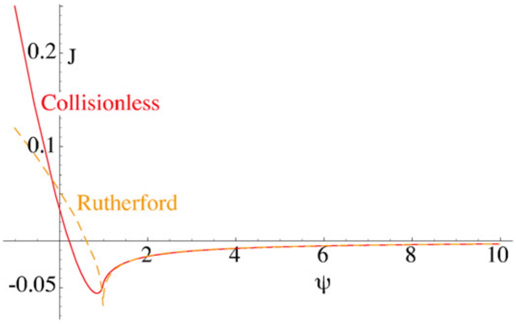

Figure 4. Comparison of the current density profiles, one determined by parallel electron momentum conservation and the other by Ohmic diffusion (dashed line).

where $\partial_x \psi^{(\infty)}(x, y)$ is to be evaluated at $x = (\psi_e^{(\infty)})^{-1}(\hat{\psi}_e, y)$ and $d\psi_e^{(\infty)}/d\psi^{(\infty)}$, being expressible as a function of $\hat{\psi}_e$ alone, can be pulled outside the integral. The final current profile follows by setting the initial and final values of $C_2(\hat{\psi}_e)$ equal at each $\hat{\psi}_e$, yielding

$$\frac{1}{\sqrt{2(\hat{\psi}_e + \eta^2)}} \frac{d\hat{\psi}_e}{d\psi^{(\infty)}} = \frac{1}{2\pi} \oint \frac{dy}{\partial_x \psi^{(\infty)}}. \tag{125}$$

Although $\psi^{(\infty)}$ is unknown, for small Δ', $\partial_x \psi^{(\infty)} \approx x \approx \sqrt{(4\psi^{(\infty)} - \cos y)/2}$. Integrating equation (125) yields,

$$\hat{\psi}_e = \iota^2(\psi^{(\infty)}) - \eta^2 . \tag{126}$$

where

$$\iota(\psi^{(\infty)}) = \frac{1}{2\pi} \int_0^{\pi} dy \sqrt{4\psi^{(\infty)} - \cos y}$$

and where we have used the fact that $\lim_{\psi \to \infty}(\hat{\psi}_e(\psi) - \psi) = 0$ to set the integration constant to zero.

Using $\psi_e^{(\infty)} = \psi^{(\infty)} + \eta^2 I(\psi^{(\infty)})$, and dropping the 'hat' and the label ∞, gives an equation for the current profile,

$$I(\psi) = -1 + \eta^{-2}\left[-\psi + \iota^2(\psi)\right]. \tag{127}$$

The function $\iota(\psi)$ is easily evaluated in terms of elliptic integrals:

$$\iota(\psi) = \begin{cases} \frac{2}{\pi}\sqrt{2\psi + 1/2}\, E(1/(2\psi + 1/2)), & \text{for } \psi > 1/4, \\ \sqrt{2}\,(E(2\psi + 1/2) + (2\psi - 1/2)K(2\psi + 1/2)), & \text{for } -1/4 < \psi < 1/4, \end{cases}$$

where K and E are the complete elliptic integral of the first and second kind.

Figure 4 shows a comparison of the above current profile with that of Rutherford [40] for resistive diffusion. Observe, the profiles are qualitatively similar. Substituting the current profile in the matching relation of (121) yields the saturation amplitude $w = \Delta' d_e^2/G$, where

$$G = 8 \int_{\psi_{\min}}^{\infty} d\psi\, I(\psi) \oint \frac{dy}{2\pi} \frac{\cos y}{\partial_x \psi} = 0.19.$$

This is close to the value $G = 0.205$ obtained by Drake and Lee [61] using a kinetic model.

The physical mechanism causing the saturation is the fact that the electron response is limited by the thermal Doppler frequency of the electrons, $k_\| v_{te} \ll \gamma$ where γ is the growth rate. In the linear regime, this results in a layer of width $\Delta' d_e^2$. As the island width reaches the layer width, however, the effective $k_\parallel$ for the particles near the island grows with w as $k_\parallel \sim k_y w/L_s$ (where w is the island width). When the island width reaches the layer width, the particles trapped in the island have a thermal Doppler frequency that exceeds the growth rate of the island. These particles experience only an ac electric field so that their averaged response vanishes.

9. Summary and conclusions

In the early sections of this paper we presented the noncanonical Hamiltonian formulation of the four-field model of [38], and showed that the associated Lie–Poisson bracket has four new independent families of Casimir invariants. These invariants led us to the discovery of variables in which the Poisson bracket has the simple form (35), and in which the system can be written in the compact form of (36)–(38).

