

## LECTURE NOTES

# An introductory guide to fluid models with anisotropic temperatures. Part 1. CGL description and collisionless fluid hierarchy

P. Hunana<sup>1,2,†</sup>, A. Tenerani<sup>3</sup>, G. P. Zank<sup>4,5</sup>, E. Khomenko<sup>1,2</sup>,  
M. L. Goldstein<sup>6</sup>, G. M. Webb<sup>4</sup>, P. S. Cally<sup>7</sup>, M. Collados<sup>1,2</sup>, M. Velli<sup>8</sup>  
and L. Adhikari<sup>4</sup>

<sup>1</sup>Instituto de Astrofísica de Canarias (IAC), La Laguna, Tenerife, 38205, Spain

<sup>2</sup>Universidad de La Laguna, La Laguna, Tenerife, 38206, Spain

<sup>3</sup>Department of Physics, The University of Texas at Austin, TX 78712, USA

<sup>4</sup>Center for Space Plasma and Aeronomic Research (CSPAR), University of Alabama, Huntsville, AL 35805, USA

<sup>5</sup>Department of Space Science, University of Alabama, Huntsville, AL 35899, USA

<sup>6</sup>Space Science Institute, Boulder, CO 80301, USA

<sup>7</sup>School of Mathematics, Monash University, Clayton, Victoria 3800, Australia

<sup>8</sup>Department of Earth, Planetary, and Space Sciences, University of California, Los Angeles, CA 90095, USA

(Received 25 January 2019; revised 27 October 2019; accepted 28 October 2019)

We present a detailed guide to advanced collisionless fluid models that incorporate kinetic effects into the fluid framework, and that are much closer to the collisionless kinetic description than traditional magnetohydrodynamics. Such fluid models are directly applicable to modelling the turbulent evolution of a vast array of astrophysical plasmas, such as the solar corona and the solar wind, the interstellar medium, as well as accretion disks and galaxy clusters. The text can be viewed as a detailed guide to Landau fluid models and it is divided into two parts. Part 1 is dedicated to fluid models that are obtained by closing the fluid hierarchy with simple (non-Landau fluid) closures. Part 2 is dedicated to Landau fluid closures. Here in Part 1, we discuss the fluid model of Chew–Goldberger–Low (CGL) in great detail, together with fluid models that contain dispersive effects introduced by the Hall term and by the finite Larmor radius corrections to the pressure tensor. We consider dispersive effects introduced by the non-gyrotropic heat flux vectors. We investigate the parallel and oblique firehose instability, and show that the non-gyrotropic heat flux strongly influences the maximum growth rate of these instabilities. Furthermore, we discuss fluid models that contain evolution equations for the gyrotropic heat flux fluctuations and that are closed at the fourth-moment level by prescribing a specific form for the distribution function. For the bi-Maxwellian distribution, such a closure is known as

† Email address for correspondence: [peter.hunana@gmail.com](mailto:peter.hunana@gmail.com)

the ‘normal’ closure. We also discuss a fluid closure for the bi-kappa distribution. Finally, by considering one-dimensional Maxwellian fluid closures at higher-order moments, we show that such fluid models are always unstable. The last possible non Landau fluid closure is therefore the ‘normal’ closure, and beyond the fourth-order moment, Landau fluid closures are required.

**Key words:** astrophysical plasmas, space plasma physics, plasma waves

## Contents

<b>1</b>	<b>Introduction</b>	<b>4</b>
<b>2</b>	<b>Pressure tensor equation</b>	<b>10</b>
2.1	Pressure tensor decomposition . . . . .	15
2.2	Parallel pressure equation . . . . .	17
2.3	Perpendicular pressure equation . . . . .	19
2.4	On the paper by Chew <i>et al.</i> (1956) . . . . .	21
2.5	Physical meaning of CGL equations . . . . .	22
2.6	Exact equations of anisotropic multi-fluid models . . . . .	23
2.7	Conservation of energy . . . . .	26
<b>3</b>	<b>CGL description</b>	<b>28</b>
3.1	Normalized equations and definitions . . . . .	31
3.2	Classical CGL model with cold electrons . . . . .	33
3.3	Linearized CGL equations . . . . .	35
3.4	Alfvén mode . . . . .	37
3.5	Slow and fast modes . . . . .	38
3.6	Slow mode can become faster than Alfvén mode . . . . .	42
3.7	CGL parallel propagation . . . . .	43
3.8	CGL with $T_{\perp} = T_{\parallel}$ is not equivalent to MHD . . . . .	44
3.9	MHD velocity eigenvector . . . . .	45
3.9.1	MHD special case $V_A = C_s$ . . . . .	47
3.10	CGL velocity eigenvector . . . . .	48
3.10.1	CGL special case $\beta_{\parallel} = 2/(4 - T_{\perp}/T_{\parallel})$ . . . . .	51
3.11	Empirical models with polytropic indices $\gamma_{\parallel}$ , $\gamma_{\perp}$ . . . . .	52
3.12	CGL and radially expanding solar wind . . . . .	56
3.13	CGL model – summary . . . . .	57
<b>4</b>	<b>Hall-CGL model</b>	<b>58</b>
4.1	Hall-CGL model with cold electrons . . . . .	59
4.2	Parallel propagation . . . . .	60
4.2.1	Firehose instability . . . . .	63
4.2.2	Generalization to causal (correct) solutions . . . . .	65
4.2.3	Polarization . . . . .	67
4.2.4	Simplest ion-cyclotron resonances . . . . .	68
4.3	Hall-CGL dispersion relation for arbitrary propagation angle . . . . .	69

<b>5</b>	<b>FLR corrections to the pressure tensor</b>	<b>71</b>
5.1	Fully nonlinear FLR corrections . . . . .	71
5.1.1	Employing a non-dispersive induction equation . . . . .	77
5.2	Simplest FLR corrections (FLR1) . . . . .	78
5.3	Hall-CGL-FLR1 fluid model . . . . .	82
5.4	CGL-FLR1 fluid model (no Hall term) . . . . .	83
5.5	Second-order FLR corrections (FLR2) . . . . .	84
5.6	Hall-CGL-FLR2 fluid model . . . . .	86
5.7	FLR corrections with gyrotropic heat flux . . . . .	88
5.8	FLR corrections with non-gyrotropic heat flux vectors (FLR3) . . . . .	90
5.9	Hall-CGL-FLR3 fluid model . . . . .	92
<b>6</b>	<b>Parallel and oblique firehose instability</b>	<b>94</b>
6.1	Parallel propagation . . . . .	94
6.2	Oblique propagation . . . . .	98
6.3	Hellinger's contours for the Hall-CGL-FLR3 model . . . . .	98
<b>7</b>	<b>Heat flux tensor equation</b>	<b>100</b>
7.1	Heat flux tensor decomposition . . . . .	104
7.2	Parallel heat flux equation . . . . .	108
7.3	Perpendicular heat flux equation . . . . .	110
7.4	Scalar heat flux equations continued . . . . .	112
7.5	Heat flux equations with 'normal' closure . . . . .	113
<b>8</b>	<b>Fourth-order fluid moment</b>	<b>114</b>
8.1	Non-gyrotropic $\mathbf{r}^{\text{ng}}$ . . . . .	118
8.2	Bi-Maxwellian distribution . . . . .	119
8.3	Bi-Maxwellian non-gyrotropic $\mathbf{r}^{\text{ng}}$ contributions . . . . .	121
<b>9</b>	<b>Bi-Maxwellian fluid model, second-order CGL (CGL2)</b>	<b>122</b>
9.1	CGL2 dispersion relation . . . . .	124
9.2	Mirror instability . . . . .	125
9.3	Hall-CGL2 dispersion relation . . . . .	127
9.4	'Static' closure – generalized isothermal closure . . . . .	128
<b>10</b>	<b>Bi-kappa fluid model (BiKappa)</b>	<b>129</b>
10.1	Bi-kappa distribution function . . . . .	129
10.2	Bi-kappa fluid closure . . . . .	134
10.3	BiKappa dispersion relation . . . . .	138
10.4	Mirror instability . . . . .	138
10.5	Hall-BiKappa dispersion relation . . . . .	139
<b>11</b>	<b>Collisionless damping in fluid models – Landau fluid models</b>	<b>140</b>
<b>12</b>	<b>Evolution equation for <math>n</math>th-order fluid moment</b>	<b>143</b>
12.1	Hierarchy of moments in a 1-D geometry (electrostatic) . . . . .	147
12.2	Impossibility of going beyond CGL2 without Landau fluid closures . . . . .	150
<b>13</b>	<b>Conclusions</b>	<b>154</b>
	<b>Appendix A. Fourier transformations</b>	<b>162</b>

<b>Appendix B. Generalized vector (cross) product</b>	<b>163</b>
<b>Appendix C. MHD dispersion relation</b>	<b>164</b>
<b>Appendix D. Non-gyrotropic heat flux tensor <math>q^{\text{ng}}</math></b>	<b>165</b>
D.1 Non-gyrotropic heat flux vector $S_{\perp}^{\parallel}$	166
D.2 Conversion of other notations	170
D.3 Non-gyrotropic heat flux vector $S_{\perp}^{\perp}$	170
D.4 Second-order heat flux vectors	172
D.5 Contributions of $q^{\text{ng}}$ to various equations	175
D.5.1 Contributions to FLR pressure tensor	175
D.5.2 Contributions to scalar heat flux equations	176
D.5.3 Contributions to scalar pressure equations	177

## 1. Introduction

Fluid models are an extremely important tool in many areas of space physics, astrophysics and laboratory plasmas. Even though many physical systems studied in these fields are almost collisionless, where a proper kinetic description should be used, traditional fluid models with isotropic scalar pressures (temperatures), such as the usual magnetohydrodynamic description (MHD) and multi-fluid models based on this description, were extremely successful in modelling, interpreting or at least offering the first insight into many space physics phenomena. For example, it was indeed the simplified fluid approach that allowed Parker (1958b) to predict the existence of the solar wind, which was surprisingly several decades after the breakthrough discoveries in quantum mechanics and relativity. In recent years, observational studies re-sparked interest in the temperature anisotropy effects, that cannot be studied with the usual MHD fluid descriptions, and we anticipate that the interest will grow even further, once data from the Parker Solar Probe and the future Solar Orbiter missions are analysed.

The correct modelling and understanding of collisionless plasmas in a fluid framework concerns not only systems with plasma temperatures far from an isotropic state. It is sometimes forgotten that while the application of fluid theory to strongly collisional systems is intuitively obvious, the approach to collisionless (or weakly collisional) systems is not. From a linear perspective, the usual MHD description does not converge to the collisionless kinetic description even in the low-frequency long-wavelength limit (even when the kinetic distribution function is considered to be an isotropic Maxwellian), since in the absence of collisions the isotropic equation of state is never correct. Additionally, Landau damping never completely vanishes (even when electrons are hot). The situation is much more complicated from a nonlinear perspective where, on average, the effect of Landau damping might be balanced by processes called stochastic plasma echoes (Meyrand *et al.* 2019).

As a linear example, the ordering of the phase speeds in MHD is always slow, Alfvén, fast, i.e.  $v_s \leq v_A \leq v_f$ . In contrast, in collisionless systems with a sufficiently high plasma beta, the real phase speed of the slow mode can become faster than the Alfvén mode, so that the ordering can become Alfvén, slow, fast, i.e.  $v_A \leq v_s \leq v_f$ . Or in another words, the phase speeds of linear eigenmodes that are present in MHD do not hold in the collisionless (or weakly collisional) regime, and this effect exists even if the temperatures are isotropic. From a fluid perspective, the main reason for this discrepancy is that, in magnetized collisionless systems, the pressure fluctuations in the

directions parallel and perpendicular to the magnetic field lines are not equivalent, and cannot be described with a single scalar pressure equation. The pressure fluctuations have to be described with two different evolution equations for  $p_{\parallel}$  and  $p_{\perp}$ , even if the mean pressure values are equal ( $p_{\parallel}^{(0)} = p_{\perp}^{(0)}$ ) and no mean temperature anisotropy exists.

The simplest collisionless fluid description is the CGL fluid model – named after Chew, Goldberger & Low (1956) – and sometimes also referred to as collisionless MHD. The CGL dispersion relation is not equivalent to the MHD dispersion relation (even in the case  $p_{\parallel}^{(0)} = p_{\perp}^{(0)}$ ), since the pressure fluctuations still remain anisotropic and the evolution equations for  $p_{\parallel}$  and  $p_{\perp}$  remain different. In other words, in collisionless systems the distribution function is free to evolve from its initial state and to become anisotropic. By ‘forcibly’ prescribing only one scalar pressure, one effectively prescribes a high-collisionality regime, even if collisions are not prescribed explicitly. Formally, the MHD equations can indeed be derived from the collisionless Vlasov equation, i.e. with no explicit collisional operator. It is therefore often stated, that the MHD description is highly collisional implicitly. Moreover, as discussed for example by Kulsrud (1983), while in the presence of a magnetic field, transverse motions can in some circumstances be described by fluid-type equations, the determination of pressures as well as longitudinal motions *a priori* requires a kinetic description.

Here we focus on collisionless fluid models. Nevertheless, weak collisions can be incorporated easily, and calculations just yield additional terms on the right-hand sides of the parallel and perpendicular pressure and heat flux equations. For the simple Bhatnagar–Gross–Krook (BGK) collisional operator (Bhatnagar, Gross & Krook 1954), see for example Snyder, Hammett & Dorland (1997). A thorough review of anisotropic fluid models, including the collisional dynamics, was presented by Barakat & Schunk (1982). A very good discussion about collisionality of various astrophysical plasmas, such as the solar wind, interstellar medium, accretion disks and galaxy clusters, can be found for example in Schekochihin *et al.* (2009). Weakly collisional fluid models also seem to be applicable for modelling the upper solar photosphere and chromosphere, where a curious situation exists and the proton–proton collisional frequency is roughly equal to the proton cyclotron frequency – see for example figure 1 in Khomenko *et al.* (2014). We note that in this guide we use the definition of a ‘collisionless fluid model’ as being a fluid model that (i) is derived from the collisionless Vlasov equation with zero right-hand side; and (ii) that has two different pressure equations. Our definition therefore differs from an alternative view of for example Zank *et al.* (2014) and Zank (2014), where models with a scattering operator that reflects charged particle scattering by electromagnetic fluctuations on the right-hand side of the Vlasov equation are also viewed as collisionless.

Fluid modelling of collisionless plasmas is an extremely attractive subject, and an enormous amount of theoretical and numerical work was done in this field in the past. The manuscript presented here has no intention of being a proper review paper in this field. In our opinion, no satisfactory easy-to-read introductory paper exists about collisionless fluid models, and this manuscript attempts to fill such a spot. Instead of just stating the major results and discussing what was done in the past and by whom, we make a significant effort to present a (hopefully consistent) derivation of basic collisionless fluid models. On one hand, the presented calculations might be considered as too detailed in many places. On the other hand, it is exactly the relatively complicated algebra of collisionless fluid models, that makes the field difficult to enter for new researchers. The primary goal of this paper is to allow

researchers new to the subject to follow the algebra easily. Instead of spending years, an interested reader should be able to become comfortable with the basics of the subject in a matter of days, or perhaps a couple of weeks. The text is separated to two parts. Part 1 is dedicated to fluid models that are obtained by closing the fluid hierarchy with simple (non-Landau fluid) closures. Part 2 is dedicated to Landau fluid closures.

Here, in Part 1, §2, we start with the detailed derivation of the pressure tensor equation. The pressure tensor  $\mathbf{p}$  is then decomposed to its gyrotropic part  $\mathbf{p}^g = p_{\parallel} \hat{\mathbf{b}}\hat{\mathbf{b}} + p_{\perp}(\mathbf{I} - \hat{\mathbf{b}}\hat{\mathbf{b}})$  (also referred to as  $\mathbf{p}^{\text{CGL}}$ ), and non-gyrotropic part  $\mathbf{\Pi}$ , the latter usually called the finite Larmor radius (FLR) corrections to the pressure tensor, or the gyroviscous stress tensor. In the highly collisional decomposition  $\mathbf{p} = \mathbf{p}^g + \mathbf{\Pi}$ , see e.g. Braginskii (1958, 1965), the quantity  $\mathbf{\Pi}$  is called the stress tensor. The decomposition procedure yields rigorously exact (even though still not closed) evolution equations for  $p_{\parallel}$  and  $p_{\perp}$ , see e.g. Oraevskii, Chodura & Feneberg (1968), Passot & Sulem (2004) and Goswami, Passot & Sulem (2005). Importantly, at the leading order (by neglecting the  $\mathbf{\Pi}$  and also the non-gyrotropic heat flux contributions  $\mathbf{q}^{\text{ng}}$ ), the equations of Chew *et al.* (1956), hereafter referred to as CGL, are recovered. We discuss the paper by Chew *et al.* (1956) and point out that the paper derived the correct form of the pressure equations with the gyrotropic heat flux contributions included, however, the quantities  $q_n, q_s$  used in that paper are related to the usual  $q_{\parallel}, q_{\perp}$  by relations  $q_n = q_{\parallel} - 3q_{\perp}$  and  $q_s = q_{\perp}$ . By further neglecting the heat flux contributions, the pressure equations can be written in conservative form. The resulting pressure equations became known as the CGL equations, and they can be interpreted as the conservation laws for the first and second adiabatic invariants (Kulsrud 1983; Gurnett & Bhattacharjee 2005).

Furthermore, we discuss the general equations of collisionless multi-fluid models. The pressure equations are rigorously exact, even though the system is not closed, since the FLR pressure tensor and the entire heat flux tensor are not specified; for a quick look see (2.89)–(2.92). Rewriting the system to a form where the usual CGL conservation laws are on the left-hand side, and all other contributions on the right-hand side, nicely represents the complicated plasma heating processes that can be encountered, and that are responsible for the breaking of the adiabatic invariants; see (2.100), (2.101). We also discuss the conservation of energy. For the case of periodic boundary conditions, the total conservation of energy has a very illuminating form, that can be found for example in Yang *et al.* (2017a,b). That formulation is obtained by considering the total ‘internal’ energy for each particle species  $E_r^{\text{int}} = \frac{1}{2} \langle \text{Tr} \mathbf{p}_r \rangle$  (where the brackets represent integration over the entire spatial domain), which only reveals the total plasma heating. Here, we split the internal energy into its parallel and perpendicular components  $E_{\parallel r}^{\text{int}} = \frac{1}{2} \langle p_{\parallel r} \rangle$ ,  $E_{\perp r}^{\text{int}} = \langle p_{\perp r} \rangle$ , and formulate the total conservation of energy with the possibility of anisotropic plasma heating, see (2.113)–(2.116).

The CGL description is analysed in great detail in §3. We derive the CGL dispersion relation, and discuss properties of the slow, Alfvén and fast modes that are present. We verify many of the classical results of Abraham-Shrauner (1967), here written in a slightly more convenient notation by using the parallel plasma beta  $\beta_{\parallel}$ , and the temperature anisotropy ratio  $a_p = T_{\perp}/T_{\parallel}$ . Collisionless plasmas cannot reach arbitrarily large values of temperature anisotropy, and the linear CGL eigenmodes can indeed become unstable, with the associated instabilities referred to as the (parallel and oblique) firehose instability and the mirror instability. Similarly to MHD, the simple CGL description does not contain any dispersive effects and is technically



scale invariant, even though valid only on the largest scales. The instability thresholds present in the CGL model therefore can be referred to as ‘hard’ thresholds, i.e. thresholds that are obtained in the long-wavelength low-frequency limit. The firehose and mirror instabilities are believed to play a crucial role in solar wind dynamics (see e.g. Hellinger *et al.* 2006; Bale *et al.* 2009), and in figure 1, they are plotted in the usual  $\beta_{\parallel} - T_{\perp}/T_{\parallel}$  plane with logarithmic scales. While the firehose threshold matches the one obtained from kinetic theory, the CGL mirror threshold contains the well-known factor of 6 error for large  $\beta_{\parallel}$  values. The factor of 6 error can be interpreted as inadequacy of the adiabatic CGL closure in the very slow-dynamics context, such as the mirror instability. This is further addressed in § 9.4, where we discuss that the ‘static’ closure, which can be viewed as a generalization of isothermal closure in the presence of temperature anisotropy and variations of magnetic field strength, reproduces the correct mirror threshold (Constantinescu 2002; Chust & Belmont 2006; Passot, Ruban & Sulem 2006).

We discuss the core differences between MHD and CGL, which can be nicely summarized with the concept of adiabatic indices  $\gamma$ , related to the number of degrees of freedom  $i$  by  $\gamma = (i + 2)/i$ . While MHD is fully three-dimensional with  $\gamma = 5/3$ , the CGL can be viewed as composed of one- and two-dimensional dynamics with  $\gamma_{\parallel} = 3$  and  $\gamma_{\perp} = 2$ . We also address the velocity and magnetic field eigenvectors. Similarly to the velocity field eigenvector in MHD (see figure 2), which shows a ‘singular’ behaviour for strictly parallel propagation with  $V_A = C_s$ , the CGL eigenvector (figure 3) shows similar singularity for  $\beta_{\parallel} = 2/(4 - T_{\perp}/T_{\parallel})$ . We also briefly discuss fluid models with empirical ‘free’ polytropic indices  $\gamma_{\parallel}, \gamma_{\perp}$ , studied for example by Hau & Sonnerup (1993), Hau *et al.* (1993) and Abraham-Shrauner (1973).

In § 4, we introduce the simplest dispersive effects by including the Hall term into the induction equation, and study dispersion relations of the Hall-CGL model. We focus on the parallel firehose instability, and show that the instability is indeed associated with the whistler mode, see figures 4 and 5. We show that negative real frequencies  $\omega_r < 0$  have to be handled carefully, and that non-causal analytic solutions (4.32)–(4.35) have to be modified to the causal form (4.43)–(4.46). We briefly discuss the simplest ion-cyclotron resonances, and compare solutions of the Hall-CGL model with solutions of linear kinetic theory, see figure 6.

In § 5, we evaluate the FLR tensor  $\Pi$  at several levels of approximation. The evaluation of the FLR tensor is cumbersome, because the tensor is described by the pressure tensor equation implicitly. We first reproduce the fully nonlinear ‘inversion’ procedure on how to obtain  $\Pi$  from expression  $(\hat{\mathbf{b}} \times \Pi) + (\hat{\mathbf{b}} \times \Pi)^T$  that can be found for example in Hsu, Hazeltine & Morrison (1986), Passot & Sulem (2004) and Ramos (2005) as a brief note. Applying this inversion procedure to the pressure tensor equation evaluates the FLR corrections correctly along the magnetic field lines, but leads to an equation for  $\Pi$  that is still implicit. Nevertheless, evaluating the resulting equation at the leading order (technically first order in frequency and wavenumber), leads to an explicit expression for  $\Pi$ . We first recover the nonlinear result of Schekochihin *et al.* (2010), derived in that paper slightly differently, without using the inversion procedure. We point out that the result can be slightly simplified, and rearranging the expression yields two different useful forms for writing the nonlinear  $\Pi$ . Finally, by using the non-dispersive (MHD) induction equation, we obtain the nonlinear result of Ramos (2005) (see also Macmahon (1965)). For further evaluation of nonlinear FLR corrections to higher orders, an advanced reader is referred to Ramos (2005). We continue with the evaluation of the FLR tensor in the linear approximation, i.e. when the magnetic field lines are not too distorted. For the

first-order tensor (here called FLR1), we recover the classical result of Yajima (1966), which is notably different from the one provided by Oraevskii *et al.* (1968). In the isotropic case, the FLR1 tensor is consistent with the one extracted from the stress tensor of Braginskii (1965), if the collisional terms are ‘ignored’. It is noteworthy that a proper collisionless limit cannot be achieved from the stress tensor of Braginskii (1965), because expressions are proportional to  $\tau$  (and also  $1/\tau$ ), where  $\tau$  is time between two collisions ( $\tau \sim 1/\nu$  where  $\nu$  is the usual collisional frequency), so the collisionless limit  $\tau \rightarrow \infty$  does not work.

We consider the Hall-CGL-FLR1 fluid model, and we provide dispersion relation for generally oblique propagation, which can be also found in Hunana & Zank (2017). For higher-order FLR corrections, we only provide analytic dispersion relations for the parallel propagating whistler and ion-cyclotron modes, as well as the perpendicular fast mode. Nevertheless, we provide linearized, normalized and Fourier transformed equations written in the  $x$ - $z$  plane for all the fluid models. To obtain the dispersion relation for an oblique propagation, the reader is encouraged to use analytic software such as Maple or Mathematica, or to solve the system numerically. The second-order FLR corrections (FLR2) are here defined as containing the Hall term and the time derivative  $\partial \Pi / \partial t$ . We also consider FLR corrections with the non-gyrotropic heat flux vectors  $S_{\perp}^{\parallel}$ ,  $S_{\perp}^{\perp}$ , that are here defined as FLR3. The precision of various FLR corrections can be compared nicely by considering the perpendicular fast mode in the long-wavelength limit. The comparison is especially meaningful, when written in the notation of Del Sarto, Pegoraro & Tenerani (2017), see (5.119)–(5.121). We proceed by showing that the FLR3 corrections (with the second-order non-gyrotropic heat flux vectors) indeed recover the fully kinetic dispersion relation for the fast mode in the long-wavelength (low-frequency) limit, a result reported by Mikhailovskii & Smolyakov (1985). Instead of expanding the pressure tensor equation, one can derive very precise linear FLR corrections from linear kinetic theory, which is not addressed here, and the reader is referred to papers by Passot & Sulem (2007) and Sulem & Passot (2015) and references therein.

In §6, we investigate the parallel and oblique firehose instability. The FLR and Hall dispersive effects are crucial for the stabilization of the firehose instability at small scales, and a comprehensive discussion can be found in Hunana & Zank (2017). That paper was essentially extracted from this guide (with many figures that we do not republish here), and an interested reader who wants to focus on the firehose instability can find further information there. Nevertheless, here we briefly investigate improvements that can be made by considering the FLR2 and FLR3 corrections, see figures 7–11. Importantly, we show that the non-gyrotropic heat flux vectors in the FLR3 model, partially reproduce the large ‘bump’ in the imaginary phase speed (growth rate normalized to the wavenumber), when the plasma is close to the long-wavelength limit ‘hard’ firehose threshold, see figure 7. The firehose instability in a fluid formalism was also investigated by Wang & Hau (2003, 2010), Schekochihin *et al.* (2010) and Rosin *et al.* (2011) and references therein.

In §7, we derive the heat flux tensor equation. The heat flux tensor is then decomposed into its gyrotropic and non-gyrotropic parts,  $\mathbf{q} = \mathbf{q}^g + \mathbf{q}^{ng}$ . The procedure yields evolution equations for the gyrotropic parallel and perpendicular heat fluxes,  $q_{\parallel}$  and  $q_{\perp}$ , that contain the tensor of the fourth-order moment  $\mathbf{r}$ . It is emphasized that, if one wants to keep the non-gyrotropic  $\Pi$  contributions in the scalar heat flux equations, one needs to keep the non-gyrotropic contributions of the fourth-order moment  $\mathbf{r}^{ng}$  as well, since there are several possible cancellations even at the linear level. The non-gyrotropic heat flux  $\mathbf{q}^{ng}$  can be further decomposed to the non-gyrotropic heat



flux vectors  $\mathbf{S}_{\perp}^{\parallel}$ ,  $\mathbf{S}_{\perp}^{\perp}$  and tensor  $\sigma$ . The detailed algebra of the non-gyrotropic heat flux vectors  $\mathbf{S}_{\perp}^{\parallel}$ ,  $\mathbf{S}_{\perp}^{\perp}$ , i.e. how to express them through lower-order moments, is presented in appendix D. The first-order expressions are obtained at the nonlinear level and the second-order expressions at the linear level. We do not address how to decompose the tensor  $\sigma$  through lower-order moments. Such calculations require a complicated ‘inversion’ procedure for a third-rank tensor  $(\hat{\mathbf{b}} \times \sigma)^S$ , and an advanced reader is referred to Ramos (2005).

In § 8, we consider the fourth-order moment  $\mathbf{r}$  which is a tensor of fourth rank,  $r_{ijkl}$ . The moment is again decomposed to its gyrotropic and non-gyrotropic parts,  $\mathbf{r} = \mathbf{r}^g + \mathbf{r}^{ng}$  and the gyrotropic part has three scalar components,  $r_{\parallel\parallel}$ ,  $r_{\parallel\perp}$  and  $r_{\perp\perp}$ . We show that for a bi-Maxwellian distribution function, the gyrotropic components can indeed be evaluated as  $r_{\parallel\parallel} = 3p_{\parallel}^2/\rho$ ,  $r_{\parallel\perp} = p_{\parallel}p_{\perp}/\rho$  and  $r_{\perp\perp} = 2p_{\perp}^2/\rho$ . This constitutes a ‘normal’ closure, a name suggested by Chust & Belmont (2006). By using a similar procedure to the one provided by Grad (1949) for dilute gases, it is possible to express the non-gyrotropic bi-Maxwellian  $\mathbf{r}^{ng}$  through a combination of  $\mathbf{p}^g$  and  $\Pi$ , see e.g. Oraevskii *et al.* (1968).

In § 9, a dispersion relation of a fluid model closed with the bi-Maxwellian ‘normal’ closure is provided for generally oblique propagation, and we call this model second-order CGL (CGL2). We specifically focus on the mirror instability, since this simple fluid model (without any Landau damping) corrects the erroneous 1/6 factor in the ‘hard’ mirror threshold found in the basic CGL description, a result also reported by Dzhililov, Kuznetsov & Staude (2011). The mirror instability is not addressed to a higher level of sophistication in this guide. However, we note that capturing the mirror growth rate (when the threshold is crossed) sufficiently well requires Landau fluid models (Snyder *et al.* 1997) and the stabilization at small scales requires FLR corrections (Passot & Sulem 2007; Sulem & Passot 2015). We also provide the dispersion relation of the Hall-CGL2 fluid model. Finally, the CGL2 model can be simplified by considering slow-dynamics regime and constructing generalized isothermal closure that is called the ‘static’ closure, yielding the simplest fluid model that captures the correct mirror threshold (Constantinescu 2002; Chust & Belmont 2006; Passot *et al.* 2006).

In § 10, we consider the bi-kappa distribution function. We show that the closure at the fourth-order moment is constructed by  $r_{\parallel\parallel} = \alpha_{\kappa} 3p_{\parallel}^2/\rho$ ,  $r_{\parallel\perp} = \alpha_{\kappa} p_{\parallel}p_{\perp}/\rho$  and  $r_{\perp\perp} = \alpha_{\kappa} 2p_{\perp}^2/\rho$ , where the coefficient  $\alpha_{\kappa} = (\kappa - 3/2)/(\kappa - 5/2)$ , and the closure is valid for  $\kappa > 5/2$ . We call this closure and the associated fluid model ‘BiKappa’, and we discuss its dispersion relations. Even though the linear modes are generally different in this fluid model than in the CGL2 fluid model, we show that the ‘hard’ thresholds for the parallel and oblique firehose instability, and for the highly oblique mirror instability, are not affected by and are independent of the  $\kappa$  value. The Hall-BiKappa fluid dispersion relation is also provided. We also provide the first-order non-gyrotropic heat flux vectors  $\mathbf{S}_{\perp}^{\parallel}$ ,  $\mathbf{S}_{\perp}^{\perp}$ . We do not calculate the  $\mathbf{r}^{ng}$  for the bi-kappa distribution, and therefore we do not provide the second-order non-gyrotropic heat flux vectors.

In § 11, we discuss the core differences between the usual fluid models and kinetic theory. Namely, we discuss why the usual fluid models do not contain collisionless damping mechanisms, such as Landau damping, regardless of the order to which the fluid hierarchy is developed. The effect of Landau damping is present in the collisionless Vlasov equation, and the crucial difference between the usual fluid hierarchy and kinetic calculations just lies in the technique of how the Vlasov equation is integrated over the velocity space. We introduce preliminary ideas as to how the

Landau fluid closures will be constructed. For example, for closures performed at the fourth-order moment instead of the ‘normal’ closure, one needs to consider perturbations around this state, and prescribe  $r_{\parallel\parallel} = 3p_{\parallel}^2/\rho + \tilde{r}_{\parallel\parallel}$ ,  $r_{\parallel\perp} = p_{\parallel}p_{\perp}/\rho + \tilde{r}_{\parallel\perp}$  and  $r_{\perp\perp} = 2p_{\perp}^2/\rho + \tilde{r}_{\perp\perp}$ . The deviations  $\tilde{r}_{\parallel\parallel}$ ,  $\tilde{r}_{\parallel\perp}$ ,  $\tilde{r}_{\perp\perp}$  will be calculated from linear kinetic theory in Part 2, by performing Landau fluid closures.

In § 12, we derive the evolution equation for a general  $n$ th-order fluid moment (a tensor with  $3^n$  components). We then consider fluid models in the simplified one-dimensional (1-D) geometry that can be viewed as an electrostatic case (or as a propagation along the mean magnetic field), and that are closed at a general  $n$ th-order level by a Maxwellian fluid closure. A dispersion relation is obtained, which for  $n > 4$  always yields some solutions that are unstable. It is therefore concluded that the last non-Landau fluid closure is the ‘normal’ closure and that for  $n > 4$ , Landau fluid closures are required. This surprising result, first reported in Hunana *et al.* (2018), serves as motivation for Part 2, which is a detailed guide to Landau fluid closures.

## 2. Pressure tensor equation

Collisionless plasmas are described by the Vlasov equation, which in CGS units reads

$$\frac{\partial f_r}{\partial t} + \mathbf{v} \cdot \nabla f_r + \frac{q_r}{m_r} \left( \mathbf{E} + \frac{1}{c} \mathbf{v} \times \mathbf{B} \right) \cdot \nabla_{\mathbf{v}} f_r = 0, \quad (2.1)$$

and which describes how a distribution function  $f_r(\mathbf{x}, \mathbf{v}, t)$  evolves in time. The  $r$  is the index of species and  $r = p$  for protons,  $r = e$  for electrons, etc. The  $q_r$  is the particle charge,  $m_r$  the particle mass,  $c$  the speed of light,  $\mathbf{E}$  the electric field vector and  $\mathbf{B}$  the magnetic field vector. The species index  $r$  can sometimes be confusing in lengthy tensor calculations with multiple indices and for clarity we will often drop it and reintroduce it when required. To derive the fluid equations, we need to integrate (perform an averaging) at each spatial point over the ‘kinetic’ velocity  $\mathbf{v}$ . It is important to realize that the distribution function just describes the probability of finding a particle with velocity  $\mathbf{v}$  at the position  $\mathbf{x}, t$  and that the ‘kinetic’ velocity  $\mathbf{v}$  entering the distribution function  $f(\mathbf{x}, \mathbf{v}, t)$  is a completely independent variable from  $\mathbf{x}, t$ , i.e.

$$\frac{\partial v_i}{\partial t} = 0; \quad \frac{\partial v_i}{\partial x_j} = 0. \quad (2.2)$$

Also, the magnetic and electric fields  $\mathbf{B}(\mathbf{x}, t)$ ,  $\mathbf{E}(\mathbf{x}, t)$  are macroscopic quantities that do not depend on  $\mathbf{v}$  and can be moved outside of velocity integrals over  $d^3v$ , or in another words  $\partial B_i / \partial v_j = 0$  and  $\partial E_i / \partial v_j = 0$ . The definitions of the fluid moments are

$$n = \int f d^3v; \quad (2.3)$$

$$n\mathbf{u} = \int \mathbf{v} f d^3v; \quad (2.4)$$

$$\mathbf{p} = m \int (\mathbf{v} - \mathbf{u})(\mathbf{v} - \mathbf{u}) f d^3v; \quad (2.5)$$

$$\mathbf{q} = m \int (\mathbf{v} - \mathbf{u})(\mathbf{v} - \mathbf{u})(\mathbf{v} - \mathbf{u}) f d^3v; \quad (2.6)$$

$$\mathbf{r} = m \int (\mathbf{v} - \mathbf{u})(\mathbf{v} - \mathbf{u})(\mathbf{v} - \mathbf{u})(\mathbf{v} - \mathbf{u}) f d^3v, \quad (2.7)$$

where we have omitted the tensor product notation that is sometimes written down explicitly as  $\mathbf{uu} = \mathbf{u} \otimes \mathbf{u}$  and in the index notation  $(\mathbf{uu})_{ij} = u_i u_j$ . If two vectors or tensors are next to each other without an operator between them, a tensor product is always assumed. The number density  $n$  is a scalar, the fluid velocity  $\mathbf{u}$  is a vector, the pressure tensor  $\mathbf{p}$  is a tensor of second rank ( $3 \times 3$  matrix), the heat flux tensor  $\mathbf{q}$  is a tensor of third rank (with  $3 \times 3 \times 3$  components), the tensor  $\mathbf{r}$  is of fourth rank (with  $3 \times 3 \times 3 \times 3$  components), etc. Directly from the definitions, it is obvious that all fluid tensors must be symmetric in all of their indices, i.e.  $p_{ij} = p_{ji}$ ,  $q_{ijk} = q_{ikj} = \dots$ . The fluctuating velocity is defined as

$$\mathbf{c} = \mathbf{v} - \mathbf{u}, \quad (2.8)$$

and should not be confused with the speed of light  $c$ .

The second important concept that is used in deriving the fluid hierarchy, is the use of the usual Gauss–Ostrogradsky (divergence) theorem. The divergence theorem is used in velocity space and, written in a form that is typically encountered when calculating the fluid hierarchy, it reads

$$\int_V \nabla_v \cdot (f \mathbf{A}) d^3v = \int_S f \mathbf{A} \cdot d\mathbf{S}. \quad (2.9)$$

The  $\mathbf{A}$  is a general  $n$ th-order tensor,  $f$  is a distribution function, the left-hand side is a 3-D integral calculated over the entire velocity volume  $V$  and the right-hand side is a surface integral calculated over a boundary of that volume  $d\mathbf{S} = \hat{\mathbf{n}} dS$ , where  $\hat{\mathbf{n}}$  is a unit normal vector to the local surface area pointing outwards. When such an integral is encountered in the fluid hierarchy, the result is always assumed to be zero. The volume integrals are from  $v = -\infty$  to  $v = \infty$  in each velocity component and the integration on the right-hand side of (2.9) is therefore performed over the velocity surface area at infinity. The necessary (but technically not sufficient) condition for the integral to be negligible is

$$\lim_{v \rightarrow \infty} f(v) = 0. \quad (2.10)$$

Since the area  $dS \sim v^2 dv$ , the sufficient condition for the integral to vanish, can be estimated more precisely as

$$\lim_{v \rightarrow \infty} f(v) A(v) v^2 = 0. \quad (2.11)$$

When calculating the fluid hierarchy, the encountered expressions are always  $A(v) \sim v^n$ , where  $n$  is a positive integer. Then, for example, for a Maxwellian distribution  $f(v) \sim e^{-v^2}$  the limit (2.11) is always zero for all  $n$ . Even for a slower converging distribution  $f(v) \sim e^{-|v|}$ , the limit is always zero. For distribution functions proportional to a power law, for example  $f(v) \sim (v^2)^{-(\kappa+1)}$  as for a kappa distribution, the situation is more restrictive, with some minimum required values of  $\kappa$ . When we will calculate the  $n$ th-order fluid moment (§ 12), we will see that two surface integrals are encountered, one with  $A(v) \sim v^n$  and one with  $A(v) \sim v^{n+1}$ . Therefore, for a kappa distribution, the strict condition for  $\lim_{v \rightarrow \infty} (v^2)^{-(\kappa+1)} v^{n+1} v^2 = 0$  yields the requirement  $\kappa > (n+1)/2$ . For example, many fluid models discussed in this guide are closed at the fourth-order moment  $\mathbf{r}$ , so  $n = 4$ , which implies the requirement  $\kappa > 5/2$ . At first sight, the required limit (2.11) can be considered only a technical

mathematical detail since, physically, one can argue that particles with enormously large energies will not be measured/observed, and some physical mechanism that is responsible for a cutoff of the distribution function at finite energies can be usually assumed. Or in another words, even observational studies that fit the data with a kappa distribution do not assume that the fit is valid all the way up to  $v \rightarrow \infty$ , and usually a cutoff is implicitly assumed. However, once we calculate the fourth-order moments for a kappa distribution, we will see that, for example  $r_{\parallel\parallel} = \alpha_\kappa 3p_{\parallel}^2/\rho$ , where  $\alpha_\kappa = (\kappa - 3/2)/(\kappa - 5/2)$ , so the restriction  $\kappa > 5/2$  returns, and has to be applied regardless. When the limit (2.11) is satisfied, the neglect of the surface integrals (2.9) is therefore based on solid theoretical principles, and it is not an approximate or an *ad hoc* choice, the surface integrals are really zero. Nevertheless, for complete clarity in the upcoming fluid hierarchy calculations, we will differentiate between expressions that are zero exactly, and between the surface integrals (2.9) that are zero asymptotically, by using  $=0$  and  $\rightarrow 0$ .

We start directly with the derivation of the pressure tensor equation, since a detailed derivation of the density and momentum equations can be found in many books. To derive the pressure tensor equation, it is possible to multiply the Vlasov equation by  $mc_i c_j$  or  $mv_i v_j$  (another possibility is  $mc_i v_j$ ). Here, we will use the first choice, which is slightly easier to present in detail, because the second choice requires the use of density and momentum equations to cancel several terms. It is useful to derive the following identities

$$\int \mathbf{c} f \, d^3v = \int (\mathbf{v} - \mathbf{u}) f \, d^3v = \int \mathbf{v} f \, d^3v - \mathbf{u} \int f \, d^3v = n\mathbf{u} - \mathbf{u}n = 0; \quad (2.12)$$

$$\begin{aligned} m \int \mathbf{c} \mathbf{v} f \, d^3v &= m \int (\mathbf{v} - \mathbf{u})(\mathbf{v} - \mathbf{u} + \mathbf{u}) f \, d^3v \\ &= m \int (\mathbf{v} - \mathbf{u})(\mathbf{v} - \mathbf{u}) f \, d^3v + m\mathbf{u} \underbrace{\int (\mathbf{v} - \mathbf{u}) f \, d^3v}_{=0} = \mathbf{p}; \end{aligned} \quad (2.13)$$

$$m \int \mathbf{c} \mathbf{c} \mathbf{v} f \, d^3v = m \int \mathbf{c} \mathbf{c} (\mathbf{v} - \mathbf{u} + \mathbf{u}) f \, d^3v = \mathbf{q} + \mathbf{p} \mathbf{u}. \quad (2.14)$$

We write down the derivatives with respect to time  $\partial/\partial t$  and velocity  $\partial/\partial v_i$  explicitly, but we abbreviate the derivative with respect to spatial coordinates as  $\partial/\partial x_i \equiv \partial_i$ . We will need

$$\frac{\partial}{\partial t} c_i = \frac{\partial}{\partial t} (v_i - u_i) = -\frac{\partial}{\partial t} u_i; \quad (2.15)$$

$$\frac{\partial}{\partial t} (c_i c_j) = -c_i \frac{\partial u_j}{\partial t} - c_j \frac{\partial u_i}{\partial t}; \quad (2.16)$$

$$\partial_k (c_i c_j) = -c_i \partial_k u_j - c_j \partial_k u_i; \quad (2.17)$$

$$\frac{\partial}{\partial v_k} c_i = \frac{\partial}{\partial v_k} (v_i - u_i) = \frac{\partial v_i}{\partial v_k} = \delta_{ik}; \quad (2.18)$$

$$\frac{\partial}{\partial v_k} (c_i c_j) = c_i \delta_{jk} + c_j \delta_{ik}; \quad (2.19)$$

$$\begin{aligned}
\frac{\partial}{\partial v_k} [c_i c_j (\mathbf{v} \times \mathbf{B})_k] &= \frac{\partial}{\partial v_k} (c_i c_j) (\mathbf{v} \times \mathbf{B})_k + c_i c_j \frac{\partial}{\partial v_k} (\epsilon_{klm} v_l B_m) \\
&= (\delta_{ik} c_j + \delta_{jk} c_i) (\mathbf{v} \times \mathbf{B})_k + c_i c_j \underbrace{\epsilon_{klm} \delta_{lk}}_{=0} B_m \\
&= c_j (\mathbf{v} \times \mathbf{B})_i + c_i (\mathbf{v} \times \mathbf{B})_j,
\end{aligned} \tag{2.20}$$

where in the last identity we used the result  $\epsilon_{klm} \delta_{lk} = 0$  since the Levi-Civita tensor  $\epsilon_{ijk}$  is antisymmetric and the Kronecker  $\delta_{ij}$  is a symmetric tensor. Integrating the first term of the Vlasov equation yields

$$\begin{aligned}
\textcircled{1} &= m \int c_i c_j \frac{\partial f}{\partial t} d^3 v = \frac{\partial}{\partial t} \underbrace{\left( m \int c_i c_j f d^3 v \right)}_{=p_{ij}} - m \int f \frac{\partial}{\partial t} (c_i c_j) d^3 v \\
&= \frac{\partial}{\partial t} p_{ij} + m \underbrace{\frac{\partial u_j}{\partial t} \int f c_i d^3 v}_{=0} + m \underbrace{\frac{\partial u_i}{\partial t} \int f c_j d^3 v}_{=0} \\
&= \frac{\partial}{\partial t} p_{ij}.
\end{aligned} \tag{2.21}$$

The second term of the Vlasov equation yields

$$\begin{aligned}
\textcircled{2} &= m \int c_i c_j \underbrace{\mathbf{v} \cdot \nabla f}_{=v_k \partial_k} d^3 v = \partial_k \underbrace{\left( m \int c_i c_j v_k f d^3 v \right)}_{=q_{ijk} + p_{ij} u_k} - m \int v_k f \partial_k (c_i c_j) d^3 v \\
&= \partial_k (q_{ijk} + p_{ij} u_k) + (\partial_k u_j) \underbrace{m \int c_i v_k f d^3 v}_{=p_{ik}} \\
&\quad + (\partial_k u_i) \underbrace{m \int c_j v_k f d^3 v}_{=p_{jk}} = \partial_k (q_{ijk} + p_{ij} u_k) + p_{ik} \partial_k u_j + p_{jk} \partial_k u_i \\
&= \partial_k (q_{kij} + u_k p_{ij}) + p_{ik} \partial_k u_j + p_{jk} \partial_k u_i \\
&= [\nabla \cdot (\mathbf{q} + \mathbf{u}\mathbf{p}) + \mathbf{p} \cdot \nabla \mathbf{u} + (\mathbf{p} \cdot \nabla \mathbf{u})^T]_{ij}.
\end{aligned} \tag{2.22}$$

To go back and forth between the index notation and the vector notation in a fully consistent manner, it is important to establish some conventions that were not required in simple fluid models. One important convention that is typically used is that the divergence of a tensor is meant to be through its first component, i.e. for a general tensor  $X_{ijk\dots n}$ , the divergence  $\nabla \cdot \mathbf{X}$  means  $\partial_i X_{ijk\dots n}$ . This convention is the reason why in the above expression we used that  $q_{ijk} = q_{kij}$  and also the ordering of  $\mathbf{u}$  and  $\mathbf{p}$  inside  $\nabla \cdot (\mathbf{u}\mathbf{p})$  reflects that in the index notation we have  $\partial_k (u_k p_{ij})$ .

The third term of the Vlasov equation calculates

$$\begin{aligned}
\textcircled{3} &= q \int c_i c_j \mathbf{E} \cdot \nabla_v f d^3 v = q \int c_i c_j E_k \frac{\partial}{\partial v_k} f d^3 v \\
&= q E_k \underbrace{\int \frac{\partial}{\partial v_k} (c_i c_j f) d^3 v}_{\rightarrow 0} - q E_k \int \frac{\partial}{\partial v_k} (c_i c_j) f d^3 v
\end{aligned}$$

$$\begin{aligned}
&= -qE_k \int \frac{\partial}{\partial v_k} (c_i c_j) f \, d^3 v = -qE_k \int (\delta_{ik} c_j + \delta_{jk} c_i) f \, d^3 v \\
&= -qE_i \underbrace{\int c_j f \, d^3 v}_{=0} - qE_j \underbrace{\int c_i f \, d^3 v}_{=0} = 0.
\end{aligned} \tag{2.23}$$

Finally, the fourth term of the Vlasov equation calculates

$$\begin{aligned}
\textcircled{4} &= \frac{q}{c} \int c_i c_j (\mathbf{v} \times \mathbf{B}) \cdot \nabla_v f \, d^3 v = \frac{q}{c} \int c_i c_j (\mathbf{v} \times \mathbf{B})_k \frac{\partial f}{\partial v_k} \, d^3 v \\
&= \frac{q}{c} \underbrace{\int \frac{\partial}{\partial v_k} [c_i c_j (\mathbf{v} \times \mathbf{B})_k f] \, d^3 v}_{\rightarrow 0} - \frac{q}{c} \int \frac{\partial}{\partial v_k} [c_i c_j (\mathbf{v} \times \mathbf{B})_k] f \, d^3 v \\
&= -\frac{q}{c} \int [c_j (\mathbf{v} \times \mathbf{B})_i + c_i (\mathbf{v} \times \mathbf{B})_j] f \, d^3 v = -\frac{q}{c} \int c_j \epsilon_{ikl} v_k B_l f \, d^3 v - \frac{q}{c} \int c_i \epsilon_{jkl} v_k B_l f \, d^3 v \\
&= -\frac{q}{c} \underbrace{\epsilon_{ikl} B_l \int c_j v_k f \, d^3 v}_{=(1/m)p_{jk}} - \frac{q}{c} \underbrace{\epsilon_{jkl} B_l \int c_i v_k f \, d^3 v}_{=(1/m)p_{ik}} = -\frac{q}{mc} [\epsilon_{ikl} B_l p_{jk} + \epsilon_{jkl} B_l p_{ik}] \\
&= \frac{q}{mc} [\epsilon_{ikl} B_l p_{kj} + \epsilon_{jkl} B_l p_{ki}] = \frac{q}{mc} [(\mathbf{B} \times \mathbf{p})_{ij} + (\mathbf{B} \times \mathbf{p})_{ji}] = \frac{q}{mc} [\mathbf{B} \times \mathbf{p} + (\mathbf{B} \times \mathbf{p})^T]_{ij}.
\end{aligned} \tag{2.24}$$

In the expression above, one encounters a vector product of a vector with a matrix,  $\mathbf{B} \times \mathbf{p}$ . This operator might appear unusual at first, but it is just a generalization of a vector product of two vectors, it is defined as  $(\mathbf{B} \times \mathbf{p})_{ij} = \epsilon_{ikl} B_k p_{lj}$  and the result is a matrix. Similarly, a vector product between a vector and a tensor of third rank is defined as  $(\mathbf{B} \times \mathbf{q})_{ijk} = \epsilon_{ilm} B_l q_{mjk}$  and for a tensor of  $n$ th rank  $(\mathbf{B} \times \mathbf{X})_{i\dots jk} = \epsilon_{ilm} B_l X_{m\dots jk}$  (i.e. the second index in  $\epsilon$  is the vector  $\mathbf{B}$  index and the third index in  $\epsilon$  is the first index of tensor  $\mathbf{X}$ ). However, a definition of a vector product where a matrix is applied on a vector reads  $(\mathbf{p} \times \mathbf{B})_{ij} = \epsilon_{jkl} p_{ik} B_l$ , and in general  $\mathbf{p} \times \mathbf{B} = -(\mathbf{B} \times \mathbf{p}^T)^T$ , but since the  $\mathbf{p}$  is here symmetric,  $\mathbf{p} \times \mathbf{B} = -(\mathbf{B} \times \mathbf{p})^T$ . The vector product (cross-product) is addressed further in appendix B. The term (2.24) is sometimes rewritten as

$$\mathbf{B} \times \mathbf{p} + (\mathbf{B} \times \mathbf{p})^T = \mathbf{B} \times \mathbf{p} - \mathbf{p} \times \mathbf{B}. \tag{2.25}$$

Combining all the results together  $\textcircled{1} + \textcircled{2} + \textcircled{3} + \textcircled{4} = 0$ , yields the entire pressure tensor equation as

$$\frac{\partial \mathbf{p}}{\partial t} + \nabla \cdot (\mathbf{u} \mathbf{p} + \mathbf{q}) + \mathbf{p} \cdot \nabla \mathbf{u} + (\mathbf{p} \cdot \nabla \mathbf{u})^T + \frac{q}{mc} [\mathbf{B} \times \mathbf{p} + (\mathbf{B} \times \mathbf{p})^T] = 0. \tag{2.26}$$

One notices that, for example, the pressure tensor equation of Kulsrud (1983), equation (57), contains a minus sign typo in the last term, which should be written as (2.25). That it is indeed a typo and not a definition of  $\mathbf{p} \times \mathbf{B}$  is evident from his subsequent equation (59). It is useful to introduce a symmetric operator that acts on a matrix  $\mathbf{A}$  according to

$$\mathbf{A}^S \equiv \mathbf{A} + \mathbf{A}^T; \quad A_{ij}^S \equiv A_{ij} + A_{ji}, \tag{2.27}$$

which yields a more compact form of the pressure tensor equation

$$\frac{\partial \mathbf{p}}{\partial t} + \nabla \cdot (\mathbf{u} \mathbf{p} + \mathbf{q}) + \left[ \mathbf{p} \cdot \nabla \mathbf{u} + \frac{q}{mc} \mathbf{B} \times \mathbf{p} \right]^S = 0. \tag{2.28}$$



## 2.1. Pressure tensor decomposition

By introducing a unit vector in the direction of the local magnetic field  $\hat{\mathbf{b}} = \mathbf{B}/|\mathbf{B}|$  and the gyrofrequency  $\Omega = qB_0/(mc)$ , the last term in the pressure tensor equation (2.26) can be rewritten as

$$\Omega \frac{|\mathbf{B}|}{B_0} [\hat{\mathbf{b}} \times \mathbf{p} + (\hat{\mathbf{b}} \times \mathbf{p})^T]. \quad (2.29)$$

At very low frequencies  $\omega \ll \Omega$  (and at very long spatial scales) this term will dominate and must be equal to zero. The simplest possibility that makes this term to be equal to zero is  $\mathbf{p}^{\text{iso}} = p\mathbf{I}$ , where  $p$  is the scalar pressure and  $\mathbf{I}$  is the unit matrix. In index notation  $p_{ij}^{\text{iso}} = p\delta_{ij}$ , and we can verify that

$$(\hat{\mathbf{b}} \times \mathbf{p}^{\text{iso}})_{ij} = \epsilon_{ikl} \hat{b}_k p_{lj}^{\text{iso}} = \epsilon_{ikl} \hat{b}_k p \delta_{lj} = p \epsilon_{ikj} \hat{b}_k = -p \epsilon_{ijk} \hat{b}_k; \quad (2.30)$$

$$[\hat{\mathbf{b}} \times \mathbf{p}^{\text{iso}} + (\hat{\mathbf{b}} \times \mathbf{p}^{\text{iso}})^T]_{ij} = -p \epsilon_{ijk} \hat{b}_k - p \epsilon_{jik} \hat{b}_k = -p \hat{b}_k (\epsilon_{ijk} + \epsilon_{jik}) = 0. \quad (2.31)$$

There is however a much more general solution that makes the term (2.29) equal to zero, which is

$$\mathbf{p}^g = p_{\parallel} \hat{\mathbf{b}} \hat{\mathbf{b}} + p_{\perp} (\mathbf{I} - \hat{\mathbf{b}} \hat{\mathbf{b}}), \quad (2.32)$$

where the ‘g’ stands for gyrotropic. For quick analytic calculations it is sometimes easier to use  $\mathbf{p}^g = (p_{\parallel} - p_{\perp}) \hat{\mathbf{b}} \hat{\mathbf{b}} + p_{\perp} \mathbf{I}$  since we have to deal with  $\hat{\mathbf{b}} \hat{\mathbf{b}}$  only once instead of twice. We need to verify that the term (2.29) indeed disappears,

$$\begin{aligned} (\hat{\mathbf{b}} \times \mathbf{p}^g)_{ij} &= \epsilon_{ikl} \hat{b}_k p_{lj}^g = \epsilon_{ikl} \hat{b}_k ((p_{\parallel} - p_{\perp}) \hat{b}_l \hat{b}_j + p_{\perp} \delta_{lj}) \\ &= (p_{\parallel} - p_{\perp}) \hat{b}_j \underbrace{\epsilon_{ikl} \hat{b}_k \hat{b}_l}_{=0} + p_{\perp} \epsilon_{ikj} \hat{b}_k = -p_{\perp} \epsilon_{ijk} \hat{b}_k; \end{aligned} \quad (2.33)$$

$$\Rightarrow [\hat{\mathbf{b}} \times \mathbf{p}^g + (\hat{\mathbf{b}} \times \mathbf{p}^g)^T]_{ij} = -p_{\perp} \hat{b}_k (\epsilon_{ijk} + \epsilon_{jik}) = 0. \quad (2.34)$$

It is easy to see that the parallel and perpendicular pressures can be extracted from the gyrotropic pressure tensor matrix by performing double contractions according to

$$p_{\parallel} = \mathbf{p}^g : \hat{\mathbf{b}} \hat{\mathbf{b}}, \quad \text{and} \quad p_{\perp} = \mathbf{p}^g : (\mathbf{I} - \hat{\mathbf{b}} \hat{\mathbf{b}})/2. \quad (2.35)$$

The double contraction (the double dot product, represented by the colon) is a very useful operator that is frequently encountered in higher-order fluid models. For two matrices it is defined as  $\mathbf{A} : \mathbf{B} = A_{ij} B_{ij}$  and yields a scalar. Sometimes, the double contraction is defined as  $\mathbf{A} : \mathbf{B} = A_{ij} B_{ji}$ , which for symmetric matrices is equivalent to the previous definition. As a reminder, the usual matrix product of two matrices is  $(\mathbf{A} \cdot \mathbf{B})_{ij} = A_{ik} B_{kj}$  and yields a matrix. It is useful to write down the following identities that involve the double contraction,

$$\hat{\mathbf{b}} \hat{\mathbf{b}} : \hat{\mathbf{b}} \hat{\mathbf{b}} = 1; \quad (2.36)$$

$$\mathbf{I} : \hat{\mathbf{b}} \hat{\mathbf{b}} = 1; \quad (2.37)$$

$$(\mathbf{I} - \hat{\mathbf{b}} \hat{\mathbf{b}}) : \hat{\mathbf{b}} \hat{\mathbf{b}} = 0; \quad (2.38)$$

$$\mathbf{I} : \mathbf{I} = 3; \quad (2.39)$$

$$(\mathbf{I} - \hat{\mathbf{b}} \hat{\mathbf{b}}) : (\mathbf{I} - \hat{\mathbf{b}} \hat{\mathbf{b}}) = 2. \quad (2.40)$$

Also interestingly, the trace of a matrix  $\mathbf{A}$  can be expressed through the double contraction with  $\mathbf{I}$ , i.e.  $\mathbf{A} : \mathbf{I} = A_{ij} \delta_{ij} = A_{ii} = \text{Tr}(\mathbf{A})$ . The pressure decomposition (2.35)

can now be verified easily. Even more revealing is to apply the decomposition (2.35) directly to the definition of the entire pressure tensor (2.5), which yields

$$\mathbf{p} : \hat{\mathbf{b}}\hat{\mathbf{b}} = m \int \mathbf{c} \mathbf{c} f d^3 v : \hat{\mathbf{b}}\hat{\mathbf{b}} = m \int (\mathbf{v} \cdot \hat{\mathbf{b}} - \mathbf{u} \cdot \hat{\mathbf{b}})^2 f d^3 v = m \int (v_{\parallel} - u_{\parallel})^2 f d^3 v \equiv p_{\parallel}; \quad (2.41)$$

$$\mathbf{p} : (\mathbf{I} - \hat{\mathbf{b}}\hat{\mathbf{b}})/2 = \frac{m}{2} \int |\mathbf{v}_{\perp} - \mathbf{u}_{\perp}|^2 f d^3 v \equiv p_{\perp}, \quad (2.42)$$

where we have used the fact that the magnitude of the 3-D velocity vector can be decomposed as  $|\mathbf{v} - \mathbf{u}|^2 = |\mathbf{v}_{\perp} - \mathbf{u}_{\perp}|^2 + (v_{\parallel} - u_{\parallel})^2$ , or equivalently  $|\mathbf{c}|^2 = |\mathbf{c}_{\perp}|^2 + c_{\parallel}^2$ .

A distribution function is called isotropic when the direction of the velocity vector  $\mathbf{v}$  does not matter, and the distribution function depends only on magnitude  $|\mathbf{v}|$ , as for example the Maxwellian distribution  $e^{-|\mathbf{v}|^2}$ . A distribution function is called gyrotropic when the direction of the perpendicular velocity vector  $\mathbf{v}_{\perp}$  does not matter, and the function depends only on  $|\mathbf{v}_{\perp}|$ , as for example the bi-Maxwellian distribution  $e^{-\alpha_{\parallel} v_{\parallel}^2} e^{-\alpha_{\perp} |\mathbf{v}_{\perp}|^2}$ . In another words, the gyrotropic distribution function is isotropic only in its transverse velocity components. The same vocabulary is used for the fluid moments, and a fluid moment is called gyrotropic when it involves integrals over  $|\mathbf{c}_{\perp}|^{2i}$ ,  $i = 0, 1, 2, 3, \dots$ . For the pressure tensor, there are only two possibilities, represented by the parallel and perpendicular pressures (2.41), (2.42).

The part of the pressure tensor extracted by the decomposition (2.35) is therefore called the gyrotropic pressure  $\mathbf{p}^g$ , sometimes also called the CGL pressure  $\mathbf{p}^{\text{CGL}}$ . The gyrotropic approximation is sufficient at very long spatial scales, where the Larmor radius (= the gyroradius) of particles gyrating around the magnetic field is small, and the non-gyrotropic contributions become negligible. However, at sufficiently small spatial scales comparable to the gyroradius, the non-gyrotropic pressure contributions become significant and we represent these through a tensor  $\mathbf{\Pi}$ , that is called the finite Larmor radius corrections to the (gyrotropic) pressure tensor, or sometimes non-gyrotropic or gyroviscous stress tensor. The entire pressure tensor is thus decomposed according to

$$\mathbf{p} = p_{\parallel} \hat{\mathbf{b}}\hat{\mathbf{b}} + p_{\perp} (\mathbf{I} - \hat{\mathbf{b}}\hat{\mathbf{b}}) + \mathbf{\Pi}, \quad (2.43)$$

and for clarity we write down the decomposition explicitly in matrix notation

$$\begin{aligned} \mathbf{p} = & p_{\parallel} \begin{pmatrix} \hat{b}_x \hat{b}_x & \hat{b}_x \hat{b}_y & \hat{b}_x \hat{b}_z \\ \hat{b}_y \hat{b}_x & \hat{b}_y \hat{b}_y & \hat{b}_y \hat{b}_z \\ \hat{b}_z \hat{b}_x & \hat{b}_z \hat{b}_y & \hat{b}_z \hat{b}_z \end{pmatrix} + p_{\perp} \begin{pmatrix} 1 - \hat{b}_x \hat{b}_x & -\hat{b}_x \hat{b}_y & -\hat{b}_x \hat{b}_z \\ -\hat{b}_y \hat{b}_x & 1 - \hat{b}_y \hat{b}_y & -\hat{b}_y \hat{b}_z \\ -\hat{b}_z \hat{b}_x & -\hat{b}_z \hat{b}_y & 1 - \hat{b}_z \hat{b}_z \end{pmatrix} \\ & + \begin{pmatrix} \Pi_{xx} & \Pi_{xy} & \Pi_{xz} \\ \Pi_{yx} & \Pi_{yy} & \Pi_{yz} \\ \Pi_{zx} & \Pi_{zy} & \Pi_{zz} \end{pmatrix}. \end{aligned} \quad (2.44)$$

In the momentum equation, the pressure tensor enters through its divergence, which is useful to break down into its components. Since  $\nabla \cdot (\hat{\mathbf{b}}\hat{\mathbf{b}}) = \hat{\mathbf{b}}(\nabla \cdot \hat{\mathbf{b}}) + (\hat{\mathbf{b}} \cdot \nabla) \hat{\mathbf{b}}$ , and  $\nabla \cdot (p_{\perp} \mathbf{I}) = \nabla p_{\perp}$ , the divergence of the pressure tensor calculates as

$$\begin{aligned} \nabla \cdot \mathbf{p} &= \nabla \cdot (p_{\perp} \mathbf{I}) + \nabla \cdot [(p_{\parallel} - p_{\perp}) \hat{\mathbf{b}}\hat{\mathbf{b}}] + \nabla \cdot \mathbf{\Pi} \\ &= \nabla p_{\perp} + (p_{\parallel} - p_{\perp}) [\hat{\mathbf{b}}(\nabla \cdot \hat{\mathbf{b}}) + (\hat{\mathbf{b}} \cdot \nabla) \hat{\mathbf{b}}] + \hat{\mathbf{b}}(\hat{\mathbf{b}} \cdot \nabla)(p_{\parallel} - p_{\perp}) + \nabla \cdot \mathbf{\Pi}. \end{aligned} \quad (2.45)$$

Importantly, we have seen in (2.41), (2.42) that the scalar parallel and perpendicular pressures are extracted by applying the double contractions to the entire pressure tensor (and not necessarily only at the gyrotropic part). The decomposition (2.43) can be rewritten as

$$\mathbf{p} = (\mathbf{p} : \hat{\mathbf{b}}\hat{\mathbf{b}})\hat{\mathbf{b}}\hat{\mathbf{b}} + (\mathbf{p} : (\mathbf{I} - \hat{\mathbf{b}}\hat{\mathbf{b}})/2)(\mathbf{I} - \hat{\mathbf{b}}\hat{\mathbf{b}}) + \Pi. \quad (2.46)$$

Applying the double contraction  $:\hat{\mathbf{b}}\hat{\mathbf{b}}$  to (2.46) yields the first requirement for the FLR tensor

$$\Pi : \hat{\mathbf{b}}\hat{\mathbf{b}} = 0. \quad (2.47)$$

Similarly, applying  $:(\mathbf{I} - \hat{\mathbf{b}}\hat{\mathbf{b}})/2$  to (2.46) yields the second requirement for the FLR tensor,

$$\Pi : (\mathbf{I} - \hat{\mathbf{b}}\hat{\mathbf{b}}) = 0. \quad (2.48)$$

By further using the first requirement, the second requirement can be rewritten as

$$\Pi : \mathbf{I} = \text{Tr}(\Pi) = 0. \quad (2.49)$$

By using the pressure decomposition (2.43) in the pressure tensor equation (2.26), and using the fact that the gyrotropic part of pressure satisfies  $\hat{\mathbf{b}} \times \mathbf{p}^g + (\hat{\mathbf{b}} \times \mathbf{p}^g)^T = 0$ , the pressure tensor equation can be rewritten as

$$\frac{\partial \mathbf{p}}{\partial t} + \nabla \cdot (\mathbf{u}\mathbf{p} + \mathbf{q}) + \mathbf{p} \cdot \nabla \mathbf{u} + (\mathbf{p} \cdot \nabla \mathbf{u})^T + \Omega \frac{|\mathbf{B}|}{B_0} [\hat{\mathbf{b}} \times \Pi + (\hat{\mathbf{b}} \times \Pi)^T] = 0. \quad (2.50)$$

We can now derive the time dependent equations for parallel and perpendicular pressures by applying the usual double contractions to this pressure tensor equation.

## 2.2. Parallel pressure equation

We calculate the double contraction with  $\hat{\mathbf{b}}\hat{\mathbf{b}}$  term by term. We will need the following identities

$$\frac{\partial}{\partial t}(\hat{\mathbf{b}}\hat{\mathbf{b}}) : \hat{\mathbf{b}}\hat{\mathbf{b}} = 0; \quad \partial_k(\hat{\mathbf{b}}\hat{\mathbf{b}}) : \hat{\mathbf{b}}\hat{\mathbf{b}} = 0. \quad (2.51)$$

The first term calculates as

$$\begin{aligned} \frac{\partial \mathbf{p}}{\partial t} : \hat{\mathbf{b}}\hat{\mathbf{b}} &= \frac{\partial}{\partial t} (p_{\parallel} \hat{b}_i \hat{b}_j + p_{\perp} (\delta_{ij} - \hat{b}_i \hat{b}_j) + \Pi_{ij}) \hat{b}_i \hat{b}_j = \frac{\partial p_{\parallel}}{\partial t} \underbrace{|\hat{\mathbf{b}}|^2 |\hat{\mathbf{b}}|^2}_{=1} \\ &\quad + p_{\parallel} \underbrace{\frac{\partial}{\partial t} (\hat{b}_i \hat{b}_j) \hat{b}_i \hat{b}_j}_{=0} + \frac{\partial p_{\perp}}{\partial t} \underbrace{(\delta_{ij} - \hat{b}_i \hat{b}_j) \hat{b}_i \hat{b}_j}_{=0} \\ &\quad + p_{\perp} \underbrace{\hat{b}_i \hat{b}_j \frac{\partial}{\partial t} (\delta_{ij} - \hat{b}_i \hat{b}_j)}_{=0} + \frac{\partial}{\partial t} (\underbrace{\Pi_{ij} \hat{b}_i \hat{b}_j}_{=0}) - \Pi_{ij} \frac{\partial}{\partial t} (\hat{b}_i \hat{b}_j) = \frac{\partial p_{\parallel}}{\partial t} - \Pi : \frac{\partial}{\partial t} (\hat{\mathbf{b}}\hat{\mathbf{b}}). \end{aligned} \quad (2.52)$$

The second term calculates as

$$\begin{aligned}\nabla \cdot (\mathbf{u}\mathbf{p}) : \hat{\mathbf{b}}\hat{\mathbf{b}} &= \partial_k(u_k p_{ij}) \hat{b}_i \hat{b}_j = \partial_k(u_k \underbrace{p_{ij} \hat{b}_i \hat{b}_j}_{=p_{\parallel}}) - u_k p_{ij} \partial_k(\hat{b}_i \hat{b}_j) = \partial_k(u_k p_{\parallel}) - u_k p_{\parallel} \underbrace{\partial_k(\hat{b}_i \hat{b}_j)}_{=0} \\ &\quad - p_{\perp} u_k \underbrace{(\delta_{ij} - \hat{b}_i \hat{b}_j) \partial_k(\hat{b}_i \hat{b}_j)}_{=0} - \Pi_{ij} u_k \partial_k(\hat{b}_i \hat{b}_j) = \nabla \cdot (p_{\parallel} \mathbf{u}) - \Pi : (\mathbf{u} \cdot \nabla(\hat{\mathbf{b}}\hat{\mathbf{b}}));\end{aligned}\quad (2.53)$$

$$(\nabla \cdot \mathbf{q}) : \hat{\mathbf{b}}\hat{\mathbf{b}} = (\nabla \cdot \mathbf{q})_{ij} \hat{b}_i \hat{b}_j = \hat{\mathbf{b}} \cdot (\nabla \cdot \mathbf{q}) \cdot \hat{\mathbf{b}}. \quad (2.54)$$

The third term calculates as

$$\begin{aligned}(\mathbf{p} \cdot \nabla \mathbf{u}) : \hat{\mathbf{b}}\hat{\mathbf{b}} &= (\mathbf{p} \cdot \nabla \mathbf{u})_{ij} \hat{b}_i \hat{b}_j = (p_{ik} \partial_k u_j) \hat{b}_i \hat{b}_j = (\partial_k u_j) p_{ik} \hat{b}_i \hat{b}_j \\ &= (\partial_k u_j) [p_{\parallel} \hat{b}_i \hat{b}_k + p_{\perp} (\delta_{ik} - \hat{b}_i \hat{b}_k) + \Pi_{ik}] \hat{b}_i \hat{b}_j \\ &= (\partial_k u_j) [p_{\parallel} \hat{b}_i \hat{b}_j + p_{\perp} \underbrace{(\hat{b}_i \hat{b}_j - \hat{b}_k \hat{b}_j)}_{=0} + \Pi_{ik} \hat{b}_i \hat{b}_j] = p_{\parallel} \hat{b}_k (\partial_k u_j) \hat{b}_j + \Pi_{ik} (\partial_k u_j) \hat{b}_i \hat{b}_j \\ &= p_{\parallel} \hat{\mathbf{b}} \cdot \nabla \mathbf{u} \cdot \hat{\mathbf{b}} + (\Pi \cdot \nabla \mathbf{u}) : \hat{\mathbf{b}}\hat{\mathbf{b}}.\end{aligned}\quad (2.55)$$

The fourth term calculates similarly as

$$\begin{aligned}(\mathbf{p} \cdot \nabla \mathbf{u})^T : \hat{\mathbf{b}}\hat{\mathbf{b}} &= (\mathbf{p} \cdot \nabla \mathbf{u})_{ji} \hat{b}_i \hat{b}_j = (p_{jk} \partial_k u_i) \hat{b}_i \hat{b}_j = (\partial_k u_i) p_{jk} \hat{b}_i \hat{b}_j \\ &= (\partial_k u_i) [p_{\parallel} \hat{b}_j \hat{b}_k + p_{\perp} (\delta_{jk} - \hat{b}_j \hat{b}_k) + \Pi_{jk}] \hat{b}_i \hat{b}_j \\ &= (\partial_k u_i) [p_{\parallel} \hat{b}_i \hat{b}_j + p_{\perp} \underbrace{(\hat{b}_i \hat{b}_j - \hat{b}_k \hat{b}_i)}_{=0} + \Pi_{jk} \hat{b}_i \hat{b}_j] = p_{\parallel} \hat{b}_k (\partial_k u_i) \hat{b}_i + \Pi_{jk} (\partial_k u_i) \hat{b}_i \hat{b}_j \\ &= p_{\parallel} \hat{\mathbf{b}} \cdot \nabla \mathbf{u} \cdot \hat{\mathbf{b}} + (\Pi \cdot \nabla \mathbf{u})_{ji} \hat{b}_i \hat{b}_j = p_{\parallel} \hat{\mathbf{b}} \cdot \nabla \mathbf{u} \cdot \hat{\mathbf{b}} + (\Pi \cdot \nabla \mathbf{u})^T : \hat{\mathbf{b}}\hat{\mathbf{b}},\end{aligned}\quad (2.56)$$

and the final fifth term becomes

$$\begin{aligned}[\hat{\mathbf{b}} \times \Pi + (\hat{\mathbf{b}} \times \Pi)^T] : \hat{\mathbf{b}}\hat{\mathbf{b}} &= [(\hat{\mathbf{b}} \times \Pi)_{ij} + (\hat{\mathbf{b}} \times \Pi)_{ji}] \hat{b}_i \hat{b}_j = [\epsilon_{ikl} \hat{b}_k \Pi_{lj} + \epsilon_{jkl} \hat{b}_k \Pi_{li}] \hat{b}_i \hat{b}_j \\ &= \hat{b}_j \Pi_{lj} \underbrace{\epsilon_{ikl} \hat{b}_i \hat{b}_k}_{=0} + \hat{b}_i \Pi_{li} \underbrace{\epsilon_{jkl} \hat{b}_j \hat{b}_k}_{=0} = 0.\end{aligned}\quad (2.57)$$

The entire equation for the parallel pressure therefore reads

$$\begin{aligned}\frac{\partial p_{\parallel}}{\partial t} + \nabla \cdot (p_{\parallel} \mathbf{u}) + 2p_{\parallel} \hat{\mathbf{b}} \cdot \nabla \mathbf{u} \cdot \hat{\mathbf{b}} + \hat{\mathbf{b}} \cdot (\nabla \cdot \mathbf{q}) \cdot \hat{\mathbf{b}} - \Pi : \left[ \left( \frac{\partial}{\partial t} + \mathbf{u} \cdot \nabla \right) (\hat{\mathbf{b}}\hat{\mathbf{b}}) \right] \\ + [\Pi \cdot \nabla \mathbf{u} + (\Pi \cdot \nabla \mathbf{u})^T] : \hat{\mathbf{b}}\hat{\mathbf{b}} = 0.\end{aligned}\quad (2.58)$$

Using the symmetric operator (2.27) and the definition for the convective derivative  $d/dt \equiv \partial/\partial t + \mathbf{u} \cdot \nabla$ , the equation is written in a simple form as

$$\frac{\partial p_{\parallel}}{\partial t} + \nabla \cdot (p_{\parallel} \mathbf{u}) + 2p_{\parallel} \hat{\mathbf{b}} \cdot \nabla \mathbf{u} \cdot \hat{\mathbf{b}} + \hat{\mathbf{b}} \cdot (\nabla \cdot \mathbf{q}) \cdot \hat{\mathbf{b}} - \Pi : \frac{d}{dt} (\hat{\mathbf{b}}\hat{\mathbf{b}}) + (\Pi \cdot \nabla \mathbf{u})^S : \hat{\mathbf{b}}\hat{\mathbf{b}} = 0, \quad (2.59)$$

which corresponds to (26) of Passot & Sulem (2004), equation (9) of Goswami *et al.* (2005) (see also Oraevskii *et al.* (1968), Passot & Sulem (2007) and Passot, Sulem & Hunana (2012)).

## 2.3. Perpendicular pressure equation

It is possible to either apply the double contraction:  $(\mathbf{I} - \hat{\mathbf{b}}\hat{\mathbf{b}})/2$  to the pressure tensor equation to directly obtain the equation for  $\partial p_{\perp}/\partial t$ , or to apply the trace (the double contraction:  $\mathbf{I}$ ) to obtain the equation for  $\partial p_{\parallel}/\partial t + 2\partial p_{\perp}/\partial t$ , and then subtract the previously obtained equation for  $\partial p_{\parallel}/\partial t$ . Both approaches are equivalent, since  $p_{\perp} = (\text{Tr}(\mathbf{p}) - p_{\parallel})/2$ . In another words, directly calculating the trace of the pressure tensor yields

$$\text{Tr} \mathbf{p} = \delta_{ij} p_{ij} = p_{ii} = p_{\parallel} \underbrace{\hat{b}_i \hat{b}_i}_{=1} + p_{\perp} \underbrace{(\delta_{ii} - \hat{b}_i \hat{b}_i)}_{=2} + \underbrace{\Pi_{ii}}_{=0} = p_{\parallel} + 2p_{\perp}, \quad (2.60)$$

where we have used that  $\text{Tr} \mathbf{\Pi} = \Pi_{ii} = 0$ . Applying the trace operator to the pressure tensor equation (2.50) term by term yields for the first term

$$\text{Tr} \frac{\partial}{\partial t} \mathbf{p} = \delta_{ij} \frac{\partial}{\partial t} p_{ij} = \frac{\partial}{\partial t} p_{ii} = \frac{\partial}{\partial t} (p_{\parallel} + 2p_{\perp}); \quad (2.61)$$

for the second term

$$\begin{aligned} \text{Tr}(\nabla \cdot (\mathbf{u} \mathbf{p})) &= \delta_{ij} \partial_k (u_k p_{ij}) = \partial_k (u_k p_{ii}) = u_k \partial_k p_{ii} + p_{ii} \partial_k u_k \\ &= \mathbf{u} \cdot \nabla (p_{\parallel} + 2p_{\perp}) + (p_{\parallel} + 2p_{\perp}) \nabla \cdot \mathbf{u}; \end{aligned} \quad (2.62)$$

$$\text{Tr}(\nabla \cdot \mathbf{q}) = \text{unchanged}; \quad (2.63)$$

the third term

$$\begin{aligned} \text{Tr}(\mathbf{p} \cdot \nabla \mathbf{u}) &= \delta_{ij} (\mathbf{p} \cdot \nabla \mathbf{u})_{ij} = p_{ik} \partial_k u_i = ((p_{\parallel} - p_{\perp}) \hat{b}_i \hat{b}_k + p_{\perp} \delta_{ik} + \Pi_{ik}) \partial_k u_i \\ &= (p_{\parallel} - p_{\perp}) \hat{b}_k (\partial_k u_i) \hat{b}_i + p_{\perp} \partial_i u_i + \Pi_{ik} \partial_k u_i \\ &= (p_{\parallel} - p_{\perp}) \hat{\mathbf{b}} \cdot \nabla \mathbf{u} \cdot \hat{\mathbf{b}} + p_{\perp} \nabla \cdot \mathbf{u} + \text{Tr}(\mathbf{\Pi} \cdot \nabla \mathbf{u}); \end{aligned} \quad (2.64)$$

the fourth term is equivalent to the third term

$$\text{Tr}(\mathbf{p} \cdot \nabla \mathbf{u})^T = (\mathbf{p} \cdot \nabla \mathbf{u})_{ii} = (p_{\parallel} - p_{\perp}) \hat{\mathbf{b}} \cdot \nabla \mathbf{u} \cdot \hat{\mathbf{b}} + p_{\perp} \nabla \cdot \mathbf{u} + \text{Tr}(\mathbf{\Pi} \cdot \nabla \mathbf{u})^T; \quad (2.65)$$

even though we kept the last expression in transpose form so that we can use the symmetric operator after we combine all the terms. Finally the fifth term

$$\text{Tr}(\hat{\mathbf{b}} \times \mathbf{\Pi}) = \delta_{ij} \epsilon_{ikl} \hat{b}_k \Pi_{lj} = \hat{b}_k \underbrace{\epsilon_{ikl} \Pi_{li}}_{=0} = 0; \quad (2.66)$$

$$\text{Tr}(\hat{\mathbf{b}} \times \mathbf{\Pi})^T = \text{Tr}(\hat{\mathbf{b}} \times \mathbf{\Pi}) = 0. \quad (2.67)$$

Combining all the terms, using the parallel pressure equation (2.59) and dividing by 2 yields the time-dependent equation for the perpendicular pressure that reads

$$\begin{aligned} \frac{\partial p_{\perp}}{\partial t} + \nabla \cdot (p_{\perp} \mathbf{u}) + p_{\perp} \nabla \cdot \mathbf{u} - p_{\perp} \hat{\mathbf{b}} \cdot \nabla \mathbf{u} \cdot \hat{\mathbf{b}} + \frac{1}{2} [\text{Tr} \nabla \cdot \mathbf{q} - \hat{\mathbf{b}} \cdot (\nabla \cdot \mathbf{q}) \cdot \hat{\mathbf{b}}] \\ + \frac{1}{2} \left[ \text{Tr}(\mathbf{\Pi} \cdot \nabla \mathbf{u})^S + \mathbf{\Pi} : \frac{d}{dt} (\hat{\mathbf{b}} \hat{\mathbf{b}}) - (\mathbf{\Pi} \cdot \nabla \mathbf{u})^S : \hat{\mathbf{b}} \hat{\mathbf{b}} \right] = 0. \end{aligned} \quad (2.68)$$

The result corresponds to, for example, equation (25) of Passot & Sulem (2004), equation (8) of Goswami *et al.* (2005). It is important to note that the parallel

and perpendicular pressure equations are exact equations. These equations are valid regardless of what kind of higher-order closure will be adopted later. It is useful to rewrite the term  $\Pi : (\text{d}/\text{d}t)(\hat{\mathbf{b}}\hat{\mathbf{b}})$  slightly so that we can directly use the induction equation and eliminate the time derivative. The term can be rewritten as

$$\begin{aligned}\Pi : \frac{\text{d}}{\text{d}t}(\hat{\mathbf{b}}\hat{\mathbf{b}}) &= \Pi_{ij} \frac{\text{d}}{\text{d}t}(\hat{b}_i \hat{b}_j) = \Pi_{ij} \left( \hat{b}_i \frac{\text{d}}{\text{d}t} \hat{b}_j + \hat{b}_j \frac{\text{d}}{\text{d}t} \hat{b}_i \right) \\ &= \hat{b}_i \Pi_{ij} \frac{\text{d}}{\text{d}t} \hat{b}_j + \hat{b}_j \underbrace{\Pi_{ij}}_{\Pi_{ji}} \frac{\text{d}}{\text{d}t} \hat{b}_i = 2\hat{\mathbf{b}} \cdot \Pi \cdot \frac{\text{d}\hat{\mathbf{b}}}{\text{d}t}.\end{aligned}\quad (2.69)$$

We will need the following identity for the time derivative of the unit vector (see later in the text),

$$\frac{\text{d}}{\text{d}t} \hat{\mathbf{b}} = \frac{1}{|\mathbf{B}|} \left[ \frac{\text{d}\mathbf{B}}{\text{d}t} - \hat{\mathbf{b}} \left( \hat{\mathbf{b}} \cdot \frac{\text{d}\mathbf{B}}{\text{d}t} \right) \right], \quad (2.70)$$

which, when used in the above expression together with  $\hat{\mathbf{b}} \cdot \Pi \cdot \hat{\mathbf{b}} = 0$ , allows us to write

$$\Pi : \frac{\text{d}}{\text{d}t}(\hat{\mathbf{b}}\hat{\mathbf{b}}) = \frac{2}{|\mathbf{B}|} \hat{\mathbf{b}} \cdot \Pi \cdot \frac{\text{d}\mathbf{B}}{\text{d}t}, \quad (2.71)$$

which is an expression of general validity since no specific form of the induction equation was assumed yet. In the next section we will calculate the decomposition of the heat flux tensor  $\mathbf{q}$  and we will show that, if only the gyrotropic heat flux components  $q_{\parallel}, q_{\perp}$  are considered, the heat flux terms entering the pressure equations read

$$\hat{\mathbf{b}} \cdot (\nabla \cdot \mathbf{q}^g) \cdot \hat{\mathbf{b}} = \nabla \cdot (q_{\parallel} \hat{\mathbf{b}}) - 2q_{\perp} \nabla \cdot \hat{\mathbf{b}}; \quad (2.72)$$

$$\frac{1}{2} [\text{Tr} \nabla \cdot \mathbf{q}^g - \hat{\mathbf{b}} \cdot (\nabla \cdot \mathbf{q}^g) \cdot \hat{\mathbf{b}}] = \nabla \cdot (q_{\perp} \hat{\mathbf{b}}) + q_{\perp} \nabla \cdot \hat{\mathbf{b}}. \quad (2.73)$$

Therefore, by splitting the heat flux into gyrotropic and non-gyrotropic parts,  $\mathbf{q} = \mathbf{q}^g + \mathbf{q}^{\text{ng}}$ , the pressure equations read

$$\begin{aligned}\frac{\partial p_{\parallel}}{\partial t} + \nabla \cdot (p_{\parallel} \mathbf{u}) + 2p_{\parallel} \hat{\mathbf{b}} \cdot \nabla \mathbf{u} \cdot \hat{\mathbf{b}} + \nabla \cdot (q_{\parallel} \hat{\mathbf{b}}) - 2q_{\perp} \nabla \cdot \hat{\mathbf{b}} + \hat{\mathbf{b}} \cdot (\nabla \cdot \mathbf{q}^{\text{ng}}) \cdot \hat{\mathbf{b}} \\ - \frac{2}{|\mathbf{B}|} \hat{\mathbf{b}} \cdot \Pi \cdot \frac{\text{d}\mathbf{B}}{\text{d}t} + (\Pi \cdot \nabla \mathbf{u})^S : \hat{\mathbf{b}}\hat{\mathbf{b}} = 0;\end{aligned}\quad (2.74)$$

$$\begin{aligned}\frac{\partial p_{\perp}}{\partial t} + \nabla \cdot (p_{\perp} \mathbf{u}) + p_{\perp} \nabla \cdot \mathbf{u} - p_{\perp} \hat{\mathbf{b}} \cdot \nabla \mathbf{u} \cdot \hat{\mathbf{b}} + \nabla \cdot (q_{\perp} \hat{\mathbf{b}}) \\ + q_{\perp} \nabla \cdot \hat{\mathbf{b}} + \frac{1}{2} [\text{Tr} \nabla \cdot \mathbf{q}^{\text{ng}} - \hat{\mathbf{b}} \cdot (\nabla \cdot \mathbf{q}^{\text{ng}}) \cdot \hat{\mathbf{b}}] \\ + \frac{1}{2} \left[ \text{Tr}(\Pi \cdot \nabla \mathbf{u})^S + \frac{2}{|\mathbf{B}|} \hat{\mathbf{b}} \cdot \Pi \cdot \frac{\text{d}\mathbf{B}}{\text{d}t} - (\Pi \cdot \nabla \mathbf{u})^S : \hat{\mathbf{b}}\hat{\mathbf{b}} \right] = 0.\end{aligned}\quad (2.75)$$

Terms with the non-gyrotropic heat flux  $\mathbf{q}^{\text{ng}}$  can be further split into non-gyrotropic heat flux vectors  $\mathbf{S}_{\perp}^{\parallel}, \mathbf{S}_{\perp}^{\perp}$  and the heat flux tensor  $\sigma$ , which is addressed in appendix D, see (D 124), (D 126). Terms containing the FLR pressure tensor  $\Pi$  are usually called



the FLR stress forces, and these forces can generate complicated plasma heating processes. The FLR stress forces can generate both parallel and perpendicular plasma heating that can be determined only from fully nonlinear numerical simulations, since all the terms disappear at the linear level. This process is also sometimes called ‘stochastic heating’. The importance of stochastic heating was shown by Laveder, Passot & Sulem (2013), who performed 1.5-dimensional numerical simulations with the sophisticated FLR Landau fluid model of Passot & Sulem (2007), Passot *et al.* (2012).

#### 2.4. On the paper by Chew *et al.* (1956)

The very influential paper by Chew *et al.* (1956) that is typically cited for the CGL equations (with zero heat fluxes) actually also derived the pressure equations when the gyrotropic heat fluxes are considered. It is important to emphasize that, even though the scalar pressures  $p_n$ ,  $p_s$  in the notation of that paper are equal to the usual pressures  $p_{\parallel}$ ,  $p_{\perp}$ , this is not the case for the heat fluxes. The gyrotropic heat flux tensor in that paper is decomposed into components  $q_n$ ,  $q_s$  according to (Chew *et al.* 1956, equation (34))

$$q_{ijk}^g = q_n \hat{b}_i \hat{b}_j \hat{b}_k + q_s (\delta_{ij} \hat{b}_k + \delta_{jk} \hat{b}_i + \delta_{ik} \hat{b}_j). \quad (2.76)$$

In contrast, the usual gyrotropic heat flux decomposition to  $q_{\parallel}$ ,  $q_{\perp}$  reads (see later in the text)

$$q_{ijk}^g = q_{\parallel} \hat{b}_i \hat{b}_j \hat{b}_k + q_{\perp} (\delta_{ij} \hat{b}_k + \delta_{jk} \hat{b}_i + \delta_{ik} \hat{b}_j - 3 \hat{b}_i \hat{b}_j \hat{b}_k), \quad (2.77)$$

i.e. the last term is not present in their decomposition. However, this does not mean that their pressure equations (Chew *et al.* 1956, equations (31), (32)) are incorrect. It only implies that their equations are written in terms of

$$q_n = q_{\parallel} - 3q_{\perp}; \quad q_s = q_{\perp}. \quad (2.78)$$

By neglecting the FLR corrections  $\Pi$  and  $\mathbf{q}^{\text{ng}}$ , the parallel and perpendicular pressure equations (2.74), (2.75) simplify to

$$\frac{dp_{\parallel}}{dt} + p_{\parallel} \nabla \cdot \mathbf{u} + 2p_{\parallel} \hat{\mathbf{b}} \cdot \nabla \mathbf{u} \cdot \hat{\mathbf{b}} + \nabla \cdot (q_{\parallel} \hat{\mathbf{b}}) - 2q_{\perp} \nabla \cdot \hat{\mathbf{b}} = 0; \quad (2.79)$$

$$\frac{dp_{\perp}}{dt} + 2p_{\perp} \nabla \cdot \mathbf{u} - p_{\perp} \hat{\mathbf{b}} \cdot \nabla \mathbf{u} \cdot \hat{\mathbf{b}} + \nabla \cdot (q_{\perp} \hat{\mathbf{b}}) + q_{\perp} \nabla \cdot \hat{\mathbf{b}} = 0. \quad (2.80)$$

By using (2.78), the heat flux contributions in the parallel pressure equation above can be rewritten as

$$\nabla \cdot (q_{\parallel} \hat{\mathbf{b}}) - 2q_{\perp} \nabla \cdot \hat{\mathbf{b}} = \nabla \cdot [(q_n + q_s) \hat{\mathbf{b}}] + 2\hat{\mathbf{b}} \cdot \nabla q_s, \quad (2.81)$$

which agrees with the parallel pressure equation obtained by Chew *et al.* (1956), equation (31). The pressure equations of Chew *et al.* (1956) are therefore correct, but it has to be remembered that  $q_n$  is not the usual parallel heat flux  $q_{\parallel} = m \int (v_{\parallel} - u_{\parallel})^3 f d^3 v$ .<sup>1</sup> Actually, the authors do not even call the quantities  $q_n$ ,  $q_s$  heat flux components, but components of a ‘pressure-transport’ tensor.

To conclude this subsection, we will go slightly ahead and use (3.51), (3.52) (without the zero right-hand sides) that are the usual CGL pressure equations and

<sup>1</sup>Sometimes the parallel heat flux is defined with the factor of 1/2 as  $q_{\parallel} = m/2 \int (v_{\parallel} - u_{\parallel})^3 f d^3 v$ , but this definition should be avoided.

valid only at long spatial scales when the Hall term in the induction equation is neglected. Under this assumption, the pressure equations (2.79), (2.80) are rewritten as

$$\frac{d}{dt} \left( \frac{p_{\parallel} |\mathbf{B}|^2}{\rho^3} \right) = - \frac{|\mathbf{B}|^2}{\rho^3} [\nabla \cdot (q_{\parallel} \hat{\mathbf{b}}) - 2q_{\perp} \nabla \cdot \hat{\mathbf{b}}]; \quad (2.82)$$

$$\frac{d}{dt} \left( \frac{p_{\perp}}{\rho |\mathbf{B}|} \right) = - \frac{1}{\rho |\mathbf{B}|} [\nabla \cdot (q_{\perp} \hat{\mathbf{b}}) + q_{\perp} \nabla \cdot \hat{\mathbf{b}}]. \quad (2.83)$$

It is often forgotten that the paper by Chew *et al.* (1956) derived the correct pressure equations with gyrotropic heat flux contributions. The authors, however, did not write the equations in the form (2.82), (2.83), and left them in the form (2.79), (2.80), using the ‘conservative’ form only after prescribing the closure  $q_{\parallel} = q_{\perp} = 0$  (their equations (35), (36)),

$$\frac{d}{dt} \left( \frac{p_{\parallel} |\mathbf{B}|^2}{\rho^3} \right) = 0; \quad (2.84)$$

$$\frac{d}{dt} \left( \frac{p_{\perp}}{\rho |\mathbf{B}|} \right) = 0. \quad (2.85)$$

Over time, equations (2.84), (2.85) became known as the CGL description. Equations (2.84), (2.85) are often replaced by the equations of state  $p_{\parallel} |\mathbf{B}|^2 / \rho^3 = \text{const.}$  and  $p_{\perp} / (\rho |\mathbf{B}|) = \text{const.}$ , see e.g. Mjølhus (2009). Additionally, similarly to the definition of (appropriately normalized) entropy in MHD,  $s = \ln(p/\rho^{\gamma})$ , it is useful to define parallel and perpendicular entropy in the CGL model, according to

$$s_{\parallel} = \frac{1}{3} \ln \left( \frac{p_{\parallel} |\mathbf{B}|^2}{\rho^3} \right); \quad s_{\perp} = \frac{2}{3} \ln \left( \frac{p_{\perp}}{\rho |\mathbf{B}|} \right). \quad (2.86)$$

The total entropy,  $s_{\parallel} + s_{\perp} = \ln(p_{\parallel}^{1/3} p_{\perp}^{2/3} \rho^{-5/3})$ , is equivalent to the MHD entropy when  $p_{\parallel} = p_{\perp}$ , see for example (9)–(11) in Abraham-Shrauner (1967).

### 2.5. Physical meaning of CGL equations

The exact derivation of the CGL equation (2.84), (2.85) will be completed in the next section. Here, we briefly discuss the physical meaning of these equations. The CGL equations (2.85), (2.84) can be interpreted as the conservation of the first and second adiabatic invariants, as is nicely discussed for example by Kulsrud (1983) and by Gurnett & Bhattacharjee (2005). Considering many particles with velocity components  $v_{\parallel}$  and  $v_{\perp}$  with respect to the mean magnetic field, the parallel and perpendicular pressures can be estimated as

$$p_{\parallel} \sim \langle v_{\parallel}^2 \rangle \rho; \quad p_{\perp} \sim \langle v_{\perp}^2 \rangle \rho, \quad (2.87)$$

where the brackets represent an average over all particles. To complete the estimate, we need expressions for  $\langle v_{\parallel}^2 \rangle$  and  $\langle v_{\perp}^2 \rangle$ , which come from the conservation of the adiabatic invariants. The first adiabatic invariant is the conservation of magnetic moment  $\mu$  of a particle that is gyrating periodically around a mean magnetic field. The second adiabatic invariant, sometimes also called the longitudinal invariant, is associated with a particle bouncing periodically between two magnetic mirror points.

The conserved quantity is the integral over the parallel momentum of a particle  $J \equiv \oint_a^b mv_{\parallel} dl$ , where the magnetic mirrors are located at ‘a’ and ‘b’ and the integral is performed along the magnetic field line. The distance between the two magnetic mirrors at ‘a’ and ‘b’ can be labelled as  $L$ . The conservation of the first and second adiabatic invariants is therefore

$$\mu \equiv \frac{mv_{\perp}^2}{2B} = \text{const.}; \quad J \equiv mv_{\parallel}L = \text{const.} \quad (2.88)$$

Since we are considering only non-relativistic particles, the particle mass ‘m’ is also a constant. The conservation of magnetic moment  $\mu$  therefore implies  $v_{\perp}^2 \sim B$ , and using this in (2.87) yields  $p_{\perp} \sim B\rho$ , recovering (2.85). The CGL equation (2.85) therefore corresponds to conservation of magnetic moment, i.e. the first adiabatic invariant. The conservation of the second adiabatic invariant  $J$  is more tricky, since we need to somehow estimate the non-intuitive length  $L$  between the two magnetic mirrors, and no magnetic mirrors are explicitly assumed, since we are just dealing with somewhat random magnetic field lines. The length  $L$  can be nevertheless estimated from two fundamental physical principles. Consider a magnetic flux tube (a deformed cylinder) with cross-sectional area  $A$  and length  $L$ . The conservation of the total magnetic flux through the area  $A$  implies  $AB = \text{const.}$ , which yields an estimate for the area  $A \sim 1/B$ . The second principle is the conservation of the total mass of particles that are completely trapped in that flux tube and that cannot escape,  $mnAL = \text{const.}$ , yielding  $L \sim 1/(\rho A)$ . The use of  $A \sim 1/B$  yields the final estimate for the non-intuitive length  $L \sim B/\rho$ . The conservation of the second adiabatic invariant  $J$  therefore implies  $v_{\parallel} \sim 1/L \sim \rho/B$ . Using this result in (2.87) implies  $p_{\parallel} \sim \rho^3/B^2$ , recovering (2.84). The CGL equation (2.84) corresponds to conservation of the second adiabatic invariant  $J$ , and the equation is valid for particles that are completely trapped and therefore cannot carry a heat flux. As we discuss later, the consideration of the Hall term, the heat flux and the FLR stress forces (the stochastic heating), leads to the breaking of these two adiabatic invariants. Nevertheless, exact conservation laws can still be derived, and the very simple CGL equations (2.84), (2.85) have to be modified with non-zero right-hand sides.

In Part 2 of our guide (see §4.1), we show that the conservation of magnetic moment is very useful for understanding the form of the perturbed distribution function  $f^{(1)}$  in the gyrotropic limit. By performing calculations in the laboratory reference frame, we show that to obtain the correct  $f^{(1)}$  in the gyrotropic limit actually requires the usual complicated kinetic integration around the unperturbed orbit (see the Appendix of Part 2) and only then the gyrotropic limit can be imposed. In contrast, performing calculations in the guiding-centre reference frame allows one to prescribe the conservation of magnetic moment from the beginning, and obtain the correct  $f^{(1)}$  in the gyrotropic limit perhaps more intuitively. Conservation of adiabatic invariants is important in many areas of space physics and astrophysics. For example, conservation of adiabatic invariants is used to construct models that describe particles trapped inside of magnetic islands and that are being accelerated during magnetic island contraction and merging (Drake, Swisdak & Fermo 2013; Zank *et al.* 2014).

## 2.6. Exact equations of anisotropic multi-fluid models

Before we discuss solutions of specific fluid models, it is beneficial to summarize the most general equations of multi-fluid models that were derived with no simplifications

at all. By reintroducing the species index  $r$ , the equations that were directly obtained by integrating the collisionless Vlasov equation read

$$\frac{\partial \rho_r}{\partial t} + \nabla \cdot (\rho_r \mathbf{u}_r) = 0; \quad (2.89)$$

$$\frac{\partial \mathbf{u}_r}{\partial t} + \mathbf{u}_r \cdot \nabla \mathbf{u}_r + \frac{1}{\rho_r} \nabla \cdot \mathbf{p}_r - \frac{q_r}{m_r} \left( \mathbf{E} + \frac{1}{c} \mathbf{u}_r \times \mathbf{B} \right) = 0; \quad (2.90)$$

$$\begin{aligned} \frac{\partial p_{\parallel r}}{\partial t} + \nabla \cdot (p_{\parallel r} \mathbf{u}_r) + 2p_{\parallel r} \hat{\mathbf{b}} \cdot \nabla \mathbf{u}_r \cdot \hat{\mathbf{b}} + \hat{\mathbf{b}} \cdot (\nabla \cdot \mathbf{q}_r) \cdot \hat{\mathbf{b}} \\ - \frac{2}{|\mathbf{B}|} \hat{\mathbf{b}} \cdot \mathbf{\Pi}_r \cdot \frac{d_r \mathbf{B}}{dt} + (\mathbf{\Pi}_r \cdot \nabla \mathbf{u}_r)^S : \hat{\mathbf{b}} \hat{\mathbf{b}} = 0; \end{aligned} \quad (2.91)$$

$$\begin{aligned} \frac{\partial p_{\perp r}}{\partial t} + \nabla \cdot (p_{\perp r} \mathbf{u}_r) + p_{\perp r} \nabla \cdot \mathbf{u}_r - p_{\perp r} \hat{\mathbf{b}} \cdot \nabla \mathbf{u}_r \cdot \hat{\mathbf{b}} + \frac{1}{2} [\text{Tr} \nabla \cdot \mathbf{q}_r - \hat{\mathbf{b}} \cdot (\nabla \cdot \mathbf{q}_r) \cdot \hat{\mathbf{b}}] \\ + \frac{1}{2} \left[ \text{Tr}(\mathbf{\Pi}_r \cdot \nabla \mathbf{u}_r)^S + \frac{2}{|\mathbf{B}|} \hat{\mathbf{b}} \cdot \mathbf{\Pi}_r \cdot \frac{d_r \mathbf{B}}{dt} - (\mathbf{\Pi}_r \cdot \nabla \mathbf{u}_r)^S : \hat{\mathbf{b}} \hat{\mathbf{b}} \right] = 0, \end{aligned} \quad (2.92)$$

where the convective derivative  $d_r/dt = \partial/\partial t + \mathbf{u}_r \cdot \nabla$ . In the above equations, the pressure tensor for each species was just decomposed into gyrotropic contributions  $p_{\parallel r}$ ,  $p_{\perp r}$  and the rest of the pressure tensor (the FLR pressure tensor  $\mathbf{\Pi}_r$ ), according to

$$\mathbf{p}_r = p_{\parallel r} \hat{\mathbf{b}} \hat{\mathbf{b}} + p_{\perp r} (\mathbf{I} - \hat{\mathbf{b}} \hat{\mathbf{b}}) + \mathbf{\Pi}_r, \quad (2.93)$$

which is a rigorous decomposition not introducing any simplifications. The heat flux  $\mathbf{q}_r$  contributions are here left at the most general level without any simplifications either. The equations are accompanied by Maxwell's equations, that for now we write down here with no simplifications to emphasize that no Maxwell equations were used in deriving the above system,

$$\nabla \cdot \mathbf{E} = 4\pi \rho_c = 4\pi \sum_r q_r n_r; \quad \nabla \cdot \mathbf{B} = 0; \quad (2.94)$$

$$\frac{\partial \mathbf{B}}{\partial t} = -c \nabla \times \mathbf{E}; \quad \mathbf{j} = \sum_r q_r n_r \mathbf{u}_r = \frac{c}{4\pi} \nabla \times \mathbf{B} - \frac{1}{4\pi} \frac{\partial \mathbf{E}}{\partial t}. \quad (2.95)$$

Equations (2.89)–(2.95) represent an exact (even though not closed) kinetic Vlasov–Maxwell system formulated in fluid variables, and the description is valid for any general distribution function and for all considered spatial and temporal scales. All advanced collisionless multi-fluid models have to be based in one way or another on these equations. One can concentrate on the proton dynamics or one can concentrate on the electron dynamics and simplify these equations accordingly for these spatial scales. Naturally, one can consider complicated multi-fluid descriptions composed of many particle species.

For a new reader who just jumped straight to this section, the symmetric operator ‘ $S$ ’ acts on a matrix  $A_{ij}$  according to  $A_{ij}^S = A_{ij} + A_{ji}$ . If one does not like the symmetric operator in the expressions above, the symmetric operator is actually not necessary, since

$$(\mathbf{\Pi} \cdot \nabla \mathbf{u})_{ij} = \Pi_{ik} \partial_k u_j; \quad (2.96)$$

$$(\mathbf{\Pi} \cdot \nabla \mathbf{u})_{ij}^S = \Pi_{ik} \partial_k u_j + \Pi_{jk} \partial_k u_i; \quad (2.97)$$

$$(\boldsymbol{\Pi} \cdot \nabla \mathbf{u})^S : \hat{\mathbf{b}}\hat{\mathbf{b}} = 2\hat{b}_i \Pi_{ik} (\partial_k u_j) \hat{b}_j = 2(\boldsymbol{\Pi} \cdot \nabla \mathbf{u}) : \hat{\mathbf{b}}\hat{\mathbf{b}} = 2\hat{\mathbf{b}} \cdot (\boldsymbol{\Pi} \cdot \nabla \mathbf{u}) \cdot \hat{\mathbf{b}}; \quad (2.98)$$

$$\text{Tr}(\boldsymbol{\Pi} \cdot \nabla \mathbf{u})^S = 2\Pi_{ik} \partial_k u_i = 2\boldsymbol{\Pi} : \nabla \mathbf{u} = 2(\boldsymbol{\Pi} \cdot \nabla) \cdot \mathbf{u}, \quad (2.99)$$

where we also got rid of the double contractions. Note that  $\Pi_{ik} = \Pi_{ki}$  and also  $p_{ik} = p_{ki}$ .