In section 5 we used the Hamiltonian formulation to obtain a variational principle that gives a set of coupled differential equations that generalize the Grad–Shafranov equilibrium equation. In the limit of vanishing electron mass ($d_e \to 0$) the equilibrium equations reduce to previously known results. This limit provides some insight into the relationship of the Casimirs to the more familiar conserved quantities of conventional low-β drift models. We have presented two solutions of the equilibrium equations, the first describing dipole-like equilibria and the second describing homogeneous equilibria that support drift-acoustic and Alfvén modes.

In section 6 we investigated the linear dispersion relation for homogeneous equilibria and described the map to the appropriate Hamiltonian forms. We also presented thresholds for spectral and energy stability. In section 6.3 we described a method for determining the energy signature of a mode. This method is of general utility and can be applied to all valid models, provided one understands their Hamiltonian structure. In obtaining reduced fluid models, there can be ambiguity about the energy for the full dynamics, and linear theory alone cannot be used to uniquely determine the correct energy of the linear dynamics. The only reliable way to determine the energy is from a Hamiltonian or action principle formulation, where the energy for the linear dynamics is obtained by expansion of a Hamiltonian associated with time translation symmetry.

In sections 7 and 8 we demonstrated the usefulness of the Hamiltonian formulation for the analysis of the linear collisionless tearing mode and its nonlinear saturation. In the case of linear stability, the Jacobi identity allows the reduction of the system to a form analogous to that of MHD but where the conductivity is replaced by a spatially varying ac conductivity. Applications of the formalism left for future work include the study of the stability of the saturated states against secondary modes and an investigation of saturation in a more general dynamical context using additional Casimirs.

Another area for future work concerns the families associated with the invariants $T_\pm$, which generalize a pair of Casimirs $G_\pm$ found for a low-β two-field model derived in [4]. A natural question is whether the invariants $T_\pm$ play a role analogous to the one played by $G_\pm$ in the two-field limit in determining the alignment of current density and vorticity along the separatrices of the magnetic field during the nonlinear evolution of the system [33]. The present model makes possible an investigation of this question along the lines carried out in [22].

Acknowledgments

ET and DG were partly supported by the Euratom Communities under the contract of Association between EURATOM/ENEA. The views and opinions expressed herein do not necessarily reflect those of the European Commission. They would also like to thank F Pegoraro for drawing their attention to this problem, T J Schep for useful discussions and would also like to acknowledge support from a Grant PRIN-2006. PJM and FLW were supported by the US Department of Energy Contract No DE-FG03-96ER-54346. PJM would like to thank W Horton for useful discussions.