Equations (2.89)–(2.92) contain two very important catches – the entire heat flux tensor  $\mathbf{q}_r$  and the non-gyrotropic pressure tensor  $\boldsymbol{\Pi}_r$  are not specified yet. In this moment, the entire class of collisionless fluid models separates to many possible classes and sub-classes, depending on how precisely one wants to evaluate the heat flux tensor  $\mathbf{q}_r$  and the FLR pressure tensor  $\boldsymbol{\Pi}_r$ . There are several complications how to correctly model these two quantities, but before we discuss this, we want to point out that by using the simple general identities (3.46), (3.47) (that are nothing more than calculating derivatives and using the density equation for each species separately), the above pressure equations can be rewritten to the following form

$$\begin{aligned} \frac{d_r}{dt} \left( \frac{p_{\parallel r} |\mathbf{B}|^2}{\rho_r^3} \right) &= \frac{|\mathbf{B}|^2}{\rho_r^3} \left\{ 2p_{\parallel r} \nabla \cdot \mathbf{u}_r - 2p_{\parallel r} \hat{\mathbf{b}} \cdot \nabla \mathbf{u}_r \cdot \hat{\mathbf{b}} + 2 \frac{p_{\parallel r}}{|\mathbf{B}|} \hat{\mathbf{b}} \cdot \frac{d_r \mathbf{B}}{dt} - \hat{\mathbf{b}} \cdot (\nabla \cdot \mathbf{q}_r) \cdot \hat{\mathbf{b}} \right. \\ &\quad \left. + \frac{2}{|\mathbf{B}|} \hat{\mathbf{b}} \cdot \boldsymbol{\Pi}_r \cdot \frac{d_r \mathbf{B}}{dt} - (\boldsymbol{\Pi}_r \cdot \nabla \mathbf{u}_r)^S : \hat{\mathbf{b}}\hat{\mathbf{b}} \right\}; \end{aligned} \quad (2.100)$$

$$\begin{aligned} \frac{d_r}{dt} \left( \frac{p_{\perp r}}{\rho_r |\mathbf{B}|} \right) &= \frac{1}{\rho_r |\mathbf{B}|} \left\{ -p_{\perp r} \nabla \cdot \mathbf{u}_r + p_{\perp r} \hat{\mathbf{b}} \cdot \nabla \mathbf{u}_r \cdot \hat{\mathbf{b}} - \frac{p_{\perp r}}{|\mathbf{B}|} \hat{\mathbf{b}} \cdot \frac{d_r \mathbf{B}}{dt} \right. \\ &\quad - \frac{1}{2} [\text{Tr} \nabla \cdot \mathbf{q}_r - \hat{\mathbf{b}} \cdot (\nabla \cdot \mathbf{q}_r) \cdot \hat{\mathbf{b}}] \\ &\quad \left. - \frac{1}{2} \left[ \text{Tr}(\boldsymbol{\Pi}_r \cdot \nabla \mathbf{u}_r)^S + \frac{2}{|\mathbf{B}|} \hat{\mathbf{b}} \cdot \boldsymbol{\Pi}_r \cdot \frac{d_r \mathbf{B}}{dt} - (\boldsymbol{\Pi}_r \cdot \nabla \mathbf{u}_r)^S : \hat{\mathbf{b}}\hat{\mathbf{b}} \right] \right\}. \end{aligned} \quad (2.101)$$

Again, no simplifications were introduced and no Maxwell equations were used yet. The equations are exact. The equations represent the complicated plasma heating processes that are responsible for anisotropic plasma heating, and that are in general very difficult to classify. The left-hand sides of (2.100)–(2.101) are the familiar CGL equations that represent conservations of the first and second adiabatic invariants, and all the expressions on the right-hand sides break these adiabatic invariants. On the right-hand side, we have the  $\partial \mathbf{B} / \partial t$  that couples various species together through Maxwell's equations, and that also introduces the Hall term responsible for the simplest dispersive effects. The heat flux tensor  $\mathbf{q}_r$  (a tensor of third rank), is decomposed into its gyrotropic and non-gyrotropic (FLR) parts  $\mathbf{q}_r = \mathbf{q}_r^g + \mathbf{q}_r^{\text{ng}}$ . At large scales, the gyrotropic heat flux can be viewed as a gateway for the simplest form of collisionless damping mechanism, known as Landau damping, that can be further separated to a 'pure' electrostatic Landau damping, and its magnetic analogue, the transit-time damping – see Part 2 of the text, subsection 'Coulomb force and mirror force (Landau damping and transit time damping)'. Nevertheless, fluid models that contain the gyrotropic heat flux fluctuations, but that do not contain any Landau damping, can be constructed as well (see CGL2 model later in the text). The FLR pressure tensor  $\boldsymbol{\Pi}_r$  introduces further dispersive effects and therefore also modifies the collisionless damping rates, as well as the heat flux  $\mathbf{q}_r$  modifies the FLR pressure corrections  $\boldsymbol{\Pi}_r$ . The quantities  $\mathbf{q}_r$  and  $\boldsymbol{\Pi}_r$  are therefore generally coupled. Importantly, the FLR pressure tensor  $\boldsymbol{\Pi}_r$  introduces complicated turbulent plasma heating processes, referred to as stochastic heating. Since these FLR stress forces completely disappear at the linear level in the above pressure equations, they can be explored only with fully

nonlinear numerical simulations ( $\Pi_r$  still enters at the linear level in the momentum equation and modifies the dispersion relations).

From a fluid perspective, there are several major difficulties how to correctly model the quantities  $\Pi_r$  and  $\mathbf{q}_r$ . The heat flux tensor  $\mathbf{q}_r$  is described by an infinite hierarchy of higher-order fluid moments, and one needs to find an appropriate way, how to close the system. The FLR pressure tensor  $\Pi_r$  is described by equations that are implicit, and the FLR corrections have to be typically expanded on temporal and spatial scales in order to obtain tractable expressions that can be used for numerical simulations. This expansion will naturally restrict the area of validity of such fluid models to those scales considered. For example, for the proton–electron plasma where the spatial and temporal scales are largely separated because of the mass ratio  $m_p/m_e = 1836$ , simple first-order FLR corrections expanded around the proton scales  $k\rho_i$ ,  $\omega/\Omega_p$  will long lose their validity at the electron scales. The most complicated Landau fluid models that use linear kinetic theory to evaluate  $\Pi$  and that contain the Bessel functions (Passot & Sulem 2007; Passot *et al.* 2012; Sulem & Passot 2015), have technically no restriction for wavenumbers  $k_\perp\rho_i$ , but the first-order frequency restriction is there nevertheless.

### 2.7. Conservation of energy

The equations (2.100)–(2.101), that correctly split the entire plasma heating into parallel and perpendicular heating, are indeed quite complicated. To gain further insight, let us briefly consider a case in which we are not interested in parallel and perpendicular heating (whose net effect can possibly be zero), and we are only interested in the heating of the entire system. Such formulations are sometimes used to interpret plasma heating in fully kinetic simulations (Yang *et al.* 2017a,b). By summing together the pressure equations (2.91), (2.92), we are interested in the evolution equation  $\partial(p_\parallel + 2p_\perp)/\partial t$ , or in another words, we are interested in  $\partial \text{Tr} \mathbf{p}/\partial t$ . Summing the equations together (actually the trace of the entire pressure tensor equation was calculated a few pages back, that is how the  $p_\perp$  equation was derived) yields

$$\frac{\partial}{\partial t}(p_{\parallel r} + 2p_{\perp r}) + \nabla \cdot ((p_{\parallel r} + 2p_{\perp r})\mathbf{u}_r) + \nabla \cdot (\text{Tr} \mathbf{q}_r) + 2(\underbrace{(\mathbf{p}_r^g + \Pi_r) \cdot \nabla}_{=\mathbf{p}_r}) \cdot \mathbf{u}_r = 0. \quad (2.102)$$

Note that  $\text{Tr} \nabla \cdot \mathbf{q}_r = \nabla \cdot (\text{Tr} \mathbf{q}_r)$ , and also  $\text{Tr} \Pi_r = 0$ . By integrating over the entire spatial volume of the plasma, one can define the internal energy (or thermal energy) for each particle species as

$$E_r^{\text{int}} = \frac{1}{2} \int \text{Tr} \mathbf{p}_r d^3x = \frac{1}{2} \int (p_{\parallel r} + 2p_{\perp r}) d^3x \equiv \frac{1}{2} \langle p_{\parallel r} + 2p_{\perp r} \rangle. \quad (2.103)$$

By considering a special case of periodic boundary conditions, i.e. typical numerical simulations of turbulence in a periodic box, one can use the Gauss–Ostrogradsky theorem (2.9), now in a spatial domain, and for periodic boundary conditions all divergence operators in (2.102) vanish, yielding

$$\frac{\partial}{\partial t} E_r^{\text{int}} = -\langle ((\mathbf{p}_r^g + \Pi_r) \cdot \nabla) \cdot \mathbf{u}_r \rangle. \quad (2.104)$$

Equations (2.102), (2.104) are equivalent to (4) and (7) of Yang *et al.* (2017b). Here the equations are just written in such a way that one can explicitly see the pressure



components  $p_{\parallel}, p_{\perp}, \mathbf{\Pi}$ , but the equations are equivalent. Following Yang *et al.* (2017b), one can also easily derive the evolution equations for the ‘kinetic’ energy and the electromagnetic energy

$$E_r^{\text{kin}} = \frac{1}{2} \langle \rho_r |\mathbf{u}_r|^2 \rangle; \quad E^{\text{mag}} = \frac{1}{8\pi} \langle (|\mathbf{B}|^2 + |\mathbf{E}|^2) \rangle. \quad (2.105)$$

For the kinetic energy, by first applying  $\mathbf{u}_r \cdot$  to the momentum equation (2.90)

$$\mathbf{u}_r \cdot \left( \rho_r \frac{\partial \mathbf{u}_r}{\partial t} + \rho_r \mathbf{u}_r \cdot \nabla \mathbf{u}_r \right) = -\mathbf{u}_r \cdot (\nabla \cdot \mathbf{p}_r) + \underbrace{q_r n_r \mathbf{u}_r \cdot \mathbf{E}}_{=j_r}, \quad (2.106)$$

which further implies

$$\begin{aligned} \frac{1}{2} \frac{\partial}{\partial t} (\rho |\mathbf{u}|^2) &= \frac{1}{2} \left( \frac{\partial \rho}{\partial t} |\mathbf{u}|^2 + \rho \frac{\partial |\mathbf{u}|^2}{\partial t} \right) = -\frac{1}{2} \nabla \cdot (\rho \mathbf{u}) |\mathbf{u}|^2 + \rho \mathbf{u} \cdot \frac{\partial \mathbf{u}}{\partial t} \\ &= -\frac{1}{2} \nabla \cdot (\rho \mathbf{u} |\mathbf{u}|^2) + \mathbf{u} \cdot \left( \rho \frac{\partial \mathbf{u}}{\partial t} + \rho \mathbf{u} \cdot \nabla \mathbf{u} \right) \\ &= -\frac{1}{2} \nabla \cdot (\rho_r \mathbf{u}_r |\mathbf{u}_r|^2) - \nabla \cdot (\mathbf{p}_r \mathbf{u}_r) + (\mathbf{p}_r \cdot \nabla) \cdot \mathbf{u}_r + \mathbf{j}_r \cdot \mathbf{E}, \end{aligned} \quad (2.107)$$

and for the electromagnetic energy

$$\begin{aligned} \frac{1}{8\pi} \frac{\partial}{\partial t} (|\mathbf{B}|^2 + |\mathbf{E}|^2) &= \frac{1}{4\pi} \frac{\partial}{\partial t} \left( \mathbf{B} \cdot \frac{\partial \mathbf{B}}{\partial t} + \mathbf{E} \cdot \frac{\partial \mathbf{E}}{\partial t} \right) \\ &= -\frac{c}{4\pi} (\mathbf{B} \cdot (\nabla \times \mathbf{E}) - \mathbf{E} \cdot (\nabla \times \mathbf{B})) - \mathbf{j} \cdot \mathbf{E} \\ &= -\frac{c}{4\pi} \nabla \cdot (\mathbf{E} \times \mathbf{B}) - \mathbf{j} \cdot \mathbf{E}. \end{aligned} \quad (2.108)$$

Therefore, by integrating over the entire volume and assuming periodic boundary conditions, the total conservation of energy can be expressed as (Yang *et al.* 2017b)

$$\frac{\partial}{\partial t} E_r^{\text{int}} = -\langle (\mathbf{p}_r \cdot \nabla) \cdot \mathbf{u}_r \rangle; \quad (2.109)$$

$$\frac{\partial}{\partial t} E_r^{\text{kin}} = +\langle (\mathbf{p}_r \cdot \nabla) \cdot \mathbf{u}_r \rangle + \langle \mathbf{j}_r \cdot \mathbf{E} \rangle; \quad (2.110)$$

$$\frac{\partial}{\partial t} E^{\text{mag}} = -\langle \mathbf{j} \cdot \mathbf{E} \rangle, \quad (2.111)$$

which beautifully clarifies how the energy can be transferred. Nevertheless, we point out that it is exactly the splitting into the parallel and perpendicular heating that is so complicated, since even with periodic boundary conditions, only a few terms in the pressure equations (2.91), (2.92) vanish. Considering anisotropic heating, and splitting the internal (thermal) energy into parallel and perpendicular parts

$$E_r^{\text{int}} = E_{\parallel r}^{\text{int}} + E_{\perp r}^{\text{int}}; \quad E_{\parallel r}^{\text{int}} = \frac{1}{2} \langle p_{\parallel r} \rangle; \quad E_{\perp r}^{\text{int}} = \langle p_{\perp r} \rangle, \quad (2.112)$$

the conservation of total energy can be expressed as

$$\begin{aligned} \frac{\partial}{\partial t} E_{\parallel r}^{\text{int}} = & -\langle p_{\parallel r} \hat{\mathbf{b}} \cdot \nabla \mathbf{u}_r \cdot \hat{\mathbf{b}} \rangle - \frac{1}{2} \langle \hat{\mathbf{b}} \cdot (\nabla \cdot \mathbf{q}_r) \cdot \hat{\mathbf{b}} \rangle \\ & + \left\langle \frac{\hat{\mathbf{b}}}{|\mathbf{B}|} \cdot \boldsymbol{\Pi}_r \cdot \frac{d_r \mathbf{B}}{dt} \right\rangle - \langle \hat{\mathbf{b}} \cdot (\boldsymbol{\Pi}_r \cdot \nabla \mathbf{u}_r) \cdot \hat{\mathbf{b}} \rangle; \end{aligned} \quad (2.113)$$

$$\begin{aligned} \frac{\partial}{\partial t} E_{\perp r}^{\text{int}} = & -\langle p_{\perp r} \nabla \cdot \mathbf{u}_r \rangle + \langle p_{\perp r} \hat{\mathbf{b}} \cdot \nabla \mathbf{u}_r \cdot \hat{\mathbf{b}} \rangle + \frac{1}{2} \langle \hat{\mathbf{b}} \cdot (\nabla \cdot \mathbf{q}_r) \cdot \hat{\mathbf{b}} \rangle \\ & - \left\langle \frac{\hat{\mathbf{b}}}{|\mathbf{B}|} \cdot \boldsymbol{\Pi}_r \cdot \frac{d_r \mathbf{B}}{dt} \right\rangle + \langle \hat{\mathbf{b}} \cdot (\boldsymbol{\Pi}_r \cdot \nabla \mathbf{u}_r) \cdot \hat{\mathbf{b}} \rangle - \langle (\boldsymbol{\Pi}_r \cdot \nabla) \cdot \mathbf{u}_r \rangle; \end{aligned} \quad (2.114)$$

$$\frac{\partial}{\partial t} E_r^{\text{kin}} = +\langle (\mathbf{p}_r \cdot \nabla) \cdot \mathbf{u}_r \rangle + \langle \mathbf{j}_r \cdot \mathbf{E} \rangle; \quad (2.115)$$

$$\frac{\partial}{\partial t} E^{\text{mag}} = -\langle \mathbf{j} \cdot \mathbf{E} \rangle, \quad (2.116)$$

where

$$\langle (\mathbf{p}_r \cdot \nabla) \cdot \mathbf{u}_r \rangle = \langle p_{\parallel r} \hat{\mathbf{b}} \cdot \nabla \mathbf{u}_r \cdot \hat{\mathbf{b}} \rangle - \langle p_{\perp r} \hat{\mathbf{b}} \cdot \nabla \mathbf{u}_r \cdot \hat{\mathbf{b}} \rangle + \langle p_{\perp r} \nabla \cdot \mathbf{u}_r \rangle + \langle (\boldsymbol{\Pi}_r \cdot \nabla) \cdot \mathbf{u}_r \rangle. \quad (2.117)$$

The anisotropic plasma heating is obviously a very complicated process.

### 3. CGL description

The CGL model is the simplest possible fluid model that incorporates anisotropic temperatures. In contrast to MHD, CGL contains two pressure equations. For very good MHD reviews, from the perspective of turbulence in the solar wind, and the use of MHD in modelling heliospheric and astrophysical plasmas, we recommend the reviews by Goldstein, Roberts & Matthaeus (1995), Tu & Marsch (1995), Zank (1999), Zhou, Matthaeus & Dmitruk (2004) and Bruno & Carbone (2013). The CGL model is obtained by using the parallel and perpendicular pressure equations (2.74), (2.75) and prescribing zero heat flux  $q_{\parallel} = q_{\perp} = 0$ ,  $\mathbf{q}^{\text{ng}} = 0$ , and by neglecting the FLR pressure tensor  $\boldsymbol{\Pi} = 0$ . The pressure equations greatly simplify to the form

$$\frac{\partial p_{\parallel}}{\partial t} + \nabla \cdot (p_{\parallel} \mathbf{u}) + 2p_{\parallel} \hat{\mathbf{b}} \cdot \nabla \mathbf{u} \cdot \hat{\mathbf{b}} = 0; \quad (3.1)$$

$$\frac{\partial p_{\perp}}{\partial t} + \nabla \cdot (p_{\perp} \mathbf{u}) + p_{\perp} \nabla \cdot \mathbf{u} - p_{\perp} \hat{\mathbf{b}} \cdot \nabla \mathbf{u} \cdot \hat{\mathbf{b}} = 0. \quad (3.2)$$

By using the notation for the convective derivative  $d/dt = \partial/\partial t + \mathbf{u} \cdot \nabla$ , the pressure equations can be rewritten as

$$\frac{dp_{\parallel}}{dt} + p_{\parallel} \nabla \cdot \mathbf{u} + 2p_{\parallel} \hat{\mathbf{b}} \cdot \nabla \mathbf{u} \cdot \hat{\mathbf{b}} = 0; \quad (3.3)$$

$$\frac{dp_{\perp}}{dt} + 2p_{\perp} \nabla \cdot \mathbf{u} - p_{\perp} \hat{\mathbf{b}} \cdot \nabla \mathbf{u} \cdot \hat{\mathbf{b}} = 0. \quad (3.4)$$

Note that the CGL pressure equations are fully nonlinear and that the nonlinearities are actually of fourth order, i.e. CGL pressure equations are more nonlinear than the usual MHD pressure equation  $dp/dt + \gamma p \nabla \cdot \mathbf{u} = 0$ , where the nonlinearity is only

of second order. At first sight, expressions  $\hat{\mathbf{b}} \cdot \nabla \mathbf{u} \cdot \hat{\mathbf{b}}$  might appear to be a little unusual when written without any brackets. From MHD, one is familiar with nonlinear expressions of the type  $\mathbf{u} \cdot \nabla \mathbf{u}$ , which are also usually written without any brackets and can be interpreted in two equivalent forms, either as  $(\mathbf{u} \cdot \nabla) \mathbf{u}$  (a scalar  $\mathbf{u} \cdot \nabla$  applied on a vector  $\mathbf{u}$ ; or  $\mathbf{u} \cdot (\nabla \mathbf{u})$  (a vector  $\mathbf{u}$  multiplied by a matrix  $\nabla \mathbf{u}$ ). In a similar fashion, the expression  $\hat{\mathbf{b}} \cdot \nabla \mathbf{u} \cdot \hat{\mathbf{b}}$  is meant to be interpreted as  $(\hat{\mathbf{b}} \cdot \nabla) \mathbf{u} \cdot \hat{\mathbf{b}}$  or  $\hat{\mathbf{b}} \cdot (\nabla \mathbf{u}) \cdot \hat{\mathbf{b}}$ ; but importantly, it is not meant to be interpreted as  $\hat{\mathbf{b}} \cdot \nabla(\mathbf{u} \cdot \hat{\mathbf{b}})$ , because the derivative is meant to act only on  $\mathbf{u}$  in this case. For direct numerical simulations, the fourth-order nonlinearity explicitly reads

$$\begin{aligned} \hat{\mathbf{b}} \cdot \nabla \mathbf{u} \cdot \hat{\mathbf{b}} &= \hat{b}_i (\nabla \mathbf{u})_{ij} \hat{b}_j = \hat{b}_i \hat{b}_j \partial_i u_j = \hat{b}_x (\hat{b}_x \partial_x u_x + \hat{b}_y \partial_x u_y + \hat{b}_z \partial_x u_z) \\ &+ \hat{b}_y (\hat{b}_x \partial_y u_x + \hat{b}_y \partial_y u_y + \hat{b}_z \partial_y u_z) + \hat{b}_z (\hat{b}_x \partial_z u_x + \hat{b}_y \partial_z u_y + \hat{b}_z \partial_z u_z), \end{aligned} \quad (3.5)$$

and is of course a scalar. So far, we have not made any assumptions about particle species, and have only integrated the Vlasov equation and derived the density, momentum and scalar pressure equations, which can be done for each species separately. By re-introducing the species index  $r$ , where  $r=p$  for protons,  $r=e$  for electrons etc., the n-fluid CGL-type equations become

$$\frac{\partial \rho_r}{\partial t} + \nabla \cdot (\rho_r \mathbf{u}_r) = 0; \quad (3.6)$$

$$\frac{\partial \mathbf{u}_r}{\partial t} + \mathbf{u}_r \cdot \nabla \mathbf{u}_r + \frac{1}{\rho_r} \nabla \cdot \mathbf{p}_r - \frac{q_r}{m_r} \left( \mathbf{E} + \frac{1}{c} \mathbf{u}_r \times \mathbf{B} \right) = 0; \quad (3.7)$$

$$\frac{\partial p_{\parallel r}}{\partial t} + \nabla \cdot (\mathbf{u}_r p_{\parallel r}) + 2p_{\parallel r} \hat{\mathbf{b}} \cdot \nabla \mathbf{u}_r \cdot \hat{\mathbf{b}} = 0; \quad (3.8)$$

$$\frac{\partial p_{\perp r}}{\partial t} + \nabla \cdot (\mathbf{u}_r p_{\perp r}) + p_{\perp r} \nabla \cdot \mathbf{u}_r - p_{\perp r} \hat{\mathbf{b}} \cdot \nabla \mathbf{u}_r \cdot \hat{\mathbf{b}} = 0. \quad (3.9)$$

The above CGL pressure equations are valid regardless of the form of the induction equation  $\partial \mathbf{B} / \partial t$  and are therefore very general and valid for a wide range of CGL-type n-fluid models. It is advisable to keep them in this form when considering direct numerical simulations or solving linear dispersion relations. The above equations are accompanied by Maxwell's equations. Specifically, (i) Gauss's law  $\nabla \cdot \mathbf{E} = 4\pi \rho_c$  where the total charge density  $\rho_c = \sum_r q_r n_r$ , (ii)  $\nabla \cdot \mathbf{B} = 0$ , (iii) Faraday's induction equation  $\nabla \times \mathbf{E} = -(1/c)(\partial \mathbf{B} / \partial t)$  and (iv) Ampère's law  $\nabla \times \mathbf{B} = (4\pi/c) \mathbf{j} + (1/c)(\partial \mathbf{E} / \partial t)$ , where the total current  $\mathbf{j} = \sum_r q_r n_r \mathbf{u}_r$ . In many areas of space physics and astrophysics it is often very useful to eliminate very high frequency effects occurring at frequencies higher than the plasma frequency. This is achieved by prescribing local charge neutrality (so called quasi-neutrality since the plasma is still ionized) by  $\rho_c = 0$  and by neglecting the second term in Ampère's law  $(1/c)(\partial \mathbf{E} / \partial t)$  known as Maxwell's displacement current (or Maxwell's correction to Ampère's law). This eliminates high-frequency effects such as for example plasma oscillations, known as Langmuir waves. This assumption is also equivalent to assuming that the Alfvén speed is much less than the speed of light  $V_A/c \ll 1$ . It is important to emphasize that Gauss's law  $\nabla \cdot \mathbf{E} = 4\pi \rho_c$  is replaced by the quasi-neutrality condition  $\rho_c = 0$  and no requirement is imposed on  $\nabla \cdot \mathbf{E}$  (other than that it is small). For further discussion concerning the charge neutrality and  $\nabla \cdot \mathbf{E}$ , see for example Braginskii (1965) page 264, Webb *et al.* (2007) equation (7) and Passot & Sulem (2007) equation (A 1). Maxwell's equations in this approximation simplify to

$$\nabla \cdot \mathbf{B} = 0; \quad (3.10)$$

$$\frac{\partial \mathbf{B}}{\partial t} = -c \nabla \times \mathbf{E}; \quad (3.11)$$

$$\mathbf{j} = \sum_r q_r n_r \mathbf{u}_r = \frac{c}{4\pi} \nabla \times \mathbf{B}, \quad (3.12)$$

and this form of Maxwell's equations is considered in most of the fluid models that we discuss, regardless of future complexities arising from FLR corrections, heat flux, Landau damping and so on. The exception is the section in Part 2 'Electron Landau damping of the Langmuir mode', where the displacement current has to be retained. Here, we further focus only on a plasma consisting of protons and electrons, with charges  $q_p = e$  and  $q_e = -e$ , but models with more species can of course be considered. Charge neutrality implies that the proton and electron number densities are equal, i.e.  $n_p = n_e = n$ . By using the definition of current  $\mathbf{j} = en\mathbf{u}_p - en\mathbf{u}_e$ , the electron velocities are related to the proton velocities by

$$\mathbf{u}_e = \mathbf{u}_p - \frac{1}{en} \mathbf{j}, \quad (3.13)$$

which after using (3.12) yields

$$\mathbf{u}_e = \mathbf{u}_p - \frac{1}{en} \frac{c}{4\pi} \nabla \times \mathbf{B}. \quad (3.14)$$

By using the momentum equation (3.7) for electrons, the electric field can be expressed as

$$\mathbf{E} = -\frac{1}{c} \mathbf{u}_e \times \mathbf{B} - \frac{1}{en} \nabla \cdot \mathbf{p}_e - \frac{m_e}{e} \left( \frac{\partial \mathbf{u}_e}{\partial t} + \mathbf{u}_e \cdot \nabla \mathbf{u}_e \right). \quad (3.15)$$

The last term represents the effects of electron inertia and is for example responsible for the effect that at small spatial scales the frequency  $\omega$  of the parallel whistler mode converges to the electron cyclotron frequency  $\Omega_e$  instead of increasing to infinity. When considering only relatively low frequencies, it is very useful to neglect the electron inertia term.<sup>2</sup> This yields somewhat simpler analytic expressions, but most importantly, it provides a great benefit for direct numerical simulations since the time step is not restricted by the requirement to fully resolve the electron motion. This of course restricts the validity of a model to frequencies that are sufficiently smaller than the electron cyclotron frequency. By neglecting the electron inertia and using (3.14), the electric field can be expressed as

$$\mathbf{E} = -\frac{1}{c} \mathbf{u}_p \times \mathbf{B} + \frac{1}{4\pi en} (\nabla \times \mathbf{B}) \times \mathbf{B} - \frac{1}{en} \nabla \cdot \mathbf{p}_e. \quad (3.17)$$

<sup>2</sup>The magnitude of terms on the right-hand side of (3.15) can be easily compared to the advection term by estimating

$$\frac{\left| \frac{1}{en} \nabla \cdot \mathbf{p}_e \right|}{\left| \frac{1}{c} \mathbf{u}_e \times \mathbf{B} \right|} \sim \frac{\rho_e}{l} \frac{v_{\text{the}}}{u_e}; \quad \frac{\left| \frac{m_e}{e} \frac{d\mathbf{u}_e}{dt} \right|}{\left| \frac{1}{c} \mathbf{u}_e \times \mathbf{B} \right|} \sim \frac{1}{\Omega_e \tau}, \quad (3.16)$$

where  $l, \tau$  are the characteristic length scale and time scale,  $\rho_e$  the electron gyroradius and  $v_{\text{the}}$  the electron thermal velocity.

Using this expression in the proton momentum equation and in the induction equation eliminates the electric field from the entire system. The system of equations therefore reads

$$\frac{\partial \rho_p}{\partial t} + \nabla \cdot (\rho_p \mathbf{u}_p) = 0; \quad (3.18)$$

$$\frac{\partial \mathbf{u}_p}{\partial t} + \mathbf{u}_p \cdot \nabla \mathbf{u}_p + \frac{1}{\rho_p} \nabla \cdot (\mathbf{p}_p + \mathbf{p}_e) - \frac{1}{4\pi\rho_p} (\nabla \times \mathbf{B}) \times \mathbf{B} = 0; \quad (3.19)$$

$$\frac{\partial \mathbf{B}}{\partial t} = \nabla \times (\mathbf{u}_p \times \mathbf{B}) - \frac{c}{4\pi e} \nabla \times \left[ \frac{1}{n} (\nabla \times \mathbf{B}) \times \mathbf{B} \right] + \frac{c}{e} \nabla \times \left[ \frac{1}{n} \nabla \cdot \mathbf{p}_e \right], \quad (3.20)$$

and is accompanied by the parallel and perpendicular pressure equations (3.8), (3.9) for the proton and electron species. In the induction equation (3.20), the first term is the familiar term from the MHD induction equation and considering only this first term yields the classical CGL model (where in addition the electrons are prescribed to be cold to eliminate  $\mathbf{p}_e$  from the momentum equation). The second term in the induction equation represents the Hall term, and fluid models containing this term are generally referred to as Hall-CGL fluid models. The third term represents the electron pressure contributions and in CGL plasmas the term is in general non-zero, since

$$\nabla \cdot \mathbf{p}_e = (p_{\parallel e} - p_{\perp e}) (\hat{\mathbf{b}} \cdot \nabla \hat{\mathbf{b}} + \hat{\mathbf{b}} \nabla \cdot \hat{\mathbf{b}}) + \hat{\mathbf{b}} \hat{\mathbf{b}} \cdot \nabla (p_{\parallel e} - p_{\perp e}) + \nabla p_{\perp e}. \quad (3.21)$$

In contrast, by prescribing the electrons to be isotropic and isothermal (a typical assumption used in hybrid simulations) with pressure  $\mathbf{p}_e = p_e \mathbf{I}$ ,  $p_e = nT_e^{(0)}$ , so that  $\nabla \cdot \mathbf{p}_e = \nabla p_e = T_e^{(0)} \nabla n$ , eliminates the pressure term since  $\nabla \times [(1/n) \nabla n] = \nabla \times [\nabla \ln(n)] = 0$ . By prescribing the electrons to be isotropic but not isothermal, with pressure  $\mathbf{p}_e = p_e \mathbf{I}$ , the term initially does not disappear and it is equal to

$$\frac{c}{e} \nabla \times \left[ \frac{1}{n} \nabla \cdot \mathbf{p}_e \right] = -\frac{c}{en^2} (\nabla n) \times (\nabla p_e). \quad (3.22)$$

The term only disappears after the equation of state is prescribed, for example for  $p_e \sim n^\gamma$ . The term (3.22) is sometimes called the ‘battery’ term (Biermann 1950; Kulsrud *et al.* 1997; Khomenko *et al.* 2017), since it is argued that for any deviations from an ideal equation of state (caused for example by shocks), the term can produce magnetic field fluctuations even if there is no magnetic field initially.

### 3.1. Normalized equations and definitions

As with MHD and Hall-MHD, it is often useful to work with normalized equations. The density, velocity, magnetic field and pressure are normalized with respect to  $\rho_0$ ,  $u_0$ ,  $B_0$ ,  $p_{\parallel}^{(0)}$ . The length is normalized with respect to (for now) arbitrary  $x_0$  and the time with respect to  $t_0 = x_0/u_0$ . The normalized quantities are then

$$\tilde{\rho} = \frac{\rho}{\rho_0}; \quad \tilde{\mathbf{u}} = \frac{\mathbf{u}}{u_0}; \quad \tilde{\mathbf{B}} = \frac{\mathbf{B}}{B_0}; \quad \tilde{p}_{\parallel} = \frac{p_{\parallel}}{p_{\parallel}^{(0)}}; \quad \tilde{p}_{\perp} = \frac{p_{\perp}}{p_{\parallel}^{(0)}}; \quad (3.23)$$

$$\tilde{\mathbf{x}} = \frac{\mathbf{x}}{x_0}; \quad \tilde{\nabla} = \nabla x_0; \quad \tilde{t} = t \frac{u_0}{x_0}; \quad \frac{\partial}{\partial \tilde{t}} = \frac{\partial}{\partial t} \frac{x_0}{u_0}. \quad (3.24)$$

Obviously  $\tilde{n} = \tilde{\rho}$ . After dropping the tilde, the normalization yields the same density and pressure equations, whereas the normalized momentum and induction equation become

$$\frac{\partial \mathbf{u}_p}{\partial t} + \mathbf{u}_p \cdot \nabla \mathbf{u}_p + \frac{p_{\parallel}^{(0)}}{\rho_0 u_0^2 \rho_p} \frac{1}{\rho_p} \nabla \cdot (\mathbf{p}_p + \mathbf{p}_e) - \frac{V_A^2}{u_0^2} \frac{1}{\rho_p} (\nabla \times \mathbf{B}) \times \mathbf{B} = 0; \quad (3.25)$$

$$\frac{\partial \mathbf{B}}{\partial t} = \nabla \times (\mathbf{u} \times \mathbf{B}) - \frac{V_A^2}{\Omega_p u_0 x_0} \nabla \times \left[ \frac{1}{n} (\nabla \times \mathbf{B}) \times \mathbf{B} \right] + \frac{p_{\parallel}^{(0)}}{\Omega_p \rho_0 u_0 x_0} \nabla \times \left[ \frac{1}{n} \nabla \cdot \mathbf{p}_e \right], \quad (3.26)$$

where the Alfvén speed  $V_A = B_0 / \sqrt{4\pi\rho_0}$ . The momentum equation implies that it is beneficial to choose the (so far unspecified) normalizing velocity  $u_0$  to be the Alfvén speed  $u_0 = V_A$ . If the second and third terms are neglected in the induction equation (i.e. when the usual MHD induction equation is used), the CGL system is independent of length scale  $x_0$ , similarly to the MHD system. If the Hall term is considered, it is beneficial to choose the (so far unspecified) normalizing length scale  $x_0$  to be the ion inertial length

$$d_i = \frac{V_A}{\Omega_p}. \quad (3.27)$$

Therefore by choosing  $x_0 = d_i$ , the normalizations that we use in all fluid models in this paper read

$$\tilde{\mathbf{u}} = \frac{\mathbf{u}}{V_A}; \quad \tilde{\mathbf{x}} = \frac{\mathbf{x}}{d_i}; \quad \tilde{t} = t \frac{V_A}{d_i} = t \Omega_p. \quad (3.28)$$

The normalization also has the advantage that, when transferred to Fourier space, one obtains the dispersion relations for normalized wavenumber  $\tilde{k}$  and frequency  $\tilde{\omega}$  as

$$\tilde{k} = \frac{k V_A}{\Omega_p}, \quad \tilde{\omega} = \frac{\omega}{\Omega_p}. \quad (3.29)$$

It is useful to define the parallel and perpendicular thermal speeds

$$v_{\text{th}\parallel} = \sqrt{\frac{2T_{\parallel}^{(0)}}{m_p}}; \quad v_{\text{th}\perp} = \sqrt{\frac{2T_{\perp}^{(0)}}{m_p}}, \quad (3.30)$$

and the definition of parallel and perpendicular temperatures are  $T_{\parallel} = p_{\parallel}/n$  and  $T_{\perp} = p_{\perp}/n$ . We use the usual convention with the Boltzmann constant  $k_B = 1$ .<sup>3</sup> The parallel and perpendicular plasma beta are defined according to

$$\beta_{\parallel} = \frac{v_{\text{th}\parallel}^2}{V_A^2} = \frac{2p_{\parallel}^{(0)}}{\rho_0 V_A^2} = \frac{p_{\parallel}^{(0)}}{B_0^2/(8\pi)}; \quad \beta_{\perp} = \frac{v_{\text{th}\perp}^2}{V_A^2} = \frac{2p_{\perp}^{(0)}}{\rho_0 V_A^2} = \frac{p_{\perp}^{(0)}}{B_0^2/(8\pi)} = \beta_{\parallel} \frac{T_{\perp}^{(0)}}{T_{\parallel}^{(0)}}, \quad (3.31)$$

and later we will often use the abbreviation  $a_p$  for the proton temperature anisotropy ratio

$$a_p = \frac{T_{\perp}^{(0)}}{T_{\parallel}^{(0)}} = \frac{p_{\perp}^{(0)}}{p_{\parallel}^{(0)}}. \quad (3.32)$$

<sup>3</sup>Without this convention the temperature is defined as  $T_{\parallel\perp} = p_{\parallel\perp}/(nk_B)$  and the thermal speeds  $v_{\text{th}\parallel\perp} = \sqrt{2k_B T_{\parallel\perp}^{(0)}/m_p}$ , so the plasma  $\beta_{\parallel}$ ,  $\beta_{\perp}$  expressions are the same as above.



The normalized momentum and induction equations therefore read (tilde are dropped)

$$\frac{\partial \mathbf{u}_p}{\partial t} + \mathbf{u}_p \cdot \nabla \mathbf{u}_p + \frac{\beta_{\parallel}}{2} \frac{1}{\rho_p} \nabla \cdot (\mathbf{p}_p + \mathbf{p}_e) - \frac{1}{\rho_p} (\nabla \times \mathbf{B}) \times \mathbf{B} = 0; \quad (3.33)$$

$$\frac{\partial \mathbf{B}}{\partial t} = \nabla \times (\mathbf{u} \times \mathbf{B}) - \nabla \times \left[ \frac{1}{n} (\nabla \times \mathbf{B}) \times \mathbf{B} \right] + \frac{\beta_{\parallel}}{2} \nabla \times \left[ \frac{1}{n} \nabla \cdot \mathbf{p}_e \right]. \quad (3.34)$$

Another useful quantity that we will use later in the text is the proton Larmor radius (the gyroradius)  $\rho_i$ , defined according to

$$\rho_i = \frac{v_{\text{th}\perp}}{\Omega_p}. \quad (3.35)$$

We note that normalizations do not have to be done with respect to mean values, and normalizations with respect to fluctuating (turbulent) quantities are used for example in the nearly incompressible models of Zank *et al.* (2017) and Zank & Matthaeus (1993).

### 3.2. Classical CGL model with cold electrons

Here we want to consider the simplest possible CGL model for the proton species. We assume the electrons to be massless and cold ( $\mathbf{p}_e = 0$ ), and we also neglect the Hall term in the induction equation. Since only proton species are present, for simplicity we can drop the proton index ‘ $p$ ’. Let us work in physical units for a moment. The classical CGL model with cold electrons therefore reads (written in physical units)

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{u}) = 0; \quad (3.36)$$

$$\frac{\partial \mathbf{u}}{\partial t} + \mathbf{u} \cdot \nabla \mathbf{u} + \frac{1}{\rho} \nabla \cdot \mathbf{p} - \frac{1}{4\pi\rho} (\nabla \times \mathbf{B}) \times \mathbf{B} = 0; \quad (3.37)$$

$$\frac{\partial p_{\parallel}}{\partial t} + \nabla \cdot (p_{\parallel} \mathbf{u}) + 2p_{\parallel} \hat{\mathbf{b}} \cdot \nabla \mathbf{u} \cdot \hat{\mathbf{b}} = 0; \quad (3.38)$$

$$\frac{\partial p_{\perp}}{\partial t} + \nabla \cdot (p_{\perp} \mathbf{u}) + p_{\perp} \nabla \cdot \mathbf{u} - p_{\perp} \hat{\mathbf{b}} \cdot \nabla \mathbf{u} \cdot \hat{\mathbf{b}} = 0; \quad (3.39)$$

$$\frac{\partial \mathbf{B}}{\partial t} = \nabla \times (\mathbf{u} \times \mathbf{B}). \quad (3.40)$$

Similarly to the usual MHD model, the CGL equations are scale invariant and do not contain any information about the physical length scale. A length scale is introduced by considering the Hall term in the induction equation, and we discuss the Hall-CGL model in the next section. Also, the absence of the Hall term allows the CGL pressure equations to be rewritten into ‘conservative’ form as

$$\frac{d}{dt} \left( \frac{p_{\parallel} |\mathbf{B}|^2}{\rho^3} \right) = 0; \quad \frac{d}{dt} \left( \frac{p_{\perp}}{\rho |\mathbf{B}|} \right) = 0, \quad (3.41)$$

where  $d/dt$  is the convective derivative. This formulation is the most often cited form of the CGL pressure equations (often the  $|\mathbf{B}|$  is abbreviated only as  $B$ , which can

lead to confusion on how to correctly calculate the derivatives). We want to show the equivalence of (3.41) and (3.38), (3.39). We need the following identity,

$$\frac{\partial}{\partial t} |\mathbf{B}| = \frac{\partial}{\partial t} (B_x^2 + B_y^2 + B_z^2)^{1/2} = \frac{1}{|\mathbf{B}|} \mathbf{B} \cdot \frac{\partial \mathbf{B}}{\partial t} = \hat{\mathbf{b}} \cdot \frac{\partial \mathbf{B}}{\partial t}, \quad (3.42)$$

and the result is of course a scalar. Similarly for the spatial and convective derivative

$$\partial_i |\mathbf{B}| = \hat{\mathbf{b}} \cdot (\partial_i \mathbf{B}); \quad \frac{d}{dt} |\mathbf{B}| = \left( \frac{\partial}{\partial t} + u_i \partial_i \right) |\mathbf{B}| = \hat{\mathbf{b}} \cdot \frac{d\mathbf{B}}{dt}. \quad (3.43)$$

Later we will need

$$\frac{d}{dt} |\mathbf{B}|^2 = \frac{d}{dt} (B_x^2 + B_y^2 + B_z^2) = 2\mathbf{B} \cdot \frac{d\mathbf{B}}{dt}. \quad (3.44)$$

Now we can directly calculate

$$\begin{aligned} \frac{d}{dt} \left( \frac{p_{\perp}}{\rho |\mathbf{B}|} \right) &= \frac{1}{\rho |\mathbf{B}|} \frac{dp_{\perp}}{dt} - \frac{p_{\perp}}{|\mathbf{B}| \rho^2} \frac{d\rho}{dt} - \frac{p_{\perp}}{\rho |\mathbf{B}|^2} \frac{d|\mathbf{B}|}{dt} \\ &= \frac{1}{\rho |\mathbf{B}|} \left( \frac{dp_{\perp}}{dt} - \frac{p_{\perp}}{\rho} \frac{d\rho}{dt} - \frac{p_{\perp}}{|\mathbf{B}|} \frac{d|\mathbf{B}|}{dt} \right), \end{aligned} \quad (3.45)$$

and, by using the density equation  $d\rho/dt = -\rho \nabla \cdot \mathbf{u}$  and the identity (3.43), we obtain

$$\frac{d}{dt} \left( \frac{p_{\perp}}{\rho |\mathbf{B}|} \right) = \frac{1}{\rho |\mathbf{B}|} \left( \frac{dp_{\perp}}{dt} + p_{\perp} \nabla \cdot \mathbf{u} - \frac{p_{\perp}}{|\mathbf{B}|} \hat{\mathbf{b}} \cdot \frac{d\mathbf{B}}{dt} \right). \quad (3.46)$$

The above equation is completely general since we did not use the induction equation so far and we will use this equation in the next section when we consider the Hall-CGL model. In a similar way one derives a completely general identity

$$\begin{aligned} \frac{d}{dt} \left( \frac{p_{\parallel} |\mathbf{B}|^2}{\rho^3} \right) &= \frac{|\mathbf{B}|^2}{\rho^3} \left( \frac{dp_{\parallel}}{dt} + \frac{p_{\parallel}}{|\mathbf{B}|^2} \frac{d|\mathbf{B}|^2}{dt} - 3 \frac{p_{\parallel}}{\rho} \frac{d\rho}{dt} \right) \\ &= \frac{|\mathbf{B}|^2}{\rho^3} \left( \frac{dp_{\parallel}}{dt} + 3p_{\parallel} \nabla \cdot \mathbf{u} + \frac{2p_{\parallel}}{|\mathbf{B}|} \hat{\mathbf{b}} \cdot \frac{d\mathbf{B}}{dt} \right). \end{aligned} \quad (3.47)$$

Now, we use the induction equation

$$\frac{\partial \mathbf{B}}{\partial t} = \nabla \times (\mathbf{u} \times \mathbf{B}) = \underbrace{\mathbf{u} \nabla \cdot \mathbf{B}}_{=0} - \mathbf{B} \nabla \cdot \mathbf{u} + \mathbf{B} \cdot \nabla \mathbf{u} - \mathbf{u} \cdot \nabla \mathbf{B}, \quad (3.48)$$

that allows us to calculate

$$\frac{d\mathbf{B}}{dt} = \frac{\partial \mathbf{B}}{\partial t} + \mathbf{u} \cdot \nabla \mathbf{B} = -\mathbf{B} \nabla \cdot \mathbf{u} + \mathbf{B} \cdot \nabla \mathbf{u}; \quad (3.49)$$

$$\frac{\hat{\mathbf{b}}}{|\mathbf{B}|} \cdot \frac{d\mathbf{B}}{dt} = -\nabla \cdot \mathbf{u} + \hat{\mathbf{b}} \cdot \nabla \mathbf{u} \cdot \hat{\mathbf{b}}, \quad (3.50)$$

and which, when used in (3.47)–(3.46) yields the final result

$$\frac{d}{dt} \left( \frac{p_{\parallel} |\mathbf{B}|^2}{\rho^3} \right) = \frac{|\mathbf{B}|^2}{\rho^3} \left( \frac{dp_{\parallel}}{dt} + p_{\parallel} \nabla \cdot \mathbf{u} + 2p_{\parallel} \hat{\mathbf{b}} \cdot \nabla \mathbf{u} \cdot \hat{\mathbf{b}} \right) = 0; \quad (3.51)$$

$$\frac{d}{dt} \left( \frac{p_{\perp}}{\rho|\mathbf{B}|} \right) = \frac{1}{\rho|\mathbf{B}|} \left( \frac{dp_{\perp}}{dt} + 2p_{\perp} \nabla \cdot \mathbf{u} - p_{\perp} \hat{\mathbf{b}} \cdot \nabla \mathbf{u} \cdot \hat{\mathbf{b}} \right) = 0. \quad (3.52)$$

Equations (3.51)–(3.52) verify that the CGL pressure equations (3.41) written in the ‘conservative’ form and the directly derived CGL equations (3.38)–(3.39) are equivalent. We will see that the right-hand side of the usual CGL equations (3.41) is non-zero if the Hall term is considered, and that it is further modified by the heat flux contributions and by the FLR stress forces.

### 3.3. Linearized CGL equations

To obtain the dispersion relations, the equations need to be linearized and transformed to Fourier space. Equations (3.36)–(3.40) are linearized with respect to mean values  $\rho_0$ ;  $\mathbf{u}^{(0)} \equiv \langle \mathbf{u} \rangle = 0$ ;  $p_{\perp}^{(0)}$ ;  $\mathbf{B}_0 = (0, 0, B_0)$ ;  $\hat{\mathbf{b}}_0 = (0, 0, 1)$  where the magnetic field  $B_0$  is assumed to be in the  $z$ -direction. We assume that the mean value of the velocity is zero (the zero mean velocity value  $\mathbf{u}^{(0)} = 0$  should not be confused with the normalizing velocity as in (3.23), which is later chosen to be  $V_A$ ). The terms  $\hat{\mathbf{b}} \cdot \nabla \mathbf{u} \cdot \hat{\mathbf{b}}$  are linearized as  $\hat{\mathbf{b}} \cdot \nabla \mathbf{u} \cdot \hat{\mathbf{b}} \stackrel{\text{lin}}{=} \partial_z u_z$  and the linearized system reads

$$\frac{\partial \rho}{\partial t} + \rho_0 \nabla \cdot \mathbf{u} = 0; \quad (3.53)$$

$$\frac{\partial \mathbf{u}}{\partial t} + \frac{1}{\rho_0} \nabla \cdot \mathbf{p} - \frac{1}{4\pi\rho_0} (\nabla \times \mathbf{B}) \times \mathbf{B}_0 = 0; \quad (3.54)$$

$$\frac{\partial p_{\parallel}}{\partial t} + p_{\parallel}^{(0)} \nabla \cdot \mathbf{u} + 2p_{\parallel}^{(0)} \partial_z u_z = 0; \quad (3.55)$$

$$\frac{\partial p_{\perp}}{\partial t} + 2p_{\perp}^{(0)} \nabla \cdot \mathbf{u} - p_{\perp}^{(0)} \partial_z u_z = 0; \quad (3.56)$$

$$\frac{\partial \mathbf{B}}{\partial t} = \nabla \times (\mathbf{u} \times \mathbf{B}_0). \quad (3.57)$$

The divergence of the pressure tensor is

$$(\nabla \cdot \mathbf{p})_i = \partial_j p_{ji} = \hat{b}_i \hat{b}_j \partial_j p_{\parallel} + (\delta_{ij} - \hat{b}_i \hat{b}_j) \partial_j p_{\perp} + (p_{\parallel} - p_{\perp}) (\hat{b}_i \partial_j \hat{b}_j + \hat{b}_j \partial_j \hat{b}_i), \quad (3.58)$$

and the first step of linearization yields, by components,

$$(\nabla \cdot \mathbf{p})_x \stackrel{\text{lin}}{=} \partial_x p_{\perp} + (p_{\parallel}^{(0)} - p_{\perp}^{(0)}) \partial_z \hat{b}_x; \quad (3.59)$$

$$(\nabla \cdot \mathbf{p})_y \stackrel{\text{lin}}{=} \partial_y p_{\perp} + (p_{\parallel}^{(0)} - p_{\perp}^{(0)}) \partial_z \hat{b}_y; \quad (3.60)$$

$$(\nabla \cdot \mathbf{p})_z \stackrel{\text{lin}}{=} \partial_z p_{\parallel} + (p_{\parallel}^{(0)} - p_{\perp}^{(0)}) (\nabla \cdot \hat{\mathbf{b}} + \partial_z \hat{b}_z), \quad (3.61)$$

which can be conveniently written in matrix notation as

$$\nabla \cdot \mathbf{p} \stackrel{\text{lin}}{=} \nabla \cdot \begin{pmatrix} p_{\perp} & 0 & 0 \\ 0 & p_{\perp} & 0 \\ 0 & 0 & p_{\parallel} \end{pmatrix} + (p_{\parallel}^{(0)} - p_{\perp}^{(0)}) \nabla \cdot \begin{pmatrix} 0 & 0 & \hat{b}_x \\ 0 & 0 & \hat{b}_y \\ \hat{b}_x & \hat{b}_y & 2\hat{b}_z \end{pmatrix} \Big|_{\text{lin}}. \quad (3.62)$$

Notice that, since  $\partial_i |\mathbf{B}| = \hat{\mathbf{b}} \cdot (\partial_i \mathbf{B})$ , the derivatives of unit vectors are calculated according to

$$\partial_i \hat{\mathbf{b}} = \partial_i \frac{\mathbf{B}}{|\mathbf{B}|} = \frac{1}{|\mathbf{B}|} \left[ \partial_i \mathbf{B} - \frac{\mathbf{B}}{|\mathbf{B}|} \partial_i |\mathbf{B}| \right] = \frac{1}{|\mathbf{B}|} [\partial_i \mathbf{B} - \hat{\mathbf{b}} (\hat{\mathbf{b}} \cdot (\partial_i \mathbf{B}))], \quad (3.63)$$

and in the linear approximation one has

$$\partial_i \hat{\mathbf{b}} \stackrel{\text{lin}}{=} \frac{1}{B_0} [\partial_i \mathbf{B} - \hat{\mathbf{b}}_0 \partial_i B_z], \quad (3.64)$$

which by components reads

$$\partial_i \hat{b}_x \stackrel{\text{lin}}{=} \frac{1}{B_0} \partial_i B_x; \quad \partial_i \hat{b}_y \stackrel{\text{lin}}{=} \frac{1}{B_0} \partial_i B_y; \quad \partial_i \hat{b}_z \stackrel{\text{lin}}{=} \frac{1}{B_0} (\partial_i B_z - \partial_z B_z), \quad (3.65)$$

where, importantly,  $\partial_z \hat{b}_z \stackrel{\text{lin}}{=} 0$ . The divergence of the pressure tensor in the linear approximation is therefore expressed as

$$\nabla \cdot \mathbf{p} \stackrel{\text{lin}}{=} \nabla \cdot \begin{pmatrix} p_\perp & 0 & 0 \\ 0 & p_\perp & 0 \\ 0 & 0 & p_\parallel \end{pmatrix} + \frac{1}{B_0} (p_\parallel^{(0)} - p_\perp^{(0)}) \nabla \cdot \begin{pmatrix} 0 & 0 & B_x \\ 0 & 0 & B_y \\ B_x & B_y & 0 \end{pmatrix}, \quad (3.66)$$

where the second term contributes if the temperature anisotropy is present, i.e. when  $p_\parallel^{(0)} \neq p_\perp^{(0)}$ . The vectors  $(\nabla \times \mathbf{B}) \times \mathbf{B}_0$  and  $\nabla \times (\mathbf{u} \times \mathbf{B}_0)$  are straightforward to calculate and are

$$(\nabla \times \mathbf{B}) \times \mathbf{B}_0 = B_0 \begin{pmatrix} -\partial_x B_z + \partial_z B_x \\ -\partial_y B_z + \partial_z B_y \\ 0 \end{pmatrix}; \quad \nabla \times (\mathbf{u} \times \mathbf{B}_0) = B_0 \begin{pmatrix} \partial_z u_x \\ \partial_z u_y \\ -\partial_x u_x - \partial_y u_y \end{pmatrix}. \quad (3.67)$$

The entire set of linearized CGL equations reads

$$\frac{\partial \rho}{\partial t} + \rho_0 \nabla \cdot \mathbf{u} = 0; \quad (3.68)$$

$$\frac{\partial u_x}{\partial t} + \frac{1}{\rho_0} \partial_x p_\perp + \frac{1}{B_0 \rho_0} (p_\parallel^{(0)} - p_\perp^{(0)}) \partial_z B_x - \frac{B_0}{4\pi \rho_0} (-\partial_x B_z + \partial_z B_x) = 0; \quad (3.69)$$

$$\frac{\partial u_y}{\partial t} + \frac{1}{\rho_0} \partial_y p_\perp + \frac{1}{B_0 \rho_0} (p_\parallel^{(0)} - p_\perp^{(0)}) \partial_z B_y - \frac{B_0}{4\pi \rho_0} (-\partial_y B_z + \partial_z B_y) = 0; \quad (3.70)$$

$$\frac{\partial u_z}{\partial t} + \frac{1}{\rho_0} \partial_z p_\parallel + \frac{1}{B_0 \rho_0} (p_\parallel^{(0)} - p_\perp^{(0)}) (\partial_x B_x + \partial_y B_y) = 0; \quad (3.71)$$

$$\frac{\partial B_x}{\partial t} = B_0 \partial_z u_x; \quad \frac{\partial B_y}{\partial t} = B_0 \partial_z u_y; \quad \frac{\partial B_z}{\partial t} = -B_0 \partial_x u_x - B_0 \partial_y u_y; \quad (3.72)$$

$$\frac{\partial p_\parallel}{\partial t} + p_\parallel^{(0)} (\partial_x u_x + \partial_y u_y) + 3p_\parallel^{(0)} \partial_z u_z = 0; \quad (3.73)$$

$$\frac{\partial p_\perp}{\partial t} + 2p_\perp^{(0)} (\partial_x u_x + \partial_y u_y) + p_\perp^{(0)} \partial_z u_z = 0. \quad (3.74)$$

Additionally, the 3 components of the induction equation are not independent, but satisfy the  $\nabla \cdot \mathbf{B} = 0$  constraint. In order to find the linear dispersion relations, without any loss of generality, we consider wave propagation in the  $x$ - $z$  plane and assume that there is no variation with respect to the  $y$ -coordinate (i.e. one assumes that all expressions containing  $\partial_y$  are zero; alternatively, one can consider propagation in the  $y$ - $z$  plane, where all expressions containing  $\partial_x$  would be zero). By exploring the above system, it is noteworthy that the density equation does not play any role in determining the dispersion relation, since no other equation contains the variable  $\rho$ .

## 3.4. Alfvén mode

On considering (3.68)–(3.74) written in the  $x$ – $z$  plane with  $\partial_y = 0$ , it is apparent that the equations for  $u_y$  and  $B_y$  decouple from the entire system and applying  $\partial/\partial t$  to the equation for  $u_y$  yields

$$\frac{\partial^2 u_y}{\partial t^2} - \frac{1}{\rho_0} \left( \frac{B_0^2}{4\pi} - p_{\parallel}^{(0)} + p_{\perp}^{(0)} \right) \partial_z^2 u_y = 0, \quad (3.75)$$

which is a wave equation describing the propagation of (generally oblique) Alfvén waves in the CGL description. The same wave equation can be obtained for the component  $B_y$ . For isotropic mean pressures  $p_{\parallel}^{(0)} = p_{\perp}^{(0)}$  the equation is equivalent to the MHD description with the usual Alfvén speed  $V_A = B_0/(4\pi\rho_0)^{1/2}$ . Since

$$\frac{\beta_{\parallel}}{2} = \frac{p_{\parallel}^{(0)}}{\rho_0 V_A^2}; \quad a_p \frac{\beta_{\parallel}}{2} = \frac{p_{\perp}^{(0)}}{\rho_0 V_A^2}, \quad (3.76)$$

where  $a_p$  is the temperature ratio (3.32), the Alfvén wave equation can be rewritten as

$$\frac{\partial^2 u_y}{\partial t^2} - V_A^2 \left[ 1 + \frac{\beta_{\parallel}}{2} (a_p - 1) \right] \partial_z^2 u_y = 0. \quad (3.77)$$

Considering a wave propagating obliquely in the  $x$ – $z$  plane with the wavevector  $\mathbf{k} = (k_{\perp}, 0, k_{\parallel})$  at an angle  $\theta$  with respect to the  $z$ -direction (the  $B_0$  direction), the parallel and perpendicular wavenumbers are defined as

$$k_{\parallel} = k \cos \theta; \quad k_{\perp} = k \sin \theta, \quad (3.78)$$

and the transformation to Fourier space is performed according to (see appendix A)

$$\frac{\partial}{\partial t} \rightarrow -i\omega; \quad \partial_z \rightarrow ik_{\parallel}; \quad \partial_x \rightarrow ik_{\perp}. \quad (3.79)$$

The wave equation (3.77) is then

$$\omega^2 u_y - V_A^2 \left[ 1 + \frac{\beta_{\parallel}}{2} (a_p - 1) \right] k^2 \cos^2 \theta u_y = 0, \quad (3.80)$$

which yields the dispersion relation for the Alfvén waves

$$\omega = \pm k \cos \theta V_A \sqrt{1 + \frac{1}{2} \beta_{\parallel} (a_p - 1)} = \pm k_{\parallel} V_A \sqrt{1 + \frac{1}{2} \beta_{\parallel} (a_p - 1)}. \quad (3.81)$$

It is important to emphasize that, similarly to MHD, the Alfvén mode propagates in all the oblique directions, with a frequency  $\omega$  that is proportional to the projection of the wavenumber  $\mathbf{k}$  in the direction of  $B_0$ , i.e.  $k_{\parallel} = k \cos \theta$ . If the expression under the square root in (3.81) becomes negative, i.e. if  $1 + \beta_{\parallel} (a_p - 1)/2 < 0$ , the frequency  $\omega$  is imaginary and the Alfvén mode becomes unstable. This instability, that affects the obliquely propagating Alfvén mode, is known as the oblique firehose instability. The necessary (but not sufficient) condition for the instability to develop

is  $a_p = T_{\perp}^{(0)}/T_{\parallel}^{(0)} < 1$ . The firehose instability criterion can be rewritten in many forms, the most common being

$$1 - \frac{\beta_{\parallel}}{2} + \frac{1}{2}\beta_{\parallel}a_p < 0; \quad \frac{T_{\perp}^{(0)}}{T_{\parallel}^{(0)}} + \frac{2}{\beta_{\parallel}} - 1 < 0; \quad \beta_{\parallel} - \beta_{\perp} > 2; \quad p_{\parallel}^{(0)} - p_{\perp}^{(0)} > \frac{B_0^2}{4\pi}. \quad (3.82)$$

The first expression directly implies that, if  $\beta_{\parallel} \leq 2$ , no firehose instability can exist. The firehose instability therefore exists only for a relatively high plasma beta  $\beta_{\parallel} > 2$  and only if the parallel temperature is higher than the perpendicular temperature. Importantly, the firehose instability criterion in the CGL model is equivalent to the one found from linear kinetic theory in the long-wavelength limit (for example, Rosenbluth (1956), Chandrasekhar, Kaufman & Watson (1958), Parker (1958a) and Vedenov & Sagdeev (1958)). Alfvén wave propagation only affects components  $u_y$  and  $B_y$  which are decoupled from the rest of the system, and yields the other (eigenvector) components  $u_x, u_z, B_x, B_z$  as zero. The  $u_y$  component does not enter the density equation and Alfvén mode fluctuations therefore do not produce any density fluctuations or pressure fluctuations. The Alfvén mode is therefore incompressible. It is also useful to define the magnetic compressibility  $\chi_b(k_{\parallel}, k_{\perp})$  in Fourier space, that measures the relative ratio of magnetic energy in the parallel component versus the total magnetic energy

$$\chi_b(k_{\parallel}, k_{\perp}) = \frac{|B_z|^2}{|B_x|^2 + |B_y|^2 + |B_z|^2}. \quad (3.83)$$

Since the Alfvén mode produces only  $B_y$  fluctuations, with  $B_x = B_z = 0$ , the magnetic compressibility for this mode is

$$\chi_b(k_{\parallel}, k_{\perp}) = 0, \quad (3.84)$$

for all the propagation directions. Sometimes the magnetic compressibility is defined as a ratio of parallel and perpendicular energies  $|B_z|^2/(|B_x|^2 + |B_y|^2)$ , this quantity however has a slight disadvantage in that the values can become very large (and actually equal to infinity), which makes plotting of this quantity problematic. In contrast, the magnetic compressibility as defined in (3.83), has a nicely bounded range of values between 0 and 1. A similar quantity is defined for the ratio of the energy in velocity fluctuations

$$\chi_u(k_{\parallel}, k_{\perp}) = \frac{|u_z|^2}{|u_x|^2 + |u_y|^2 + |u_z|^2}, \quad (3.85)$$

and the Alfvén mode has  $\chi_u = 0$  for all propagation directions, since only the perpendicular velocity component  $u_y$  is non-zero.

### 3.5. Slow and fast modes

By applying  $\partial/\partial t$  to the equations for  $\partial u_x/\partial t$  and  $\partial u_z/\partial t$  and using the evolution equations for  $B_x, B_z, p_{\parallel}, p_{\perp}$  one obtains

$$\frac{\partial^2 u_x}{\partial t^2} - \frac{p_{\perp}^{(0)}}{\rho_0} (2\partial_x^2 u_x + \partial_x \partial_z u_z) + \left( \frac{p_{\parallel}^{(0)}}{\rho_0} - \frac{p_{\perp}^{(0)}}{\rho_0} - V_A^2 \right) \partial_z^2 u_x - V_A^2 \partial_x^2 u_x = 0; \quad (3.86)$$

$$\frac{\partial^2 u_z}{\partial t^2} - \frac{3p_{\parallel}^{(0)}}{\rho_0} \partial_z^2 u_z - \frac{p_{\perp}^{(0)}}{\rho_0} \partial_x \partial_z u_x = 0. \quad (3.87)$$

The easiest way to solve this coupled system is to transform to Fourier space, which yields

$$\begin{pmatrix} \omega^2 - \left( \frac{2p_{\perp}^{(0)}}{\rho_0} + V_A^2 \right) k_{\perp}^2 + \left( \frac{p_{\parallel}^{(0)}}{\rho_0} - \frac{p_{\perp}^{(0)}}{\rho_0} - V_A^2 \right) k_{\parallel}^2; & -\frac{p_{\perp}^{(0)}}{\rho_0} k_{\perp} k_{\parallel} \\ -\frac{p_{\perp}^{(0)}}{\rho_0} k_{\perp} k_{\parallel}; & \omega^2 - \frac{3p_{\parallel}^{(0)}}{\rho_0} k_{\parallel}^2 \end{pmatrix} \begin{pmatrix} u_x \\ u_z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \quad (3.88)$$

The system has non-trivial solutions only if the determinant of the matrix is zero, yielding the CGL dispersion relation at fourth order in frequency  $\omega$  that contains the slow and fast modes

$$\begin{aligned} \omega^4 - \omega^2 \left[ \left( \frac{2p_{\perp}^{(0)}}{\rho_0} + V_A^2 \right) k_{\perp}^2 + \left( \frac{2p_{\parallel}^{(0)}}{\rho_0} + \frac{p_{\perp}^{(0)}}{\rho_0} + V_A^2 \right) k_{\parallel}^2 \right] \\ + 3k_{\parallel}^2 \frac{p_{\parallel}^{(0)}}{\rho_0} \left[ \left( \frac{2p_{\perp}^{(0)}}{\rho_0} - \frac{p_{\perp}^{(0)2}}{3p_{\parallel}^{(0)}\rho_0} + V_A^2 \right) k_{\perp}^2 - \left( \frac{p_{\parallel}^{(0)}}{\rho_0} - \frac{p_{\perp}^{(0)}}{\rho_0} - V_A^2 \right) k_{\parallel}^2 \right] = 0. \end{aligned} \quad (3.89)$$

One can proceed in several ways with the above dispersion relation.

In the specific cases of strictly parallel ( $k_{\perp} = 0$ ) and strictly perpendicular ( $k_{\parallel} = 0$ ) propagation directions it is actually easier to work directly with the system (3.88), because the off-diagonal components in the matrix are zero and the dynamics in the  $u_x$  and  $u_z$  components decouples. For strictly parallel propagation, one solution is in the  $u_x$  component, with the dispersion relation identical to the Alfvén wave (3.81). Since the magnetic and velocity fluctuations are only in the  $u_x$ ,  $b_x$  components, the compressibility  $\chi_b = 0$  and  $\chi_u = 0$  for this mode. The second solution for parallel propagation is in the  $u_z$  component, and it is called the sound mode, or the ion-acoustic mode, with dispersion relation  $\omega = \pm k_{\parallel} \sqrt{3p_{\parallel}^{(0)}/\rho_0} = \pm k_{\parallel} V_A \sqrt{3\beta_{\parallel}/2}$ . The parallel ion-acoustic mode has  $\chi_u = 1$  and  $\chi_b = 0$ . Considering strictly perpendicular propagation, the waves in the  $u_z$  component vanish with the solution  $\omega = 0$ . The waves in the  $u_x$  component are fast magnetosonic waves with dispersion relation  $\omega = \pm k_{\perp} \sqrt{V_A^2 + 2p_{\perp}^{(0)}/\rho_0} = \pm k_{\perp} V_A \sqrt{1 + a_p \beta_{\parallel}}$ , and  $\chi_u = 0$  and  $\chi_b = 1$ . It is also possible to define the parallel and perpendicular sound speeds  $C_{\parallel} = \sqrt{3p_{\parallel}^{(0)}/\rho_0}$ ;  $C_{\perp} = \sqrt{2p_{\perp}^{(0)}/\rho_0}$ , which yields the parallel acoustic mode dispersion  $\omega = \pm k_{\parallel} C_{\parallel}$  and the perpendicular fast mode dispersion  $\omega = \pm k_{\perp} \sqrt{V_A^2 + C_{\perp}^2}$ .<sup>4</sup>

<sup>4</sup>Sometimes the parallel and perpendicular sound speeds are defined as  $C_{\parallel} = \sqrt{p_{\parallel}^{(0)}/\rho_0}$ ;  $C_{\perp} = \sqrt{p_{\perp}^{(0)}/\rho_0}$ , which yields the parallel acoustic mode dispersion  $\omega = \pm k_{\parallel} \sqrt{3} C_{\parallel}$  and the perpendicular fast mode dispersion  $\omega = \pm k_{\perp} \sqrt{V_A^2 + 2C_{\perp}^2}$ .



In the general case of oblique propagation, we choose to use the normalized frequencies and wavenumbers (3.29), and rewrite the dispersion relation as

$$\left. \begin{aligned} \tilde{\omega}^4 - A_2 \tilde{\omega}^2 + A_0 &= 0; \\ A_2 &= \tilde{k}_\perp^2 (1 + a_p \beta_\parallel) + \tilde{k}_\parallel^2 (1 + \beta_\parallel + \tfrac{1}{2} a_p \beta_\parallel); \\ A_0 &= \tfrac{3}{2} \tilde{k}_\parallel^2 \beta_\parallel \left[ \tilde{k}_\parallel^2 (1 - \tfrac{1}{2} \beta_\parallel + \tfrac{1}{2} a_p \beta_\parallel) + \tilde{k}_\perp^2 (1 + a_p \beta_\parallel - \tfrac{1}{6} a_p^2 \beta_\parallel) \right]. \end{aligned} \right\} \quad (3.90)$$

It is sometimes useful to introduce the parallel Alfvén phase speed  $v_{A\parallel}$ , and its normalization  $\tilde{v}_{A\parallel} \equiv v_{A\parallel}/V_A$ , so that

$$v_{A\parallel} = V_A \sqrt{1 + \frac{\beta_\parallel}{2} (a_p - 1)}; \quad \tilde{v}_{A\parallel}^2 = 1 + \frac{\beta_\parallel}{2} (a_p - 1), \quad (3.91)$$

since  $\tilde{v}_{A\parallel}^2$  can be used to write the CGL dispersion relation in a shorter form, as done later in (3.193). The solution of the polynomial (3.90) is simply  $\tilde{\omega}^2 = (A_2 \pm \sqrt{A_2^2 - 4A_0})/2$ , and it is obvious that the coefficient  $A_2$  is always positive, since both the plasma beta and the temperature anisotropy ratio must always be positive. It is also possible to show that the discriminant under the square root  $A_2^2 - 4A_0 \geq 0$  (proof shown a few lines below). The solution with the minus sign is called the slow mode and the solution with the plus sign is called the fast mode. For the slow mode,  $\tilde{\omega}^2$  becomes negative (and  $\tilde{\omega}$  complex) when  $A_0 < 0$ , implying that the slow mode becomes unstable when

$$[k_\parallel^2 (1 - \tfrac{1}{2} \beta_\parallel + \tfrac{1}{2} a_p \beta_\parallel) + k_\perp^2 (1 + a_p \beta_\parallel - \tfrac{1}{6} a_p^2 \beta_\parallel)] < 0, \quad (3.92)$$

where we suppress the tildes on the wavenumbers since the above condition is valid for both  $\tilde{k}$  and  $k$  forms of wavenumbers. The above condition implies that there are two competing instabilities. The highly parallel propagation with  $k_\parallel \gg k_\perp$  yields an instability threshold equivalent to (3.82) and the associated instability is called the parallel firehose instability. The highly oblique propagation with  $k_\perp \gg k_\parallel$  yields the mirror instability and the threshold in the CGL description reads<sup>5</sup>

$$1 + a_p \beta_\parallel - \frac{1}{6} a_p^2 \beta_\parallel < 0; \quad \frac{1}{6} \frac{T_\perp^{(0)}}{T_\parallel^{(0)}} - 1 - \frac{1}{\beta_\perp} > 0; \quad \frac{1}{6} \frac{\beta_\perp^2}{\beta_\parallel} > 1 + \beta_\perp. \quad (3.93)$$

It is well documented that the mirror instability threshold in the CGL description is not equivalent to the one found in linear kinetic theory. In particular, the CGL mirror threshold contains a factor of 6 error (Abraham-Shrauner 1967; Tajiri 1967; Kulsrud 1983; Ferrière & André 2002), since the correct threshold obtained from kinetic theory reads

$$1 + a_p \beta_\parallel - a_p^2 \beta_\parallel < 0; \quad \frac{T_\perp^{(0)}}{T_\parallel^{(0)}} - 1 - \frac{1}{\beta_\perp} > 0; \quad \frac{\beta_\perp^2}{\beta_\parallel} > 1 + \beta_\perp. \quad (3.94)$$

The threshold (3.93) implies that the necessary condition for the mirror instability to exist in the CGL description is  $a_p = T_\perp^{(0)}/T_\parallel^{(0)} > 6$ , whereas in kinetic theory the necessary condition is  $a_p > 1$ . A very good discussion about the physical mechanism

<sup>5</sup>To obtain the mirror instability threshold, it is not possible to consider strictly perpendicular propagation with  $k_\parallel = 0$  because the coefficient  $A_0$  would be exactly zero. The mirror threshold is revealed only in the highly oblique limit  $k_\perp \gg k_\parallel$ .

of the mirror instability and the role of resonant particles can be found in Southwood & Kivelson (1993) and Kivelson & Southwood (1996). The correct mirror instability threshold is obtained when using the normal closure discussed later on in the text or, more simply, its quasi-static limit (Constantinescu 2002; Chust & Belmont 2006; Passot *et al.* 2006). In the opposite case where ions are cold, a similar electron temperature anisotropy instability is obtained which is called the field-swelling instability, see e.g. Basu & Coppi (1984).

Returning to the oblique dispersion relation, we show that indeed  $A_2^2 - 4A_0 \geq 0$ . By using (3.78) the wavenumber  $k$  can be pulled out of the  $A_0, A_2$  coefficients and the dispersion relation can be rewritten in terms of a phase speed as

$$(\tilde{\omega}/\tilde{k})^4 - A_2(\tilde{\omega}/\tilde{k})^2 + A_0 = 0; \quad (3.95)$$

$$A_2 = 1 + a_p \beta_{\parallel} \left(1 - \frac{1}{2} \cos^2 \theta\right) + \beta_{\parallel} \cos^2 \theta; \quad (3.96)$$

$$A_0 = \frac{3}{2} \beta_{\parallel} \cos^2 \theta \left[1 + a_p \beta_{\parallel} \left(1 - \frac{1}{2} \cos^2 \theta\right) - \frac{1}{2} \beta_{\parallel} \cos^2 \theta - \frac{1}{6} a_p^2 \beta_{\parallel} \sin^2 \theta\right], \quad (3.97)$$

and (after staring for a sufficiently long time) it can be shown that

$$A_2^2 - 4A_0 = \left[1 + a_p \beta_{\parallel} \left(1 - \frac{1}{2} \cos^2 \theta\right) - 2\beta_{\parallel} \cos^2 \theta\right]^2 + a_p^2 \beta_{\parallel}^2 \sin^2 \theta \cos^2 \theta, \quad (3.98)$$

which is obviously always non-negative. The dispersion relation for the slow and fast waves in the CGL description can be written in the compact form

$$\begin{aligned} \left(\frac{\tilde{\omega}}{\tilde{k}}\right)_{sf}^2 &= \left(\frac{\omega}{kV_A}\right)_{sf}^2 = \frac{1}{2} \left[1 + a_p \beta_{\parallel} \left(1 - \frac{1}{2} \cos^2 \theta\right) + \beta_{\parallel} \cos^2 \theta\right] \\ &\pm \frac{1}{2} \sqrt{\left[1 + a_p \beta_{\parallel} \left(1 - \frac{1}{2} \cos^2 \theta\right) - 2\beta_{\parallel} \cos^2 \theta\right]^2 + a_p^2 \beta_{\parallel}^2 \sin^2 \theta \cos^2 \theta}, \end{aligned} \quad (3.99)$$

which is equivalent to the dispersion relation of Abraham-Shrauner (1967), equation (30), here written in the convenient units of  $\beta_{\parallel}$  and  $a_p = T_{\perp}^{(0)}/T_{\parallel}^{(0)}$  (for a direct comparison with the units used in that paper, the relations are  $\beta_{\parallel} = 2S_{\parallel}$  and  $a_p \beta_{\parallel} = 2S_{\perp}$ ). The magnetic compressibility of the slow and fast modes can be calculated easily directly from the induction equations (3.72). The slow and fast modes generate only components  $B_x$  and  $B_z$  with the component  $B_y = 0$ . The transformation of (3.72) to Fourier space yields  $-\omega B_x = B_0 k_{\parallel} u_x$ , and  $-\omega B_z = -B_0 k_{\perp} u_x$ , and by applying the magnitude operator yields the magnetic compressibility

$$\chi_b(k_{\parallel}, k_{\perp}) = \frac{k_{\perp}^2}{k_{\parallel}^2 + k_{\perp}^2} = \sin^2 \theta. \quad (3.100)$$

For quasi-parallel propagation angles the magnetic compressibility  $\chi_b$  is therefore not a good indicator for mode recognition, since all three modes have  $\chi_b$  close to zero. The magnetic compressibility is, however, an excellent tool for oblique propagation angles, because the Alfvén mode has  $\chi_b = 0$  regardless of the propagation direction, whereas for the slow and fast modes  $\chi_b$  increases with  $\theta$  towards  $\chi_b = 1$ . We will leave discussion about the CGL velocity eigenvector (the ratio  $\chi_u$ ) to a later subsection since the discussion is a bit technical, and we instead focus now on an interesting effect that can be directly obtained by considering the CGL dispersion relations.

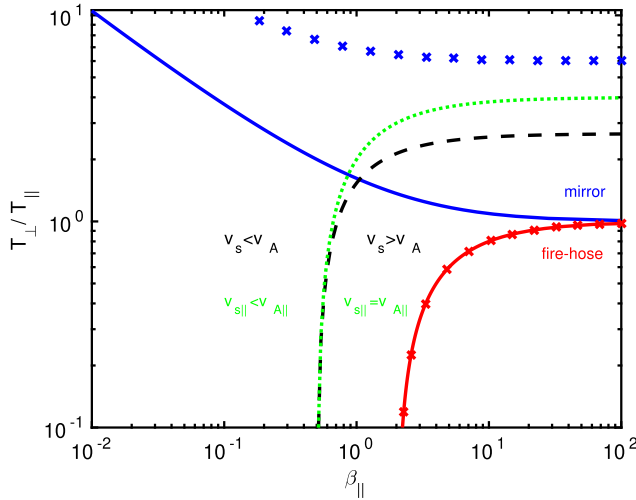


FIGURE 1. Analytic ‘hard’ thresholds of the mirror instability (blue) and the firehose instability (red) at long wavelengths. Solid lines are solutions of the linear kinetic theory (bi-Maxwellian) and crosses are solutions of the linear CGL model. The firehose instability is described correctly (both the oblique firehose and the parallel firehose), whereas the mirror instability contains an asymptotic factor of 6 error for large  $\beta_{\parallel}$  values. The black dashed line is the CGL threshold (3.103), where for all oblique directions the slow mode speed  $v_s$  matches the Alfvén mode speed  $v_A$ , and which separates two regions where  $v_s < v_A$  and  $v_s > v_A$ . The green dotted curve is a special case for parallel propagation (3.108) where  $v_{s\parallel} = v_{A\parallel}$ . For parallel propagation, the slow mode speed cannot exceed the Alfvén mode speed, but can only match it, and the green dotted curve separates two regions where  $v_{s\parallel} < v_{A\parallel}$  and  $v_{s\parallel} = v_{A\parallel}$ . For a similar plot in linear scale, see figure 1 of Abraham-Shrauner (1967).

### 3.6. Slow mode can become faster than Alfvén mode

There is a very interesting phenomenon present in the CGL description that is also observed in the kinetic description but that is absent in the MHD description. For a sufficiently high plasma beta, the slow mode phase speed can be larger than the Alfvén mode phase speed. To obtain the critical plasma beta when this happens, one can write the dispersion relation for the Alfvén wave as  $(\tilde{\omega}/k)_A^2 = (1 - \beta_{\parallel}/2 + a_p\beta_{\parallel}/2) \cos^2 \theta \equiv A_A$  and using this expression with the dispersion relation (3.95), the critical condition when the slow mode speed matches the Alfvén speed is  $(A_2 - \sqrt{A_2^2 - 4A_0})/2 = A_A$ , and since  $A_2^2 - 4A_0 \geq 0$ , this further yields  $A_A^2 - A_2A_A + A_0 = 0$ . It takes quite a few algebraic operations to simplify the result, the easiest way is to collect terms with different powers of  $\beta_{\parallel}$  and then simplify. A very valuable approach available now is to use mathematical software such as Maple or Mathematica for guidance, and for example the Maple command `simplify(expr,trig)` immediately simplifies the result, which can then be of course verified with pencil and paper. The result is

$$A_A^2 - A_2A_A + A_0 = -\sin^2 \theta \cos^2 \theta \left[ 1 + \beta_{\parallel} \left( \frac{3}{2}a_p - 2 \right) + \beta_{\parallel}^2 a_p \left( \frac{3}{4}a_p - 2 \right) \right] = 0. \quad (3.101)$$

The expression is naturally satisfied for cases of strictly parallel and strictly perpendicular propagation. For  $\theta = \pi/2$ , both the slow mode and the Alfvén mode

have zero frequency. For  $\theta = 0$ , the slow mode speed can be only smaller or equal to the speed of the Alfvén mode (never higher), depending on another condition that will be easier to clarify later when specifically considering the parallel propagation limit. Now, on which side of the threshold (3.101) does the speed of the slow mode become faster than the speed of the Alfvén mode? Working with inequalities is a bit trickier. By assuming that both modes have real frequencies (and are not firehose or mirror unstable), the condition is derived to be

$$-\sin^2 \theta \cos^2 \theta \left[ 1 + \beta_{\parallel} \left( \frac{3}{2} a_p - 2 \right) + \beta_{\parallel}^2 a_p \left( \frac{3}{4} a_p - 2 \right) \right] \geq 0, \quad (3.102)$$

which agrees with the result obtained by Abraham-Shrauner (1967), equation (42). The expression (3.102) reveals that there is a condition that can be satisfied for all directions of propagation and that reads

$$1 + \beta_{\parallel} \left( \frac{3}{2} a_p - 2 \right) + \beta_{\parallel}^2 a_p \left( \frac{3}{4} a_p - 2 \right) \leq 0. \quad (3.103)$$

This is a quadratic equation in  $\beta_{\parallel}$  which can be solved easily for a given  $a_p$  (or solved for  $a_p$  for a given  $\beta_{\parallel}$ ). To visualize this equation, the threshold is plotted in figure 1 as a black dashed line. A particularly useful result is obtained for isotropic temperatures ( $a_p = 1$ ). In this case the inequality (3.103) becomes  $1 - \beta_{\parallel}/2 - \beta_{\parallel}^2/4 \leq 0$  and the critical plasma beta at which the slow mode speed becomes larger than the Alfvén mode speed is

$$\beta_{\parallel}^{\text{crit}} \geq \frac{\sqrt{21} - 1}{5} \approx 0.72. \quad (3.104)$$

Thus, this effect exists in the CGL description even when the temperatures are isotropic.

### 3.7. CGL parallel propagation

When numerically solving the CGL dispersion relations for finite (non-zero) propagation angles, the slow and fast mode expressions always get the correct ordering with  $\omega_s \leq \omega_f$ . It is sometimes confusing about how the dispersion relations hold in the limit of parallel propagation  $\theta \rightarrow 0$  which we will address here. In this limit, the dispersion relations (3.99) for the slow and fast modes can be directly evaluated as

$$\left( \frac{\tilde{\omega}}{\tilde{k}} \right)_{sf}^2 = \frac{1}{2} \left[ 1 + \frac{1}{2} a_p \beta_{\parallel} + \beta_{\parallel} \right] \pm \frac{1}{2} \sqrt{\left[ 1 + \frac{1}{2} a_p \beta_{\parallel} - 2\beta_{\parallel} \right]^2}. \quad (3.105)$$

Now, because  $\sqrt{a^2} = |a|$ , for further calculation one needs to determine the sign of the expression under the square root. For  $1 + \frac{1}{2} a_p \beta_{\parallel} - 2\beta_{\parallel} > 0$ , i.e. when  $\beta_{\parallel} < 2/(4 - a_p)$ , the above equations yield

$$(\tilde{\omega}/\tilde{k})_s^2 = \frac{3}{2} \beta_{\parallel}; \quad (\tilde{\omega}/\tilde{k})_f^2 = 1 + \frac{\beta_{\parallel}}{2} (a_p - 1) = (\tilde{\omega}/\tilde{k})_A^2, \quad (3.106)$$

so the slow mode is the sound/acoustic mode and the fast mode coincides with the Alfvén mode. However, for  $1 + \frac{1}{2} a_p \beta_{\parallel} - 2\beta_{\parallel} < 0$ , i.e. when  $\beta_{\parallel} > 2/(4 - a_p)$ , the dispersion relations yield

$$(\tilde{\omega}/\tilde{k})_s^2 = 1 + \frac{\beta_{\parallel}}{2} (a_p - 1) = (\tilde{\omega}/\tilde{k})_A^2; \quad (\tilde{\omega}/\tilde{k})_f^2 = \frac{3}{2} \beta_{\parallel}, \quad (3.107)$$

so the slow mode coincides with the Alfvén mode and the fast mode is the sound/acoustic mode. In this region, the slow mode phase speed will always stay equal to the Alfvén mode phase speed, regardless of further increases of  $\beta_{\parallel}$ . In the special case when  $1 + \frac{1}{2}a_p\beta_{\parallel} - 2\beta_{\parallel} = 0$ , which is equivalent to

$$\beta_{\parallel} = \frac{2}{4 - a_p}; \quad a_p = 4 - \frac{2}{\beta_{\parallel}}, \quad (3.108)$$

all three modes propagate with the same phase speed  $(\tilde{\omega}/k)_{Asf}^2 = \frac{3}{2}\beta_{\parallel}$ . The criterion (3.108) can be satisfied only for  $\beta_{\parallel} \geq 1/2$  and in the limit  $\beta_{\parallel} \rightarrow \infty$  the asymptote is  $a_p = 4$ , so the possible range of  $a_p$  is  $[0, 4]$ . This curve is plotted in figure 1 as a green dotted line. The curve separates two regions. In the first region  $\beta_{\parallel} < 2/(4 - a_p)$  and  $v_{s\parallel} < v_{A\parallel}$ . In the second region  $\beta_{\parallel} > 2/(4 - a_p)$  and  $v_{s\parallel} = v_{A\parallel}$ . The interesting result to remember is that, for isotropic temperatures ( $a_p = 1$ ) and parallel propagation, the critical plasma beta at which the slow mode speed matches the Alfvén mode speed is

$$\beta_{\parallel}^{\text{crit}}(0) \geq \frac{2}{3}. \quad (3.109)$$

### 3.8. CGL with $T_{\perp} = T_{\parallel}$ is not equivalent to MHD

It is important to emphasize that, even for isotropic temperatures  $a_p = 1$ , the CGL dispersion (3.99) is not equivalent to the MHD dispersion for the slow and fast waves. The MHD dispersion is usually written in the form (see (C 20) in appendix C)

$$\left(\frac{\omega}{k}\right)^2 = \frac{1}{2}(V_A^2 + C_s^2) \pm \frac{1}{2}\sqrt{(V_A^2 + C_s^2)^2 - 4V_A^2C_s^2\cos^2\theta}, \quad (3.110)$$

or alternatively

$$\left(\frac{\omega}{k}\right)^2 = \frac{1}{2}(V_A^2 + C_s^2) \pm \frac{1}{2}\sqrt{(V_A^2 - C_s^2)^2 + 4V_A^2C_s^2\sin^2\theta}, \quad (3.111)$$

which nicely shows that the expression under the square root is always greater or equal to zero. The MHD sound speed is defined as  $C_s^2 = \gamma p^{(0)}/\rho_0$ , where  $\gamma = 5/3$ . Rewritten with the parameters we use here  $C_s^2/V_A^2 = \gamma\beta_{\parallel}/2 = 5\beta_{\parallel}/6$ , the MHD dispersion for the slow and fast waves reads

$$\left(\frac{\omega}{kV_A}\right)_{\text{MHD}}^2 = \frac{1}{2}\left(1 + \frac{\gamma}{2}\beta_{\parallel}\right) \pm \frac{1}{2}\sqrt{\left(1 + \frac{\gamma}{2}\beta_{\parallel}\right)^2 - 2\gamma\beta_{\parallel}\cos^2\theta}, \quad (3.112)$$

or alternatively

$$\left(\frac{\omega}{kV_A}\right)_{\text{MHD}}^2 = \frac{1}{2}\left(1 + \frac{\gamma}{2}\beta_{\parallel}\right) \pm \frac{1}{2}\sqrt{\left(1 - \frac{\gamma}{2}\beta_{\parallel}\right)^2 + 2\gamma\beta_{\parallel}\sin^2\theta}. \quad (3.113)$$

In the CGL description with isotropic temperatures ( $a_p = 1$ ) the dispersion relation (3.99) for slow and fast waves reads

$$\begin{aligned} \left(\frac{\omega}{kV_A}\right)_{\text{CGL}}^2 &= \frac{1}{2}\left[1 + \beta_{\parallel}\left(1 + \frac{1}{2}\cos^2\theta\right)\right] \\ &\pm \frac{1}{2}\sqrt{\left[1 + \beta_{\parallel}\left(1 - \frac{5}{2}\cos^2\theta\right)\right]^2 + \beta_{\parallel}^2\sin^2\theta\cos^2\theta}, \end{aligned} \quad (3.114)$$

or alternatively

$$\begin{aligned} \left(\frac{\omega}{kV_A}\right)_{\text{CGL}}^2 &= \frac{1}{2} \left[ 1 + \beta_{\parallel} \left( 1 + \frac{1}{2} \cos^2 \theta \right) \right] \\ &\pm \frac{1}{2} \sqrt{\left[ 1 + \beta_{\parallel} \left( 1 + \frac{1}{2} \cos^2 \theta \right) \right]^2 - 6\beta_{\parallel} \cos^2 \theta \left( 1 + \frac{5}{6} \beta_{\parallel} \sin^2 \theta \right)}. \end{aligned} \quad (3.115)$$

We wrote down all major possibilities on purpose, since we wanted to clearly show that, regardless of the choice of  $\beta_{\parallel}$  and  $\gamma$ , the MHD dispersion relation cannot match the CGL dispersion relation simultaneously for all propagation directions  $\theta$ . One could even play with an abstract idea to make the MHD  $\gamma$  dependent on the angle  $\theta$ , and possibly  $\beta_{\parallel}$ . Even then, there is no obvious way to modify the MHD dispersion relation so that it behaves like the CGL model. Additionally, the CGL description leads to the development of pressure anisotropy even when the initial pressure is isotropic.

### 3.9. MHD velocity eigenvector

Before we proceed with the discussion of the CGL velocity eigenvector (which can appear to be a bit tricky), for the sake of clarity it is useful to consider the simpler MHD case first. By following the steps outlined above, it is easy to show that again the Alfvén mode dynamics in the  $u_y$  decouples from the dynamics in the other directions and the velocity eigenvector matrix for the slow and fast waves in MHD is

$$\begin{pmatrix} \omega^2 - (V_A^2 k^2 + C_s^2 k_{\perp}^2) & -C_s^2 k_{\parallel} k_{\perp} \\ -C_s^2 k_{\parallel} k_{\perp} & \omega^2 - C_s^2 k_{\parallel}^2 \end{pmatrix} \begin{pmatrix} u_x \\ u_z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad (3.116)$$

or alternatively

$$\begin{pmatrix} (\omega/k)^2 - (V_A^2 + C_s^2 \sin^2 \theta) & -C_s^2 \sin \theta \cos \theta \\ -C_s^2 \sin \theta \cos \theta & (\omega/k)^2 - C_s^2 \cos^2 \theta \end{pmatrix} \begin{pmatrix} u_x \\ u_z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \quad (3.117)$$

This yields the MHD dispersion relations for the slow and fast modes in the form of (3.110) or alternatively (3.111). It is possible to either express the ratio  $u_x/u_z$  or  $u_z/u_x$ . We want to calculate the velocity ratio  $\chi_u$  and we choose the first possibility that allows us to express that for the slow and fast modes (since  $u_y = 0$ )

$$\chi_u(k_{\parallel}, k_{\perp}) = \frac{|u_z|^2}{|u_x|^2 + |u_z|^2} = \frac{1}{\left| \frac{u_x}{u_z} \right|^2 + 1}. \quad (3.118)$$

The ratio  $u_x/u_z$  can be found for example from the second row of the matrix as

$$\frac{u_x}{u_z} = \frac{(\omega/k)^2 - C_s^2 \cos^2 \theta}{C_s^2 \sin \theta \cos \theta}, \quad (3.119)$$

or equivalently from the first row of the matrix as

$$\frac{u_x}{u_z} = \frac{C_s^2 \sin \theta \cos \theta}{(\omega/k)^2 - (V_A^2 + C_s^2 \sin^2 \theta)}. \quad (3.120)$$

The possible range of the ratio  $u_x/u_z$  is from  $-\infty$  to  $+\infty$  and the possible range for  $\chi_u$  is from 0 to 1. Solutions can be easily divided into two categories – depending on the behaviour for parallel propagation – and are plotted in figure 2. For the sake of clarity, we only consider propagation with positive wavenumbers  $k_{\parallel}$  and  $k_{\perp}$ , so the possible range of angle  $\theta$  is from 0 to  $\pi/2$  and  $\sin \theta \geq 0$ ,  $\cos \theta \geq 0$ . Consider first the case  $C_s < V_A$ . For parallel propagation, the dispersion relation (3.111) directly yields that for the fast mode  $(\omega/k)_f^2|_{\theta=0} = V_A^2$  and for the slow mode  $(\omega/k)_s^2|_{\theta=0} = C_s^2$ . Using the expression (3.119) then yields that for the fast mode  $(u_x/u_z)_f|_{\theta=0} = +\infty$  and for the slow mode  $(u_x/u_z)_s|_{\theta=0} = 0$ . This implies that for  $C_s < V_A$ , all the  $\chi_u(\theta)$  curves start at

$$C_s < V_A: \quad \chi_u^f(\theta=0) = 0; \quad \chi_u^s(\theta=0) = 1, \quad (3.121)$$

regardless of the actual value of  $C_s/V_A < 1$ . The second case  $C_s > V_A$  is straightforward, since for parallel propagation everything is just turned around; the fast mode dispersion relation is  $(\omega/k)_f^2|_{\theta=0} = C_s^2$  and the slow mode dispersion relation is  $(\omega/k)_s^2|_{\theta=0} = V_A^2$ . Implying that for  $C_s > V_A$ , all the  $\chi_u(\theta)$  curves start at

$$C_s > V_A: \quad \chi_u^f(\theta=0) = 1; \quad \chi_u^s(\theta=0) = 0, \quad (3.122)$$

regardless of the actual value of  $C_s/V_A > 1$ . The perpendicular propagation limit is even easier since the split into two categories is not required and for both cases the dispersion (3.110) yields  $(\omega/k)_f^2|_{\theta=\pi/2} = V_A^2 + C_s^2$  and  $(\omega/k)_s^2|_{\theta=\pi/2} = 0$ , implying that for the fast mode  $(u_x/u_z)_f|_{\theta=\pi/2} = +\infty$  and that for the slow mode  $(u_x/u_z)_s|_{\theta=\pi/2} = 0$ . All the  $\chi_u(\theta)$  curves therefore ‘end’ at

$$\chi_u^f(\theta = \pi/2) = 0; \quad \chi_u^s(\theta = \pi/2) = 1, \quad (3.123)$$

whether the ratio  $C_s/V_A < 1$  or  $C_s/V_A > 1$ . The behaviour of the  $\chi_u(\theta)$  curves is plotted in figure 2.

In MHD, the ratio  $C_s^2/V_A^2$  is sometimes used as the definition of plasma beta, which we will call here  $\beta_{\text{MHD}} \equiv C_s^2/V_A^2$ . As can be seen in figure 2, for  $\beta_{\text{MHD}}$  approaching 1, the curves display a ‘singular’ behaviour, and it is difficult to guess what will happen when  $\beta_{\text{MHD}}$  reaches the value of 1 exactly. This special case is addressed in the next subsection. There are also two special cases, one for  $C_s \ll V_A$  or  $\beta_{\text{MHD}} \ll 1$  and one for  $C_s \gg V_A$  or  $\beta_{\text{MHD}} \gg 1$ . Both limits  $C_s \ll V_A$  and  $C_s \gg V_A$  are easily obtained by rewriting the dispersion relation (3.110) to a form

$$\left(\frac{\omega}{k}\right)^2 = \frac{1}{2}(V_A^2 + C_s^2) \pm \frac{1}{2}(V_A^2 + C_s^2) \sqrt{1 - \frac{4V_A^2 C_s^2}{(V_A^2 + C_s^2)^2} \cos^2 \theta}. \quad (3.124)$$

In both limits the second term under the square root is small and the expression can be expanded as  $\sqrt{1 - \epsilon} = 1 - \epsilon/2$ , yielding the slow and fast mode dispersion relations

$$C_s \ll V_A \text{ or } C_s \gg V_A: \quad \left(\frac{\omega}{k}\right)_f^2 = V_A^2 + C_s^2 - \frac{V_A^2 C_s^2}{V_A^2 + C_s^2} \cos^2 \theta; \quad \left(\frac{\omega}{k}\right)_s^2 = \frac{V_A^2 C_s^2}{V_A^2 + C_s^2} \cos^2 \theta, \quad (3.125)$$



which further yield dispersion relations (3.126), (3.128). Note that the limit  $C_s \ll V_A$  can be easily obtained directly from the velocity matrix (3.117) by neglecting the off-diagonal terms that decouple the two modes. In contrast, the limit  $C_s \gg V_A$  is not that obvious and one needs to calculate the determinant.

In the limit  $\beta_{\text{MHD}} \ll 1$ , the MHD dispersion relation simplifies to

$$\beta_{\text{MHD}} \ll 1: \quad (\omega/k)_f^2 = V_A^2 + C_s^2 \sin^2 \theta; \quad (\omega/k)_s^2 = C_s^2 \cos^2 \theta. \quad (3.126)$$

The expression (3.119) yields for the fast mode  $(u_x/u_z)_f(\theta) = \infty$  and the expression (3.120) yields for the slow mode  $(u_x/u_z)_s(\theta) = 0$ , implying

$$\beta_{\text{MHD}} \ll 1: \quad \chi_u^f(\theta) = 0; \quad \chi_u^s(\theta) = 1. \quad (3.127)$$

In the opposite limit  $\beta_{\text{MHD}} \gg 1$ , the MHD dispersions simplify to

$$\beta_{\text{MHD}} \gg 1: \quad (\omega/k)_f^2 = C_s^2 + V_A^2 \sin^2 \theta; \quad (\omega/k)_s^2 = V_A^2 \cos^2 \theta. \quad (3.128)$$

In this limit, the expression (3.119) yields  $(u_x/u_z)_f(\theta) = (1 + V_A^2/C_s^2) \tan \theta$ , that describes the behaviour of the fast mode for large beta values and the complete limit is  $(u_x/u_z)_f(\theta) = \tan \theta$ . For the slow mode the expression (3.119) yields  $(u_x/u_z)_s(\theta) = -(1 - V_A^2/C_s^2) \cot \theta$  and the complete limit is  $(u_x/u_z)_s(\theta) = -\cot \theta$ . The  $\chi_u(\theta)$  solutions for large beta values therefore converge to

$$\beta_{\text{MHD}} \gg 1: \quad \chi_u^f(\theta) = \cos^2 \theta; \quad \chi_u^s(\theta) = \sin^2 \theta, \quad (3.129)$$

and these two curves are plotted in figure 2 as dashed lines.

### 3.9.1. MHD special case $V_A = C_s$

Consider the very peculiar case of parallel propagation  $\theta = 0$  with  $V_A = C_s$ . The velocity eigenvector matrix (3.117), when evaluated exactly at these values reads

$$\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} u_x \\ u_z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \quad (3.130)$$

It is not possible to determine the ratio  $u_x/u_z$  and the eigenvector must be found by taking the limit  $\theta \rightarrow 0$ . For  $V_A = C_s$  the dispersion relation for the slow and fast modes (3.111) simplifies to  $(\omega/k)^2 = V_A^2 \pm V_A^2 |\sin \theta|$ . As before, for the sake of simplicity we consider only positive wavenumbers, so the propagation angle  $\theta$  is in the range  $[0, \pi/2]$  and  $\sin \theta \geq 0$ ,  $\cos \theta \geq 0$ . Evaluating the velocity matrix (3.117) at  $V_A = C_s$  only, it is possible to pull the  $\sin \theta$  out of the system so that

$$\begin{pmatrix} \pm 1 - \sin \theta & -\cos \theta \\ -\cos \theta & \pm 1 + \sin \theta \end{pmatrix} V_A^2 \sin \theta \begin{pmatrix} u_x \\ u_z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad (3.131)$$

which nicely demonstrates that for  $\theta \rightarrow 0$  the system approaches (3.130). Nevertheless, we can now calculate the velocity ratio (for example from the second row) as

$$\frac{u_x}{u_z} = \frac{\pm 1 + \sin \theta}{\cos \theta}, \quad (3.132)$$

which can be evaluated at  $\theta = 0$ . For the fast mode  $(u_x/u_z)_f(0) = 1$  and for the slow mode  $(u_x/u_z)_s(0) = -1$ . This implies that for the peculiar case of  $V_A = C_s$ , the velocity

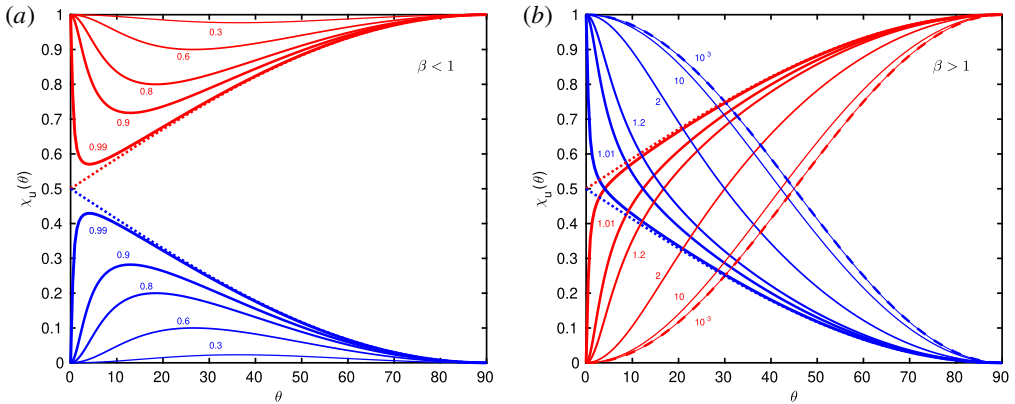


FIGURE 2. MHD velocity ratio  $\chi_u(\theta) = |u_{\parallel}|^2/|u|^2$  for the slow mode (red lines) and the fast mode (blue lines) for different ratios of  $C_s^2/V_A^2 = \beta_{\text{MHD}}$ . Panel (a) is for  $\beta_{\text{MHD}} < 1$ , and the thickness of the lines increases with  $\beta_{\text{MHD}}$  when approaching the critical value of  $\beta_{\text{MHD}} = 1$ . The plotted lines have  $\beta_{\text{MHD}} = 0.3; 0.6; 0.8; 0.9; 0.99$ . Panel (b) is for  $\beta_{\text{MHD}} > 1$  and the thickness of the lines decreases with  $\beta_{\text{MHD}}$  when going away from the critical value of  $\beta_{\text{MHD}} = 1$ . The plotted lines have  $\beta_{\text{MHD}} = 1.01; 1.1; 1.2; 2.0; 10; 10^3$ . In both figures, the dotted lines are the critical case  $\beta_{\text{MHD}} = 1$  with solutions (3.134). In (b) the dashed lines represent the limit  $\beta_{\text{MHD}} \gg 1$  with solutions (3.129) and the dashed lines nicely overlap with a thin solid line solutions obtained for  $\beta_{\text{MHD}} = 10^3$ . The Alfvén mode has  $\chi_u(\theta) = 0$  regardless of the plasma beta and is not plotted.

ratio  $\chi_u^{\text{sf}}(0) = 1/2$  for both the fast and slow modes. The expression (3.132) can of course be directly obtained by using the result  $(\omega/k)^2 = V_A^2 \pm V_A^2 \sin \theta$  in the expression (3.119) and by cancelling the  $\sin \theta$  in the nominator and denominator. The expression (3.132) is valid for any angle  $\theta$ . However, for the slow mode it becomes singular at  $\theta = \pi/2$ , and has to be rearranged by multiplying and dividing with  $1 + \sin \theta$  (or equivalently, use the first row of the matrix). To have useful expression for the entire range of  $\theta$ , we therefore split the expressions for the slow and fast modes as

$$(u_x/u_z)^s = -\frac{\cos \theta}{1 + \sin \theta}; \quad (u_x/u_z)^f = \frac{1 + \sin \theta}{\cos \theta}, \quad (3.133)$$

which yields, in this peculiar case of  $V_A = C_s$ , that the ratio  $\chi_u$  for the slow and fast modes can be expressed as

$$V_A = C_s: \quad \chi_u^s(\theta) = \frac{1}{2}(1 + \sin \theta); \quad \chi_u^f(\theta) = \frac{1}{2}(1 - \sin \theta). \quad (3.134)$$

These solutions are plotted in figure 2 as dotted lines.