References

[1] Rogers B N, Denton R E, Drake J F and Shay M A 2001 Role of dispersive waves in collisionless magnetic reconnection *Phys. Rev. Lett.* **87** 195004
[2] Priest E R and Forbes T G 2000 *Magnetic Reconnection* (Cambridge: Cambridge University Press)
[3] Biskamp D 2000 *Magnetic Reconnection in Plasmas* (Cambridge: Cambridge University Press)
[4] Schep T J, Pegoraro F and Kuvshinov B N 1994 Generalized two-fluid theory of nonlinear magnetic structures *Phys. Plasmas* **1** 2843–51
[5] Dastgeeer S, Mahajan S and Weiland J 2002 Zonal flows and transport in ion temperature gradient turbulence *Phys. Plasmas* **9** 4911–6
[6] Krommes J A and Kolesnikov R A 2004 Hamiltonian description of convective-cell generation *Phys. Plasmas* **11** L29–L32
[7] Kolesnikov R A and Krommes J A 2005 Transition to collisionless ion-temperature-gradient-driven plasma turbulence: a dynamical systems approach *Phys. Rev. Lett.* **94** 235002
[8] Kolesnikov R A and Krommes J A 2005 Bifurcation theory of the transition to collisionless ion-temperature-gradient-driven plasma turbulence *Phys. Plasmas* **12** 122302
[9] Xu X Q, Cohen R H, Rognlien T D and Myra J R 2000 Low-to-high confinement transition simulations in divertor geometry *Phys. Plasmas* **7** 1951–8
[10] Rogers B N, Drake J F and Zeiler A 1998 Phase space of tokamak edge turbulence, the L–H transition, and the formation of the edge pedestal *Phys. Rev. Lett.* **81** 4396–9
[11] Guzdar P N, Mahajan S M and Yoshida Z 2005 A theory for the pressure pedestal in high (H) mode tokamak discharges *Phys. Plasmas* **12** 032502
[12] Scott B D 2007 Tokamak edge turbulence: background theory and computation *Plasma Phys. Control. Fusion* **49** S25–S41
[13] Scott B D 2005 Free-energy conservation in local gyrofluid models *Phys. Plasmas* **12** 102307
[14] Morrison P J 1982 Poisson brackets for fluids and plasmas *Mathematical Methods in Hydrodynamics and Integrability in Dynamical Systems (American Institute of Physics Conference Proceedings* vol 88*)* ed M Tabor and Y Treve (New York: American Institute of Physics) pp 13–45
[15] Morrison P J 1998 Hamiltonian description of the ideal fluid *Rev. Mod. Phys.* **70** 467–521
[16] Morrison P J 2005 Hamiltonian and action principle formulations of plasma physics *Phys. Plasmas* **12** 058102–1–058102–13
[17] Morrison P J and Greene J M 1980 Noncanonical Hamiltonian density formulation of hydrodynamics and ideal magnetohydrodynamics *Phys. Rev. Lett.* **45** 790–3
Morrison P J and Greene J M 1982 *Phys. Rev. Lett.* **48** 569 (erratum)
[18] Morrison P J and Hazeltine R D 1984 Hamiltonian formulation of reduced magnetohydrodynamics *Phys. Fluids* **27** 886–97
[19] Marsden J E and Morrison P J 1984 Noncanonical Hamiltonian field theory and reduced MHD *Contemp. Math.* **28** 133–50
[20] Hazeltine R D, Hsu C T and Morrison P J 1987 Hamiltonian four-field model for nonlinear tokamak dynamics *Phys. Fluids* **30** 3204–11
[21] Kuvshinov B N, Pegoraro F and Schep T J 1994 Hamiltonian formulation of low-frequency, nonlinear plasma dynamics *Phys. Lett.* A **191** 296–300
[22] Grasso D, Califano F, Pegoraro F and Porcelli F 2001 Phase mixing and saturation in Hamiltonian reconnection *Phys. Rev. Lett.* **86** 5051–4
[23] Grasso D, Califano F, Pegoraro F and Porcelli F 1999 Hamiltonian magnetic reconnection *Plasma Phys. Control. Fusion* **41** 1497–515

[24] Waelbroeck F L, Morrison P J and Horton W 2004 Hamiltonian formulation and coherent structures in electrostatic turbulence *Plasma Phys. Control. Fusion* **46** 1331–50

[25] Hazeltine R D and Meiss J D 1985 Shear-alfven dynamics of toroidally confined plasmas *Phys. Rep.* **121** 1–164

[26] Porcelli F, Borgogno D, Califano F, Grasso D, Ottaviani M and Pegoraro F 2002 Recent advances in collisionless magnetic reconnection *Plasma Phys. Control. Fusion* **44** B389–B405

[27] Kuvshinov B N, Lakhin V P, Pegoraro F and Schep T J 1998 Hamiltonian vortices and reconnection in a magnetized plasma *J. Plasma Phys.* **59** 727–36

[28] Aydemir A Y 1992 Nonlinear studies of $m = 1$ modes in high-temperature plasmas *Phys. Fluids* B **4** 3469–72

[29] Kleva R G, Drake J F and Waelbroeck F L 1995 Fast reconnection in high temperature plasmas *Phys. Plasmas* **2** 23–34

[30] Hazeltine R D, Kotschenreuther M and Morrison P J 1985 A four-field model for tokamak plasma dynamics *Phys. Fluids* **28** 2466–77

[31] Hazeltine R D, Kotschenreuther M and Morrison P J 1986 *Phys. Fluids* **29** 341–1 (erratum)

[32] Ottaviani M and Porcelli F 1993 Nonlinear collisionless magnetic reconnection *Phys. Rev. Lett.* **71** 3802–5

[33] Cafaro E, Grasso D, Pegoraro F, Porcelli F and Saluzzi A 1998 Invariants and geometric structures in nonlinear Hamiltonian magnetic reconnection *Phys. Rev. Lett.* **80** 4430–3

[34] Del Sarto D, Califano F and Pegoraro F 2007 Electron parallel compressibility in the nonlinear development of two-dimensional collisionless magnetohydrodynamic reconnection *Mod. Phys. Lett.* B **20** 931–61

[35] Vekstein G and Bian N H 2006 Hall assisted forced magnetic reconnection *Phys. Plasmas* **13** 122105

[36] Bian N and Vekstein G 2007 On the two-fluid modification of the resistive tearing instability *Phys. Plasmas* **14** 072107