### 3.10. CGL velocity eigenvector

To find the ratio for velocity components  $\chi_u$  in the CGL description, it is useful for a moment to work with abbreviated symbols, and we write the matrix (3.88) for the  $u_x$  and  $u_z$  components as

$$\begin{pmatrix} (\tilde{\omega}/\tilde{k})^2 - a; & -b \\ -b & (\tilde{\omega}/\tilde{k})^2 - c \end{pmatrix} \begin{pmatrix} u_x \\ u_z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad (3.135)$$

where the coefficients are

$$a = 1 + a_p \beta_{\parallel} \left( 1 - \frac{1}{2} \cos^2 \theta \right) - \frac{\beta_{\parallel}}{2} \cos^2 \theta; \quad b = \frac{\beta_{\parallel}}{2} a_p \sin \theta \cos \theta; \quad c = \frac{3}{2} \beta_{\parallel} \cos^2 \theta. \quad (3.136)$$

The determinant must be zero, which implies  $(\tilde{\omega}/\tilde{k})^4 - (\tilde{\omega}/\tilde{k})^2(a+c) + ac - b^2 = 0$ , and has a solution

$$(\tilde{\omega}/\tilde{k})^2 = \frac{1}{2}(a+c) \pm \frac{1}{2}\sqrt{(a+c)^2 - 4ac + 4b^2} = \frac{1}{2}(a+c) \pm \frac{1}{2}\sqrt{(a-c)^2 + 4b^2}. \quad (3.137)$$

The last step was actually used in obtaining (3.98) which is otherwise quite difficult to see. Using this result in the above matrix, the eigenvector matrix becomes

$$\begin{pmatrix} -\frac{1}{2}(a-c) \pm \frac{1}{2}\sqrt{(a-c)^2 + 4b^2}; & -b \\ -b & \frac{1}{2}(a-c) \pm \frac{1}{2}\sqrt{(a-c)^2 + 4b^2} \end{pmatrix} \begin{pmatrix} u_x \\ u_z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad (3.138)$$

where

$$a - c = 1 + a_p \beta_{\parallel} \left( 1 - \frac{1}{2} \cos^2 \theta \right) - 2\beta_{\parallel} \cos^2 \theta. \quad (3.139)$$

The ratio of  $u_x/u_z$  can be expressed (for example from the second row) as

$$\begin{aligned} \frac{u_x}{u_z} &= \frac{1}{b}[(\tilde{\omega}/\tilde{k})^2 - c] = \frac{1}{2b}[a - c \pm \sqrt{(a-c)^2 + 4b^2}] \\ &= \frac{1}{\beta_{\parallel} a_p \sin \theta \cos \theta} \left[ 1 + a_p \beta_{\parallel} \left( 1 - \frac{1}{2} \cos^2 \theta \right) - 2\beta_{\parallel} \cos^2 \theta \right. \\ &\quad \left. \pm \sqrt{\left( 1 + a_p \beta_{\parallel} \left( 1 - \frac{1}{2} \cos^2 \theta \right) - 2\beta_{\parallel} \cos^2 \theta \right)^2 + a_p^2 \beta_{\parallel}^2 \sin^2 \theta \cos^2 \theta} \right], \quad (3.140) \end{aligned}$$

or alternatively (from the first row) as

$$\begin{aligned} \frac{u_x}{u_z} &= \frac{b}{(\tilde{\omega}/\tilde{k})^2 - a} = \frac{2b}{-(a-c) \pm \sqrt{(a-c)^2 + 4b^2}} \\ &= \beta_{\parallel} a_p \sin \theta \cos \theta \left[ -1 - a_p \beta_{\parallel} \left( 1 - \frac{1}{2} \cos^2 \theta \right) + 2\beta_{\parallel} \cos^2 \theta \right. \\ &\quad \left. \pm \sqrt{\left( 1 + a_p \beta_{\parallel} \left( 1 - \frac{1}{2} \cos^2 \theta \right) - 2\beta_{\parallel} \cos^2 \theta \right)^2 + a_p^2 \beta_{\parallel}^2 \sin^2 \theta \cos^2 \theta} \right]^{-1}. \quad (3.141) \end{aligned}$$

The behaviour of these solutions is plotted in figure 3, where we plot solutions for isotropic temperatures  $a_p = 1$ . Similarly to MHD, two classes of solutions can be characterized based upon the behaviour for parallel propagation  $\theta = 0$ , depending if the quantity  $a - c$  (3.139) evaluated at  $\theta = 0$  is positive or negative. The quantity  $(a - c)|_{\theta=0} = 1 + \frac{1}{2} a_p \beta_{\parallel} - 2\beta_{\parallel}$  was discussed in the section for parallel propagation and

this quantity represents the green dotted curve in figure 1 that separates the system into two regions. For  $(a - c)|_{\theta=0} > 0$ , i.e. in the region  $\beta_{\parallel} < 2/(4 - a_p)$ , for the slow mode (use (3.141)) the ratio  $(u_x/u_z)^s(0) = 0$  and for the fast mode (use (3.140)) the ratio  $(u_x/u_z)^f(0) = +\infty$ . In this region, all the  $\chi_u(\theta)$  curves start at

$$\beta_{\parallel} < 2/(4 - a_p) : \quad \chi_u^s(0) = 1; \quad \chi_u^f(0) = 0. \quad (3.142)$$

For  $(a - c)|_{\theta=0} < 0$ , i.e. in the region  $\beta_{\parallel} > 2/(4 - a_p)$ , for the slow mode (use (3.140)) the ratio  $(u_x/u_z)^s(0) = +\infty$  and for the fast mode (use (3.141)) the ratio  $(u_x/u_z)^f(0) = 0$ . In this region, all the  $\chi_u(\theta)$  curves start at

$$\beta_{\parallel} > 2/(4 - a_p) : \quad \chi_u^s(0) = 0; \quad \chi_u^f(0) = 1. \quad (3.143)$$

For perpendicular propagation  $\theta = \pi/2$  the quantity  $(a - c)|_{\theta=\pi/2} = 1 + a_p\beta_{\parallel}$  is always positive and the separation into two regions is not necessary. For the slow mode  $(u_x/u_z)^s(\pi/2) = 0$  and for the fast mode  $(u_x/u_z)^f(\pi/2) = +\infty$ . All the  $\chi_u(\theta)$  curves, regardless of the value of  $\beta_{\parallel}$  therefore end at

$$\chi_u^s(\pi/2) = 1; \quad \chi_u^f(\pi/2) = 0. \quad (3.144)$$

This summarizes the general behaviour of the solutions. However, as can be seen from figure 3, there is a special case when the parameters approach  $\beta_{\parallel} = 2/(4 - a_p)$ , i.e. when  $(a - c)|_{\theta=0} = 0$ , that will be discussed in the next subsection. One can again consider the two limiting cases when  $\beta_{\parallel}$  is small or large. The  $\beta_{\parallel} \ll 1$  case yields dispersion relations<sup>6</sup>

$$\beta_{\parallel} \ll 1 : \quad (\tilde{\omega}/\tilde{k})_f^2 = 1 + a_p\beta_{\parallel} \left(1 - \frac{1}{2} \cos^2 \theta\right) - \frac{\beta_{\parallel}}{2} \cos^2 \theta; \quad (3.146)$$

$$(\tilde{\omega}/\tilde{k})_s^2 = \frac{3}{2}\beta_{\parallel} \cos^2 \theta. \quad (3.147)$$

Using expression (3.141) for the slow mode yields  $(u_x/u_z)^s(\theta) = 0$  and expression (3.140) for the fast mode yields  $(u_x/u_z)^f(\theta) = \infty$ , implying that, in low beta limit, the  $\chi_u(\theta)$  solutions are

$$\beta_{\parallel} \ll 1 : \quad \chi_u^s(\theta) = 1; \quad \chi_u^f(\theta) = 0, \quad (3.148)$$

and the constant solutions approaching this limit are visible in the plot of figure 3. The opposite  $\beta_{\parallel} \gg 1$  limit in the CGL description is not particularly revealing since there is no obvious way to simplify the expansion, even for isotropic temperatures. For  $a_p = 1$ , the  $\beta_{\parallel} \gg 1$  expansion yields

$$\beta_{\parallel} \gg 1 : \quad (\tilde{\omega}/\tilde{k})_{sf}^2 = \frac{1}{2} \left(1 + \beta_{\parallel} \left(1 + \frac{1}{2} \cos^2 \theta\right)\right) \pm \frac{1}{2} \left[\beta_{\parallel} \sqrt{A} + \frac{1 - \frac{5}{2} \cos^2 \theta}{\sqrt{A}}\right], \quad (3.149)$$

$$A = \left(1 - \frac{5}{2} \cos^2 \theta\right)^2 + \sin^2 \theta \cos^2 \theta, \quad (3.150)$$

<sup>6</sup>One can directly work with the dispersion relation (3.137) and for  $\beta_{\parallel} \ll 1$  and  $a_p$  bounded and less than an extremely large values, the parameter  $(a - c)$  is positive and  $(a - c) \sim 1$ , on the other hand the parameters  $b \ll 1$  and  $c \ll 1$ , so an expansion of the square root is possible as

$$\frac{1}{2}(a + c) \pm \frac{1}{2}(a - c) \sqrt{1 + \frac{4b^2}{(a - c)^2}} \simeq \frac{1}{2}(a + c) \pm \frac{1}{2}(a - c) \left(1 + \frac{2b^2}{(a - c)^2}\right) \simeq \frac{1}{2}(a + c) \pm \frac{1}{2} \left[(a - c) + \frac{2b^2}{a}\right]. \quad (3.145)$$

The fast mode  $(\tilde{\omega}/\tilde{k})_f^2 \simeq a + (b^2/a) \simeq a$ , and the slow mode  $(\tilde{\omega}/\tilde{k})_s^2 \simeq c - (b^2/a) \simeq c$ , which yields low  $\beta_{\parallel}$  dispersion relations. With ‘less confidence’ the same result can be directly obtained from the velocity matrix (3.135) by neglecting the off-diagonal terms.

where we neglected a term proportional to  $1/\beta_{\parallel}$ . The expression results in the correct parallel limit  $(\tilde{\omega}/\tilde{k})_f^2(0) = \frac{3}{2}\beta_{\parallel}$ ,  $(\tilde{\omega}/\tilde{k})_s^2(0) = 1$  and the correct perpendicular limit  $(\tilde{\omega}/\tilde{k})_f^2(\pi/2) = 1 + \beta_{\parallel}$ ,  $(\tilde{\omega}/\tilde{k})_s^2(\pi/2) = 0$ . If the perpendicular fast mode limit is sufficient to be  $(\tilde{\omega}/\tilde{k})_f^2(\pi/2) = \beta_{\parallel}$  (which is acceptable since  $\beta_{\parallel} \gg 1$ ), only the terms proportional to  $\beta_{\parallel}$  can be retained and the fast mode dispersion relation can be simplified to

$$\beta_{\parallel} \gg 1: \quad (\tilde{\omega}/\tilde{k})_f^2 = \frac{1}{2}\beta_{\parallel} \left[ 1 + \frac{1}{2} \cos^2 \theta + \sqrt{\left(1 - \frac{5}{2} \cos^2 \theta\right)^2 + \sin^2 \theta \cos^2 \theta} \right]. \quad (3.151)$$

The expression (3.140) yields the velocity ratio of the fast mode as

$$\beta_{\parallel} \gg 1: \quad \left( \frac{u_x}{u_z} \right)^f = \frac{1 - \frac{5}{2} \cos^2 \theta + \sqrt{\left(1 - \frac{5}{2} \cos^2 \theta\right)^2 + \sin^2 \theta \cos^2 \theta}}{\sin \theta \cos \theta}, \quad (3.152)$$

that is independent of  $\beta_{\parallel}$ . For the slow mode wave the dispersion relation cannot be further simplified because if we collect only terms proportional to  $\beta_{\parallel}$ , the resulting expression will be non-zero everywhere except exactly at  $\theta = 0$  (and also  $\theta = \pi/2$ ). For  $\theta = 0$ , the entire contribution of  $(\tilde{\omega}/\tilde{k})_s^2(0) = 1$  comes from terms in (3.149) that are not proportional to  $\beta_{\parallel}$ . Nevertheless, the analytic velocity ratio (3.152) for the fast mode is sufficiently simple and to obtain analytic  $\chi_u^s(\theta)$  for the slow mode, it is actually possible to cheat a little by realizing that the slow and fast mode curves are always symmetric and that  $\chi_u^s(\theta) + \chi_u^f(\theta) = 1$ . Our final large beta results therefore are

$$\chi_u^f(\theta) = \left[ 1 + \left( \frac{1 - \frac{5}{2} \cos^2 \theta + \sqrt{\left(1 - \frac{5}{2} \cos^2 \theta\right)^2 + \sin^2 \theta \cos^2 \theta}}{\sin \theta \cos \theta} \right)^2 \right]^{-1}; \quad (3.153)$$

$$\chi_u^s(\theta) = 1 - \chi_u^f(\theta). \quad (3.154)$$

Solutions are plotted in figure 3 as dashed lines and fit the solutions of the full dispersion relations with  $\beta_{\parallel} = 10^3$  very nicely.

### 3.10.1. CGL special case $\beta_{\parallel} = 2/(4 - T_{\perp}/T_{\parallel})$

In this peculiar case when all 3 modes propagate in the parallel direction with the same phase speed, one again faces complications when calculating the velocity eigenvector, as with the MHD peculiar case when  $V_A = C_s$ . This CGL special case can be encountered only for  $\beta_{\parallel} \geq 1/2$  because the temperature anisotropy must be  $a_p \geq 0$ . We start with the velocity eigenvector matrix (3.138) and for the special case considered  $1 + \frac{1}{2}a_p\beta_{\parallel} - 2\beta_{\parallel} = 0$ , the expression (3.139) simplifies to  $a - c = (4\beta_{\parallel} - 1) \sin^2 \theta$ . One can already see that for strictly parallel propagation the entire matrix will be zero and it will not be possible to evaluate the eigenvector. Similarly to the MHD case, the  $\sin \theta$  can be pulled out of the matrix or alternatively one can use (3.140), (3.141) and calculate the limit. For the special case considered here, the parameter  $b = \frac{1}{2}(4\beta_{\parallel} - 2) \sin \theta \cos \theta$ . Considering again for simplicity only positive wavenumbers, the expression (3.140) yields in this particular case

$$\frac{u_x}{u_z} = \frac{4\beta_{\parallel} - 1}{4\beta_{\parallel} - 2} \tan \theta \pm \sqrt{\left( \frac{4\beta_{\parallel} - 1}{4\beta_{\parallel} - 2} \right)^2 \tan^2 \theta + 1}. \quad (3.155)$$

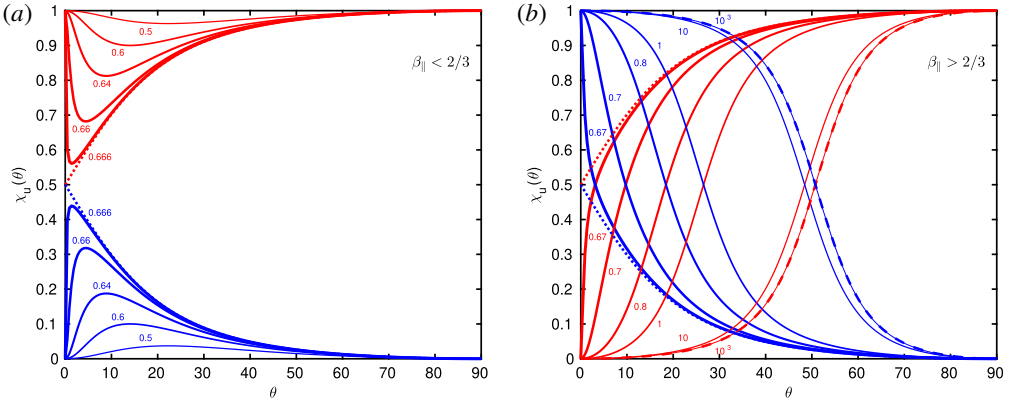


FIGURE 3. CGL velocity ratio  $\chi_u(\theta) = |u_{||}|^2/|u|^2$  for the slow mode (red lines) and the fast mode (blue lines) for different values of  $\beta_{||}$ , with isotropic temperature  $T_{\perp} = T_{||}$  (or the temperature anisotropy ratio  $a_p = 1$ ). Panel (a) is for  $\beta_{||} < 2/3$ , and the thickness of the lines increases with  $\beta_{||}$  when approaching the critical value of  $\beta_{||} = 2/3$ . The plotted lines have  $\beta_{||} = 0.5; 0.6; 0.64; 0.66; 0.666$ . Panel (b) is for  $\beta_{||} > 2/3$  and the thickness of the lines decreases with  $\beta_{||}$  when going away from the critical value of  $\beta_{||} = 2/3$ . The plotted lines have  $\beta_{||} = 0.67; 0.7; 0.8; 1.0; 10; 10^3$ . In both figures, the dotted lines are the critical case  $\beta_{||} = 2/3$  with solutions (3.155). In (b) the dashed lines represent the limit  $\beta_{||} \gg 1$  with solutions (3.153), (3.154) and the dashed lines nicely overlap with a thin solid line solution obtained for  $\beta_{||} = 10^3$ . The Alfvén mode has  $\chi_u(\theta) = 0$  regardless of the plasma beta and is not plotted.

The expression can be evaluated at  $\theta = 0$ , which yields for the slow mode  $(u_x/u_z)^s(0) = -1$  and for the fast mode  $(u_x/u_z)^f(0) = +1$ , implying that both modes start at

$$\beta_{||} = 2/(4 - a_p) : \quad \chi_u^s(0) = \frac{1}{2}; \quad \chi_u^f(0) = \frac{1}{2}. \quad (3.156)$$

For the perpendicular propagation the discussion is unchanged from the previous general case and yields  $\chi_u^s(\pi/2) = 1$  and  $\chi_u^f(\pi/2) = 0$ .

### 3.11. Empirical models with polytropic indices $\gamma_{||}$ , $\gamma_{\perp}$

The differences between MHD and CGL can be further clarified with polytropic indices. The linearized MHD pressure equation reads

$$\frac{\partial p}{\partial t} + \gamma p^{(0)} (\partial_x u_x + \partial_y u_y + \partial_z u_z) = 0. \quad (3.157)$$

One can introduce the concept of polytropic indices  $\gamma_{||}$  and  $\gamma_{\perp}$  to the CGL pressure equations (3.73), (3.74) and write

$$\frac{\partial p_{||}}{\partial t} + p_{||}^{(0)} (\partial_x u_x + \partial_y u_y) + \gamma_{||} p_{||}^{(0)} \partial_z u_z = 0; \quad (3.158)$$

$$\frac{\partial p_{\perp}}{\partial t} + \gamma_{\perp} p_{\perp}^{(0)} (\partial_x u_x + \partial_y u_y) + p_{\perp}^{(0)} \partial_z u_z = 0, \quad (3.159)$$

where for the CGL model  $\gamma_{\parallel} = 3$  and  $\gamma_{\perp} = 2$ .<sup>7</sup> The reasoning for such a construction is to enable a smooth transition between two extreme cases – the CGL model on one side, and the isothermal model  $\gamma_{\parallel} = 1$  and  $\gamma_{\perp} = 1$  on the other side. The important observation is that in the parallel pressure equation the  $\gamma_{\parallel}$  acts only on the parallel velocity component  $u_z$ , and in the perpendicular pressure equation the  $\gamma_{\perp}$  acts only on the perpendicular velocity components  $u_x, u_y$ . In MHD, the  $\gamma$  acts on all three components  $u_x, u_y, u_z$ . The values of polytropic indices are directly related to the number of degrees of freedom  $i$ , through the well-known relation  $\gamma = (i + 2)/i$ . In the MHD description the number of degrees of freedom is  $i = 3$  and so  $\gamma = 5/3$ . In contrast, the CGL description can be interpreted as being composed of a (strongly coupled) mixture of one-dimensional and two-dimensional dynamics, where for the parallel direction  $i = 1$  and  $\gamma_{\parallel} = 3$  and for the perpendicular direction  $i = 2$  and  $\gamma_{\perp} = 2$ . The apparently subtle, but nevertheless fundamental difference between (3.157) and (3.158), (3.159) is the core of many differences between the MHD and CGL descriptions, such as for example the effect of the slow mode becoming faster than the Alfvén mode. The two systems remain different even if the mean CGL pressures (or temperatures) are prescribed to be isotropic  $p_{\parallel}^{(0)} = p_{\perp}^{(0)}$  and equal to the mean MHD pressure  $p^{(0)}$ . This means that the MHD does not match the CGL description even if the distribution function is isotropic, as for example Maxwellian.

We briefly consider models with polytropic indices  $\gamma_{\parallel}, \gamma_{\perp}$  that are based on the linearized pressure equations (3.158), (3.159). If these pressure equations are ‘un-linearized’, i.e. if one performs a ‘reverse engineering’ procedure to obtain nonlinear equations (a procedure that is possible to do because of the known guidance by the CGL model), the nonlinear equations read

$$\frac{dp_{\parallel}}{dt} + p_{\parallel} \nabla \cdot \mathbf{u} + (\gamma_{\parallel} - 1) p_{\parallel} \hat{\mathbf{b}} \cdot \nabla \mathbf{u} \cdot \hat{\mathbf{b}} = 0; \quad (3.160)$$

$$\frac{dp_{\perp}}{dt} + \gamma_{\perp} p_{\perp} \nabla \cdot \mathbf{u} - (\gamma_{\perp} - 1) p_{\perp} \hat{\mathbf{b}} \cdot \nabla \mathbf{u} \cdot \hat{\mathbf{b}} = 0, \quad (3.161)$$

which can be rewritten as

$$\frac{d}{dt} \left( \frac{p_{\parallel} |\mathbf{B}|^{\gamma_{\parallel}-1}}{\rho^{\gamma_{\parallel}}} \right) = 0; \quad \frac{d}{dt} \left( \frac{p_{\perp}}{\rho |\mathbf{B}|^{\gamma_{\perp}-1}} \right) = 0. \quad (3.162)$$

Empirical fluid models of this type were studied for example by Hau & Sonnerup (1993) and Hau *et al.* (1993) with the motivation that spacecraft observations in the Earth’s magnetosheath often show behaviour that cannot be explained by the CGL values of  $\gamma_{\parallel} = 3, \gamma_{\perp} = 2$  or by the isothermal values of  $\gamma_{\parallel} = 1, \gamma_{\perp} = 1$ , but can be fitted by the values that lie somewhere in between (and actually quite close to the isothermal case). For a direct comparison with the CGL model, the equations (3.162) can be rewritten to a form

$$\frac{d}{dt} \left( \frac{p_{\parallel} |\mathbf{B}|^2}{\rho^3} \right) = -(\gamma_{\parallel} - 3) \left( \frac{p_{\parallel} |\mathbf{B}|^2}{\rho^3} \right) \frac{\rho}{|\mathbf{B}|} \frac{d}{dt} \left( \frac{|\mathbf{B}|}{\rho} \right); \quad (3.163)$$

<sup>7</sup>One can then define the parallel and perpendicular sound speeds as  $C_{\parallel} = \sqrt{\gamma_{\parallel} p_{\parallel}^{(0)} / \rho_0}$  and  $C_{\perp} = \sqrt{\gamma_{\perp} p_{\perp}^{(0)} / \rho_0}$ , since in this model the parallel acoustic mode will have a dispersion  $\omega = \pm C_{\parallel} k_{\parallel}$  and the perpendicular fast mode  $\omega = \pm k_{\perp} \sqrt{V_A^2 + C_{\perp}^2}$ , as can be easily verified.



$$\frac{d}{dt} \left( \frac{p_{\perp}}{\rho |\mathbf{B}|} \right) = (\gamma_{\perp} - 2) \left( \frac{p_{\perp}}{\rho |\mathbf{B}|} \right) \frac{1}{|\mathbf{B}|} \frac{d}{dt} |\mathbf{B}|, \quad (3.164)$$

or to an interesting alternative form

$$\frac{d}{dt} \ln \left( \frac{p_{\parallel} |\mathbf{B}|^2}{\rho^3} \right) = -(\gamma_{\parallel} - 3) \frac{d}{dt} \ln \left( \frac{|\mathbf{B}|}{\rho} \right); \quad (3.165)$$

$$\frac{d}{dt} \ln \left( \frac{p_{\perp}}{\rho |\mathbf{B}|} \right) = (\gamma_{\perp} - 2) \frac{d}{dt} \ln |\mathbf{B}|. \quad (3.166)$$

Obviously, for CGL values of  $\gamma_{\parallel} = 3$ ,  $\gamma_{\perp} = 2$  the right-hand sides disappear and one obtains the classical CGL expressions. The right-hand sides of the above expressions can be compared with equations of Chew *et al.* (1956) equations (2.82), (2.83) which contain the gyrotropic heat flux contributions. Therefore, the deviations from the CGL values of  $\gamma_{\parallel} = 3$ ,  $\gamma_{\perp} = 2$  can be considered to offer a very simple empirical model (perhaps too simple) that shows how the heat flux contributions modify the CGL dynamics. The dispersion relations for this fluid model can be found in Hau *et al.* (1993) and also in Abraham-Shrauner (1973), who studied an even more general empirical model with four polytropic indices (essentially, the  $\gamma_{\parallel}$ ,  $\gamma_{\perp}$  acting on  $|\mathbf{B}|$  and  $\rho$  in (3.162) can be further split to two independent parameters). Interestingly, complex polytropic indices have been suggested to model Landau damping (Belmont & Mazelle 1992). Considering the linearized CGL equations (3.68)–(3.72) and linearized polytropic pressure equations (3.158)–(3.159) written in the  $x$ – $z$  plane yields the dispersion relations for this model. The Alfvén wave is still decoupled from the rest of the system and the Alfvén wave dispersion relation is unaffected by the different pressure equations, therefore also yielding the same oblique firehose instability threshold. What is modified is the dispersion relation of the slow and fast waves. The velocity equations read

$$\frac{\partial^2 u_x}{\partial t^2} - \frac{p_{\perp}^{(0)}}{\rho_0} (\gamma_{\perp} \partial_x^2 u_x + \partial_x \partial_z u_z) + \left( \frac{p_{\parallel}^{(0)}}{\rho_0} - \frac{p_{\perp}^{(0)}}{\rho_0} - V_A^2 \right) \partial_z^2 u_x - V_A^2 \partial_x^2 u_x = 0; \quad (3.167)$$

$$\frac{\partial^2 u_z}{\partial t^2} - \gamma_{\parallel} \frac{p_{\parallel}^{(0)}}{\rho_0} \partial_z^2 u_z - \frac{p_{\perp}^{(0)}}{\rho_0} \partial_x \partial_z u_x = 0, \quad (3.168)$$

and the transformation to Fourier space yields the velocity matrix

$$\begin{pmatrix} \omega^2 - \left( \gamma_{\perp} \frac{p_{\perp}^{(0)}}{\rho_0} + V_A^2 \right) k_{\perp}^2 + \left( \frac{p_{\parallel}^{(0)}}{\rho_0} - \frac{p_{\perp}^{(0)}}{\rho_0} - V_A^2 \right) k_{\parallel}^2; & -\frac{p_{\perp}^{(0)}}{\rho_0} k_{\perp} k_{\parallel} \\ -\frac{p_{\perp}^{(0)}}{\rho_0} k_{\perp} k_{\parallel}; & \omega^2 - \gamma_{\parallel} \frac{p_{\parallel}^{(0)}}{\rho_0} k_{\parallel}^2 \end{pmatrix} \begin{pmatrix} u_x \\ u_z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad (3.169)$$

or alternatively

$$\begin{pmatrix} \tilde{\omega}^2 - \left( \gamma_{\perp} a_p \frac{\beta_{\parallel}}{2} + 1 \right) \tilde{k}_{\perp}^2 - \left( 1 + \frac{\beta_{\parallel}}{2} (a_p - 1) \right) \tilde{k}_{\parallel}^2; & -\frac{1}{2} a_p \beta_{\parallel} \tilde{k}_{\perp} \tilde{k}_{\parallel} \\ -\frac{1}{2} a_p \beta_{\parallel} \tilde{k}_{\perp} \tilde{k}_{\parallel}; & \tilde{\omega}^2 - \gamma_{\parallel} \frac{\beta_{\parallel}}{2} \tilde{k}_{\parallel}^2 \end{pmatrix} \begin{pmatrix} u_x \\ u_z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \quad (3.170)$$

Calculating the determinant yields a dispersion relation for the slow and fast modes in the form

$$\tilde{\omega}^4 - A_2 \tilde{\omega}^2 + A_0 = 0; \quad (3.171)$$

$$A_2 = \tilde{k}_\perp^2 \left( 1 + \frac{\gamma_\perp}{2} a_p \beta_\parallel \right) + \tilde{k}_\parallel^2 \left( 1 + \frac{1}{2} \beta_\parallel (\gamma_\parallel - 1) + \frac{1}{2} a_p \beta_\parallel \right); \quad (3.172)$$

$$A_0 = \frac{\gamma_\parallel}{2} \tilde{k}_\parallel^2 \beta_\parallel \left[ \tilde{k}_\parallel^2 \left( 1 - \frac{1}{2} \beta_\parallel + \frac{1}{2} a_p \beta_\parallel \right) + \tilde{k}_\perp^2 \left( 1 + \frac{\gamma_\perp}{2} a_p \beta_\parallel - \frac{1}{2\gamma_\parallel} a_p^2 \beta_\parallel \right) \right]. \quad (3.173)$$

Solutions for parallel propagation ( $k_\perp = 0$ ) are the usual CGL Alfvén mode and the ion-acoustic mode  $\tilde{\omega} = \pm \sqrt{(\gamma_\parallel/2)\beta_\parallel} \tilde{k}_\parallel$ , and for perpendicular propagation ( $k_\parallel = 0$ ) the fast mode  $\tilde{\omega} = \pm \sqrt{1 + (\gamma_\perp/2)a_p \beta_\parallel} \tilde{k}_\perp$ . Similarly to the CGL description it can be shown that  $A_2^2 - 4A_0 \geq 0$  because of the trick (3.137), implying that the slow mode becomes unstable when  $A_0 < 0$ , i.e. when

$$\left[ k_\parallel^2 \left( 1 - \frac{1}{2} \beta_\parallel + \frac{1}{2} a_p \beta_\parallel \right) + k_\perp^2 \left( 1 + \frac{\gamma_\perp}{2} a_p \beta_\parallel - \frac{1}{2\gamma_\parallel} a_p^2 \beta_\parallel \right) \right] < 0. \quad (3.174)$$

The two competing instabilities are again the parallel firehose instability and the highly oblique mirror instability. The parallel firehose instability has the same threshold as obtained previously. However, the mirror instability threshold reads

$$1 + \frac{\gamma_\perp}{2} a_p \beta_\parallel - \frac{1}{2\gamma_\parallel} a_p^2 \beta_\parallel < 0, \quad (3.175)$$

which is consistent with the result of Hau & Sonnerup (1993), their equation (14), written in the form<sup>8</sup>

$$\gamma_\parallel \beta_\parallel < \frac{\beta_\perp^2}{2 + \gamma_\perp \beta_\perp}. \quad (3.176)$$

For  $\gamma_\parallel = 3$ ,  $\gamma_\perp = 2$ , the condition (3.175) is naturally equivalent to the CGL mirror threshold. However, the condition offers a temptingly simple result to keep the perpendicular polytropic index unchanged  $\gamma_\perp = 2$  and modifies the parallel polytropic index to  $\gamma_\parallel = 1/2$ . In this way, the mirror threshold matches the correct kinetic threshold (3.94). Nevertheless, this is just an interesting empirical concept, it fixes the highly oblique mirror threshold but at the same time moves the slow and fast mode dispersion relations in other directions, for example the parallel sound speed in such a model is  $C_\parallel^2 = \frac{1}{2} p_\parallel^{(0)} / \rho_0$ , which appears to be unrealistically low. Such a ‘tuning’ of models with free parameters – only in one particular direction and only at the level of the dispersion relations – can be sometimes useful for interpretation of observational data, however, such models will usually lead to unphysical situations. We will not consider models with polytropic indices further in this text. Instead, to obtain a better match with kinetic description, it is advisable to use the correct CGL values of  $\gamma_\parallel = 3$ ,  $\gamma_\perp = 2$  and focus on fluid models that derive the correct heat flux contributions  $q_\parallel$  and  $q_\perp$ .

<sup>8</sup>Note that the equation in the cited paper has the opposite sign, which is not a typo, the authors just call the mirror threshold a condition when a mode is stable – opposite to the more usual condition when a mode becomes unstable. Similarly, the authors write the condition for the firehose-stable region as  $\beta_\parallel \leq 2 + \beta_\perp$ .

## 3.12. CGL and radially expanding solar wind

Similarly to MHD, the CGL fluid model can be very useful for understanding the evolution of temperature in the radially expanding solar wind. A good discussion can be found for example in Matteini *et al.* (2007), Matteini *et al.* (2013) and references therein. Let us first consider the MHD case. Combining the pressure equation  $dp/dt + \gamma \nabla \cdot \mathbf{u} = 0$  with the density equation  $d\rho/dt + \rho \nabla \cdot \mathbf{u} = 0$  yields

$$\text{MHD: } \frac{d}{dt} \left( \frac{p}{\rho^\gamma} \right) = 0; \quad \Rightarrow \quad \frac{d}{dt} \left( \frac{T}{\rho^{\gamma-1}} \right) = 0, \quad (3.177)$$

which in a steady state (or co-moving with the convective derivative) implies  $T \sim \rho^{\gamma-1}$ . By considering simple radially expanding solar wind, one prescribes that the density  $\rho$  evolves with the heliocentric distance  $r$  as  $\rho \sim 1/r^2$ , and by using  $\gamma = 5/3$ , the ideal MHD model yields (in the absence of any dissipation and associated turbulent heating)

$$T \sim r^{-4/3}. \quad (3.178)$$

In contrast, the CGL equations (3.41) are rewritten as

$$\text{CGL: } \frac{d}{dt} \left( \frac{T_{\parallel} |\mathbf{B}|^2}{\rho^2} \right) = 0; \quad \frac{d}{dt} \left( \frac{T_{\perp}}{|\mathbf{B}|} \right) = 0, \quad (3.179)$$

which in a steady state implies

$$T_{\parallel} \sim \frac{\rho^2}{|\mathbf{B}|^2}; \quad T_{\perp} \sim |\mathbf{B}|; \quad \frac{T_{\perp}}{T_{\parallel}} = \frac{|\mathbf{B}|^3}{\rho^2}. \quad (3.180)$$

For the magnetic field it is possible to consider profiles of roughly  $|\mathbf{B}| \sim 1/r^2$ , or  $|\mathbf{B}| \sim 1/r$  (one can be more precise and consider the Parker spiral profile). The first magnetic field profile yields

$$|\mathbf{B}| \sim 1/r^2; \quad \Rightarrow \quad T_{\parallel} \sim \text{const.}; \quad T_{\perp} \sim 1/r^2; \quad \frac{T_{\perp}}{T_{\parallel}} \sim 1/r^2. \quad (3.181)$$

The second magnetic field profile yields

$$|\mathbf{B}| \sim 1/r; \quad \Rightarrow \quad T_{\parallel} \sim 1/r^2; \quad T_{\perp} \sim 1/r; \quad \frac{T_{\perp}}{T_{\parallel}} \sim r. \quad (3.182)$$

There is also a curious case when the magnetic field profile is prescribed to be  $|\mathbf{B}| \sim r^{-4/3}$ , i.e. a profile between two cases of  $1/r$  and  $1/r^2$ , that yields

$$|\mathbf{B}| \sim r^{-4/3}; \quad \Rightarrow \quad T_{\parallel} \sim r^{-4/3}; \quad T_{\perp} \sim r^{-4/3}; \quad \frac{T_{\perp}}{T_{\parallel}} \sim \text{const.}, \quad (3.183)$$

and the temperature anisotropy stays constant. The estimations are of course valid only until a firehose threshold or a mirror threshold is reached. For example, evolution according to (3.181) will lead the system to a firehose threshold, and evolution according to (3.182) will lead the system to a mirror threshold. In general, if we fix the density profile to  $\rho \sim 1/r^2$ , any magnetic field profile steeper than  $r^{-4/3}$  will evolve the system towards a firehose threshold, and a magnetic field profile shallower

than  $r^{-4/3}$  will evolve the system towards a mirror threshold. The estimations do not reveal the presence of the firehose instability or the mirror instability in the CGL model, since the solutions are written in a steady state, and concern only evolution of mean temperatures (i.e. the solutions should be written with  $T_{\parallel}^{(0)}, T_{\perp}^{(0)}$ ). To find an instability, one of course needs to consider evolution equations for fluctuating variables and to analyse the associated dispersion relations, since it is the fluctuations that become unstable. Further discussion can be found in Hunana & Zank (2017).

From a linear perspective, the CGL model has to be accompanied by the Hall term and FLR corrections, so that the instabilities are stabilized at small spatial scales. However, considering nonlinear evolution with finite amplitudes of fluctuations, even the non-dispersive CGL model is stabilized. For example, considering parallel propagation, Tenerani, Velli & Hellinger (2017) has shown that the firehose instability criterion reads (see (9) in that paper)

$$V_A^2 + \frac{1}{\rho_0} \frac{p_{\perp}^{(0)} - p_{\parallel}^{(0)}}{1 + \left| \frac{\mathbf{B}_{\perp}}{B_0} \right|^2} < 0; \quad \Rightarrow \quad 1 + \frac{\beta_{\parallel}}{2} \frac{a_p - 1}{1 + \left| \frac{\mathbf{B}_{\perp}}{B_0} \right|^2} < 0, \quad (3.184)$$

i.e. for infinitely small amplitudes of the magnetic field, the usual firehose criterion is obtained. However, when the CGL system reaches firehose-unstable regime and the amplitude of the fluctuations starts to grow, the firehose criterion is modified according to (3.184). Therefore, the amplitude of the fluctuations cannot grow without bounds, and beyond some critical amplitude, the system becomes again stabilized. The condition (3.184) nicely demonstrates how the CGL system is stabilized during nonlinear evolution, once the firehose instability is reached. For further discussion, see Tenerani *et al.* (2017), Tenerani & Velli (2018).

To finish this subsection, the heuristically generalized CGL model with polytropic indices  $\gamma_{\parallel}, \gamma_{\perp}$  yields that, in the steady state, the mean temperatures evolve according to

$$T_{\parallel} = \left( \frac{\rho}{|\mathbf{B}|} \right)^{\gamma_{\parallel}-1}; \quad T_{\perp} = |\mathbf{B}|^{\gamma_{\perp}-1}; \quad \frac{T_{\perp}}{T_{\parallel}} = \frac{|\mathbf{B}|^{\gamma_{\parallel}+\gamma_{\perp}-2}}{\rho^{\gamma_{\parallel}-1}}, \quad (3.185)$$

and such heuristic models can yield a wide array of possible temperature profiles.

### 3.13. CGL model – summary

The usual ‘combining’ of evolution equations (by performing  $\partial/\partial t$  and substituting one equation into another) is very useful to gain insight into the linear eigenmodes of a considered system. However, for more advanced fluid models, such a procedure can become very analytically involved, and it is much easier to calculate the determinant of the entire system with analytic software, such as Maple or Mathematica. Therefore, for more advanced fluid models we often do not write down the final dispersion relation for arbitrary propagation angle (which might be uneconomically large to write down), and we only provide normalized, linearized in the  $x$ – $z$  plane and Fourier transformed equations. For easy comparison with more advanced fluid models, it is beneficial to write the final CGL equations in the same form. It is useful to work in normalized units, and we drop writing the tilde. The normalized, linearized in the  $x$ – $z$  plane, and Fourier transformed CGL equations read

$$-\omega\rho + k_{\perp}u_x + k_{\parallel}u_z = 0; \quad (3.186)$$

$$-\omega u_x + \frac{\beta_{\parallel}}{2}k_{\perp}p_{\perp} - v_{A\parallel}^2 k_{\parallel}B_x + k_{\perp}B_z = 0; \quad (3.187)$$

$$-\omega u_y - v_{A\parallel}^2 k_{\parallel}B_y = 0; \quad (3.188)$$

$$-\omega u_z + \frac{\beta_{\parallel}}{2}k_{\parallel}p_{\parallel} + \frac{\beta_{\parallel}}{2}(1 - a_p)k_{\perp}B_x = 0; \quad (3.189)$$

$$-\omega B_x - k_{\parallel}u_x = 0; \quad -\omega B_y - k_{\parallel}u_y = 0; \quad -\omega B_z + k_{\perp}u_x = 0; \quad (3.190)$$

$$-\omega p_{\parallel} + k_{\perp}u_x + 3k_{\parallel}u_z = 0; \quad (3.191)$$

$$-\omega p_{\perp} + 2a_p k_{\perp}u_x + a_p k_{\parallel}u_z = 0, \quad (3.192)$$

where we have introduced the normalized parallel Alfvén speed  $v_{A\parallel}^2 \equiv 1 + (\beta_{\parallel}/2)(a_p - 1)$  (the tilde on  $v_{A\parallel}$  is dropped too). The determinant can be easily calculated and factorized in Maple, yielding the CGL dispersion relation in normalized units (tilde are dropped)

$$(\omega^2 - k_{\parallel}^2 v_{A\parallel}^2)(\omega^4 - A_2 \omega^2 + A_0) = 0; \quad (3.193)$$

$$A_2 = k_{\perp}^2 (1 + a_p \beta_{\parallel}) + k_{\parallel}^2 (v_{A\parallel}^2 + \frac{3}{2} \beta_{\parallel}); \quad (3.194)$$

$$A_0 = \frac{3}{2} k_{\perp}^2 \beta_{\parallel} [k_{\parallel}^2 v_{A\parallel}^2 + k_{\perp}^2 (1 + a_p \beta_{\parallel} - \frac{1}{6} a_p^2 \beta_{\parallel})], \quad (3.195)$$

which of course agrees with the previously obtained CGL solutions for the Alfvén mode (3.81), and the slow and fast modes (3.90).

#### 4. Hall-CGL model

Here we consider a slightly generalized CGL model when the Hall term is included in the induction equation. Concerning the isotropic Hall-MHD model, a good comparison with solutions of kinetic theory can be found for example in Howes (2009). Before we proceed with the discussion of dispersion relations, we want to point out how the Hall term and the electron pressure contributions and the electron inertia contributions modify the usual ‘CGL conservation equations’ for protons. The general electric field was already written down in (3.15). For a moment it is beneficial to consider this most general form including the electron inertia and the electron pressure contributions and separate the electric field into two parts

$$\mathbf{E} = -\frac{1}{c} \mathbf{u}_p \times \mathbf{B} + \mathbf{E}_H \quad (4.1)$$

$$\mathbf{E}_H = \frac{1}{4\pi en} (\nabla \times \mathbf{B}) \times \mathbf{B} - \frac{1}{en} \nabla \cdot \mathbf{p}_e - \frac{m_e}{e} \left( \frac{\partial \mathbf{u}_e}{\partial t} + \mathbf{u}_e \cdot \nabla \mathbf{u}_e \right). \quad (4.2)$$

The notation  $\mathbf{E}_H$  is perhaps slightly misleading since only the first term in (4.2) is usually called the Hall electric field. Note that the first term in  $\mathbf{E}_H$  still has to be used in the proton momentum equation even for the simplest CGL (or MHD) models. The induction equation then reads

$$\frac{\partial \mathbf{B}}{\partial t} = \nabla \times (\mathbf{u}_p \times \mathbf{B}) - c \nabla \times \mathbf{E}_H, \quad (4.3)$$

which is then rewritten with the use of the convective derivative as

$$\frac{d\mathbf{B}}{dt} = -\mathbf{B} \nabla \cdot \mathbf{u}_p + \mathbf{B} \cdot \nabla \mathbf{u}_p - c \nabla \times \mathbf{E}_H; \quad (4.4)$$

$$\frac{\hat{\mathbf{b}}}{|\mathbf{B}|} \cdot \frac{d\mathbf{B}}{dt} = -\nabla \cdot \mathbf{u}_p + \hat{\mathbf{b}} \cdot \nabla \mathbf{u}_p \cdot \hat{\mathbf{b}} - \frac{c\hat{\mathbf{b}}}{|\mathbf{B}|} \cdot (\nabla \times \mathbf{E}_H). \quad (4.5)$$

The last expression is then used in the completely general equations (3.46), (3.47), which yields

$$\frac{d}{dt} \left( \frac{p_{\parallel p} |\mathbf{B}|^2}{\rho_p^3} \right) = \frac{|\mathbf{B}|^2}{\rho_p^3} \left( \frac{dp_{\parallel p}}{dt} + p_{\parallel p} \nabla \cdot \mathbf{u}_p + 2p_{\parallel p} \hat{\mathbf{b}} \cdot \nabla \mathbf{u}_p \cdot \hat{\mathbf{b}} - 2p_{\parallel p} \frac{c\hat{\mathbf{b}}}{|\mathbf{B}|} \cdot (\nabla \times \mathbf{E}_H) \right); \quad (4.6)$$

$$\frac{d}{dt} \left( \frac{p_{\perp p}}{\rho_p |\mathbf{B}|} \right) = \frac{1}{\rho_p |\mathbf{B}|} \left( \frac{dp_{\perp p}}{dt} + 2p_{\perp p} \nabla \cdot \mathbf{u}_p - p_{\perp p} \hat{\mathbf{b}} \cdot \nabla \mathbf{u}_p \cdot \hat{\mathbf{b}} + p_{\perp p} \frac{c\hat{\mathbf{b}}}{|\mathbf{B}|} \cdot (\nabla \times \mathbf{E}_H) \right). \quad (4.7)$$

By cancelling pressure equations (3.3), (3.4), that are also the pressure equations of the Hall-CGL model, one obtains that for a general Hall-CGL model the electric field  $\mathbf{E}_H$  modifies the conservation laws according to

$$\frac{d}{dt} \left( \frac{p_{\parallel p} |\mathbf{B}|^2}{\rho_p^3} \right) = \frac{p_{\parallel p} |\mathbf{B}|^2}{\rho_p^3} \left[ -2 \frac{c\hat{\mathbf{b}}}{|\mathbf{B}|} \cdot (\nabla \times \mathbf{E}_H) \right]; \quad (4.8)$$

$$\frac{d}{dt} \left( \frac{p_{\perp p}}{\rho_p |\mathbf{B}|} \right) = \frac{p_{\perp p}}{\rho_p |\mathbf{B}|} \left[ \frac{c\hat{\mathbf{b}}}{|\mathbf{B}|} \cdot (\nabla \times \mathbf{E}_H) \right], \quad (4.9)$$

implying the Hall term breaks the first and second adiabatic invariants.

#### 4.1. Hall-CGL model with cold electrons

Here, we want to study the simplest Hall-CGL model for the proton species. We neglect the electron inertia and we make the electrons cold with  $\mathbf{p}_e = 0$ . The full model is described by the usual CGL equations (3.36)–(3.39), but the induction equation now reads

$$\frac{\partial \mathbf{B}}{\partial t} = \nabla \times (\mathbf{u}_p \times \mathbf{B}) - \frac{c}{4\pi e} \nabla \times \left( \frac{1}{n} (\nabla \times \mathbf{B}) \times \mathbf{B} \right). \quad (4.10)$$

The induction equation is linearized as

$$\frac{\partial \mathbf{B}}{\partial t} \stackrel{\text{lin}}{=} \nabla \times (\mathbf{u}_p \times \mathbf{B}_0) - \frac{c}{4\pi e n_0} \nabla \times ((\nabla \times \mathbf{B}) \times \mathbf{B}_0). \quad (4.11)$$

By using (3.67), direct calculation of the second term yields

$$\nabla \times ((\nabla \times \mathbf{B}) \times \mathbf{B}_0) = B_0 \begin{pmatrix} \partial_y \partial_z B_z - \partial_z^2 B_y \\ -\partial_x \partial_z B_z + \partial_z^2 B_x \\ \partial_x \partial_z B_y - \partial_y \partial_z B_x \end{pmatrix}, \quad (4.12)$$

and by using  $c/(4\pi e n_0) = V_A^2/(\Omega_p B_0)$ , the Hall-term contributions are

$$c \nabla \times \mathbf{E}_H \stackrel{\text{lin}}{=} \frac{V_A^2}{\Omega_p} \begin{pmatrix} \partial_y \partial_z B_z - \partial_z^2 B_y \\ -\partial_x \partial_z B_z + \partial_z^2 B_x \\ \partial_x \partial_z B_y - \partial_y \partial_z B_x \end{pmatrix}. \quad (4.13)$$

The entire linearized induction equation reads

$$\frac{\partial \mathbf{B}}{\partial t} \stackrel{\text{lin}}{=} B_0 \begin{pmatrix} \partial_z u_x \\ \partial_z u_y \\ -\partial_x u_x - \partial_y u_y \end{pmatrix} - \frac{V_A^2}{\Omega_p} \begin{pmatrix} \partial_y \partial_z B_z - \partial_z^2 B_y \\ -\partial_x \partial_z B_z + \partial_z^2 B_x \\ \partial_x \partial_z B_y - \partial_y \partial_z B_x \end{pmatrix}, \quad (4.14)$$

which when written in the  $x$ - $z$  plane (all  $\partial_y = 0$ ) yields

$$\frac{\partial B_x}{\partial t} = B_0 \partial_z u_x + \frac{V_A^2}{\Omega_p} \partial_z^2 B_y; \quad (4.15)$$

$$\frac{\partial B_y}{\partial t} = B_0 \partial_z u_y + \frac{V_A^2}{\Omega_p} (\partial_x \partial_z B_z - \partial_z^2 B_x); \quad (4.16)$$

$$\frac{\partial B_z}{\partial t} = -B_0 \partial_x u_x - \frac{V_A^2}{\Omega_p} \partial_x \partial_z B_y. \quad (4.17)$$

These equations replace the  $\partial B/\partial t$  equations in the linearized CGL system written in the  $x$ - $z$  plane. And the trouble is immediately apparent. In contrast to the CGL system, we are no longer able to separate the Alfvén mode which was before exclusively in the  $B_y$ ,  $u_y$  components. Now, all the components are coupled and all three modes, the Alfvén mode, the slow mode and the fast mode are coupled. This is not surprising, since the same situation is in the Hall-MHD. In this case, we have no other choice and we have to calculate the dispersion relation for all three modes. The dispersion relation can be solved analytically only for special cases and, in general, the dispersion relation has to be solved numerically. Let us quickly consider the two special cases of parallel and perpendicular propagation that can be solved analytically.

#### 4.2. Parallel propagation

For propagation parallel to the mean magnetic field, all  $\partial_x = 0$  and the familiar sound mode/ion-acoustic mode decouples in the  $u_z$ ,  $p_{\parallel}$  components as

$$\frac{\partial^2 u_z}{\partial t^2} - \frac{3p_{\parallel}^{(0)}}{\rho_0} \partial_z^2 u_z = 0, \quad (4.18)$$

with the familiar dispersion relation  $\omega^2 = C_{\parallel}^2 k_{\parallel}^2$ . The dispersion relation of this mode does not depend on the Hall term. However, the other two modes are now coupled through

$$\frac{\partial u_x}{\partial t} - \frac{1}{B_0} \left( V_A^2 - \frac{p_{\parallel}^{(0)}}{\rho_0} + \frac{p_{\perp}^{(0)}}{\rho_0} \right) \partial_z B_x = 0; \quad (4.19)$$

$$\frac{\partial u_y}{\partial t} - \frac{1}{B_0} \left( V_A^2 - \frac{p_{\parallel}^{(0)}}{\rho_0} + \frac{p_{\perp}^{(0)}}{\rho_0} \right) \partial_z B_y = 0; \quad (4.20)$$

$$\frac{\partial B_x}{\partial t} = B_0 \partial_z u_x + \frac{V_A^2}{\Omega_p} \partial_z^2 B_y; \quad (4.21)$$

$$\frac{\partial B_y}{\partial t} = B_0 \partial_z u_y - \frac{V_A^2}{\Omega_p} \partial_z^2 B_x. \quad (4.22)$$



If the last terms in the  $B$ -field equations are neglected, i.e. if the Hall term is neglected, the two modes decouple and one naturally recovers the CGL parallel propagating Alfvén modes, both propagating with the phase speed

$$v_{A\parallel}^2 = V_A^2 - \frac{p_{\parallel}^{(0)}}{\rho_0} + \frac{p_{\perp}^{(0)}}{\rho_0} = V_A^2 \left( 1 + \frac{\beta_{\parallel}}{2} (a_p - 1) \right). \quad (4.23)$$

Applying  $\partial_t$  to the  $B$ -field equations yields

$$\frac{\partial^2 B_x}{\partial t^2} = v_{A\parallel}^2 \partial_z^2 B_x + \frac{V_A^2}{\Omega_p^2} \partial_z^2 \frac{\partial B_y}{\partial t}; \quad (4.24)$$

$$\frac{\partial^2 B_y}{\partial t^2} = v_{A\parallel}^2 \partial_z^2 B_y - \frac{V_A^2}{\Omega_p^2} \partial_z^2 \frac{\partial B_x}{\partial t}, \quad (4.25)$$

which when transformed to Fourier space read

$$\begin{pmatrix} \omega^2 - v_{A\parallel}^2 k_{\parallel}^2; & +i \frac{V_A^2}{\Omega_p^2} k_{\parallel}^2 \omega \\ -i \frac{V_A^2}{\Omega_p^2} k_{\parallel}^2 \omega; & \omega^2 - v_{A\parallel}^2 k_{\parallel}^2 \end{pmatrix} \begin{pmatrix} B_x \\ B_y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad (4.26)$$

and the zero determinant requirement implies

$$\omega^4 - \omega^2 k_{\parallel}^2 \left( 2v_{A\parallel}^2 + k_{\parallel}^2 \frac{V_A^4}{\Omega_p^2} \right) + k_{\parallel}^4 v_{A\parallel}^4 = 0. \quad (4.27)$$

The solutions of the dispersion relation (quadratic polynomial in  $\omega^2$ ) is simply

$$\omega_w^2 = k_{\parallel}^2 V_A^2 \left[ 1 + \frac{\beta_{\parallel}}{2} (a_p - 1) + \frac{k_{\parallel}^2 V_A^2}{2\Omega_p^2} + \frac{|k_{\parallel}| V_A}{\Omega_p} \sqrt{1 + \frac{\beta_{\parallel}}{2} (a_p - 1) + \frac{k_{\parallel}^2 V_A^2}{4\Omega_p^2}} \right]; \quad (4.28)$$

$$\omega_{IC}^2 = k_{\parallel}^2 V_A^2 \left[ 1 + \frac{\beta_{\parallel}}{2} (a_p - 1) + \frac{k_{\parallel}^2 V_A^2}{2\Omega_p^2} - \frac{|k_{\parallel}| V_A}{\Omega_p} \sqrt{1 + \frac{\beta_{\parallel}}{2} (a_p - 1) + \frac{k_{\parallel}^2 V_A^2}{4\Omega_p^2}} \right], \quad (4.29)$$

where, importantly,  $\sqrt{k_{\parallel}^2} = |k_{\parallel}|$  was used. The solution (4.28) is the whistler mode, and the solution (4.29) is the ion-cyclotron mode. For isotropic temperatures ( $a_p = 1$ ), the result is equal to the Hall-MHD dispersion relation. At long wavelengths ( $k_{\parallel} \rightarrow 0$ ) the solutions become non-dispersive and connect to the usual CGL Alfvén modes. For parallel propagation, the Hall term is therefore responsible for splitting the two modes that otherwise would propagate at the same speed, while it does not influence the ion-acoustic mode. Or in another words, using the vocabulary of slow, Alfvén, fast, depending on plasma  $\beta_{\parallel}$  and temperature anisotropy  $a_p$ , the Hall term either splits the parallel propagating Alfvén and fast mode, or the slow and Alfvén mode. At very short spatial scales ( $k_{\parallel} \gg 1$ ), the limits are<sup>9</sup>

$$k_{\parallel} \rightarrow \pm\infty: \quad \omega_{IC}^2 = \Omega_p^2 \left( 1 + \frac{\beta_{\parallel}}{2} (a_p - 1) \right)^2; \quad \omega_w^2 = k_{\parallel}^4 \frac{V_A^4}{\Omega_p^2}. \quad (4.30)$$

<sup>9</sup>Note that the limit for the ion-cyclotron mode has to be calculated with a trick  $a - b = (a^2 - b^2)/(a + b)$  since both  $a, b$  go to  $\infty$ .

For isotropic temperatures  $a_p = 1$  the frequency of the ion-cyclotron mode (that should be actually called the proton-cyclotron mode) converges towards the proton-cyclotron frequency  $\pm\Omega_p$ . For anisotropic temperatures, the frequency of the mode starts to be  $\beta_{\parallel}$  dependent and will converge to frequencies that can be higher or lower than  $\Omega_p$ . The frequency of the whistler mode is not bounded and technically goes to infinity, which is a consequence of neglecting the electron inertia. Importantly, both modes are stable for high wavenumbers. The solutions (4.28), (4.29) are analytically correct, but similarly to the Hall-MHD, the solutions can be rewritten to a more useful form, by realizing that the dispersion equation (4.27) can be decomposed as<sup>10</sup>

$$\left(\omega^2 + \frac{k_{\parallel}^2 V_A^2}{\Omega_p} \omega - k_{\parallel}^2 v_{A\parallel}^2\right) \left(\omega^2 - \frac{k_{\parallel}^2 V_A^2}{\Omega_p} \omega - k_{\parallel}^2 v_{A\parallel}^2\right) = 0. \quad (4.31)$$

The first bracket yields two solutions

$$\omega_{IC,1} = -\frac{k_{\parallel}^2 V_A^2}{2\Omega_p} + |k_{\parallel}| V_A \sqrt{1 + \frac{\beta_{\parallel}}{2}(a_p - 1) + \frac{k_{\parallel}^2 V_A^2}{4\Omega_p^2}}; \quad \text{—Green line;} \quad (4.32)$$

$$\omega_{W,1} = -\frac{k_{\parallel}^2 V_A^2}{2\Omega_p} - |k_{\parallel}| V_A \sqrt{1 + \frac{\beta_{\parallel}}{2}(a_p - 1) + \frac{k_{\parallel}^2 V_A^2}{4\Omega_p^2}}; \quad \text{—dashed Magenta line,} \quad (4.33)$$

and the second bracket yields two solutions

$$\omega_{W,2} = +\frac{k_{\parallel}^2 V_A^2}{2\Omega_p} + |k_{\parallel}| V_A \sqrt{1 + \frac{\beta_{\parallel}}{2}(a_p - 1) + \frac{k_{\parallel}^2 V_A^2}{4\Omega_p^2}}; \quad \text{—Blue line;} \quad (4.34)$$

$$\omega_{IC,2} = +\frac{k_{\parallel}^2 V_A^2}{2\Omega_p} - |k_{\parallel}| V_A \sqrt{1 + \frac{\beta_{\parallel}}{2}(a_p - 1) + \frac{k_{\parallel}^2 V_A^2}{4\Omega_p^2}}; \quad \text{—dashed Cyan line.} \quad (4.35)$$

It is straightforward to check that  $(\omega - \omega_{W,1})(\omega - \omega_{W,2}) = 0$  yields the whistler mode solution (4.28), and  $(\omega - \omega_{IC,1})(\omega - \omega_{IC,2}) = 0$  yields the ion-cyclotron mode solution (4.29). Often, solutions (4.32)–(4.35) are written in an alternative form where the absolute value on  $k_{\parallel}$  is removed (which obviously can be done), and the solutions are just slightly re-arranged. Additionally, for more complicated fluid models, we want to write solution in one line instead of four lines. In this example, we just write a shortcut that two solutions are

$$\omega = \pm \frac{k_{\parallel}^2 V_A^2}{2\Omega_p} + k_{\parallel} V_A \sqrt{1 + \frac{\beta_{\parallel}}{2}(a_p - 1) + \frac{k_{\parallel}^2 V_A^2}{4\Omega_p^2}}, \quad (4.36)$$

with another two solutions obtained by substituting  $\omega$  with  $-\omega$ .

<sup>10</sup>One can double check the result by working with quantities  $B^{\pm} = iB_x \pm B_y$ , yielding  $(\omega^2 \pm (k_{\parallel}^2 V_A^2 / \Omega_p) \omega - k_{\parallel}^2 v_{A\parallel}^2) B^{\mp} = 0$ .

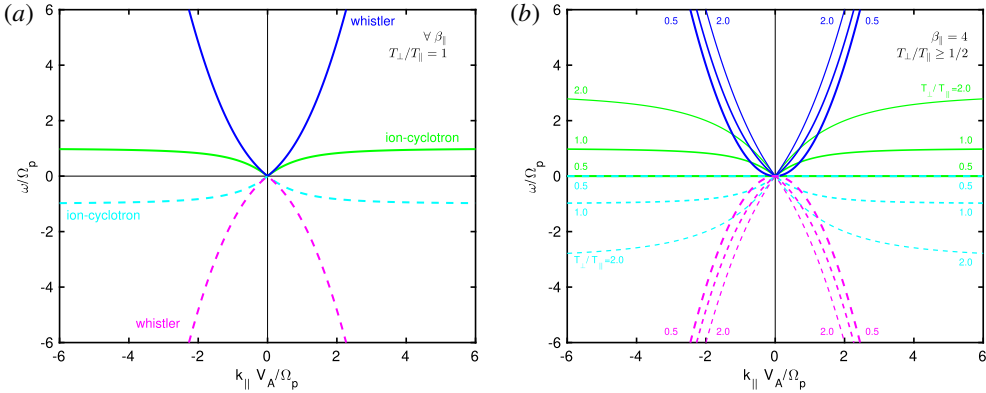


FIGURE 4. Dispersion relations for the Hall-CGL model (with cold electrons), for parallel propagation  $\theta = 0$ . The colour coding of lines is according to (4.32)–(4.35). (a) Isotropic temperatures  $a_p = T_{\perp}/T_{\parallel} = 1.0$ . The dispersion relations do not depend on the value of  $\beta_{\parallel}$ , which is a consequence of neglecting the FLR pressure corrections, etc. The figure is actually equivalent to the Hall-MHD model. (b) the value of  $\beta_{\parallel}$  is fixed with the value  $\beta_{\parallel} = 4.0$  and the temperature anisotropy is varied as  $a_p = 2.0; 1.0; 0.5$ . The thickness of the curves increases with decreasing  $a_p$  as we approach the firehose threshold at  $a_p = 0.5$ . All the curves have purely real frequency  $\omega$ .

#### 4.2.1. Firehose instability

Now we can discuss the firehose instability. If the quantity  $1 + (\beta_{\parallel}/2)(a_p - 1) > 0$ , then all the solutions (4.32)–(4.35) have frequencies that are purely real for the entire range of wavenumbers. For isotropic temperatures  $a_p = 1$ , the dispersion relations do not depend on the value of  $\beta_{\parallel}$ , which is a consequence of neglecting the FLR pressure corrections, as will be discussed later, and all solutions are equivalent to the Hall-MHD dispersion relations. Solutions (4.32)–(4.35) with  $a_p = 1$  are plotted in figure 4(a), where we plotted both positive and negative wavenumbers and positive and negative frequencies to clearly understand how the modes are connected to each other. The magenta and cyan lines are plotted as dashed lines, so that the growth rates in the firehose-unstable regime (plotted in figure 5) are clearly visible. For anisotropic temperatures  $a_p \neq 1$ , the solutions are naturally  $\beta_{\parallel}$  dependent. To show solutions in a firehose unstable regime, one has to choose  $\beta_{\parallel} > 2$  since the instability does not exist for lower values of  $\beta_{\parallel}$ . In figure 4(b), we have chosen a value of  $\beta_{\parallel} = 4$  and varied the temperature anisotropy from  $a_p = 2$  down to a critical value of  $a_p = 0.5$ . In this regime, all the modes have purely real frequencies.

Importantly, if the temperature anisotropy parameter  $a_p$  is further decreased, the frequencies of all the modes will become complex numbers (with a real and imaginary part) in the region where

$$1 + \frac{\beta_{\parallel}}{2}(a_p - 1) + \frac{k_{\parallel}^2 V_A^2}{4\Omega_p^2} < 0, \quad (4.37)$$

which can be viewed as a modified Hall-CGL firehose instability threshold that is now length-scale dependent. At long spatial scales ( $k_{\parallel} \rightarrow 0$ ), the above criterion is naturally equivalent to the CGL firehose criterion. The term proportional to  $k_{\parallel}^2$  in (4.37) is always positive and as  $k_{\parallel}$  increases, after some critical wavenumber the frequency

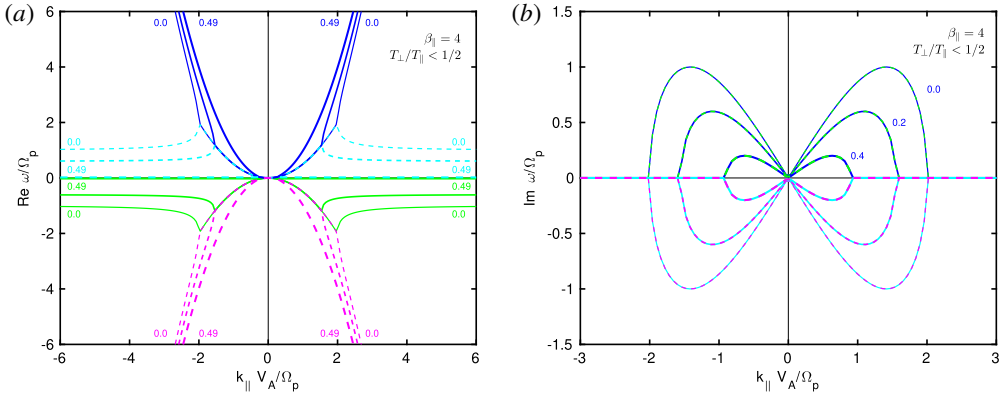


FIGURE 5. Dispersion relations for the Hall-CGL model (with cold electrons), for parallel propagation  $\theta = 0$ , with fixed  $\beta_{||} = 4$ , after crossing the firehose threshold  $a_p = T_{\perp}/T_{||} < 0.5$ . The colour coding of lines is according to (4.32)–(4.35). The thickness of the lines decreases as going away from the threshold  $a_p = 0.5$ . (a) Real frequency and the curves have  $a_p = 0.49; 0.2; 0.0$ . (b) Imaginary frequency and the curves have  $a_p = 0.4; 0.2; 0.0$ . The mode with  $a_p = 0.49$ , shown in (a), has an imaginary part of the frequency which is very close to zero for the entire range of  $k_{||}$ . For this reason this mode is not plotted in (b), where instead a mode with  $a_p = 0.4$  is plotted. A mode that has  $\text{Im } \omega > 0$  is unstable and growing in time. A mode that has  $\text{Im } \omega < 0$  is stable and damped. According to the figure, whistler modes with  $\omega_r > 0$  are unstable, and ion-cyclotron modes with  $\omega_r < 0$  are unstable. It is assumed that  $\sqrt{-1} = +i$ . The solutions for  $\omega_r < 0$  are here non-causal, and (b) should be re-plotted with causal solutions (4.43)–(4.46), so that whistler modes are always unstable, and ion-cyclotron modes stable.

of modes will become purely real. The Hall term is therefore responsible for the stabilization of the firehose instability at sufficiently high wavenumbers (small spatial scales). The solutions in this firehose-unstable regime are plotted in figure 5, where again  $\beta_{||} = 4$  and the temperature anisotropy is varied from  $a_p = 0.49$  to the lowest possible value of  $a_p = 0.0$ . Figure 5(a) shows real frequencies and the (b) shows imaginary frequencies.

It is of interest to determine which mode becomes unstable. According to our Fourier transform (see appendix A), a mode with positive imaginary frequency is unstable and growing, and a mode with negative imaginary frequency is stable and damped. If the firehose threshold (4.37) is satisfied, for some range of  $k_{||}$  the second terms in solutions (4.32)–(4.35) become purely imaginary and the first terms stay purely real. Consider one of the four possible quadrants in figure 5(a), for example the quadrant with  $k_{||} > 0; \omega_r > 0$ . Up to some critical  $k_{||}$  where the Hall-CGL firehose criterion is satisfied, both the blue mode and the cyan mode propagate with the same real frequency  $\omega_r = +k_{||}^2 V_A^2/(2\Omega_p)$ . The differences are in the imaginary frequency, see figure 5(b), where in this quadrant the blue mode is unstable and the cyan mode is stable (and damped). After a critical  $k_{||}$  (for example for  $a_p = 0$  it is  $k_{||} V_A/\Omega_p = 2$ ), both modes become purely real and evolve with different real frequencies, here the blue mode is obviously the whistler mode and the cyan mode is the ion-cyclotron mode. The same situation is in the quadrant  $k_{||} < 0; \omega_r > 0$ . We can therefore conclude that for positive real frequencies,  $\omega_r > 0$ , the whistler mode is unstable and the ion-cyclotron mode is stable and damped.

Now we could say that considering both positive and negative wavenumbers and frequencies is redundant, and that we are finished. However, from a perspective of this guide, we consider useful to show that one might easily come into contradictions, if negative frequencies are not handled with sufficient care. Importantly, according to figure 5, for negative real frequencies  $\omega_r < 0$ , the whistler mode is stable and the ion-cyclotron mode is unstable. This is indeed contradictory. Keeping the  $k_{\parallel} > 0$ , a change from positive to negative  $\omega_r$  represents a change of direction of propagation along  $B_0$ . One can argue, that in a homogeneous system (e.g. excluding propagation in a stratified fluid), a change of direction of propagation should not yield a change of stability. Additionally, kinetic theory and simulations show, that it is the whistler mode that is firehose unstable for both  $\omega_r > 0$  and  $\omega_r < 0$ , and the ion-cyclotron mode is stable. It is important to emphasize that, from a perspective of solutions (4.32)–(4.35), at the range of wavenumbers where the firehose instability is present, the whistler and ion-cyclotron modes propagate (in a given quadrant) with the same real frequency. Moreover, both modes have the same polarization of electric and magnetic field (see § 4.2.3). Both modes are completely degenerate, and not distinguishable. Thus, how the modes are continued for higher wavenumbers once the firehose instability disappears, is not obvious (since  $\sqrt{-1}$  can be possibly both  $+i$  or  $-i$ ). One should consider all the possible solutions and verify, which solutions are physically plausible. Therefore, after the firehose threshold is crossed, the analytic solutions (4.32)–(4.35) are not correctly separated. Obviously, more ‘physical information’ is needed to distinguish between the ion-cyclotron and whistler modes in the firehose-unstable regime, and to determine which solutions are causal and which are non-causal.

#### 4.2.2. Generalization to causal (correct) solutions

The mind boggling puzzle as why the Hall-CGL solutions (4.32)–(4.35) are in contradiction to kinetic theory and to common sense (a discrepancy found by one of the referees of this text), was beautifully solved by Paul Cally. The easiest way to obtain correct solutions is to introduce a very small ‘causal’ dissipation into the system, and remove it later. Alternatively, it should be possible to obtain the same result by using Laplace–Fourier transforms on temporal-spatial scales and considering an initial value problem.

We introduce causality, by adding a tiny amount of dissipation on the right-hand sides of momentum equations (4.19)–(4.20), in the following form

$$\frac{\partial u_x}{\partial t} - v_{A\parallel}^2 \partial_z \frac{B_x}{B_0} = +\epsilon \frac{2V_A^2}{\Omega_p} \partial_z^2 u_x; \quad (4.38)$$

$$\frac{\partial u_y}{\partial t} - v_{A\parallel}^2 \partial_z \frac{B_y}{B_0} = +\epsilon \frac{2V_A^2}{\Omega_p} \partial_z^2 u_y, \quad (4.39)$$

where  $\epsilon$  is a dimensionless real small positive number, for example  $\epsilon = 10^{-8}$ . The dissipation will be later completely removed by the limit  $\lim_{\epsilon \rightarrow 0^+}$ . The  $V_A^2/\Omega_p$  is there so that the parameter  $\epsilon$  is indeed dimensionless (since here we usually normalize with respect to  $V_A$  and  $\Omega_p$ , but other normalizations can be of course considered). The factor of 2 is there for pure convenience, so that the final dispersion relations (4.43)–(4.46) are of the utmost beauty, and that we do not have to later redefine  $\epsilon$ . The momentum equations are accompanied by the induction equations (4.21)–(4.22). The reader is encouraged to verify the calculations by switching to normalized variables  $\tilde{k}_{\parallel} = k_{\parallel} V_A/\Omega_p$ ,  $\tilde{\omega} = \omega/\Omega_p$ ,  $\tilde{v}_{A\parallel}^2 = 1 + (\beta_{\parallel}/2)(a_p - 1)$ , we continue in physical units.

By adding the dissipation, the dispersion relation (4.31) is modified to the following form

$$\left[ \omega^2 + \frac{k_{\parallel}^2 V_A^2}{\Omega_p^2} \omega(1 + 2i\epsilon) - k_{\parallel}^2 v_{A\parallel}^2 + 2i\epsilon \frac{k_{\parallel}^4 V_A^4}{\Omega_p^2} \right] (iB_x - B_y) = 0; \quad (4.40)$$

$$\left[ \omega^2 - \frac{k_{\parallel}^2 V_A^2}{\Omega_p^2} \omega(1 - 2i\epsilon) - k_{\parallel}^2 v_{A\parallel}^2 - 2i\epsilon \frac{k_{\parallel}^4 V_A^4}{\Omega_p^2} \right] (iB_x + B_y) = 0. \quad (4.41)$$

It is useful to define the Hall-CGL firehose threshold

$$\Delta \equiv 1 + \frac{\beta_{\parallel}}{2}(a_p - 1) + \frac{k_{\parallel}^2 V_A^2}{4\Omega_p^2}, \quad (4.42)$$

and similar generalizations can be done for models with FLR corrections. Then, since  $\epsilon$  is small, it is easy to show that the non-causal solutions (4.32)–(4.35) are ‘corrected’ by the tiny dissipation to the following causal form

$$\omega_{IC,1} = -\frac{k_{\parallel}^2 V_A^2}{2\Omega_p^2} + |k_{\parallel}| V_A \lim_{\epsilon \rightarrow 0^+} \sqrt{\Delta - i\epsilon}; \quad \text{—Green line;} \quad (4.43)$$

$$\omega_{W,1} = -\frac{k_{\parallel}^2 V_A^2}{2\Omega_p^2} - |k_{\parallel}| V_A \lim_{\epsilon \rightarrow 0^+} \sqrt{\Delta - i\epsilon}; \quad \text{—dashed Magenta line;} \quad (4.44)$$

$$\omega_{W,2} = +\frac{k_{\parallel}^2 V_A^2}{2\Omega_p^2} + |k_{\parallel}| V_A \lim_{\epsilon \rightarrow 0^+} \sqrt{\Delta + i\epsilon}; \quad \text{—Blue line;} \quad (4.45)$$

$$\omega_{IC,2} = +\frac{k_{\parallel}^2 V_A^2}{2\Omega_p^2} - |k_{\parallel}| V_A \lim_{\epsilon \rightarrow 0^+} \sqrt{\Delta + i\epsilon}; \quad \text{—dashed Cyan line.} \quad (4.46)$$

The reader is encouraged to plot these solutions, for example with  $\epsilon = 10^{-8}$ , and verify that figure 5(b), is indeed corrected, i.e. that the whistler is unstable for both  $\omega_r > 0$  and  $\omega_r < 0$ . We want to further address the limits. For clarity, let us first consider the following simplest case

$$\lim_{\epsilon \rightarrow 0^+} \sqrt{-1 \pm i\epsilon} = \frac{\epsilon}{2} \pm i, \quad (4.47)$$

where one can further suppress the small real  $\epsilon/2$  on the right-hand side. The limit (4.47) provides an answer, when  $\sqrt{-1}$  is equal to  $+i$  or  $-i$ . The result (4.47) can be easily generalized, and limits in expressions (4.43)–(4.46) are calculated according to

$$\lim_{\epsilon \rightarrow 0^+} \sqrt{\Delta \pm i\epsilon} = \begin{cases} \sqrt{\Delta}, & \Delta > 0; \\ 0, & \Delta = 0; \\ \pm i\sqrt{-\Delta}, & \Delta < 0. \end{cases} \quad (4.48)$$

Obviously, the whistler solutions (4.44), (4.45) are unstable, and the ion-cyclotron solutions (4.43), (4.46) are damped. Only now we can conclude with confidence, that the parallel firehose instability as an instability of the whistler mode.

We note that if instead of the momentum equations, we added the dissipation (diffusivity) to the induction equations, the solutions (4.43)–(4.46) are recovered with  $\epsilon \rightarrow -\epsilon$ . In other words, by adding dissipation to the induction equations, the whistler mode would be always damped and the ion-cyclotron mode unstable. The result makes sense with respect to kinetic theory, since the ion-cyclotron resonances enter through the velocity moment (by integration over the perturbation of the distribution function  $f^{(1)}$ ), or more precisely through the current  $\mathbf{j} = \sum_r q_r n_r \mathbf{u}_r$ , see Part 2.

### 4.2.3. Polarization

The polarization of the magnetic field can be easily checked, actually without plugging any dispersion relations into the matrix (4.26), since the matrix can be represented simply as

$$\begin{pmatrix} a & +ib \\ -ib & a \end{pmatrix} \begin{pmatrix} B_x \\ B_y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad (4.49)$$

where  $a = \omega^2 - v_{A\parallel}^2 k_{\parallel}^2$  and  $b = V_A^2 k_{\parallel}^2 \omega / \Omega_p$ . It is useful to define the polarization as the angle between the complex  $B_x$  and  $B_y$  components, which is calculated as  $\text{Arg}(B_y/B_x)$ . If the result is smaller than zero, the wave is left polarized. If the result is larger than zero, the wave is right polarized. The determinant of the matrix (4.49) is  $a^2 - b^2 = 0$ , so  $a = \pm b$ . The case  $a = -b$  (which is directly equal to the first bracket in (4.31) with solutions (4.32), (4.33)) yields that  $B_y = -iB_x$ , so that

$$\frac{B_y}{B_x} = -i = e^{-i(\pi/2)}; \quad \Rightarrow \quad \text{Arg} \frac{B_y}{B_x} = -\frac{\pi}{2} < 0 \quad (\text{left polarized}). \quad (4.50)$$

The second case  $a = +b$  (which is equal to the second bracket in (4.31) with solutions (4.34), (4.35)) yields that  $B_y = +iB_x$ , so that

$$\frac{B_y}{B_x} = +i = e^{+i(\pi/2)}; \quad \Rightarrow \quad \text{Arg} \frac{B_y}{B_x} = +\frac{\pi}{2} > 0 \quad (\text{right polarized}). \quad (4.51)$$

Similarly, one can calculate the polarization of electric field  $\mathbf{E}$ , and (when not specified that the  $\mathbf{B}$  field is considered), the words left/right polarized are usually meant for the  $\mathbf{E}$  field. The ratio  $E_y/E_x$  can be easily calculated, for example from equations (4.74), (4.75) (written there in the normalized form), which when multiplied by  $\omega$  and use of momentum equations, yields for the special case of parallel propagation considered here

$$\frac{E_y}{E_x} = \frac{-\tilde{v}_{A\parallel}^2 + i\tilde{\omega} \frac{B_y}{B_x}}{\tilde{v}_{A\parallel}^2 \frac{B_y}{B_x} + i\tilde{\omega}}. \quad (4.52)$$

When plotting many solutions, it is useful to additionally divide the angle by  $\pi$ , and define the  $\mathbf{B}$  and  $\mathbf{E}$  field polarizations as

$$\mathcal{P}_B = \frac{1}{\pi} \text{Arg} \frac{B_y}{B_x}; \quad \mathcal{P}_E = \frac{1}{\pi} \text{Arg} \frac{E_y}{E_x}. \quad (4.53)$$

Electric field polarization defined this way was used for example by Hunana *et al.* (2013) to investigate properties of (highly oblique) kinetic Alfvén waves in various fluid models and the kinetic description; and also by Camporeale & Burgess (2017), who compared various linear modes in hybrid, gyrokinetic and fully kinetic descriptions.<sup>11</sup> When  $\mathcal{P}_E < 0$  the wave is called left polarized and when  $\mathcal{P}_E > 0$ , the wave is called right polarized. For the first group (4.50) this yields  $E_y/E_x = -i$  and

<sup>11</sup>Similar definition was used by Sahraoui, Belmont & Goldstein (2012), who however plot a somewhat confusing  $\text{Arg}(E_y)/\pi$ , since in the WHAMP code the electric field is normalized so that  $\text{Arg}(E_x) = 0$ .



$\mathcal{P}_E = -1/2 < 0$  (left polarized wave), and for the second group (4.51)  $E_y/E_x = +i$  and  $\mathcal{P}_E = +1/2 > 0$  (right polarized wave).

Finally, it is of interest to note another possible definition of polarization, that is defined with respect to positive real frequency, i.e. the electric field polarization (4.53) is redefined to

$$\mathcal{P}_E = \frac{\text{sign}(\omega_r)}{\pi} \text{Arg} \frac{E_y}{E_x}. \quad (4.54)$$

A similar definition of polarization with  $\text{sign}(\omega_r)$  is used for example in the book by Gary (1993), page 91. With such a definition, the whistler mode is always right polarized, both in the firehose-stable and firehose-unstable regimes, since for the whistler mode ((4.34), blue line) the real frequency remains  $\omega_r > 0$ , and for the whistler mode ((4.33), magenta line)  $\omega_r < 0$ . The situation is more confusing for the ion-cyclotron mode, where the firehose-stable and firehose-unstable regimes have to be considered separately. (i) In the firehose-stable regime, both ion-cyclotron modes are left polarized, since for the mode ((4.32), green line)  $\omega_r > 0$ , and for the mode ((4.35), cyan line)  $\omega_r < 0$ . (ii) In contrast, in the firehose-unstable regime, both ion-cyclotron modes are right polarized, since for the mode ((4.32), green line)  $\omega_r < 0$ , and for the mode ((4.35), cyan line)  $\omega_r > 0$ .

#### 4.2.4. Simplest ion-cyclotron resonances

When talking about ‘resonances’, one often expects to see expressions with a denominator that becomes zero. Generally speaking, the simplest ‘singular’ expressions representing resonances just arise from a technique used in deriving the dispersion relations. If one eliminates the electric field  $\mathbf{E}$  from the beginning and expresses the fluid model through a matrix multiplied by a vector  $(\rho, u_x, u_y, u_z, b_x, b_y, b_z, p_{\parallel}, p_{\perp})$ , leads to dispersion relations that are not singular explicitly. In contrast, expressing the model through a matrix multiplied by a vector  $(E_x, E_y, E_z)$ , yields equivalent dispersion relations where the resonances are shown explicitly. The dispersion equation (4.31) can be easily rewritten as

$$\frac{k_{\parallel}^2}{\omega^2} = -\frac{\Omega_p/V_A^2}{\omega - \Omega_p \left(1 + \frac{\beta_{\parallel}}{2}(a_p - 1)\right)}; \quad \frac{k_{\parallel}^2}{\omega^2} = +\frac{\Omega_p/V_A^2}{\omega + \Omega_p \left(1 + \frac{\beta_{\parallel}}{2}(a_p - 1)\right)}, \quad (4.55)$$

which, in the cold plasma limit ( $\beta_{\parallel} = 0$ ) or in this case also for isotropic temperatures ( $a_p = 1$ ), simplifies to

$$\frac{k_{\parallel}^2}{\omega^2} = -\frac{\Omega_p/V_A^2}{\omega - \Omega_p}; \quad \frac{k_{\parallel}^2}{\omega^2} = +\frac{\Omega_p/V_A^2}{\omega + \Omega_p}, \quad (4.56)$$

making it more evident that we have the simplest ion-cyclotron resonances here too. Indeed, for  $k \rightarrow \infty$  the frequency  $\omega \rightarrow \Omega_p$  (first case) and  $\omega \rightarrow -\Omega_p$  (second case). For a plasma with finite  $\beta_{\parallel}$  and non-isotropic temperatures, the  $k \rightarrow \infty$  limit yields more general ion-cyclotron resonances (4.30).

To better understand what effects are present and what effects are absent in the Hall-CGL model, we plot fully kinetic solutions obtained with the WHAMP code in figure 6. The proton temperature is prescribed to be isotropic  $T_{\perp}/T_{\parallel} = 1$  and the electrons are prescribed to be cold with  $T_e/T_p = 0$  (in the WHAMP code, the value is chosen to be  $10^{-8}$ ). In the WHAMP code it is necessary to prescribe the ratio of



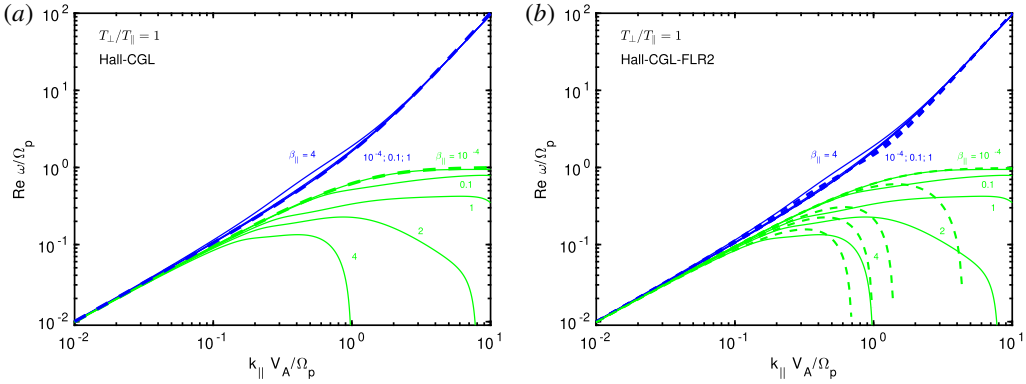


FIGURE 6. (a) Comparison of solutions of full kinetic theory obtained with the WHAMP code (solid lines) and the simple dispersion of the Hall-CGL model (dashed lines) for the whistler mode (blue) and the ion-cyclotron mode (green). The proton temperatures are isotropic  $T_{\perp}/T_{\parallel} = 1$ , electrons are cold and the angle of propagation  $\theta = 0$ . Real frequency is plotted. The  $\beta_{\parallel}$  is varied as  $\beta_{\parallel} = 10^{-4}; 0.1; 1; 2; 4$ . The kinetic solutions show that while the frequency of whistler mode is not very  $\beta_{\parallel}$  dependent (only the case  $\beta_{\parallel} = 4$  is slightly different), the frequency of the ion-cyclotron mode is strongly  $\beta_{\parallel}$  dependent since the mode experiences strong damping. The Hall-CGL model has only the simplest ion-cyclotron resonance, and the Hall-CGL ion-cyclotron mode matches the kinetic ion-cyclotron mode only for the lowest  $\beta_{\parallel} = 10^{-4}$ . (b) Solutions of the Hall-CGL-FLR2 model are plotted. Note that the real frequency of the ion-cyclotron mode is crudely reproduced.

the parallel proton thermal speed to the speed of light, and we choose  $v_{th\parallel}/c = 10^{-4}$ . The solid lines are kinetic solutions with different  $\beta_{\parallel} = 10^{-4}; 0.1; 1; 2; 4$ . The blue curve is the whistler mode and the green curve is the ion-cyclotron mode. Solutions of the Hall-CGL model are shown in figure 6(a) and, for isotropic temperatures, the solutions are completely  $\beta_{\parallel}$  independent and are represented with two dashed curves. It is shown that the kinetic solutions for the whistler mode are almost  $\beta_{\parallel}$  independent and only the mode with  $\beta_{\parallel} = 4$  deviates from the other whistler solutions, and only between  $kd_i = 0.1 - 2.0$ . For comparison, in figure 6(b) we plot solutions of the Hall-CGL-FLR2 model (the model is discussed later in the text).

#### 4.3. Hall-CGL dispersion relation for arbitrary propagation angle

To make everything as clear as possible, let us write down one more time the normalized equations we are solving, including the linearized and Fourier transformed equations in the  $x$ - $z$  plane. The nonlinear Hall-CGL equations with cold electrons are (dropping the tilde)

$$\frac{\partial \rho_p}{\partial t} + \nabla \cdot (\rho_p \mathbf{u}_p) = 0; \quad (4.57)$$

$$\frac{\partial \mathbf{u}_p}{\partial t} + \mathbf{u}_p \cdot \nabla \mathbf{u}_p + \frac{\beta_{\parallel}}{2} \frac{1}{\rho_p} \nabla \cdot \mathbf{p}_p - \frac{1}{\rho_p} (\nabla \times \mathbf{B}) \times \mathbf{B} = 0; \quad (4.58)$$

$$\frac{\partial \mathbf{B}}{\partial t} - \nabla \times (\mathbf{u}_p \times \mathbf{B}) + \nabla \times \left[ \frac{1}{\rho_p} (\nabla \times \mathbf{B}) \times \mathbf{B} \right] = 0; \quad (4.59)$$

$$\frac{\partial p_{\parallel p}}{\partial t} + \nabla \cdot (p_{\parallel p} \mathbf{u}_p) + 2p_{\parallel p} \hat{\mathbf{b}} \cdot \nabla \mathbf{u}_p \cdot \hat{\mathbf{b}} = 0; \quad (4.60)$$

$$\frac{\partial p_{\perp p}}{\partial t} + \nabla \cdot (p_{\perp p} \mathbf{u}_p) + p_{\perp p} \nabla \cdot \mathbf{u}_p - p_{\perp p} \hat{\mathbf{b}} \cdot \nabla \mathbf{u}_p \cdot \hat{\mathbf{b}} = 0. \quad (4.61)$$

Linearizing in normalized units is even easier, because (temporarily introducing back the tilde for clarity)  $\tilde{p}_{\parallel p}^{(0)} = 1$ ;  $\tilde{p}_{\perp p}^{(0)} = a_p$ ;  $\tilde{\rho}_0 = 1$ ;  $\tilde{B}_0 = 1$ . Linearizing these equations, writing them in the  $x$ - $z$  plane and Fourier transforming yields (dropping also the proton index ' $p$ ')

$$-\omega \rho + k_{\perp} u_x + k_{\parallel} u_z = 0; \quad (4.62)$$

$$-\omega u_x + \frac{\beta_{\parallel}}{2} k_{\perp} p_{\perp} - v_{A\parallel}^2 k_{\parallel} B_x + k_{\perp} B_z = 0; \quad (4.63)$$

$$-\omega u_y - v_{A\parallel}^2 k_{\parallel} B_y = 0; \quad (4.64)$$

$$-\omega u_z + \frac{\beta_{\parallel}}{2} k_{\parallel} p_{\parallel} + \frac{\beta_{\parallel}}{2} (1 - a_p) k_{\perp} B_x = 0; \quad (4.65)$$

$$-\omega B_x - k_{\parallel} u_x - i k_{\parallel}^2 B_y = 0; \quad (4.66)$$

$$-\omega B_y - k_{\parallel} u_y - i k_{\parallel} k_{\perp} B_z + i k_{\parallel}^2 B_x = 0; \quad (4.67)$$

$$-\omega B_z + k_{\perp} u_x + i k_{\parallel} k_{\perp} B_y = 0; \quad (4.68)$$

$$-\omega p_{\parallel} + k_{\perp} u_x + 3 k_{\parallel} u_z = 0; \quad (4.69)$$

$$-\omega p_{\perp} + 2 a_p k_{\perp} u_x + a_p k_{\parallel} u_z = 0, \quad (4.70)$$

where the normalized  $v_{A\parallel}^2 = 1 + (\beta_{\parallel}/2)(a_p - 1)$ . Alternatively, it is sometimes illuminating to work with the general induction equation where the electric field is used

$$-\omega B_x - k_{\parallel} E_y = 0; \quad (4.71)$$

$$-\omega B_y - k_{\perp} E_z + k_{\parallel} E_x = 0; \quad (4.72)$$

$$-\omega B_z + k_{\perp} E_y = 0, \quad (4.73)$$

and in this case with cold electrons the electric field components are

$$E_x = -u_y - i k_{\perp} B_z + i k_{\parallel} B_x; \quad (4.74)$$

$$E_y = u_x + i k_{\parallel} B_y; \quad (4.75)$$

$$E_z = 0. \quad (4.76)$$

The induction equation (4.71)–(4.73) with the above electric field components is of course equivalent to the induction equation (4.66)–(4.68). All equations are normalized, the tilde are dropped, and the electric field  $\tilde{E} = E/E_0$ , where  $E_0 = V_A B_0/c$ . The Hall-CGL dispersion relation reads

$$\omega^6 - A_4 \omega^4 + A_2 \omega^2 - A_0 = 0; \quad (4.77)$$

$$A_4 = k_{\parallel}^2 \left( \frac{3}{2} \beta_{\parallel} + 2 v_{A\parallel}^2 \right) + k_{\perp}^2 (1 + a_p \beta_{\parallel}) + k^2 k_{\parallel}^2; \quad (4.78)$$

$$A_2 = k_{\parallel}^2 \left[ k_{\parallel}^2 v_{A\parallel}^2 (v_{A\parallel}^2 + 3 \beta_{\parallel}) + k_{\perp}^2 (v_{A\parallel}^2 (1 + a_p \beta_{\parallel}) + \frac{3}{2} \beta_{\parallel} (1 + a_p \beta_{\parallel} - \frac{1}{6} a_p^2 \beta_{\parallel})) \right] + k^2 k_{\parallel}^2 \beta_{\parallel} \left( \frac{3}{2} k_{\parallel}^2 + a_p k_{\perp}^2 \right); \quad (4.79)$$

$$A_0 = \frac{3}{2} \beta_{\parallel} k_{\parallel}^4 v_{A\parallel}^2 \left[ k_{\parallel}^2 v_{A\parallel}^2 + k_{\perp}^2 (1 + a_p \beta_{\parallel} - \frac{1}{6} a_p^2 \beta_{\parallel}) \right] + k^2 k_{\parallel}^4 k_{\perp}^2 \frac{5}{4} a_p \beta_{\parallel}^2, \quad (4.80)$$

where the wavenumber  $k^2 = k_{\parallel}^2 + k_{\perp}^2$ . Perhaps the nicest form is to keep the CGL contributions on the left-hand side (so the terms can be factorized to form the CGL dispersion relation) and put the Hall term contributions on the right-hand side, which yields the Hall-CGL dispersion relation

$$(\omega^2 - k_{\parallel}^2 v_{A\parallel}^2)(\omega^4 - A_2 \omega^2 + A_0) = k^2 k_{\parallel}^2 \left[ \omega^4 - \beta_{\parallel} \left( \frac{3}{2} k_{\parallel}^2 + a_p k_{\perp}^2 \right) \omega^2 + k_{\parallel}^2 k_{\perp}^2 \frac{5}{4} a_p \beta_{\parallel}^2 \right]; \quad (4.81)$$

$$A_2 = k_{\perp}^2 (1 + a_p \beta_{\parallel}) + k_{\parallel}^2 (v_{A\parallel}^2 + \frac{3}{2} \beta_{\parallel}); \quad (4.82)$$

$$A_0 = \frac{3}{2} k_{\parallel}^2 \beta_{\parallel} \left[ k_{\parallel}^2 v_{A\parallel}^2 + k_{\perp}^2 \left( 1 + a_p \beta_{\parallel} - \frac{1}{6} a_p^2 \beta_{\parallel} \right) \right]. \quad (4.83)$$

Solutions for the parallel propagation ( $k_{\perp} = 0$ ) are of course the whistler and the ion-cyclotron mode

$$\omega = \pm \frac{k_{\parallel}^2}{2} + k_{\parallel} \sqrt{v_{A\parallel}^2 + \frac{k_{\parallel}^2}{4}}, \quad (4.84)$$

with other two solutions obtained by exchanging  $\omega$  with  $-\omega$ . By bringing back the tilde to the normalized solution (4.84), the result (4.36) in physical units is easily recovered. The other parallel solution is the ion-acoustic mode  $\omega = \pm \sqrt{3\beta_{\parallel}/2} k_{\parallel}$ . For perpendicular propagation ( $k_{\parallel} = 0$ ), the fast mode solution is  $\omega = \pm k_{\perp} \sqrt{1 + a_p \beta_{\parallel}}$ , which is equivalent to CGL since the Hall term disappears.

We investigated solutions of the Hall-CGL model in much greater detail in our paper ‘On the Parallel and Oblique Firehose Instability in Fluid Models’ (Hunana & Zank 2017), a paper which was essentially pulled out of this guide. We do not want to repeat the discussion with many associated figures here, and a reader who is further interested in the Hall-CGL firehose instability can find much more information in that paper.

## 5. FLR corrections to the pressure tensor

The finite Larmor radius corrections to the pressure tensor can be derived at several levels of approximation. The pressure tensor equation (2.50) can be rewritten as

$$\hat{\mathbf{b}} \times \boldsymbol{\Pi} + (\hat{\mathbf{b}} \times \boldsymbol{\Pi})^{\text{T}} = -\frac{1}{\Omega} \frac{B_0}{|\mathbf{B}|} \left[ \frac{\partial \mathbf{p}}{\partial t} + \nabla \cdot (\mathbf{u} \mathbf{p} + \mathbf{q}) + \mathbf{p} \cdot \nabla \mathbf{u} + (\mathbf{p} \cdot \nabla \mathbf{u})^{\text{T}} \right]. \quad (5.1)$$

This equation is exact and since  $\mathbf{p} = \mathbf{p}^{\text{g}} + \boldsymbol{\Pi}$ , the FLR tensor  $\boldsymbol{\Pi}$  is described implicitly.

### 5.1. Fully nonlinear FLR corrections

If preservation of all nonlinearities is desired, this implicit equation can be further rewritten by the following procedure that can be found for example in Hsu *et al.* (1986), Passot & Sulem (2004) and Ramos (2005) as a brief note. By applying  $\hat{\mathbf{b}} \times$  to the left-hand side, the first term calculates as

$$\begin{aligned} [\hat{\mathbf{b}} \times (\hat{\mathbf{b}} \times \boldsymbol{\Pi})]_{ij} &= \epsilon_{ikl} \hat{b}_k (\hat{\mathbf{b}} \times \boldsymbol{\Pi})_{lj} = \epsilon_{ikl} \hat{b}_k \epsilon_{lrs} \hat{b}_r \Pi_{sj} = (\delta_{ir} \delta_{ks} - \delta_{is} \delta_{kr}) \hat{b}_k \hat{b}_r \Pi_{sj} \\ &= \hat{b}_k \hat{b}_i \Pi_{kj} - \underbrace{\hat{b}_k \hat{b}_k}_{=1} \Pi_{ij} = \hat{b}_i \hat{b}_k \Pi_{kj} - \Pi_{ij}, \end{aligned} \quad (5.2)$$

where we have used that  $\epsilon_{ikl} \epsilon_{lrs} = \delta_{ir} \delta_{ks} - \delta_{is} \delta_{kr}$ . The second term on the left-hand side of (5.1) calculates quite differently and we will need an identity

$$\epsilon_{ikl} \epsilon_{jrs} = \delta_{ij} (\delta_{kr} \delta_{ls} - \delta_{ks} \delta_{lr}) - \delta_{ir} (\delta_{kj} \delta_{ls} - \delta_{ks} \delta_{lj}) + \delta_{is} (\delta_{kj} \delta_{lr} - \delta_{kr} \delta_{lj}). \quad (5.3)$$

Let us also remind ourselves that in the index notation  $\Pi_{ss} = \text{Tr } \Pi = 0$  and  $\hat{b}_k \hat{b}_l \Pi_{kl} = \Pi : \hat{b}\hat{b} = 0$ . The second term calculates as

$$\begin{aligned} [\hat{b} \times (\hat{b} \times \Pi)]_{ij} &= \epsilon_{ikl} \hat{b}_k (\hat{b} \times \Pi)_{ij}^T = \epsilon_{ikl} \hat{b}_k (\hat{b} \times \Pi)_{jl} = \epsilon_{ikl} \hat{b}_k \epsilon_{jrs} \hat{b}_r \Pi_{sl} = \epsilon_{ikl} \epsilon_{jrs} \hat{b}_k \hat{b}_r \Pi_{sl} \\ &= \delta_{ij} (\hat{b}_k \hat{b}_k \underbrace{\Pi_{ss}}_{=0} - \underbrace{\hat{b}_k \hat{b}_l \Pi_{kl}}_{=0}) - (\hat{b}_j \hat{b}_i \underbrace{\Pi_{ss}}_{=0} - \hat{b}_k \hat{b}_i \Pi_{kj}) + (\hat{b}_j \hat{b}_l \Pi_{il} - \hat{b}_k \hat{b}_k \Pi_{ij}) \\ &= \hat{b}_i \hat{b}_k \Pi_{kj} + \hat{b}_j \Pi_{il} \hat{b}_l - \Pi_{ij}. \end{aligned} \quad (5.4)$$

The entire left-hand side therefore reads

$$[\hat{b} \times (\hat{b} \times \Pi) + \hat{b} \times (\hat{b} \times \Pi)^T]_{ij} = -2\Pi_{ij} + 2\hat{b}_i \hat{b}_l \Pi_{lj} + \hat{b}_j \Pi_{il} \hat{b}_l. \quad (5.5)$$

Performing a usual matrix (single) dot product of this result with matrix  $\hat{b}\hat{b}$  yields

$$\begin{aligned} [[\hat{b} \times (\hat{b} \times \Pi) + \hat{b} \times (\hat{b} \times \Pi)^T] \cdot (\hat{b}\hat{b})]_{ij} &= [\hat{b} \times (\hat{b} \times \Pi) + \hat{b} \times (\hat{b} \times \Pi)^T]_{ik} (\hat{b}\hat{b})_{kj} \\ &= [-2\Pi_{ik} + 2\hat{b}_i \hat{b}_l \Pi_{lk} + \hat{b}_k \Pi_{il} \hat{b}_l] \hat{b}_k \hat{b}_j = -2\Pi_{ik} \hat{b}_k \hat{b}_j + 2\hat{b}_i \hat{b}_j \underbrace{\Pi_{lk} \hat{b}_l \hat{b}_k}_{=0} + \Pi_{il} \hat{b}_l \hat{b}_j \underbrace{\hat{b}_k \hat{b}_k}_{=1} \\ &= -2\Pi_{ik} \hat{b}_k \hat{b}_j + \Pi_{il} \hat{b}_l \hat{b}_j = -2\Pi_{ik} \hat{b}_k \hat{b}_j + \Pi_{ik} \hat{b}_k \hat{b}_j = -\Pi_{ik} \hat{b}_k \hat{b}_j. \end{aligned} \quad (5.6)$$

This further yields that

$$\begin{aligned} [[\hat{b} \times (\hat{b} \times \Pi) + \hat{b} \times (\hat{b} \times \Pi)^T] \cdot (I + 3\hat{b}\hat{b})]_{ij} &= -2\Pi_{ij} + 2\hat{b}_i \hat{b}_k \Pi_{kj} + \hat{b}_j \Pi_{ik} \hat{b}_k - 3\Pi_{ik} \hat{b}_k \hat{b}_j \\ &= -2\Pi_{ij} + 2\hat{b}_i \hat{b}_k \Pi_{kj} - 2\Pi_{ik} \hat{b}_k \hat{b}_j, \end{aligned} \quad (5.7)$$

and adding this result together with its transpose implies

$$\begin{aligned} [[\hat{b} \times (\hat{b} \times \Pi) + \hat{b} \times (\hat{b} \times \Pi)^T] \cdot (I + 3\hat{b}\hat{b})]_{ij}^S \\ = -2\Pi_{ij} + 2\hat{b}_i \hat{b}_k \Pi_{kj} - 2\Pi_{ik} \hat{b}_k \hat{b}_j - 2\Pi_{ji} + 2\hat{b}_j \hat{b}_k \Pi_{ki} - 2\Pi_{jk} \hat{b}_k \hat{b}_i = -4\Pi_{ij}, \end{aligned} \quad (5.8)$$

where we have used that the  $\Pi_{ij} = \Pi_{ji}$ . One therefore derives that the FLR tensor can be extracted from the left-hand side of (5.1) by performing an ‘inversion’ procedure

$$\Pi = -\frac{1}{4} [[\hat{b} \times (\hat{b} \times \Pi) + \hat{b} \times (\hat{b} \times \Pi)^T] \cdot (I + 3\hat{b}\hat{b})]^S. \quad (5.9)$$

We can introduce matrix  $\kappa$  that will represent the right-hand side of (5.1) as

$$\hat{b} \times \Pi + (\hat{b} \times \Pi)^T = -\kappa; \quad (5.10)$$

$$\kappa = \frac{1}{\Omega} \frac{B_0}{|B|} \left[ \frac{\partial p}{\partial t} + \nabla \cdot (up + q) + p \cdot \nabla u + (p \cdot \nabla u)^T \right]. \quad (5.11)$$

By performing operations (5.9) in this equation yields the FLR tensor in the following form

$$\Pi = \frac{1}{4} [(\hat{b} \times \kappa) \cdot (I + 3\hat{b}\hat{b})]^S. \quad (5.12)$$

The expression (5.12) is completely general since no simplifications were introduced, the expression is exact. However, the equation is still implicit, since on the right-hand side the tensor  $\kappa$  contains the full pressure tensor  $p = p^g + \Pi$ . Nevertheless, the

equation is extremely useful, once the expansion of the right-hand side is performed, as we will do below. Note that the brackets are not unique, since

$$(\hat{\mathbf{b}} \times \boldsymbol{\kappa}) \cdot (\mathbf{I} + 3\hat{\mathbf{b}}\hat{\mathbf{b}}) = \hat{\mathbf{b}} \times [\boldsymbol{\kappa} \cdot (\mathbf{I} + 3\hat{\mathbf{b}}\hat{\mathbf{b}})] = \hat{\mathbf{b}} \times \boldsymbol{\kappa} \cdot (\mathbf{I} + 3\hat{\mathbf{b}}\hat{\mathbf{b}}), \quad (5.13)$$

where the last choice leaves the brackets unspecified (since both options are allowed), similarly to  $\hat{\mathbf{b}} \cdot \nabla \mathbf{u}$ . Alternatively, one can get rid of the operator  $\cdot \mathbf{I}$  (which acting on a matrix yields the same matrix), and split the result into two parts

$$\boldsymbol{\Pi} = \frac{1}{4}(\hat{\mathbf{b}} \times \boldsymbol{\kappa})^S + \frac{3}{4}(\hat{\mathbf{b}} \times \boldsymbol{\kappa} \cdot \hat{\mathbf{b}}\hat{\mathbf{b}})^S. \quad (5.14)$$

Another possible form can be obtained by subtracting the evolution equations for the gyrotropic pressure components  $\partial p_{\parallel}/\partial t$  and  $\partial p_{\perp}/\partial t$ . The scalar pressure equations were derived by applying  $:\hat{\mathbf{b}}\hat{\mathbf{b}}$  and  $:(\mathbf{I} - \hat{\mathbf{b}}\hat{\mathbf{b}})/2$  to the pressure tensor equation. The subtraction can be formally represented by introducing ‘overbar’ projection operator  $\bar{\mathbf{a}}$  that projects any  $(3 \times 3)$  matrix  $\mathbf{a}$  to

$$\bar{\mathbf{a}} = \mathbf{a} - [\mathbf{a} : \hat{\mathbf{b}}\hat{\mathbf{b}}]\hat{\mathbf{b}}\hat{\mathbf{b}} - [\mathbf{a} : (\mathbf{I} - \hat{\mathbf{b}}\hat{\mathbf{b}})/2](\mathbf{I} - \hat{\mathbf{b}}\hat{\mathbf{b}}). \quad (5.15)$$

The definition is of course motivated by the pressure decomposition (2.46). By applying this operator to (5.10), the left-hand side remains the same, yielding  $(\hat{\mathbf{b}} \times \boldsymbol{\Pi})^S = -\bar{\boldsymbol{\kappa}}$ . The solution (5.12) for the FLR tensor therefore rewrites

$$\boldsymbol{\Pi} = \frac{1}{4}[(\hat{\mathbf{b}} \times \bar{\boldsymbol{\kappa}}) \cdot (\mathbf{I} + 3\hat{\mathbf{b}}\hat{\mathbf{b}})]^S. \quad (5.16)$$

The subtraction of scalar pressure equations is motivated by the observation that when working in the linear approximation directly with (5.10), i.e. without performing the inversion procedure, the scalar pressure equations have to be subtracted at the end. However, by performing the inversion procedure, the scalar pressure equations are subtracted ‘automatically’ during calculations, and it is actually easier to work directly with (5.12) instead of (5.16).