[37] Rogers B N, Kobayashi S, Ricci P, Dorland W, Drake J and Tatsuno T 2007 Gyrokinetic simulations of collisionless magnetic reconnection *Phys. Plasmas* **14** 092110

[38] Fitzpatrick R and Porcelli F 2004 Collisionless magnetic reconnection with arbitrary guide field *Phys. Plasmas* **11** 4713–18

Fitzpatrick R and Porcelli F 2007 *Phys. Plasmas* **14** 049902 (erratum)

[39] Tassi E, Morrison P J and Grasso D 2007 Hamiltonian structure of a collisionless reconnection model valid for high and low β plasmas *Collective Phenomena in Macroscopic Systems* ed G Bertin *et al* (Singapore: World Scientific) pp 197–206

[40] Rutherford P H 1973 Nonlinear growth of the tearing mode *Phys. Fluids* **16** 1903–8

[41] Thiffeault J L and Morrison P J 2000 Classification and Casimir invariants of Lie–Poisson brackets *Physica* D **136** 205–44

[42] Yoshida Z and Mahajan S M 2002 Variational principles and self-organization in two-fluid plasmas *Phys. Rev. Lett.* **88** 095001

[43] Yoshida Z, Mahajan S M and Ohsaki S 2004 Scale hierarchy created in plasma flow *Phys. Plasmas* **11** 3660–4

[44] Goedbloed J P 2004 Variational principles for stationary one- and two-fluid equilibria of axisymmetric laboratory and astrophysical plasmas *Phys. Plasmas* **11** L81–L84

[45] Hirota M, Yoshida Z and Hameiri E 2006 Variational principle for linear stability of flowing plasmas in hall magnetohydrodynamics *Phys. Plasmas* **13** 022107

[46] Eliezer Hameiri and Torasso R 2004 Linear stability of static equilibrium states in the hall-magnetohydrodynamics model *Phys. Plasmas* **11** 4934–45

[47] Atsushi Ito, Jesús J Ramos and Noriyoshi Nakajima 2007 Ellipticity of axisymmetric equilibria with flow and pressure anisotropy in single-fluid and hall magnetohydrodynamics *Phys. Plasmas* **14** 062502

[48] Steinhauer L C and Ishida A 2006 Nearby-fluids equilibria: I. Formalism and transition to single-fluid magnetohydrodynamics *Phys. Plasmas* **13** 052513

[49] Steinhauer L C and Guo H Y 2006 Nearby-fluids equilibria: II. Zonal flows in a high-beta, self-organized plasma experiment *Phys. Plasmas* **13** 052514

[50] Gardner C S 1971 Korteweg–de Vries equation and generalizations: IV. The Korteweg–de Vries equations as a Hamiltonian system *J. Math. Phys.* **12** 1548–51

[51] Williamson J 1936 On an algebraic problem concerning the normal forms of linear dynamical systems *Am. J. Math.* **58** 141–63

[52] Laub A J and Meyer K 1974 Canonical forms for symplectic and Hamiltonian matrices *Celest. Mech.* **9** 213–38

[53] Morrison P J and Kotschenreuther M 1990 The free energy principle, negative energy modes, and stability *Nonlinear World: IV Int. Workshop on Nonlinear and Turbulent Processes in Physics* ed V G Baryakhtar *et al* (Singapore: World Scientific)

[54] Moser J 1958 New aspects in the theory of stability of Hamiltonian systems *Commun. Pure Appl. Math.* **11** 81–114

[55] Coppi B, Rosenbluth M N and Sudan R N 1969 Non-linear interactions of positive and negative energy modes *Ann. Phys.* **55** 248–70

[56] Kueny C S and Morrison P J 1995 Nonlinear instability and chaos in plasma wave-wave interactions: I. Introduction *Phys. Plasmas* **2** 1926–40

[57] Mirnov V V, Hegna C C and Prager S C 2004 Two-fluid tearing instability in force-free magnetic configuration *Phys. Plasmas* **11** 4468–82

[58] Drake J F and Lee Y C 1977 Kinetic theory of tearing instabilities *Phys. Fluids* **20** 1341–53

[59] Grasso D, Ottaviani M and Porcelli F 2002 Growth and stabilization of drift-tearing modes in weakly collisional plasmas *Nucl. Fusion* **42** 1067–74

[60] Waelbroeck F L 1989 Current sheets and nonlinear growth of the $m = 1$ kink-tearing mode *Phys. Fluids* B **1** 2372–80

[61] Drake J F and Lee Y C 1977 Nonlinear evolution of collisionless and semicollisional tearing modes *Phys. Rev. Lett.* **39** 453–6