Therefore, let us work with (5.12). The leading-order nonlinear FLR corrections are obtained by neglecting the non-gyrotropic contributions  $\boldsymbol{\Pi}$  and  $\mathbf{q}^g$  on the right-hand side of (5.12), i.e. by prescribing

$$\boldsymbol{\Pi} = \frac{1}{4}[(\hat{\mathbf{b}} \times \boldsymbol{\kappa}^{(1)}) \cdot (\mathbf{I} + 3\hat{\mathbf{b}}\hat{\mathbf{b}})]^S; \quad (5.17)$$

$$\boldsymbol{\kappa}^{(1)} = \frac{1}{\Omega} \frac{B_0}{|\mathbf{B}|} \left[ \frac{\partial \mathbf{p}^g}{\partial t} + \nabla \cdot (\mathbf{u} \mathbf{p}^g + \mathbf{q}^g) + \mathbf{p}^g \cdot \nabla \mathbf{u} + (\mathbf{p}^g \cdot \nabla \mathbf{u})^T \right]. \quad (5.18)$$

Therefore, the only non-gyrotropic quantities that were retained on the right-hand side of (5.12), are the perpendicular components of the velocity  $\mathbf{u}$ . The slightly unappealing factors  $B_0/|\mathbf{B}|$  in all the expressions can be removed, by redefining the cyclotron frequency with  $|\mathbf{B}|$  instead of  $B_0$ , so that  $\Omega_{|\mathbf{B}|} = e|\mathbf{B}|/(mc)$ . Let us calculate (5.17) step by step. For the brevity of the calculations, it is useful to introduce convective derivative  $d/dt = \partial/\partial t + \mathbf{u} \cdot \nabla$  and work with

$$\boldsymbol{\kappa}^{(1)} = \frac{1}{\Omega_{|\mathbf{B}|}} \left[ \frac{d\mathbf{p}^g}{dt} + \mathbf{p}^g \nabla \cdot \mathbf{u} + \nabla \cdot \mathbf{q}^g + \mathbf{p}^g \cdot \nabla \mathbf{u} + (\mathbf{p}^g \cdot \nabla \mathbf{u})^T \right]. \quad (5.19)$$

We want to present complete derivation with the least amount of algebra involved, and we first want to obtain the form of Schekochihin *et al.* (2010). Let us first write each term in  $\kappa^{(1)}$  in the index notation. By using

$$p_{ij}^g = (p_{\parallel} - p_{\perp})\hat{b}_i\hat{b}_j + p_{\perp}\delta_{ij}; \quad (5.20)$$

$$q_{ijk}^g = (q_{\parallel} - 3q_{\perp})\hat{b}_i\hat{b}_j\hat{b}_k + q_{\perp}(\delta_{ij}\hat{b}_k + \delta_{jk}\hat{b}_i + \delta_{ik}\hat{b}_j), \quad (5.21)$$

in the following form

$$\left(\frac{d\mathbf{p}^g}{dt}\right)_{ij} = \hat{b}_i\hat{b}_j\frac{d}{dt}(p_{\parallel} - p_{\perp}) + (p_{\parallel} - p_{\perp})\frac{d}{dt}(\hat{b}_i\hat{b}_j) + \delta_{ij}\frac{dp_{\perp}}{dt}; \quad (5.22)$$

$$(\mathbf{p}^g \nabla \cdot \mathbf{u})_{ij} = ((p_{\parallel} - p_{\perp})\hat{b}_i\hat{b}_j + \delta_{ij}p_{\perp})\nabla \cdot \mathbf{u}; \quad (5.23)$$

$$(\mathbf{p}^g \cdot \nabla \mathbf{u})_{ij} = (p_{\parallel} - p_{\perp})\hat{b}_i\hat{\mathbf{b}} \cdot \nabla u_j + p_{\perp}\partial_i u_j; \quad (5.24)$$

$$(\mathbf{p}^g \cdot \nabla \mathbf{u})_{ij}^T = (p_{\parallel} - p_{\perp})\hat{b}_j\hat{\mathbf{b}} \cdot \nabla u_i + p_{\perp}\partial_j u_i; \quad (5.25)$$

$$\begin{aligned} (\nabla \cdot \mathbf{q}^g)_{ij} &= \hat{b}_i\hat{b}_j\nabla \cdot ((q_{\parallel} - 3q_{\perp})\hat{\mathbf{b}}) + (q_{\parallel} - 3q_{\perp})\hat{\mathbf{b}} \cdot \nabla(\hat{b}_i\hat{b}_j) \\ &\quad + \delta_{ij}\nabla \cdot (q_{\perp}\hat{\mathbf{b}}) + \partial_i(q_{\perp}\hat{b}_j) + \partial_j(q_{\perp}\hat{b}_i) \end{aligned} \quad (5.26)$$

and grouping similar terms in  $\kappa^{(1)}$  together yields

$$\begin{aligned} \kappa_{ij}^{(1)} &= \frac{1}{\Omega_{|B|}} \left\{ p_{\perp}(\partial_i u_j + \partial_j u_i) + \partial_i(q_{\perp}\hat{b}_j) + \partial_j(q_{\perp}\hat{b}_i) \right. \\ &\quad + (p_{\parallel} - p_{\perp}) \left[ \frac{d}{dt}(\hat{b}_i\hat{b}_j) + \hat{b}_i\hat{\mathbf{b}} \cdot \nabla u_j + \hat{b}_j\hat{\mathbf{b}} \cdot \nabla u_i \right] + (q_{\parallel} - 3q_{\perp})\hat{\mathbf{b}} \cdot \nabla(\hat{b}_i\hat{b}_j) \\ &\quad + \hat{b}_i\hat{b}_j \left[ \frac{d}{dt}(p_{\parallel} - p_{\perp}) + (p_{\parallel} - p_{\perp})\nabla \cdot \mathbf{u} + \nabla \cdot ((q_{\parallel} - 3q_{\perp})\hat{\mathbf{b}}) \right] \\ &\quad \left. + \delta_{ij} \left[ \frac{dp_{\perp}}{dt} + p_{\perp}\nabla \cdot \mathbf{u} + \nabla \cdot (q_{\perp}\hat{\mathbf{b}}) \right] \right\}. \end{aligned} \quad (5.27)$$

The first line in the above expression is left unchanged, since it is the final result, defined below as a matrix  $W_{ij} = p_{\perp}(\partial_i u_j + \partial_j u_i) + \partial_j(q_{\perp}\hat{b}_i) + \partial_i(q_{\perp}\hat{b}_j)$ , that can be written as  $\mathbf{W} = p_{\perp}\nabla \mathbf{u} + \nabla(q_{\perp}\hat{\mathbf{b}}) + (p_{\perp}\nabla \mathbf{u} + \nabla(q_{\perp}\hat{\mathbf{b}}))^T = (p_{\perp}\nabla \mathbf{u} + \nabla(q_{\perp}\hat{\mathbf{b}}))^S$ . However, calculating the rest of the expression according to (5.17) simplifies several terms, and for example all terms in the third line of (5.27) vanish. The terms in the third line containing  $\hat{b}_i\hat{b}_j$  vanish after applying  $\hat{\mathbf{b}} \times$ , since  $(\hat{\mathbf{b}} \times (\hat{\mathbf{b}}\hat{\mathbf{b}}))_{ij} = \epsilon_{ikl}\hat{b}_k\hat{b}_l\hat{b}_j = 0$ . The terms in the third line containing  $\delta_{ij}$  vanish after applying the symmetric operator since

$$\begin{aligned} [(\hat{\mathbf{b}} \times \mathbf{I}) \cdot (\mathbf{I} + 3\hat{\mathbf{b}}\hat{\mathbf{b}})]_{ij}^S &= [(\hat{\mathbf{b}} \times \mathbf{I})_{im}(\mathbf{I} + 3\hat{\mathbf{b}}\hat{\mathbf{b}})_{mj}]^S = [\epsilon_{ikl}\hat{b}_k\delta_{lm}(\delta_{mj} + 3\hat{b}_m\hat{b}_j)]^S \\ &= [\epsilon_{ikj}\hat{b}_k + 3\underbrace{\epsilon_{ikl}\hat{b}_k\hat{b}_l\hat{b}_j}_{=0}]^S = \epsilon_{ikj}\hat{b}_k + \epsilon_{jki}\hat{b}_k = \epsilon_{ikj}\hat{b}_k - \epsilon_{ikj}\hat{b}_k = 0. \end{aligned} \quad (5.28)$$

We therefore need to concentrate only on the second line of (5.27), and slightly rewrite that one. For example, the first term in the second line of (5.27), the term  $d/dt(\hat{b}_i\hat{b}_j)$ , calculates as

$$\begin{aligned}
\left[ \left( \hat{\mathbf{b}} \times \frac{d}{dt}(\hat{\mathbf{b}}\hat{\mathbf{b}}) \right) \cdot (\mathbf{I} + 3\hat{\mathbf{b}}\hat{\mathbf{b}}) \right]_{ij}^S &= \left[ \left( \hat{\mathbf{b}} \times \frac{d}{dt}(\hat{\mathbf{b}}\hat{\mathbf{b}}) \right)_{im} (\mathbf{I} + 3\hat{\mathbf{b}}\hat{\mathbf{b}})_{mj} \right]^S \\
&= \left[ \epsilon_{ikl} \hat{b}_k \frac{d}{dt}(\hat{b}_l \hat{b}_m) (\delta_{mj} + 3\hat{b}_m \hat{b}_j) \right]^S \\
&= \left[ \epsilon_{ikl} \hat{b}_k \left( \cancel{\hat{b}_l} \frac{d\hat{b}_m}{dt} + \hat{b}_m \frac{d\hat{b}_l}{dt} \right) (\delta_{mj} + 3\hat{b}_m \hat{b}_j) \right]^S \\
&= \left[ \epsilon_{ikl} \hat{b}_k \frac{d\hat{b}_l}{dt} (\hat{b}_j + 3\hat{b}_j) \right]^S \\
&= 4 \left[ \left( \hat{\mathbf{b}} \times \frac{d\hat{\mathbf{b}}}{dt} \right)_i \hat{b}_j \right]^S = 4 \left[ \left( \hat{\mathbf{b}} \times \frac{d\hat{\mathbf{b}}}{dt} \right) \hat{\mathbf{b}} \right]_{ij}^S. \quad (5.29)
\end{aligned}$$

The second term in the second line of (5.27), the term  $\hat{b}_i \hat{\mathbf{b}} \cdot \nabla u_j$ , vanishes after applying  $\hat{\mathbf{b}} \times$ , since

$$[\hat{\mathbf{b}} \times (\hat{\mathbf{b}}(\hat{\mathbf{b}} \cdot \nabla u))]_{ij} = \epsilon_{ikl} \hat{b}_k (\hat{\mathbf{b}}(\hat{\mathbf{b}} \cdot \nabla u))_{lj} = \underbrace{\epsilon_{ikl} \hat{b}_k \hat{b}_l}_{=0} (\hat{\mathbf{b}} \cdot \nabla u)_j = 0. \quad (5.30)$$

The third term in the second line of (5.27), the term  $\hat{b}_j \hat{\mathbf{b}} \cdot \nabla u_i$  however does not vanish. Note that  $\hat{b}_j \hat{\mathbf{b}} \cdot \nabla u_i = ((\hat{\mathbf{b}} \cdot \nabla u) \hat{\mathbf{b}})_{ij}$  and the third term calculates as

$$\begin{aligned}
[(\hat{\mathbf{b}} \times ((\hat{\mathbf{b}} \cdot \nabla u) \hat{\mathbf{b}})) \cdot (\mathbf{I} + 3\hat{\mathbf{b}}\hat{\mathbf{b}})]_{ij}^S &= [(\hat{\mathbf{b}} \times ((\hat{\mathbf{b}} \cdot \nabla u) \hat{\mathbf{b}}))_{im} (\mathbf{I} + 3\hat{\mathbf{b}}\hat{\mathbf{b}})_{mj}]^S \\
&= [\epsilon_{ikl} \hat{b}_k ((\hat{\mathbf{b}} \cdot \nabla u) \hat{\mathbf{b}})_{lm} (\delta_{mj} + 3\hat{b}_m \hat{b}_j)]^S = [\epsilon_{ikl} \hat{b}_k (\hat{\mathbf{b}} \cdot \nabla u)_l \hat{b}_m (\delta_{mj} + 3\hat{b}_m \hat{b}_j)]^S \\
&= 4[\epsilon_{ikl} \hat{b}_k (\hat{\mathbf{b}} \cdot \nabla u)_l \hat{b}_j]^S = 4[(\hat{\mathbf{b}} \times (\hat{\mathbf{b}} \cdot \nabla u))_i \hat{b}_j]^S = 4[(\hat{\mathbf{b}} \times (\hat{\mathbf{b}} \cdot \nabla u)) \hat{\mathbf{b}}]_{ij}^S. \quad (5.31)
\end{aligned}$$

The fourth term in the second line of (5.27), calculates in the same fashion than the second and third terms since  $\hat{\mathbf{b}} \cdot \nabla (\hat{b}_i \hat{b}_j) = \hat{b}_i (\hat{\mathbf{b}} \cdot \nabla \hat{b}_j) + (\hat{\mathbf{b}} \cdot \nabla \hat{b}_i) \hat{b}_j$ , yielding

$$[(\hat{\mathbf{b}} \times (\hat{\mathbf{b}} \cdot \nabla (\hat{\mathbf{b}}\hat{\mathbf{b}}))) \cdot (\mathbf{I} + 3\hat{\mathbf{b}}\hat{\mathbf{b}})]_{ij}^S = 4[(\hat{\mathbf{b}} \times (\hat{\mathbf{b}} \cdot \nabla \hat{\mathbf{b}})) \hat{\mathbf{b}}]_{ij}^S. \quad (5.32)$$

Collecting all the results together yields the FLR tensor

$$\Pi = \frac{1}{4\Omega_{|B|}} [(\hat{\mathbf{b}} \times \mathbf{W}) \cdot (\mathbf{I} + 3\hat{\mathbf{b}}\hat{\mathbf{b}})]^S + \frac{1}{\Omega_{|B|}} [(\hat{\mathbf{b}} \times \mathbf{w}) \hat{\mathbf{b}}]^S; \quad (5.33)$$

$$\mathbf{W} = [p_{\perp} \nabla \mathbf{u} + \nabla (q_{\perp} \hat{\mathbf{b}})]^S; \quad (5.34)$$

$$\mathbf{w} = (p_{\parallel} - p_{\perp}) \left( \frac{d\hat{\mathbf{b}}}{dt} + \hat{\mathbf{b}} \cdot \nabla \mathbf{u} \right) + (q_{\parallel} - 3q_{\perp}) \hat{\mathbf{b}} \cdot \nabla \hat{\mathbf{b}}, \quad (5.35)$$

where  $\mathbf{W}$  is a matrix and  $\mathbf{w}$  is a vector. The result is equivalent to (5)–(8) of Schekochihin *et al.* (2010), derived in the Appendix of that paper in a somewhat simpler way. In Schekochihin *et al.* (2010), instead of our notation for matrix  $\mathbf{W}$  and vector  $\mathbf{w}$ , the symbols  $\mathbf{S}$  and  $\boldsymbol{\sigma}$  are used, where  $\mathbf{S} = \mathbf{W}$  and  $\boldsymbol{\sigma} = -\mathbf{w}$ . We wanted to avoid the use of  $\mathbf{S}$  and  $\boldsymbol{\sigma}$ , since these symbols are typically used when decomposing

the heat flux tensor. For a reader who just jumped to this result, the symmetric operator is defined as  $A_{ij}^S = A_{ij} + A_{ji}$ . For clarity, for any vector  $\mathbf{w}$  and any matrix  $\mathbf{W}$ , the following identities hold

$$[(\hat{\mathbf{b}} \times \mathbf{w})\hat{\mathbf{b}}]^S = (\hat{\mathbf{b}} \times \mathbf{w})\hat{\mathbf{b}} + \hat{\mathbf{b}}(\hat{\mathbf{b}} \times \mathbf{w}); \quad (5.36)$$

$$[\hat{\mathbf{b}} \times \mathbf{W} \cdot (\mathbf{I} + 3\hat{\mathbf{b}}\hat{\mathbf{b}})]^S = \hat{\mathbf{b}} \times \mathbf{W} \cdot (\mathbf{I} + 3\hat{\mathbf{b}}\hat{\mathbf{b}}) - (\mathbf{I} + 3\hat{\mathbf{b}}\hat{\mathbf{b}}) \cdot \mathbf{W}^T \times \hat{\mathbf{b}}. \quad (5.37)$$

The last identity follows from an identity for general matrix  $\mathbf{A}$ , that reads  $(\hat{\mathbf{b}} \times \mathbf{A})^T = -\mathbf{A}^T \times \hat{\mathbf{b}}$ , and for a matrix  $\mathbf{A} = \mathbf{B} \cdot \mathbf{C}$  the  $\mathbf{A}^T = \mathbf{C}^T \cdot \mathbf{B}^T$ , implying  $(\hat{\mathbf{b}} \times \mathbf{B} \cdot \mathbf{C})^T = -\mathbf{C}^T \cdot \mathbf{B}^T \times \hat{\mathbf{b}}$ .

The result (5.33)–(5.35) can be slightly simplified. The matrix  $\mathbf{W}$  in the index notation reads

$$W_{ij} = p_{\perp}(\partial_i u_j + \partial_j u_i) + q_{\perp}(\partial_i \hat{b}_j + \partial_j \hat{b}_i) + (\partial_i q_{\perp})\hat{b}_j + (\partial_j q_{\perp})\hat{b}_i, \quad (5.38)$$

and by applying  $\hat{\mathbf{b}} \times$  eliminates only the last term

$$(\hat{\mathbf{b}} \times \mathbf{W})_{ij} = \epsilon_{ikl}\hat{b}_k W_{lj} = \epsilon_{ikl}\hat{b}_k [p_{\perp}(\partial_l u_j + \partial_j u_l) + q_{\perp}(\partial_l \hat{b}_j + \partial_j \hat{b}_l) + (\partial_l q_{\perp})\hat{b}_j + (\partial_j q_{\perp})\hat{b}_l], \quad (5.39)$$

implying that the matrix  $\mathbf{W}$  could possibly be redefined to

$$\mathbf{W} = [p_{\perp} \nabla \mathbf{u} + q_{\perp} \nabla \hat{\mathbf{b}}]^S + (\nabla q_{\perp})\hat{\mathbf{b}}. \quad (5.40)$$

However, now the matrix is not symmetric. Note that the brackets in the expression (5.33) are not unique, and one can for example pull the  $\hat{\mathbf{b}} \times$  out, so that the expression reads

$$\boldsymbol{\Pi} = \frac{1}{\Omega_{|B|}} \left\{ \hat{\mathbf{b}} \times \left[ \frac{1}{4} \mathbf{W} \cdot (\mathbf{I} + 3\hat{\mathbf{b}}\hat{\mathbf{b}}) + \mathbf{w}\hat{\mathbf{b}} \right] \right\}^S. \quad (5.41)$$

Let us multiply the last term in (5.40) by  $\cdot(\mathbf{I} + 3\hat{\mathbf{b}}\hat{\mathbf{b}})$ , it calculates as

$$((\nabla q_{\perp})\hat{\mathbf{b}}) \cdot (\mathbf{I} + 3\hat{\mathbf{b}}\hat{\mathbf{b}}) = 4(\nabla q_{\perp})\hat{\mathbf{b}}. \quad (5.42)$$

Obviously, it is beneficial to move this term to the vector  $\mathbf{w}$ . Therefore, redefining  $\mathbf{W}$  and  $\mathbf{w}$ , the solution reads

$$\boldsymbol{\Pi} = \frac{1}{4\Omega_{|B|}} [(\hat{\mathbf{b}} \times \mathbf{W}) \cdot (\mathbf{I} + 3\hat{\mathbf{b}}\hat{\mathbf{b}})]^S + \frac{1}{\Omega_{|B|}} [(\hat{\mathbf{b}} \times \mathbf{w})\hat{\mathbf{b}}]^S; \quad (5.43)$$

$$\mathbf{W} = [p_{\perp} \nabla \mathbf{u} + q_{\perp} \nabla \hat{\mathbf{b}}]^S; \quad (5.44)$$

$$\mathbf{w} = (p_{\parallel} - p_{\perp}) \left( \frac{d\hat{\mathbf{b}}}{dt} + \hat{\mathbf{b}} \cdot \nabla \mathbf{u} \right) + \nabla q_{\perp} + (q_{\parallel} - 3q_{\perp})\hat{\mathbf{b}} \cdot \nabla \hat{\mathbf{b}}. \quad (5.45)$$

Now the matrix  $\mathbf{W}$  is again symmetric, and  $\mathbf{w}$  is just a vector. This solution with matrix (5.44) is slightly nicer than the solution with matrix (5.34). Alternatively, one can write the solution in its full form



$$\begin{aligned} \boldsymbol{\Pi} = \frac{1}{\Omega_{|B|}} \left\{ \hat{\mathbf{b}} \times \left[ \frac{1}{4} [p_{\perp} \nabla \mathbf{u} + q_{\perp} \nabla \hat{\mathbf{b}}]^S \cdot (\mathbf{I} + 3\hat{\mathbf{b}}\hat{\mathbf{b}}) + (p_{\parallel} - p_{\perp}) \left( \frac{d\hat{\mathbf{b}}}{dt} + \hat{\mathbf{b}} \cdot \nabla \mathbf{u} \right) \hat{\mathbf{b}} \right. \right. \\ \left. \left. + (\nabla q_{\perp}) \hat{\mathbf{b}} + (q_{\parallel} - 3q_{\perp})(\hat{\mathbf{b}} \cdot \nabla \hat{\mathbf{b}}) \hat{\mathbf{b}} \right] \right\}^S. \end{aligned} \quad (5.46)$$

The nonlinear solution for the FLR tensor  $\boldsymbol{\Pi}$  can be further re-arranged. For example, to obtain the form reported by Ramos (2005), one needs to separate a part that is obtained by performing  $\boldsymbol{\Pi} \cdot \hat{\mathbf{b}}$ . The solution (5.43) in the index notation reads

$$\begin{aligned} \Pi_{ij} = \frac{1}{\Omega_{|B|}} \left\{ \epsilon_{ikl} \hat{b}_k \left[ \frac{1}{4} W_{lm} (\delta_{mj} + 3\hat{b}_m \hat{b}_j) + w_l \hat{b}_j \right] \right. \\ \left. + \epsilon_{jkl} \hat{b}_k \left[ \frac{1}{4} W_{lm} (\delta_{mi} + 3\hat{b}_m \hat{b}_i) + w_l \hat{b}_i \right] \right\}, \end{aligned} \quad (5.47)$$

and direct calculation yields

$$\Pi_{ij} \hat{b}_j = \frac{1}{\Omega_{|B|}} \left\{ \epsilon_{ikl} \hat{b}_k \left[ \frac{1}{4} W_{lm} (\delta_{mj} + 3\hat{b}_m \hat{b}_j) + w_l \hat{b}_j \right] \hat{b}_j \right\} = \frac{1}{\Omega_{|B|}} \epsilon_{ikl} \hat{b}_k (W_{lm} \hat{b}_m + w_l); \quad (5.48)$$

$$\boldsymbol{\Pi} \cdot \hat{\mathbf{b}} = \frac{1}{\Omega_{|B|}} \hat{\mathbf{b}} \times (\mathbf{W} \cdot \hat{\mathbf{b}} + \mathbf{w}) \equiv \frac{1}{\Omega_{|B|}} \hat{\mathbf{b}} \times \mathbf{h}, \quad (5.49)$$

where it was useful to define the new vector  $\mathbf{h} \equiv \mathbf{W} \cdot \hat{\mathbf{b}} + \mathbf{w}$ , by adding the quantity

$$\mathbf{W} \cdot \hat{\mathbf{b}} = p_{\perp} ((\nabla \mathbf{u}) \cdot \hat{\mathbf{b}} + \hat{\mathbf{b}} \cdot \nabla \mathbf{u}) + q_{\perp} \hat{\mathbf{b}} \cdot \nabla \hat{\mathbf{b}} \quad (5.50)$$

to the vector  $\mathbf{w}$ . By separating the  $\boldsymbol{\Pi} \cdot \hat{\mathbf{b}}$  part, the solution (5.43) for  $\boldsymbol{\Pi}$  is therefore re-arranged to the following form

$$\boldsymbol{\Pi} = \frac{1}{4\Omega_{|B|}} [(\hat{\mathbf{b}} \times \mathbf{W}) \cdot (\mathbf{I} - \hat{\mathbf{b}}\hat{\mathbf{b}})]^S + \frac{1}{\Omega_{|B|}} [(\hat{\mathbf{b}} \times \mathbf{h}) \hat{\mathbf{b}}]^S; \quad (5.51)$$

$$\mathbf{W} = [p_{\perp} \nabla \mathbf{u} + q_{\perp} \nabla \hat{\mathbf{b}}]^S; \quad (5.52)$$

$$\mathbf{h} = p_{\parallel} \left( \frac{d\hat{\mathbf{b}}}{dt} + \hat{\mathbf{b}} \cdot \nabla \mathbf{u} \right) - p_{\perp} \left( \frac{d\hat{\mathbf{b}}}{dt} - (\nabla \mathbf{u}) \cdot \hat{\mathbf{b}} \right) + \nabla q_{\perp} + (q_{\parallel} - 2q_{\perp}) \hat{\mathbf{b}} \cdot \nabla \hat{\mathbf{b}}. \quad (5.53)$$

The solution (5.51) has a nice advantage in that it is decomposed with respect to the magnetic field lines. In other words, projecting  $\boldsymbol{\Pi}$  along the field lines by performing  $\boldsymbol{\Pi} \cdot \hat{\mathbf{b}}$  cancels the first term, and directly yields vector  $\hat{\mathbf{b}} \times \mathbf{h} / \Omega_{|B|}$ , that could be defined as vector  $\tilde{\boldsymbol{\Pi}}_{\perp}$ .

### 5.1.1. Employing a non-dispersive induction equation

Considering the long-wavelength low-frequency limit, the solution (5.45) for vector  $\mathbf{w}$  and the solution (5.53) for vector  $\mathbf{h}$  can be further simplified, by using the usual

non-dispersive CGL/MHD induction equation, that can be written in the following form (see later in the text)

$$\frac{d\hat{\mathbf{b}}}{dt} = \hat{\mathbf{b}} \cdot \nabla \mathbf{u} - \hat{\mathbf{b}}[\hat{\mathbf{b}} \cdot (\nabla \mathbf{u}) \cdot \hat{\mathbf{b}}]. \quad (5.54)$$

By further applying  $\hat{\mathbf{b}} \times$  to the induction equation, the last term on the right-hand side disappears

$$\hat{\mathbf{b}} \times \frac{d\hat{\mathbf{b}}}{dt} = \hat{\mathbf{b}} \times (\hat{\mathbf{b}} \cdot \nabla \mathbf{u}), \quad (5.55)$$

and vectors (5.45), (5.53) simplify to

$$\mathbf{w} = 2(p_{\parallel} - p_{\perp})\hat{\mathbf{b}} \cdot \nabla \mathbf{u} + \nabla q_{\perp} + (q_{\parallel} - 3q_{\perp})\hat{\mathbf{b}} \cdot \nabla \hat{\mathbf{b}}; \quad (5.56)$$

$$\mathbf{h} = 2p_{\parallel}\hat{\mathbf{b}} \cdot \nabla \mathbf{u} - p_{\perp}(\hat{\mathbf{b}} \cdot \nabla \mathbf{u} - (\nabla \mathbf{u}) \cdot \hat{\mathbf{b}}) + \nabla q_{\perp} + (q_{\parallel} - 2q_{\perp})\hat{\mathbf{b}} \cdot \nabla \hat{\mathbf{b}}. \quad (5.57)$$

The result (5.57) can be further rewritten with vorticity  $\boldsymbol{\omega} = \nabla \times \mathbf{u}$ , by using identity  $(\nabla \mathbf{u}) \cdot \hat{\mathbf{b}} - \hat{\mathbf{b}} \cdot \nabla \mathbf{u} = \hat{\mathbf{b}} \times (\nabla \times \mathbf{u}) = \hat{\mathbf{b}} \times \boldsymbol{\omega}$ , finally yielding

$$\mathbf{h} = 2p_{\parallel}\hat{\mathbf{b}} \cdot \nabla \mathbf{u} + p_{\perp}\hat{\mathbf{b}} \times \boldsymbol{\omega} + \nabla q_{\perp} + (q_{\parallel} - 2q_{\perp})\hat{\mathbf{b}} \cdot \nabla \hat{\mathbf{b}}. \quad (5.58)$$

Results (5.51), (5.52), (5.58) are equivalent to equations (49)–(51) of Ramos (2005) (see also (11)–(13) of Macmahon (1965)). Further evaluation of nonlinear FLR corrections to higher precision is out of the scope of this simple guide, and the interested reader is referred to Ramos (2005). We will continue in a much simpler linear approximation.

## 5.2. Simplest FLR corrections (FLR1)

We could continue by evaluating the obtained nonlinear results in the linear approximation. However, the nonlinear FLR calculations were quite complicated and it is very useful to clarify how we derive the simplest FLR corrections directly in the linear approximation. Let us start again, and work with the general equation (5.1). To derive the simplest finite Larmor radius corrections to the pressure tensor, one has to cancel the  $\Pi$  contributions on the right-hand side of equation (5.1) and also neglect the heat flux contributions, which yields

$$\hat{\mathbf{b}} \times \Pi + (\hat{\mathbf{b}} \times \Pi)^T = -\frac{1}{\Omega} \frac{B_0}{|\mathbf{B}|} \left[ \frac{\partial \mathbf{p}^g}{\partial t} + \mathbf{u} \cdot \nabla \mathbf{p}^g + \mathbf{p}^g \nabla \cdot \mathbf{u} + \mathbf{p}^g \cdot \nabla \mathbf{u} + (\mathbf{p}^g \cdot \nabla \mathbf{u})^T \right]. \quad (5.59)$$

Considering complete linearization (so that we can easily go to Fourier space), this is equivalent to performing expansion to the first order in frequency  $\omega/\Omega$  and to the first order in wavenumber  $r_L k$ . We do not want to perform complete linearization just yet. However, to move further, we introduce simplification by not evaluating the FLR corrections along the magnetic field lines, but along the ambient magnetic field  $B_0$  that is prescribed to be in the  $z$ -direction, and therefore  $\hat{\mathbf{b}} = (0, 0, 1) \equiv \hat{\mathbf{b}}_0$ . This simplification, typically used in numerical simulations, is appropriate if the magnetic field lines are not too distorted with respect to the mean magnetic field. Here, we will distinguish between complete linearization and evaluation along  $\hat{\mathbf{b}}_0$ . However, it is important to note that the procedure of evaluating some terms with  $\hat{\mathbf{b}}_0$ , while

other nonlinearities are kept, is of course not a fully consistent procedure. With the approach presented here, the only fully consistent procedure that simplifies the general equation (5.1), is complete linearization. Additionally, as discussed by Passot *et al.* (2014), neglecting magnetic field line distortion may in some instances lead to spurious instabilities.

Let us evaluate (5.59) along  $\hat{\mathbf{b}}_0$ . On the left-hand side we calculate each component in the matrix separately. Since  $(\hat{\mathbf{b}} \times \boldsymbol{\Pi})_{ij} = \epsilon_{ikl} \hat{b}_k \Pi_{lj}$ , it is obvious that non-zero contributions are obtained only if  $\hat{b}_k = \hat{b}_z = 1$ , so the index  $k = z$ , and  $(\hat{\mathbf{b}} \times \boldsymbol{\Pi})_{ij} = \epsilon_{izl} \Pi_{lj}$ . Now, for example, the 'xx' component calculates as  $(\hat{\mathbf{b}} \times \boldsymbol{\Pi})_{xx} = \epsilon_{xzl} \Pi_{lx}$  and the only non-zero contribution is obtained for index  $l = y$ , which yields  $(\hat{\mathbf{b}} \times \boldsymbol{\Pi})_{xx} = \epsilon_{xyz} \Pi_{yx} = -\epsilon_{xyz} \Pi_{yx} = -\Pi_{yx} = -\Pi_{xy}$ , where in the last step we used that the FLR tensor is symmetric. The entire left-hand side  $[\hat{\mathbf{b}} \times \boldsymbol{\Pi} + (\hat{\mathbf{b}} \times \boldsymbol{\Pi})^T]_{ij} = \epsilon_{izl} \Pi_{lj} + \epsilon_{jzl} \Pi_{li}$  and it is straightforward to calculate each component, which yields

$$\hat{\mathbf{b}} \times \boldsymbol{\Pi} + (\hat{\mathbf{b}} \times \boldsymbol{\Pi})^T|_{\hat{\mathbf{b}}=\hat{\mathbf{b}}_0} = \begin{pmatrix} -2\Pi_{xy}, & \Pi_{xx} - \Pi_{yy}, & -\Pi_{yz} \\ \Pi_{xx} - \Pi_{yy}, & +2\Pi_{xy}, & \Pi_{xz} \\ -\Pi_{yz}, & \Pi_{xz}, & 0 \end{pmatrix}. \quad (5.60)$$

The gyrotropic pressure tensor is of course

$$\mathbf{p}^g|_{\hat{\mathbf{b}}=\hat{\mathbf{b}}_0} = \begin{pmatrix} p_{\perp} & 0 & 0 \\ 0 & p_{\perp} & 0 \\ 0 & 0 & p_{\parallel} \end{pmatrix}. \quad (5.61)$$

The first term on the right-hand side of (5.59) calculates as

$$\frac{\partial}{\partial t} p_{ij}^g = \hat{b}_i \hat{b}_j \frac{\partial}{\partial t} p_{\parallel} + (\delta_{ij} - \hat{b}_i \hat{b}_j) \frac{\partial}{\partial t} p_{\perp} + (p_{\parallel} - p_{\perp}) \left( \hat{b}_i \frac{\partial}{\partial t} \hat{b}_j + \hat{b}_j \frac{\partial}{\partial t} \hat{b}_i \right), \quad (5.62)$$

and using the matrix notation

$$\left( \frac{\partial}{\partial t} \mathbf{p}^g \right) \Big|_{\hat{\mathbf{b}}=\hat{\mathbf{b}}_0} = \frac{\partial}{\partial t} \begin{pmatrix} p_{\perp} & 0 & 0 \\ 0 & p_{\perp} & 0 \\ 0 & 0 & p_{\parallel} \end{pmatrix} + (p_{\parallel} - p_{\perp}) \begin{pmatrix} 0 & 0 & \frac{\partial \hat{b}_x}{\partial t} \\ 0 & 0 & \frac{\partial \hat{b}_y}{\partial t} \\ \frac{\partial \hat{b}_x}{\partial t} & \frac{\partial \hat{b}_y}{\partial t} & 2 \frac{\partial \hat{b}_z}{\partial t} \end{pmatrix} \Big|_{\hat{\mathbf{b}}=\hat{\mathbf{b}}_0}. \quad (5.63)$$

The second term in (5.63) contributes because of the induction equation  $\partial \mathbf{B} / \partial t = -c \nabla \times \mathbf{E}$ . For now, we will not use the induction equation and keep our FLR calculations in the above form. Let us continue with the second term on the right-hand side of (5.59). When completely linearized this term does not contribute, but for now we will keep it and combine it with the first term into a convective derivative  $\partial \mathbf{p}^g / \partial t + \mathbf{u} \cdot \nabla \mathbf{p}^g = d\mathbf{p}^g / dt$ , so that  $\partial / \partial t$  in (5.63) is replaced by  $d/dt$ . The remaining terms on the right-hand side of (5.59) are straightforward to evaluate since no derivatives of  $\hat{\mathbf{b}}$  are encountered, and the third term  $\mathbf{p}^g \nabla \cdot \mathbf{u}|_{\hat{\mathbf{b}}=\hat{\mathbf{b}}_0}$  is equal to matrix (5.61) multiplied by  $\nabla \cdot \mathbf{u}$ . The last two terms of (5.59) are together evaluated as

$$\mathbf{p}^g \cdot \nabla \mathbf{u} + (\mathbf{p}^g \cdot \nabla \mathbf{u})^T|_{\hat{\mathbf{b}}=\hat{\mathbf{b}}_0} = \begin{pmatrix} 2p_{\perp} \partial_x u_x; & p_{\perp} (\partial_x u_y + \partial_y u_x); & p_{\perp} \partial_x u_z + p_{\parallel} \partial_z u_x \\ p_{\perp} (\partial_y u_x + \partial_x u_y); & 2p_{\perp} \partial_y u_y; & p_{\perp} \partial_y u_z + p_{\parallel} \partial_z u_y \\ p_{\parallel} \partial_z u_x + p_{\perp} \partial_x u_z; & p_{\parallel} \partial_z u_y + p_{\perp} \partial_y u_z; & 2p_{\parallel} \partial_z u_z \end{pmatrix}. \quad (5.64)$$

Now we need to subtract the scalar CGL pressure equations obtained previously. Since  $\hat{\mathbf{b}} \cdot \nabla \mathbf{u} \cdot \hat{\mathbf{b}}|_{\hat{\mathbf{b}}=\hat{\mathbf{b}}_0} = \partial_z u_z$ , the scalar pressure equations read

$$\frac{dp_{\parallel}}{dt} + p_{\parallel} \nabla \cdot \mathbf{u} + 2p_{\parallel} \partial_z u_z = 0; \quad \frac{dp_{\perp}}{dt} + 2p_{\perp} \nabla \cdot \mathbf{u} - p_{\perp} \partial_z u_z = 0. \quad (5.65)$$

By cancelling terms coming from these two equations (affecting only the diagonal components), the entire equation (5.59) rewrites

$$\begin{pmatrix} -2\Pi_{xy}, & \Pi_{xx} - \Pi_{yy}, & -\Pi_{yz} \\ \Pi_{xx} - \Pi_{yy}, & +2\Pi_{xy}, & \Pi_{xz} \\ -\Pi_{yz}, & \Pi_{xz}, & 0 \end{pmatrix} = -\frac{1}{\Omega} (p_{\parallel} - p_{\perp}) \begin{pmatrix} 0 & 0 & \frac{d\hat{b}_x}{dt} \\ 0 & 0 & \frac{d\hat{b}_y}{dt} \\ \frac{d\hat{b}_x}{dt} & \frac{d\hat{b}_y}{dt} & \frac{d\hat{b}_z}{dt} \end{pmatrix} \bigg|_{\hat{\mathbf{b}}=\hat{\mathbf{b}}_0} \\ - \frac{1}{\Omega} \begin{pmatrix} p_{\perp}(\partial_x u_x - \partial_y u_y); & p_{\perp}(\partial_x u_y + \partial_y u_x); & p_{\perp} \partial_x u_z + p_{\parallel} \partial_z u_x \\ p_{\perp}(\partial_y u_x + \partial_x u_y); & p_{\perp}(\partial_y u_y - \partial_x u_x); & p_{\perp} \partial_y u_z + p_{\parallel} \partial_z u_y \\ p_{\parallel} \partial_z u_x + p_{\perp} \partial_x u_z; & p_{\parallel} \partial_z u_y + p_{\perp} \partial_y u_z; & 0 \end{pmatrix}. \quad (5.66)$$

The identity  $\boldsymbol{\Pi} : \hat{\mathbf{b}}\hat{\mathbf{b}} = 0$  evaluated for  $\hat{\mathbf{b}} = \hat{\mathbf{b}}_0$  implies that  $\Pi_{zz} = 0$  and the identity  $\text{Tr} \boldsymbol{\Pi} = 0$  further implies that  $\Pi_{xx} = -\Pi_{yy}$ . The components of the pressure tensor therefore read

$$\left. \begin{aligned} \Pi_{xx} = -\Pi_{yy} &= -\frac{p_{\perp}}{2\Omega} (\partial_x u_y + \partial_y u_x); & \Pi_{xy} &= \frac{p_{\perp}}{2\Omega} (\partial_x u_x - \partial_y u_y); & \Pi_{zz} &= 0; \\ \Pi_{xz} &= -\frac{1}{\Omega} \left( (p_{\parallel} - p_{\perp}) \frac{d\hat{b}_y}{dt} + p_{\parallel} \partial_z u_y + p_{\perp} \partial_y u_z \right); \\ \Pi_{yz} &= \frac{1}{\Omega} \left( (p_{\parallel} - p_{\perp}) \frac{d\hat{b}_x}{dt} + p_{\parallel} \partial_z u_x + p_{\perp} \partial_x u_z \right). \end{aligned} \right\} \quad (5.67)$$

Finally, the time derivative of the unit vector

$$\frac{d\hat{\mathbf{b}}}{dt} = \frac{d\mathbf{B}}{dt} \frac{1}{|\mathbf{B}|} = \frac{1}{|\mathbf{B}|} \frac{d\mathbf{B}}{dt} - \frac{\mathbf{B}}{|\mathbf{B}|^2} \frac{d|\mathbf{B}|}{dt} = \frac{1}{|\mathbf{B}|} \left[ \frac{d\mathbf{B}}{dt} - \hat{\mathbf{b}} \frac{d|\mathbf{B}|}{dt} \right] = \frac{1}{|\mathbf{B}|} \left[ \frac{d\mathbf{B}}{dt} - \hat{\mathbf{b}} \left( \hat{\mathbf{b}} \cdot \frac{d\mathbf{B}}{dt} \right) \right], \quad (5.68)$$

which evaluated for  $\hat{\mathbf{b}} = \hat{\mathbf{b}}_0 = (0, 0, 1)$ ,  $|\mathbf{B}| = B_0$  yields

$$\left. \begin{aligned} \frac{d\hat{b}_x}{dt} \bigg|_{\hat{\mathbf{b}}=\hat{\mathbf{b}}_0} &= \frac{1}{B_0} \frac{dB_x}{dt} \bigg|_{\hat{\mathbf{b}}=\hat{\mathbf{b}}_0}; & \frac{d\hat{b}_y}{dt} \bigg|_{\hat{\mathbf{b}}=\hat{\mathbf{b}}_0} &= \frac{1}{B_0} \frac{dB_y}{dt} \bigg|_{\hat{\mathbf{b}}=\hat{\mathbf{b}}_0}; \\ \frac{d\hat{b}_z}{dt} \bigg|_{\hat{\mathbf{b}}=\hat{\mathbf{b}}_0} &= \frac{1}{B_0} \left( \frac{dB_z}{dt} - \frac{dB_z}{dt} \right) = 0. \end{aligned} \right\} \quad (5.69)$$

The general induction equation reads

$$\frac{d\mathbf{B}}{dt} = -\mathbf{B} \nabla \cdot \mathbf{u} + \mathbf{B} \cdot \nabla \mathbf{u} - c \nabla \times \mathbf{E}_H, \quad (5.70)$$

which evaluates in the  $\hat{\mathbf{b}} = \hat{\mathbf{b}}_0$  approximation as

$$\left. \begin{aligned} \frac{dB_x}{dt} \Big|_{\hat{\mathbf{b}}=\hat{\mathbf{b}}_0} &= B_0 \partial_z u_x - c(\nabla \times \mathbf{E}_H)_x; & \frac{dB_y}{dt} \Big|_{\hat{\mathbf{b}}=\hat{\mathbf{b}}_0} &= B_0 \partial_z u_y - c(\nabla \times \mathbf{E}_H)_y; \\ \frac{dB_z}{dt} \Big|_{\hat{\mathbf{b}}=\hat{\mathbf{b}}_0} &= -B_0 \nabla \cdot \mathbf{u} + B_0 \partial_z u_z - c(\nabla \times \mathbf{E}_H)_z. \end{aligned} \right\} \quad (5.71)$$

These expressions are used in the FLR tensor components  $\Pi_{xz}$  and  $\Pi_{yz}$ , and the FLR tensor reads

$$\left. \begin{aligned} \Pi_{xx} &= -\Pi_{yy} = -\frac{p_{\perp}}{2\Omega} (\partial_x u_y + \partial_y u_x); & \Pi_{xy} &= \frac{p_{\perp}}{2\Omega} (\partial_x u_x - \partial_y u_y); & \Pi_{zz} &= 0; \\ \Pi_{xz} &= -\frac{1}{\Omega} \left[ (p_{\parallel} - p_{\perp}) \left( \partial_z u_y - \frac{c}{B_0} (\nabla \times \mathbf{E}_H)_y \right) + p_{\parallel} \partial_z u_y + p_{\perp} \partial_y u_z \right]; \\ \Pi_{yz} &= \frac{1}{\Omega} \left[ (p_{\parallel} - p_{\perp}) \left( \partial_z u_x - \frac{c}{B_0} (\nabla \times \mathbf{E}_H)_x \right) + p_{\parallel} \partial_z u_x + p_{\perp} \partial_x u_z \right]. \end{aligned} \right\} \quad (5.72)$$

However, since the FLR tensor was derived with a precision to only first order in wavenumber  $k$ , to make everything consistent, the Hall contributions in the induction equation are usually (but not always) neglected. We therefore have  $d\hat{\mathbf{b}}_x/dt = \partial_z u_x$  and  $d\hat{\mathbf{b}}_y/dt = \partial_z u_y$  which yields the leading-order FLR pressure tensor evaluated along the magnetic field  $\hat{\mathbf{b}} = (0, 0, 1)$  in the form

$$\left. \begin{aligned} \Pi_{xx} &= -\Pi_{yy} = -\frac{p_{\perp}}{2\Omega} (\partial_x u_y + \partial_y u_x); & \Pi_{xy} &= \frac{p_{\perp}}{2\Omega} (\partial_x u_x - \partial_y u_y); & \Pi_{zz} &= 0; \\ \Pi_{xz} &= -\frac{1}{\Omega} ((2p_{\parallel} - p_{\perp}) \partial_z u_y + p_{\perp} \partial_y u_z); & \Pi_{yz} &= \frac{1}{\Omega} ((2p_{\parallel} - p_{\perp}) \partial_z u_x + p_{\perp} \partial_x u_z), \end{aligned} \right\} \quad (5.73)$$

which is equivalent to equation (5) of Yajima (1966). Perhaps surprisingly, the result is also equivalent to the highly collisional equations (2.20)–(2.24) of Braginskii (1965), after one prescribes  $p_{\parallel} = p_{\perp}$  and after terms containing the collisional time  $\tau$  are ‘ignored’ ( $\tau \sim 1/\nu$  where  $\nu$  is the usual collisional frequency). A proper collisionless limit  $\tau \rightarrow \infty$  cannot be obtained from the expressions of Braginskii (1965), since the results are proportional to  $\tau$  (and also  $1/\tau$ ).

Furthermore, evaluation along  $\hat{\mathbf{b}} = (0, 0, 1)$  basically means that the magnetic field terms of the system were linearized and other parts of the system were not. The expansion procedure preserved some nonlinearities, but other nonlinearities that are of the same order were implicitly neglected. Therefore, the only fully consistent procedure is to evaluate the FLR tensor in the linear approximation, yielding

$$\left. \begin{aligned} \Pi_{xx}^{(1)} &= -\Pi_{yy}^{(1)} = -\frac{p_{\perp}^{(0)}}{2\Omega} (\partial_x u_y + \partial_y u_x); & \Pi_{xy}^{(1)} &= \frac{p_{\perp}^{(0)}}{2\Omega} (\partial_x u_x - \partial_y u_y); & \Pi_{zz}^{(1)} &= 0; \\ \Pi_{xz}^{(1)} &= -\frac{1}{\Omega} ((2p_{\parallel}^{(0)} - p_{\perp}^{(0)}) \partial_z u_y + p_{\perp}^{(0)} \partial_y u_z); & \Pi_{yz}^{(1)} &= \frac{1}{\Omega} ((2p_{\parallel}^{(0)} - p_{\perp}^{(0)}) \partial_z u_x + p_{\perp}^{(0)} \partial_x u_z). \end{aligned} \right\} \quad (5.74)$$

Notably, the tensor is very different from (3.10)–(3.13) of Oraevskii *et al.* (1968). The FLR corrections to the pressure tensor (these ones or more complicated ones) were used in a number of direct numerical simulations, see for example Borgogno *et al.* (2009), Hunana *et al.* (2011), Kobayashi *et al.* (2017) and Perrone *et al.* (2018), or for the isotropic case see for example Ghosh & Parashar (2015), and references therein.

## 5.3. Hall-CGL-FLR1 fluid model

The normalized, linearized, Fourier transformed system of equations written in the  $x$ - $z$  plane reads (dropping the tilde everywhere)

$$-\omega\rho + k_{\perp}u_x + k_{\parallel}u_z = 0; \quad (5.75)$$

$$-\omega u_x + \frac{\beta_{\parallel}}{2}k_{\perp}p_{\perp} - v_{A\parallel}^2 k_{\parallel}B_x + k_{\perp}B_z + \frac{\beta_{\parallel}}{2}(k_{\perp}\Pi_{xx} + k_{\parallel}\Pi_{xz}) = 0; \quad (5.76)$$

$$-\omega u_y - v_{A\parallel}^2 k_{\parallel}B_y + \frac{\beta_{\parallel}}{2}(k_{\perp}\Pi_{xy} + k_{\parallel}\Pi_{yz}) = 0; \quad (5.77)$$

$$-\omega u_z + \frac{\beta_{\parallel}}{2}k_{\parallel}p_{\parallel} + \frac{\beta_{\parallel}}{2}(1 - a_p)k_{\perp}B_x + \frac{\beta_{\parallel}}{2}k_{\perp}\Pi_{xz} = 0; \quad (5.78)$$

$$-\omega B_x - k_{\parallel}u_x - ik_{\parallel}^2 B_y = 0; \quad (5.79)$$

$$-\omega B_y - k_{\parallel}u_y - ik_{\parallel}k_{\perp}B_z + ik_{\parallel}^2 B_x = 0; \quad (5.80)$$

$$-\omega B_z + k_{\perp}u_x + ik_{\parallel}k_{\perp}B_y = 0; \quad (5.81)$$

$$-\omega p_{\parallel} + k_{\perp}u_x + 3k_{\parallel}u_z = 0; \quad (5.82)$$

$$-\omega p_{\perp} + 2a_p k_{\perp}u_x + a_p k_{\parallel}u_z = 0, \quad (5.83)$$

where  $v_{A\parallel}^2 = 1 + (\beta_{\parallel}/2)(a_p - 1)$ , and the first-order FLR tensor  $\Pi = \Pi^{(1)}$  is

$$\left. \begin{aligned} \Pi_{xx}^{(1)} &= -\frac{a_p}{2}ik_{\perp}u_y; & \Pi_{xy}^{(1)} &= \frac{a_p}{2}ik_{\perp}u_x; & \Pi_{zz}^{(1)} &= 0; \\ \Pi_{xz}^{(1)} &= -(2 - a_p)ik_{\parallel}u_y; & \Pi_{yz}^{(1)} &= (2 - a_p)ik_{\parallel}u_x + a_p ik_{\perp}u_z. \end{aligned} \right\} \quad (5.84)$$

The dispersion relation for the Hall-CGL-FLR1 fluid model can be written as

$$(\omega^2 - k_{\parallel}^2 v_{A\parallel}^2)(\omega^4 - A_2 \omega^2 + A_0) = k_{\perp}^2 k_{\parallel}^2 [\omega^4 - \beta_{\parallel} \left( \frac{3}{2} k_{\parallel}^2 + a_p k_{\perp}^2 \right) \omega^2 + k_{\parallel}^2 k_{\perp}^2 \frac{5}{4} a_p \beta_{\parallel}^2] + \mathcal{P}^{\text{FLR}}; \quad (5.85)$$

$$A_2 = k_{\perp}^2 (1 + a_p \beta_{\parallel}) + k_{\parallel}^2 (v_{A\parallel}^2 + \frac{3}{2} \beta_{\parallel}); \quad (5.86)$$

$$A_0 = \frac{3}{2} k_{\parallel}^2 \beta_{\parallel} [k_{\parallel}^2 v_{A\parallel}^2 + k_{\perp}^2 (1 + a_p \beta_{\parallel} - \frac{1}{6} a_p^2 \beta_{\parallel})], \quad (5.87)$$

where the polynomial on the right-hand side

$$\mathcal{P}^{\text{FLR}} = A'_4 \omega^4 - A'_2 \omega^2 + A'_0; \quad (5.88)$$

$$A'_4 = \beta_{\parallel}^2 \left[ k_{\parallel}^4 \left( 1 - \frac{a_p}{2} \right)^2 + k_{\perp}^4 \frac{a_p^2}{16} + k_{\parallel}^2 k_{\perp}^2 a_p \left( 1 - \frac{a_p}{2} \right) \right]; \quad (5.89)$$

$$\begin{aligned} A'_2 = & \beta_{\parallel} k_{\parallel}^2 \left\{ k_{\parallel}^6 \beta_{\parallel} \left( 1 - \frac{a_p}{2} \right)^2 + k_{\parallel}^4 \left[ \frac{3}{2} \beta_{\parallel}^2 \left( 1 - \frac{a_p}{2} \right)^2 + k_{\perp}^2 \beta_{\parallel} \left( 1 - \frac{a_p^2}{4} \right) + v_{A\parallel}^2 (a_p - 2) \right] \right. \\ & + k_{\parallel}^2 k_{\perp}^2 \left[ -1 + \frac{a_p}{2} + \beta_{\parallel} \frac{a_p}{4} \left( 1 - a_p + \beta_{\parallel} \left( 1 - \frac{a_p}{2} \right) \right) \right. \\ & \left. \left. - v_{A\parallel}^2 \left( 1 - \beta_{\parallel} \frac{a_p}{2} \left( 1 - \frac{a_p}{2} \right) \right) + k_{\perp}^2 a_p \beta_{\parallel} \left( 1 - \frac{7}{16} a_p \right) \right] \right. \\ & \left. + k_{\perp}^4 \frac{a_p}{4} \left[ -1 + 2\beta_{\parallel} - v_{A\parallel}^2 - a_p \beta_{\parallel} \left( 1 + \beta_{\parallel} \left( a_p - \frac{15}{8} \right) - \frac{1}{4} k_{\perp}^2 \right) \right] \right\}; \quad (5.90) \end{aligned}$$

$$\begin{aligned}
A'_0 = & \beta_{\parallel}^2 k_{\parallel}^4 \left\{ k_{\parallel}^6 \frac{3}{2} \beta_{\parallel} \left(1 - \frac{a_p}{2}\right)^2 + k_{\parallel}^4 \left(1 - \frac{a_p}{2}\right) \left[ -3v_{A\parallel}^2 + k_{\perp}^2 \beta_{\parallel} \frac{3}{2} \left(1 - \frac{a_p}{2} + \frac{a_p^2}{6}\right) \right] \right. \\
& + k_{\parallel}^2 k_{\perp}^2 \left[ k_{\perp}^2 \beta_{\parallel} a_p^2 \left( \frac{21}{32} - \frac{5}{16} a_p \right) - v_{A\parallel}^2 \frac{3}{2} \left(1 - \frac{a_p}{2} + \frac{1}{6} a_p^2\right) \right. \\
& \left. \left. - \left(1 - \frac{a_p}{2}\right) \frac{3}{2} \left(1 + \frac{1}{6} a_p \beta_{\parallel} - \frac{1}{6} a_p^2 \beta_{\parallel}\right) \right] \right. \\
& \left. + k_{\perp}^4 a_p \left[ k_{\perp}^2 a_p \beta_{\parallel} \left( \frac{13}{32} - \frac{3}{16} a_p \right) + \frac{v_{A\parallel}^2}{8} (1 - 2a_p) - \frac{1}{8} + \frac{3}{16} a_p \beta_{\parallel} (1 - a_p) \right] \right\}. \quad (5.91)
\end{aligned}$$

The solution for strictly parallel propagation is

$$\omega = \pm \frac{k_{\parallel}^2}{2} \left(1 + \beta_{\parallel} \left(1 - \frac{a_p}{2}\right)\right) + k_{\parallel} \sqrt{1 + \frac{\beta_{\parallel}}{2} (a_p - 1) + \frac{k_{\parallel}^2}{4} \left(1 - \beta_{\parallel} \left(1 - \frac{a_p}{2}\right)\right)^2}. \quad (5.92)$$

For strictly perpendicular propagation the solution is

$$\omega = \pm k_{\perp} \sqrt{1 + a_p \beta_{\parallel} + \frac{k_{\perp}^2 \beta_{\parallel}^2 a_p^2}{16}}, \quad (5.93)$$

and the result is consistent with equation (17) of Yajima (1966), after neglecting the electron inertia in that paper (by  $\omega_0 \rightarrow \infty$ ). The strictly perpendicular propagation is an excellent example to point out the deficiencies of the Hall-CGL-FLR1 fluid model, or actually the CGL-FLR1 model, since the Hall contributions are zero for  $k_{\parallel} = 0$ . The result (5.93) is also equal to equation (34) of Del Sarto *et al.* (2017), who discuss in detail that, compared to solutions of a truncated kinetic Vlasov system, and also of a fluid model with more precise FLR corrections, the solution (5.93) actually has the opposite sign in front of the FLR term  $\sim k_{\perp}^2$ . For further refinement with FLR2, see (5.117) and the subsequent discussion.

#### 5.4. CGL-FLR1 fluid model (no Hall term)

If the Hall term is neglected, the general dispersion relation of the CGL-FLR1 fluid model reads

$$(\omega^2 - k_{\parallel}^2 v_{A\parallel}^2)(\omega^4 - A_2 \omega^2 + A_0) = \mathcal{P}^{\text{FLR-only}}, \quad (5.94)$$

where the CGL parameters  $A_2, A_0$  are (5.87) and the FLR1 polynomial on the right-hand side has the form

$$\mathcal{P}^{\text{FLR-only}} = \omega^2 \beta_{\parallel}^2 (A_2'' \omega^2 - A_0''); \quad (5.95)$$

$$A_2'' = k_{\parallel}^4 \left(1 - \frac{a_p}{2}\right)^2 + k_{\perp}^4 \frac{a_p^2}{16} + k_{\parallel}^2 k_{\perp}^2 a_p \left(1 - \frac{a_p}{2}\right); \quad (5.96)$$

$$\begin{aligned}
A_0'' = & k_{\parallel}^2 \left\{ k_{\parallel}^4 \frac{3}{2} \beta_{\parallel} \left(1 - \frac{a_p}{2}\right)^2 + k_{\parallel}^2 k_{\perp}^2 \frac{a_p}{2} \left( v_{A\parallel}^2 + \frac{\beta_{\parallel}}{2} \right) \left(1 - \frac{a_p}{2}\right) \right. \\
& \left. + k_{\perp}^4 \frac{a_p}{2} \left[ 1 - \frac{a_p}{2} + a_p \beta_{\parallel} \left( \frac{15}{16} - \frac{a_p}{2} \right) \right] \right\}, \quad (5.97)
\end{aligned}$$

where  $v_{A\parallel}^2 = 1 + (\beta_{\parallel}/2)(a_p - 1)$ . For strictly parallel propagation, the solution for the whistler and ion-cyclotron waves reads

$$\omega = \pm k_{\parallel}^2 \frac{\beta_{\parallel}}{2} \left(1 - \frac{a_p}{2}\right) + k_{\parallel} \sqrt{v_{A\parallel}^2 + \frac{k_{\parallel}^2 \beta_{\parallel}^2}{4} \left(1 - \frac{a_p}{2}\right)^2}. \quad (5.98)$$

For strictly perpendicular propagation the solution is equivalent to (5.93) since the Hall contributions vanish.

### 5.5. Second-order FLR corrections (FLR2)

If more precise FLR corrections are desired, it is relatively easy to increase the precision to the second order in frequency  $\omega/\Omega$  while keeping the precision in wavenumber  $kr_L$  at first order. The pressure tensor equation (5.1) is simplified to

$$\hat{\mathbf{b}} \times \boldsymbol{\Pi} + (\hat{\mathbf{b}} \times \boldsymbol{\Pi})^T = -\frac{1}{\Omega} \frac{B_0}{|\mathbf{B}|} \left[ \frac{\partial}{\partial t} (\mathbf{p}^g + \boldsymbol{\Pi}) + \nabla \cdot (\mathbf{u} \mathbf{p}^g + \mathbf{q}^g) + \mathbf{p}^g \cdot \nabla \mathbf{u} + (\mathbf{p}^g \cdot \nabla \mathbf{u})^T \right]. \quad (5.99)$$

For clarity, let us first consider the case with the heat flux  $\mathbf{q} = 0$ . Compared to the first-order FLR corrections in the previous section, we just have one more matrix present,  $(\partial/\partial t)\boldsymbol{\Pi}$ . As before, the FLR tensor must be symmetric and also satisfy conditions  $\Pi_{zz} = 0$  and  $\Pi_{yy} = -\Pi_{xx}$ . Following the previous derivation, the system (5.66) now includes also the time derivative of the FLR tensor on the right-hand side. Let us write down the system directly in the linear approximation. The system reads

$$\begin{pmatrix} -2\Pi_{xy}, & 2\Pi_{xx}, & -\Pi_{yz} \\ 2\Pi_{xx}, & 2\Pi_{xy}, & \Pi_{xz} \\ -\Pi_{yz}, & \Pi_{xz}, & 0 \end{pmatrix} = -\frac{1}{\Omega} \frac{\partial}{\partial t} \begin{pmatrix} \Pi_{xx} & \Pi_{xy} & \Pi_{xz} \\ \Pi_{xy} & -\Pi_{xx} & \Pi_{yz} \\ \Pi_{xz} & \Pi_{yz} & 0 \end{pmatrix} \\ - \frac{1}{\Omega} (p_{\parallel}^{(0)} - p_{\perp}^{(0)}) \begin{pmatrix} 0 & 0 & \frac{\partial \hat{b}_x}{\partial t} \\ 0 & 0 & \frac{\partial \hat{b}_y}{\partial t} \\ \frac{\partial \hat{b}_x}{\partial t} & \frac{\partial \hat{b}_y}{\partial t} & 0 \end{pmatrix} \\ - \frac{1}{\Omega} \begin{pmatrix} p_{\perp}^{(0)} (\partial_x u_x - \partial_y u_y); & p_{\perp}^{(0)} (\partial_x u_y + \partial_y u_x); & p_{\perp}^{(0)} \partial_x u_z + p_{\parallel}^{(0)} \partial_z u_x \\ p_{\perp}^{(0)} (\partial_y u_x + \partial_x u_y); & p_{\perp}^{(0)} (\partial_y u_y - \partial_x u_x); & p_{\perp}^{(0)} \partial_y u_z + p_{\parallel}^{(0)} \partial_z u_y \\ p_{\parallel}^{(0)} \partial_z u_x + p_{\perp}^{(0)} \partial_x u_z; & p_{\parallel}^{(0)} \partial_z u_y + p_{\perp}^{(0)} \partial_y u_z; & 0 \end{pmatrix}. \quad (5.100)$$

There are only four FLR components that have to be considered and which can be rewritten separately as

$$\frac{\partial}{\partial t} \Pi_{xx} - 2\Omega \Pi_{xy} + p_{\perp}^{(0)} (\partial_x u_x - \partial_y u_y) = 0; \quad (5.101)$$

$$\frac{\partial}{\partial t} \Pi_{xy} + 2\Omega \Pi_{xx} + p_{\perp}^{(0)} (\partial_x u_y + \partial_y u_x) = 0; \quad (5.102)$$

$$\frac{\partial}{\partial t} \Pi_{xz} - \Omega \Pi_{yz} + p_{\perp}^{(0)} \partial_x u_z + p_{\parallel}^{(0)} \partial_z u_x + (p_{\parallel}^{(0)} - p_{\perp}^{(0)}) \frac{\partial \hat{b}_x}{\partial t} = 0; \quad (5.103)$$

$$\frac{\partial}{\partial t} \Pi_{yz} + \Omega \Pi_{xz} + p_{\perp}^{(0)} \partial_y u_z + p_{\parallel}^{(0)} \partial_z u_y + (p_{\parallel}^{(0)} - p_{\perp}^{(0)}) \frac{\partial \hat{b}_y}{\partial t} = 0. \quad (5.104)$$



The FLR components could be considered as independent fluid quantities described by the time-dependent equations as written above. However, from a linear perspective this introduces additional linear modes, and from a direct numerical simulation perspective it is preferred not to introduce an additional four time-dependent equations (nevertheless, numerical simulations with the full pressure tensor equation are sometimes performed, see e.g. Wang *et al.* (2015), Del Sarto, Pegoraro & Califano (2016) and Ng *et al.* (2017)). The above equations are therefore typically simplified by expanding the FLR tensor to the first and second orders as  $\Pi = \Pi^{(1)} + \Pi^{(2)}$  and the time derivative of the second-order tensor is neglected, i.e.  $(\partial/\partial t)\Pi^{(2)} = 0$  and the system reads

$$\frac{\partial}{\partial t}\Pi_{xx}^{(1)} - 2\Omega(\Pi_{xy}^{(1)} + \Pi_{xy}^{(2)}) + p_{\perp}^{(0)}(\partial_x u_x - \partial_y u_y) = 0; \quad (5.105)$$

$$\frac{\partial}{\partial t}\Pi_{xy}^{(1)} + 2\Omega(\Pi_{xx}^{(1)} + \Pi_{xx}^{(2)}) + p_{\perp}^{(0)}(\partial_x u_y + \partial_y u_x) = 0; \quad (5.106)$$

$$\begin{aligned} \frac{\partial}{\partial t}\Pi_{xz}^{(1)} - \Omega(\Pi_{yz}^{(1)} + \Pi_{yz}^{(2)}) + p_{\perp}^{(0)}\partial_x u_z + p_{\parallel}^{(0)}\partial_z u_x \\ + (p_{\parallel}^{(0)} - p_{\perp}^{(0)})\left(\partial_z u_x - \frac{c}{B_0}(\nabla \times \mathbf{E}_H)_x\right) = 0; \end{aligned} \quad (5.107)$$

$$\begin{aligned} \frac{\partial}{\partial t}\Pi_{yz}^{(1)} + \Omega(\Pi_{xz}^{(1)} + \Pi_{xz}^{(2)}) + p_{\perp}^{(0)}\partial_y u_z + p_{\parallel}^{(0)}\partial_z u_y \\ + (p_{\parallel}^{(0)} - p_{\perp}^{(0)})\left(\partial_z u_y - \frac{c}{B_0}(\nabla \times \mathbf{E}_H)_y\right) = 0, \end{aligned} \quad (5.108)$$

where we also used the linearized induction equation in the last two equations. Now, there are two possibilities to handle the Hall-term contributions coming from the induction equation. It is possible to either define  $\Pi^{(1)}$  to be equal to set (5.74) and move the Hall term to  $\Pi^{(2)}$ , or it is possible to keep the Hall term in  $\Pi^{(1)}$  that is equal to linearized set (5.72). The first choice, i.e. if  $\Pi^{(1)}$  is set, (5.74) yields the second-order components

$$\left. \begin{aligned} \Pi_{xx}^{(2)} &= -\frac{1}{2\Omega}\frac{\partial}{\partial t}\Pi_{xy}^{(1)}; & \Pi_{xy}^{(2)} &= \frac{1}{2\Omega}\frac{\partial}{\partial t}\Pi_{xx}^{(1)}; \\ \Pi_{xz}^{(2)} &= -\frac{1}{\Omega}\frac{\partial}{\partial t}\Pi_{yz}^{(1)} + \frac{1}{\Omega}(p_{\parallel}^{(0)} - p_{\perp}^{(0)})\frac{c}{B_0}(\nabla \times \mathbf{E}_H)_y; \\ \Pi_{yz}^{(2)} &= \frac{1}{\Omega}\frac{\partial}{\partial t}\Pi_{xz}^{(1)} - \frac{1}{\Omega}(p_{\parallel}^{(0)} - p_{\perp}^{(0)})\frac{c}{B_0}(\nabla \times \mathbf{E}_H)_x. \end{aligned} \right\} \quad (5.109)$$

The second choice, i.e. when  $\Pi^{(1)}$  is equal to linearized set (5.72), yields

$$\Pi_{xx}^{(2)} = -\frac{1}{2\Omega}\frac{\partial}{\partial t}\Pi_{xy}^{(1)}; \quad \Pi_{xy}^{(2)} = \frac{1}{2\Omega}\frac{\partial}{\partial t}\Pi_{xx}^{(1)}; \quad \Pi_{xz}^{(2)} = -\frac{1}{\Omega}\frac{\partial}{\partial t}\Pi_{yz}^{(1)}; \quad \Pi_{yz}^{(2)} = \frac{1}{\Omega}\frac{\partial}{\partial t}\Pi_{xz}^{(1)}. \quad (5.110)$$

The difference between the two approaches therefore is that the first approach neglects the time derivative of the Hall term. It is noted that the second approach is not necessarily better or more accurate than the first approach. For example, for parallel propagation, the frequency of the whistler mode seems to be better described by the first approach and the solutions are closer to kinetic theory.

## 5.6. Hall-CGL-FLR2 fluid model

The second-order FLR tensor (5.109) is normalized and written in Fourier space according to

$$\left. \begin{aligned} \Pi_{xx}^{(2)} &= +\frac{i}{2}\omega\Pi_{xy}^{(1)}; & \Pi_{xy}^{(2)} &= -\frac{i}{2}\omega\Pi_{xx}^{(1)}; & \Pi_{zz}^{(2)} &= 0; \\ \Pi_{xz}^{(2)} &= +i\omega\Pi_{yz}^{(1)} + (1-a_p)(\nabla \times \mathbf{E}_H)_y; & \Pi_{yz}^{(2)} &= -i\omega\Pi_{xz}^{(1)} - (1-a_p)(\nabla \times \mathbf{E}_H)_x, \end{aligned} \right\} \quad (5.111)$$

where for cold (and massless) electrons in the  $x$ - $z$  plane

$$(\nabla \times \mathbf{E}_H)_x = +k_{\parallel}^2 B_y; \quad (\nabla \times \mathbf{E}_H)_y = k_{\parallel} k_{\perp} B_z - k_{\parallel}^2 B_x; \quad (\nabla \times \mathbf{E}_H)_z = -k_{\parallel} k_{\perp} B_y. \quad (5.112)$$

We purposely wrote the equations with  $\mathbf{E}_H$  so that the generalization to a more elaborate electric field will be straightforward. Note that the Hall contributions (5.112) in the FLR tensor completely vanish for isotropic temperatures  $a_p = 1$ , regardless of the form of  $\mathbf{E}_H$ , i.e. regardless if the electrons are cold or not.

To write down the solution for parallel propagation ( $k_{\perp} = 0$ ) it is useful for brevity to introduce the following quantity

$$v_b \equiv \beta_{\parallel} \left(1 - \frac{a_p}{2}\right), \quad (5.113)$$

and the parallel propagating whistler and ion-cyclotron modes factorize as

$$\omega^2(1 + v_b k_{\parallel}^2) \mp \omega k_{\parallel}^2(1 + v_b(1 + k_{\parallel}^2)) - k_{\parallel}^2 \left(v_{A\parallel}^2 - \frac{\beta_{\parallel}}{2} k_{\parallel}^2\right) = 0, \quad (5.114)$$

and the two solutions are

$$\omega = \frac{1}{1 + v_b k_{\parallel}^2} \left[ \pm \frac{k_{\parallel}^2}{2} (1 + v_b(1 + k_{\parallel}^2)) + k_{\parallel} \sqrt{\left(1 + \frac{\beta_{\parallel}}{2} (a_p - 1) - \frac{\beta_{\parallel}}{2} k_{\parallel}^2\right) (1 + v_b k_{\parallel}^2) + \frac{k_{\parallel}^2}{4} (1 + v_b(1 + k_{\parallel}^2))^2} \right], \quad (5.115)$$

with two further solutions corresponding to  $-\omega$  on the left-hand side. It is perhaps useful to rewrite expression (5.114) to the ‘ion-cyclotron’ resonance form

$$\frac{k_{\parallel}^2}{\omega^2} = \frac{1 + v_b k_{\parallel}^2}{\pm \omega(1 + v_b(1 + k_{\parallel}^2)) + \left(v_{A\parallel}^2 - \frac{\beta_{\parallel}}{2} k_{\parallel}^2\right)}, \quad (5.116)$$

that clearly shows that the ion-cyclotron resonances (in the parallel direction) become more and more complicated. For strictly perpendicular propagation ( $k_{\parallel} = 0$ ), where notably the Hall contributions (5.112) vanish, the solution of the model is

$$\omega = \pm \frac{k_{\perp}}{1 + k_{\perp}^2 \frac{a_p}{8} \beta_{\parallel}} \sqrt{1 + a_p \beta_{\parallel} + k_{\perp}^2 \frac{a_p}{8} \beta_{\parallel} \left(1 + \frac{3}{2} a_p \beta_{\parallel}\right)}. \quad (5.117)$$

It is useful to compare this result with the solution obtained for FLR1 corrections, equation (5.93). For direct comparison, expanding the solution (5.117) for small  $|k_\perp| \ll 1$  yields

$$\omega = \pm k_\perp \sqrt{1 + a_p \beta_\parallel - k_\perp^2 \frac{a_p}{8} \beta_\parallel \left(1 + \frac{a_p}{2} \beta_\parallel\right)}. \quad (5.118)$$

Importantly, in comparison to the FLR1 solution (5.93), the FLR2 solution (5.118) has the opposite sign in front of the FLR term  $\sim k_\perp^2$ . The equation (5.118) is equivalent to (32) of Del Sarto *et al.* (2017), obtained by solving the pressure tensor equation without the expansion to  $\Pi^{(1)}$  and  $\Pi^{(2)}$ , and by expanding for small wavenumber afterwards. The difference between the FLR1 solution (5.93) and the FLR2 solution (5.118) is perhaps even more clear, when written in physical units for  $\omega$  and  $k$ , and by using the gyroradius  $\rho_i = v_{th\perp}/\Omega_p$ , in the following form

$$\text{FLR1:} \quad \omega^2 = k_\perp^2 \left[ V_A^2 + v_{th\perp}^2 \left(1 + \frac{k_\perp^2 \rho_i^2}{16}\right) \right]; \quad (5.119)$$

$$\text{FLR2:} \quad \omega^2 = k_\perp^2 \left[ V_A^2 \left(1 - \frac{k_\perp^2 \rho_i^2}{8}\right) + v_{th\perp}^2 \left(1 - \frac{k_\perp^2 \rho_i^2}{16}\right) \right]; \quad (5.120)$$

$$\text{Kinetic:} \quad \omega^2 = k_\perp^2 \left[ V_A^2 \left(1 - \frac{k_\perp^2 \rho_i^2}{8}\right) + v_{th\perp}^2 \left(1 - \frac{5}{16} k_\perp^2 \rho_i^2\right) \right]. \quad (5.121)$$

Equations (5.119), (5.120), (5.121) are equations (34), (32), (56) of Del Sarto *et al.* (2017). As discussed in that paper, the FLR1 model not only fails to capture the correction to the Alfvén velocity, it also introduces a correction to the thermal velocity with the incorrect sign. The kinetic result can be obtained for example from (2.10) of Mikhailovskii & Smolyakov (1985), when written for cold electrons ( $\beta_e = 0$ ), and rewritten to our notation (thermal speed in that paper does not have a factor of 2, so  $(\rho_i^2)^{(MS)} = \rho_i^2/2$ ).

Coming back to the generally oblique propagation, the second approach to writing the FLR2 tensor, i.e. when the time derivative of the Hall term is not neglected and kept in the first-order tensor  $\Pi^{(1)}$  with components

$$\left. \begin{aligned} \Pi_{xx}^{(1)} &= -\frac{a_p}{2} i k_\perp u_y; & \Pi_{xy}^{(1)} &= \frac{a_p}{2} i k_\perp u_x; & \Pi_{zz}^{(1)} &= 0; \\ \Pi_{xz}^{(1)} &= -(2 - a_p) i k_\parallel u_y + (1 - a_p) (\nabla \times E_H)_y; \\ \Pi_{yz}^{(1)} &= (2 - a_p) i k_\parallel u_x + a_p i k_\perp u_z - (1 - a_p) (\nabla \times E_H)_x, \end{aligned} \right\} \quad (5.122)$$

yields the second-order tensor

$$\left. \begin{aligned} \Pi_{xx}^{(2)} &= +\frac{i}{2} \omega \Pi_{xy}^{(1)}; & \Pi_{xy}^{(2)} &= -\frac{i}{2} \omega \Pi_{xx}^{(1)}; \\ \Pi_{xz}^{(2)} &= +i \omega \Pi_{yz}^{(1)}; & \Pi_{yz}^{(2)} &= -i \omega \Pi_{xz}^{(1)}; & \Pi_{zz}^{(2)} &= 0. \end{aligned} \right\} \quad (5.123)$$

The resulting dispersion relation for strictly parallel propagation is slightly different

$$\omega^2 (1 + v_b k_\parallel^2) \mp \omega k_\parallel^2 \left(1 + v_b + \frac{\beta_\parallel}{2} k_\parallel^2\right) - k_\parallel^2 \left(v_{A\parallel}^2 - \frac{\beta_\parallel}{2} k_\parallel^2\right) = 0, \quad (5.124)$$

with explicit solutions

$$\omega = \frac{1}{1 + v_b k_{\parallel}^2} \left[ \pm \frac{k_{\parallel}^2}{2} \left( 1 + v_b + \frac{\beta_{\parallel}}{2} k_{\parallel}^2 \right) + k_{\parallel} \sqrt{\left( 1 + \frac{\beta_{\parallel}}{2} (a_p - 1) - \frac{\beta_{\parallel}}{2} k_{\parallel}^2 \right) (1 + v_b k_{\parallel}^2) + \frac{k_{\parallel}^2}{4} \left( 1 + v_b + \frac{\beta_{\parallel}}{2} k_{\parallel}^2 \right)^2} \right]. \quad (5.125)$$

For isotropic temperatures ( $a_p = 1$ ) the result is equivalent to (5.115) since the Hall-term contributions for the FLR tensor disappear for all propagation directions. Also, for perpendicular propagation, the dispersion relation is naturally equal to (5.117).

### 5.7. FLR corrections with gyrotropic heat flux

The equation (5.99) contains the divergence of the gyrotropic heat flux,  $\nabla \cdot \mathbf{q}^g$ , which in the index notation calculates as

$$\begin{aligned} (\nabla \cdot \mathbf{q}^g)_{ij} &= \partial_k q_{kij}^g = \partial_k [q_{\parallel} \hat{b}_i \hat{b}_j \hat{b}_k + q_{\perp} (\delta_{ij} \hat{b}_k + \delta_{jk} \hat{b}_i + \delta_{ki} \hat{b}_j - 3 \hat{b}_i \hat{b}_j \hat{b}_k)] \\ &= \partial_k [(q_{\parallel} - 3q_{\perp}) \hat{b}_i \hat{b}_j \hat{b}_k + q_{\perp} (\delta_{ij} \hat{b}_k + \delta_{jk} \hat{b}_i + \delta_{ki} \hat{b}_j)] \\ &= \hat{b}_i \hat{b}_j \hat{b}_k \partial_k (q_{\parallel} - 3q_{\perp}) + (q_{\parallel} - 3q_{\perp}) (\hat{b}_j \hat{b}_k \partial_k \hat{b}_i + \hat{b}_i \hat{b}_k \partial_k \hat{b}_j + \hat{b}_i \hat{b}_j \partial_k \hat{b}_k) \\ &\quad + (\delta_{ij} \hat{b}_k \partial_k + \hat{b}_i \partial_j + \hat{b}_j \partial_i) q_{\perp} + q_{\perp} (\delta_{ij} \partial_k \hat{b}_k + \partial_j \hat{b}_i + \partial_i \hat{b}_j). \end{aligned} \quad (5.126)$$

Evaluation along magnetic field  $\hat{\mathbf{b}} = \hat{\mathbf{b}}_0 = (0, 0, 1)$  yields  $\hat{b}_k \partial_k \rightarrow \partial_z$  and the matrix is evaluated as

$$(\nabla \cdot \mathbf{q}^g)_{xx} = \partial_z q_{\perp} + q_{\perp} (\nabla \cdot \hat{\mathbf{b}} + 2 \partial_x \hat{b}_x); \quad (5.127)$$

$$(\nabla \cdot \mathbf{q}^g)_{xy} = q_{\perp} (\partial_y \hat{b}_x + \partial_x \hat{b}_y) = (\nabla \cdot \mathbf{q}^g)_{yx}; \quad (5.128)$$

$$(\nabla \cdot \mathbf{q}^g)_{xz} = (q_{\parallel} - 2q_{\perp}) \partial_z \hat{b}_x + \partial_x q_{\perp} + q_{\perp} \partial_x \hat{b}_z = (\nabla \cdot \mathbf{q}^g)_{zx}; \quad (5.129)$$

$$(\nabla \cdot \mathbf{q}^g)_{yz} = (q_{\parallel} - 2q_{\perp}) \partial_z \hat{b}_y + \partial_y q_{\perp} + q_{\perp} \partial_y \hat{b}_z = (\nabla \cdot \mathbf{q}^g)_{zy}; \quad (5.130)$$

$$(\nabla \cdot \mathbf{q}^g)_{yy} = \partial_z q_{\perp} + q_{\perp} (\nabla \cdot \hat{\mathbf{b}} + 2 \partial_y \hat{b}_y); \quad (5.131)$$

$$(\nabla \cdot \mathbf{q}^g)_{zz} = \partial_z q_{\parallel} + (q_{\parallel} - 2q_{\perp}) \nabla \cdot \hat{\mathbf{b}}, \quad (5.132)$$

where in the last ‘zz’ component we also used  $\partial_z \hat{b}_z = 0$ , but for brevity we left the divergence  $\nabla \cdot \hat{\mathbf{b}} = \partial_x \hat{b}_x + \partial_y \hat{b}_y$  intact in all expressions. Note that  $(\nabla \cdot \mathbf{q}^g)_{yy} \neq -(\nabla \cdot \mathbf{q}^g)_{xx}$ , which is not a problem as will be described below. The pressure tensor equation is here approximated as

$$\frac{\partial(\mathbf{p}^g + \mathbf{\Pi})}{\partial t} + \nabla \cdot (\mathbf{u} \mathbf{p}^g + \mathbf{q}^g) + \mathbf{p}^g \cdot \nabla \mathbf{u} + (\mathbf{p}^g \cdot \nabla \mathbf{u})^T + \Omega [\hat{\mathbf{b}} \times \mathbf{\Pi} + (\hat{\mathbf{b}} \times \mathbf{\Pi})^T] = 0, \quad (5.133)$$

and evaluated with respect to  $\hat{\mathbf{b}} = \hat{\mathbf{b}}_0 = (0, 0, 1)$ . All components were already evaluated and the calculation is straightforward, however, we need to subtract from diagonal components the scalar pressure equations that contain the gyrotropic heat flux contributions and that evaluated along  $\hat{\mathbf{b}}_0$  read

$$\frac{dp_{\parallel}}{dt} + p_{\parallel} \nabla \cdot \mathbf{u} + 2p_{\parallel} \partial_z u_z + \partial_z q_{\parallel} + (q_{\parallel} - 2q_{\perp}) \nabla \cdot \hat{\mathbf{b}} = 0; \quad (5.134)$$

$$\frac{dp_{\perp}}{dt} + 2p_{\perp} \nabla \cdot \mathbf{u} - p_{\perp} \partial_z u_z + \partial_z q_{\perp} + 2q_{\perp} \nabla \cdot \hat{\mathbf{b}} = 0. \quad (5.135)$$

The requirement again is that  $\Pi_{zz} = 0$  and  $\Pi_{yy} = -\Pi_{xx}$ . Direct calculation yields

$$\frac{\partial}{\partial t} \Pi_{xx} - 2\Omega \Pi_{xy} + p_{\perp} (\partial_x u_x - \partial_y u_y) + q_{\perp} (\partial_x \hat{b}_x - \partial_y \hat{b}_y) = 0; \quad (5.136)$$

$$\frac{\partial}{\partial t} \Pi_{xy} + 2\Omega \Pi_{xx} + p_{\perp} (\partial_x u_y + \partial_y u_x) + q_{\perp} (\partial_x \hat{b}_y + \partial_y \hat{b}_x) = 0; \quad (5.137)$$

$$\begin{aligned} \frac{\partial}{\partial t} \Pi_{xz} - \Omega \Pi_{yz} + p_{\perp} \partial_x u_z + p_{\parallel} \partial_z u_x + (p_{\parallel} - p_{\perp}) \frac{d\hat{b}_x}{dt} \\ + (q_{\parallel} - 2q_{\perp}) \partial_z \hat{b}_x + q_{\perp} \partial_x \hat{b}_z + \partial_x q_{\perp} = 0; \end{aligned} \quad (5.138)$$

$$\begin{aligned} \frac{\partial}{\partial t} \Pi_{yz} + \Omega \Pi_{xz} + p_{\perp} \partial_y u_z + p_{\parallel} \partial_z u_y + (p_{\parallel} - p_{\perp}) \frac{d\hat{b}_y}{dt} \\ + (q_{\parallel} - 2q_{\perp}) \partial_z \hat{b}_y + q_{\perp} \partial_y \hat{b}_z + \partial_y q_{\perp} = 0. \end{aligned} \quad (5.139)$$

Note that after subtraction of the perpendicular pressure equation, the 'xx' and 'yy' components are again anti-symmetric, i.e.  $(\nabla \cdot \mathbf{q}^g)_{yy} - \partial_z q_{\perp} - 2q_{\perp} \nabla \cdot \hat{\mathbf{b}} = -q_{\perp} (\partial_x \hat{b}_x - \partial_y \hat{b}_y)$ , and the system indeed satisfies that  $\Pi_{yy} = -\Pi_{xx}$ . Also, considering the 'zz' component, after subtraction of the parallel pressure equation all the terms cancel, i.e.  $(\nabla \cdot \mathbf{q}^g)_{zz} - \partial_z q_{\parallel} - (q_{\parallel} - 2q_{\perp}) \nabla \cdot \hat{\mathbf{b}} = 0$  and the system indeed satisfies  $\Pi_{zz} = 0$ .

Finally, expressing the above system in the linear approximation, with the assumption that mean heat flux values are zero, i.e.  $q_{\parallel}^{(0)} = 0$ ,  $q_{\perp}^{(0)} = 0$ , the only gyroscopic heat flux contributions that remain are  $\partial_x q_{\perp}$  and  $\partial_y q_{\perp}$ , in equations (5.138), (5.139). The FLR tensor is again separated to  $\Pi^{(1)} + \Pi^{(2)}$  and similarly to the Hall-term contributions, it is again a matter of choice if the heat flux contributions are pushed to  $\Pi^{(2)}$  or kept in  $\Pi^{(1)}$ . We prefer the first choice, i.e. when  $\Pi^{(1)}$  is equivalent to (5.74), which yields the second-order tensor

$$\left. \begin{aligned} \Pi_{xx}^{(2)} &= -\frac{1}{2\Omega} \frac{\partial}{\partial t} \Pi_{xy}^{(1)}; & \Pi_{xy}^{(2)} &= \frac{1}{2\Omega} \frac{\partial}{\partial t} \Pi_{xx}^{(1)}; \\ \Pi_{xz}^{(2)} &= -\frac{1}{\Omega} \frac{\partial}{\partial t} \Pi_{yz}^{(1)} + \frac{1}{\Omega} (p_{\parallel}^{(0)} - p_{\perp}^{(0)}) \frac{c}{B_0} (\nabla \times \mathbf{E}_H)_y - \frac{1}{\Omega} \partial_y q_{\perp}; \\ \Pi_{yz}^{(2)} &= \frac{1}{\Omega} \frac{\partial}{\partial t} \Pi_{xz}^{(1)} - \frac{1}{\Omega} (p_{\parallel}^{(0)} - p_{\perp}^{(0)}) \frac{c}{B_0} (\nabla \times \mathbf{E}_H)_x + \frac{1}{\Omega} \partial_x q_{\perp}. \end{aligned} \right\} \quad (5.140)$$

Note that we could have derived the contributions  $\partial_x q_{\perp}$  and  $\partial_y q_{\perp}$  in a much quicker way, if the expression (5.126) was linearized from the beginning. However, we wanted to demonstrate that if nonlinear FLR corrections are considered, the heat flux nonlinearities can significantly complicate the dynamics, even if simplified evaluation along  $\hat{\mathbf{b}}_0$  is performed. Right now, we cannot use these FLR corrections to obtain a dispersion relation, since we have no means to determine a closure for the gyroscopic heat flux  $q_{\perp}$ . We would have to consider CGL2 fluid model or more complicated Landau fluid models, that contains evolution equations for  $q_{\parallel}$  and  $q_{\perp}$ . Additionally, the perpendicular propagating fast mode, as well as the parallel propagating ion-cyclotron and whistler modes, are not influenced by the gyroscopic heat flux. Instead, as the last step, we will consider contributions of the non-gyroscopic heat flux.

### 5.8. FLR corrections with non-gyrotropic heat flux vectors (FLR3)

It is possible to further increase precision of the FLR corrections by considering non-gyrotropic heat flux contributions. Detailed calculations with the non-gyrotropic heat flux vectors  $\mathbf{S}_\perp^\parallel$  and  $\mathbf{S}_\perp^\perp$  are presented in appendix D. At the linear level, these non-gyrotropic heat flux vectors contribute to the linearized pressure equations

$$\frac{\partial p_\parallel}{\partial t} + p_\parallel^{(0)} \nabla \cdot \mathbf{u} + 2p_\parallel^{(0)} \partial_z u_z + \partial_z q_\parallel + \partial_x (S_\perp^\parallel)_x + \partial_y (S_\perp^\parallel)_y = 0; \quad (5.141)$$

$$\frac{\partial p_\perp}{\partial t} + 2p_\perp^{(0)} \nabla \cdot \mathbf{u} - p_\perp^{(0)} \partial_z u_z + \partial_z q_\perp + \partial_x (S_\perp^\perp)_x + \partial_y (S_\perp^\perp)_y = 0, \quad (5.142)$$

and also to the linearized equations for the FLR tensor

$$\frac{\partial}{\partial t} \Pi_{xx} - 2\Omega \Pi_{xy} + p_\perp^{(0)} (\partial_x u_x - \partial_y u_y) + \frac{1}{2} [\partial_x (S_\perp^\perp)_x - \partial_y (S_\perp^\perp)_y] = 0; \quad (5.143)$$

$$\frac{\partial}{\partial t} \Pi_{xy} + 2\Omega \Pi_{xx} + p_\perp^{(0)} (\partial_x u_y + \partial_y u_x) + \frac{1}{2} [\partial_y (S_\perp^\perp)_x + \partial_x (S_\perp^\perp)_y] = 0; \quad (5.144)$$

$$\frac{\partial}{\partial t} \Pi_{xz} - \Omega \Pi_{yz} + p_\perp^{(0)} \partial_x u_z + p_\parallel^{(0)} \partial_z u_x + (p_\parallel^{(0)} - p_\perp^{(0)}) \frac{\partial \hat{b}_x}{\partial t} + \partial_x q_\perp + \partial_z (S_\perp^\parallel)_x = 0; \quad (5.145)$$

$$\frac{\partial}{\partial t} \Pi_{yz} + \Omega \Pi_{xz} + p_\perp^{(0)} \partial_y u_z + p_\parallel^{(0)} \partial_z u_y + (p_\parallel^{(0)} - p_\perp^{(0)}) \frac{\partial \hat{b}_y}{\partial t} + \partial_y q_\perp + \partial_z (S_\perp^\parallel)_y = 0. \quad (5.146)$$

These equations are equivalent to (18)–(21) of Goswami *et al.* (2005) (by noting that the vector  $\mathbf{S}^\perp = q_\perp \hat{\mathbf{b}} + \mathbf{S}_\perp^\perp$ , so at the linear level  $S_z^\perp = q_\perp$ ). The non-gyrotropic heat flux vectors are separated to the first and second order

$$\mathbf{S}_\perp^\parallel = \mathbf{S}_\perp^{\parallel(1)} + \mathbf{S}_\perp^{\parallel(2)}; \quad \mathbf{S}_\perp^\perp = \mathbf{S}_\perp^{\perp(1)} + \mathbf{S}_\perp^{\perp(2)}. \quad (5.147)$$

By using the nonlinear results for the first-order vectors derived in appendix D, (D 37), (D 60), yields that, at the linear level,

$$\mathbf{S}_\perp^{\parallel(1)} = \frac{1}{\Omega} \hat{\mathbf{b}}_0 \times \left[ p_\perp^{(0)} \nabla \left( \frac{p_\parallel}{\rho} \right) + 2 \frac{p_\parallel^{(0)}}{\rho_0} (p_\parallel^{(0)} - p_\perp^{(0)}) \partial_z \hat{\mathbf{b}} \right]; \quad (5.148)$$

$$\mathbf{S}_\perp^{\perp(1)} = \frac{1}{\Omega} \hat{\mathbf{b}}_0 \times \left[ 2p_\perp^{(0)} \nabla \left( \frac{p_\perp}{\rho} \right) \right], \quad (5.149)$$

where the gradients are meant to be further linearized. Nevertheless, the above form is useful to point out that the non-gyrotropic heat fluxes are proportional to the gradients of the temperature. For clarity, partially linearized expressions are also written down by components in appendix D, see (D 24), (D 25) and (D 61), (D 62). It is important to emphasize that the expressions (5.148), (5.149) were derived for a bi-Maxwellian distribution function. As discussed later in the text, this is achieved by prescribing closures for the gyrotropic fourth-order moment in the form  $r_{\parallel\parallel} = 3p_\parallel^2/\rho + \tilde{r}_{\parallel\parallel}$ ,  $r_{\parallel\perp} = p_\parallel p_\perp/\rho + \tilde{r}_{\parallel\perp}$ , and  $r_{\perp\perp} = 2p_\perp^2/\rho + \tilde{r}_{\perp\perp}$ . In (5.148), (5.149), we additionally neglected the perturbations  $\tilde{r}$  (since right now we do not want to consider models where these perturbations are evaluated from linear kinetic theory). Therefore, the FLR corrections with the non-gyrotropic heat flux, that we call here FLR3, are distribution function specific. Nevertheless, as discussed later in the text, very similar closures can be also

obtained for the bi-kappa distribution function, just with additional coefficients  $\alpha_\kappa = (\kappa - 3/2)/(\kappa - 5/2)$ , and expressions similar to (5.148), (5.149) can be derived.

The second-order vectors are also derived in appendix D, this time directly in the linear approximation, equations (D 88), (D 89) and the vectors read

$$(\mathbf{S}_\perp^\parallel)^{(2)} = \frac{1}{\Omega} \hat{\mathbf{b}}_0 \times \left[ \frac{\partial}{\partial t} (\mathbf{S}_\perp^\parallel)^{(1)} + 2 \frac{p_\parallel^{(0)}}{\rho_0} \partial_z \Pi_{kz}^{(1)} \right]; \quad (5.150)$$

$$(\mathbf{S}_\perp^\perp)^{(2)} = \frac{1}{\Omega} \hat{\mathbf{b}}_0 \times \left[ \frac{\partial}{\partial t} (\mathbf{S}_\perp^\perp)^{(1)} + \frac{p_\perp^{(0)}}{\rho_0} \left( \partial_x \Pi_{xk}^{(1)} + \partial_y \Pi_{yk}^{(1)} \right) \right]. \quad (5.151)$$

As a reminder,  $\Pi_{ij} = \Pi_{ji}$ . For clarity, we wrote the expressions above explicitly in the index notation. The non-gyrotropic heat flux vectors are perpendicular to  $\hat{\mathbf{b}}_0 = (0, 0, 1)$ , and the non-zero components are for index  $k = x, y$ . The expressions can be written in a more elegant vector form, by defining vector  $\vec{\Pi}_z \equiv (\Pi_{xz}, \Pi_{yz}, \Pi_{zz} = 0)$ , vector  $\nabla_\perp = (\partial_x, \partial_y, 0)$  and matrix

$$\boldsymbol{\Pi}_\perp \equiv \begin{pmatrix} \Pi_{xx} & \Pi_{xy} & 0 \\ \Pi_{yx} & \Pi_{yy} & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad (5.152)$$

so that  $\nabla_\perp \cdot \boldsymbol{\Pi}_\perp = (\partial_x \Pi_{xx} + \partial_y \Pi_{yx}, \partial_x \Pi_{xy} + \partial_y \Pi_{yy})$ , yielding

$$\mathbf{S}_\perp^{\parallel(2)} = \frac{1}{\Omega} \hat{\mathbf{b}}_0 \times \left[ \frac{\partial}{\partial t} \mathbf{S}_\perp^{\parallel(1)} + 2 \frac{p_\parallel^{(0)}}{\rho_0} \partial_z \vec{\Pi}_z^{(1)} \right]; \quad (5.153)$$

$$\mathbf{S}_\perp^{\perp(2)} = \frac{1}{\Omega} \hat{\mathbf{b}}_0 \times \left[ \frac{\partial}{\partial t} \mathbf{S}_\perp^{\perp(1)} + \frac{p_\perp^{(0)}}{\rho_0} \nabla_\perp \cdot \boldsymbol{\Pi}_\perp^{(1)} \right], \quad (5.154)$$

which are equations (53) and (54) of Passot *et al.* (2012).

Normalizing the equations (dropping the tilde), Fourier transforming, and writing them in the  $x$ - $z$  plane (with  $\partial_y = 0$ ) yields

$$-i\omega \Pi_{xx} - 2\Pi_{xy} + a_p i k_\perp u_x + \frac{1}{2} i k_\perp (S_\perp^\perp)_x = 0; \quad (5.155)$$

$$-i\omega \Pi_{xy} + 2\Pi_{xx} + a_p i k_\perp u_y + \frac{1}{2} i k_\perp (S_\perp^\perp)_y = 0; \quad (5.156)$$

$$\begin{aligned} -i\omega \Pi_{xz} - \Pi_{yz} + a_p i k_\perp u_z + i k_\parallel u_x + (1 - a_p)(i k_\parallel u_x - (\nabla \times \mathbf{E}_H)_x) \\ + i k_\perp q_\perp + i k_\parallel (S_\perp^\parallel)_x = 0; \end{aligned} \quad (5.157)$$

$$-i\omega \Pi_{yz} + \Pi_{xz} + i k_\parallel u_y + (1 - a_p)(i k_\parallel u_y - (\nabla \times \mathbf{E}_H)_y) + i k_\parallel (S_\perp^\parallel)_y = 0, \quad (5.158)$$

where the terms with Hall electric field components are specified in (5.112). At this moment we do not consider higher-order fluid models with gyrotropic heat fluxes such as CGL2 or Landau fluids (see later in the text) and to have a closed model, here we prescribe  $q_\perp = 0$ ,  $q_\parallel = 0$ . It is useful to briefly explore what fluid models are obtained, if we decide to keep only the first-order vectors  $\mathbf{S}_\perp^{\parallel(1)}$ ,  $\mathbf{S}_\perp^{\perp(1)}$ , and ignore the second-order corrections  $\mathbf{S}_\perp^{\parallel(2)}$ ,  $\mathbf{S}_\perp^{\perp(2)}$ . By considering only the  $\mathbf{S}_\perp^{\parallel(1)}$ ,  $\mathbf{S}_\perp^{\perp(1)}$  contributions, we can

either put them into  $\Pi^{(1)}$  or  $\Pi^{(2)}$ . However, it can be shown that both options yield dispersion relation for the perpendicular fast mode (long-wavelength, in physical units)

$$\omega^2 = k_{\perp}^2 \left[ V_A^2 \left( 1 - \frac{k_{\perp}^2 \rho_i^2}{8} \right) + v_{\text{th}\perp}^2 \left( 1 + \frac{k_{\perp}^2 \rho_i^2}{16} \right) \right]. \quad (5.159)$$

Rather amusingly, in comparison to the FLR2 result (5.120), the wrong sign in the correction of the thermal speed is back!

### 5.9. Hall-CGL-FLR3 fluid model

Obviously, if matching with kinetic theory for the fast perpendicular mode is of upmost importance (for small wavenumbers), we have no other choice and we have to keep the second-order non-gyrotropic heat flux contributions  $\mathbf{S}_{\perp}^{\parallel(2)}$ ,  $\mathbf{S}_{\perp}^{\perp(2)}$ . We call this FLR pressure tensor as FLR3, and as stated previously, the FLR3 pressure tensor is distribution function specific. This model is very important to us and especially for the following discussion of the firehose instability. In order to be completely clear on which equations we are solving and also so that our results can be easily reproduced, let us state the entire model in all of its beauty. The linearized, normalized, Fourier transformed equations written in the  $x$ - $z$  plane read

$$\left. \begin{aligned} -\omega\rho + k_{\perp}u_x + k_{\parallel}u_z &= 0; \\ -\omega u_x + \frac{\beta_{\parallel}}{2}k_{\perp}p_{\perp} - v_{A\parallel}^2 k_{\parallel}B_x + k_{\perp}B_z + \frac{\beta_{\parallel}}{2}(k_{\perp}\Pi_{xx} + k_{\parallel}\Pi_{xz}) &= 0; \\ -\omega u_y - v_{A\parallel}^2 k_{\parallel}B_y + \frac{\beta_{\parallel}}{2}(k_{\perp}\Pi_{xy} + k_{\parallel}\Pi_{yz}) &= 0; \\ -\omega u_z + \frac{\beta_{\parallel}}{2}k_{\parallel}p_{\parallel} + \frac{\beta_{\parallel}}{2}(1 - a_p)k_{\perp}B_x + \frac{\beta_{\parallel}}{2}k_{\perp}\Pi_{xz} &= 0; \\ -\omega B_x - k_{\parallel}u_x - ik_{\parallel}^2 B_y &= 0; \\ -\omega B_y - k_{\parallel}u_y - ik_{\parallel}k_{\perp}B_z + ik_{\parallel}^2 B_x &= 0; \\ -\omega B_z + k_{\perp}u_x + ik_{\parallel}k_{\perp}B_y &= 0; \\ -\omega p_{\parallel} + k_{\perp}u_x + 3k_{\parallel}u_z + k_{\parallel}\mathcal{Q}_{\parallel} + k_{\perp}(\mathbf{S}_{\perp}^{\parallel})_x &= 0; \\ -\omega p_{\perp} + 2a_p k_{\perp}u_x + a_p k_{\parallel}u_z + k_{\parallel}\mathcal{Q}_{\perp} + k_{\perp}(\mathbf{S}_{\perp}^{\perp})_x &= 0, \end{aligned} \right\} \quad (5.160)$$

where each component of the (FLR) pressure tensor  $\Pi$ , and each component of the (FLR) heat flux vectors  $\mathbf{S}_{\perp}^{\parallel}$ ,  $\mathbf{S}_{\perp}^{\perp}$ , is separated to

$$\Pi = \Pi^{(1)} + \Pi^{(2)}; \quad \mathbf{S}_{\perp}^{\parallel} = \mathbf{S}_{\perp}^{\parallel(1)} + \mathbf{S}_{\perp}^{\parallel(2)}; \quad \mathbf{S}_{\perp}^{\perp} = \mathbf{S}_{\perp}^{\perp(1)} + \mathbf{S}_{\perp}^{\perp(2)}. \quad (5.161)$$

The components of the  $\Pi$  tensor are given by

$$\left. \begin{aligned} \Pi_{xx}^{(1)} &= -\frac{a_p}{2}ik_{\perp}u_y; & \Pi_{xx}^{(2)} &= +i\frac{\omega}{2}\Pi_{xy}^{(1)} - i\frac{k_{\perp}}{4}(\mathbf{S}_{\perp}^{\perp})_y^{(1)+(2)}; \\ \Pi_{xy}^{(1)} &= \frac{a_p}{2}ik_{\perp}u_x; & \Pi_{xy}^{(2)} &= -i\frac{\omega}{2}\Pi_{xx}^{(1)} + i\frac{k_{\perp}}{4}(\mathbf{S}_{\perp}^{\perp})_x^{(2)}; \\ \Pi_{xz}^{(1)} &= -(2 - a_p)ik_{\parallel}u_y; & \Pi_{xz}^{(2)} &= +i\omega\Pi_{yz}^{(1)} + (1 - a_p)(\nabla \times \mathbf{E}_H)_y - ik_{\parallel}(\mathbf{S}_{\perp}^{\parallel})_y^{(1)+(2)}; \\ & & \Pi_{yz}^{(1)} &= (2 - a_p)ik_{\parallel}u_x + a_p ik_{\perp}u_z; \\ \Pi_{yz}^{(2)} &= -i\omega\Pi_{xz}^{(1)} - (1 - a_p)(\nabla \times \mathbf{E}_H)_x + ik_{\perp}\mathcal{Q}_{\perp} + ik_{\parallel}(\mathbf{S}_{\perp}^{\parallel})_x^{(1)+(2)}, \end{aligned} \right\} \quad (5.162)$$



where the Hall electric field contributions for cold massless electrons read

$$(\nabla \times \mathbf{E}_H)_x = +k_{\parallel}^2 B_y; \quad (\nabla \times \mathbf{E}_H)_y = k_{\parallel} k_{\perp} B_z - k_{\parallel}^2 B_x, \quad (5.163)$$

and the components of the non-gyrotropic (FLR) heat fluxes are given by

$$\left. \begin{aligned} (S_{\perp}^{\parallel})_x^{(1)} &= -\beta_{\parallel}(1 - a_p)ik_{\parallel}B_y; & (S_{\perp}^{\parallel})_x^{(2)} &= +i\omega(S_{\perp}^{\parallel})_y^{(1)} - \beta_{\parallel}ik_{\parallel}\Pi_{yz}^{(1)}; \\ (S_{\perp}^{\parallel})_y^{(1)} &= a_p\frac{\beta_{\parallel}}{2}ik_{\perp}(p_{\parallel} - \rho) + \beta_{\parallel}(1 - a_p)ik_{\parallel}B_x; & (S_{\perp}^{\parallel})_y^{(2)} &= -i\omega(S_{\perp}^{\parallel})_x^{(1)} + \beta_{\parallel}ik_{\parallel}\Pi_{xz}^{(1)}; \\ (S_{\perp}^{\perp})_x^{(1)} &= 0; & (S_{\perp}^{\perp})_x^{(2)} &= +i\omega(S_{\perp}^{\perp})_y^{(1)} - a_p\frac{\beta_{\parallel}}{2}ik_{\perp}\Pi_{xy}^{(1)}; \\ (S_{\perp}^{\perp})_y^{(1)} &= \beta_{\parallel}a_pik_{\perp}(p_{\perp} - a_p\rho); & (S_{\perp}^{\perp})_y^{(2)} &= -i\omega(S_{\perp}^{\perp})_x^{(1)} + a_p\frac{\beta_{\parallel}}{2}ik_{\perp}\Pi_{xx}^{(1)}. \end{aligned} \right\} \quad (5.164)$$

Importantly, the  $\Pi^{(2)}$  expressions contain both the first- and second-order contributions from the heat flux vectors  $\mathbf{S}_{\perp}^{\parallel}$ ,  $\mathbf{S}_{\perp}^{\perp}$ . To explicitly see where the closure for the gyrotropic heat fluxes  $q_{\parallel}=0$ ,  $q_{\perp}=0$  was performed, we scratched the terms involving these quantities. Additionally, we also scratched contributions from  $(S_{\perp}^{\perp})_x^{(1)}$ , which are here zero at the linear level.

Let us check the solution of the Hall-CGL-FLR3 fluid model for the perpendicular propagation ( $k_{\parallel}=0$ ), which in this case reads

$$\omega^2 = \frac{k_{\perp}^2}{\left(1 + k_{\perp}^2 \frac{a_p}{8} \beta_{\parallel}\right) (1 + a_p \beta_{\parallel} k_{\perp}^2)} \left[ 1 + a_p \beta_{\parallel} + k_{\perp}^2 a_p \beta_{\parallel} \left( 1 + \frac{13}{16} a_p \beta_{\parallel} + \frac{9}{128} a_p^2 \beta_{\parallel}^2 k_{\perp}^2 \right) \right]. \quad (5.165)$$

As a ‘sanity check’, since the solution stays always positive for all the wavenumbers and we do not have any instability, the model appears to be good. For small wavenumbers, the expansion yields

$$\omega^2 = k_{\perp}^2 \left[ 1 + a_p \beta_{\parallel} - k_{\perp}^2 a_p \beta_{\parallel} \left( \frac{1}{8} + \frac{5}{16} a_p \beta_{\parallel} \right) + \dots \right], \quad (5.166)$$

and in physical units

$$\text{FLR3: } \omega^2 = k_{\perp}^2 \left[ V_A^2 \left( 1 - \frac{k_{\perp}^2 \rho_i^2}{8} \right) + v_{th\perp}^2 \left( 1 - \frac{5}{16} k_{\perp}^2 \rho_i^2 \right) \right]. \quad (5.167)$$

Therefore, the Hall-CGL-FLR3 fluid model finally matches the analytic result from kinetic theory! We note that if the  $\Pi$  contributions are neglected in the second-order heat flux expressions (5.164), instead of the correct  $-5/16$  FLR correction to the thermal speed, one obtains  $-7/16$ . The necessity to keep the second-order heat flux contributions to recover the kinetic result for the perpendicular fast mode was discovered by Mikhailovskii & Smolyakov (1985).

Checking the solution for the parallel propagating ( $k_{\perp}=0$ ) ion-cyclotron and whistler modes yields the following dispersion relation

$$\begin{aligned} \omega^2(1 + v_b k_{\parallel}^2) \mp \omega k_{\parallel}^2 \left[ 1 + v_b(1 + k_{\parallel}^2) + k_{\parallel}^2 \beta_{\parallel}^2 \left( \frac{3}{2} - a_p \right) \right] \\ - k_{\parallel}^2 \left[ v_{A\parallel}^2 - \frac{\beta_{\parallel}}{2} k_{\parallel}^2 (1 + \beta_{\parallel}(1 - a_p)) + \frac{\beta_{\parallel}^2}{2} (a_p - 2) k_{\parallel}^4 \right] = 0, \end{aligned} \quad (5.168)$$

where the quantities  $v_b = \beta_{\parallel}(1 - a_p/2)$  and  $v_{A\parallel}^2 = 1 + (\beta_{\parallel}/2)(a_p - 1)$ . We will use this dispersion relation to study the firehose instability. For a general oblique propagation direction, the dispersion relation is obviously way too large to write down, and we recommend using analytic software such as Maple or Mathematica.

### Moving the Hall term to $\Pi^{(1)}$

For completeness, just in case we want to investigate later the influence of the Hall term, moving it to  $\Pi^{(1)}$  yields the following dispersion relation for the parallel propagation

$$\begin{aligned} \omega^2(1 + v_b k_{\parallel}^2) \mp \omega k_{\parallel}^2 \left[ 1 + v_b + k_{\parallel}^2 \frac{\beta_{\parallel}}{2} + k_{\parallel}^2 \beta_{\parallel}^2 \left( \frac{3}{2} - a_p \right) \right] \\ - k_{\parallel}^2 \left[ v_{A\parallel}^2 - \frac{\beta_{\parallel}}{2} k_{\parallel}^2 (1 + \beta_{\parallel}(1 - a_p)) - \frac{\beta_{\parallel}^2}{2} k_{\parallel}^4 \right] = 0. \end{aligned} \quad (5.169)$$

## 6. Parallel and oblique firehose instability

### 6.1. Parallel propagation

For clarity, it is useful to summarize all 3 major models that describe the parallel propagating ion-cyclotron and whistler modes. By prescribing  $k_{\perp} = 0$  in the equations of the Hall-CGL-FLR3 fluid model, the model greatly simplifies. The parallel magnetic field  $B_z = 0$ , the ion-acoustic mode decouples, and both the first- and second-order contributions to  $\Pi_{xx} = 0$ ,  $\Pi_{xy} = 0$ ,  $S_{\perp}^{\perp} = 0$ . The normalized, Fourier transformed equations written in the  $x$ - $z$  plane read

$$\left. \begin{aligned} -\omega u_x - v_{A\parallel}^2 k_{\parallel} B_x + \frac{\beta_{\parallel}}{2} k_{\parallel} \Pi_{xz} &= 0; \\ -\omega u_y - v_{A\parallel}^2 k_{\parallel} B_y + \frac{\beta_{\parallel}}{2} k_{\parallel} \Pi_{yz} &= 0; \\ -\omega B_x - k_{\parallel} u_x - i k_{\parallel}^2 B_y &= 0; \\ -\omega B_y - k_{\parallel} u_y + i k_{\parallel}^2 B_x &= 0, \end{aligned} \right\} \quad (6.1)$$

where the components of the non-gyrotropic (FLR) pressure tensor  $\Pi$  are given by

$$\left. \begin{aligned} \Pi_{xz}^{(1)} &= -(2 - a_p) i k_{\parallel} u_y; & \Pi_{xz}^{(2)} &= +i \omega \Pi_{yz}^{(1)} - (1 - a_p) k_{\parallel}^2 B_x - i k_{\parallel} (S_{\perp}^{\parallel})_y^{(1)+(2)}; \\ \Pi_{yz}^{(1)} &= +(2 - a_p) i k_{\parallel} u_x; & \Pi_{yz}^{(2)} &= -i \omega \Pi_{xz}^{(1)} - (1 - a_p) k_{\parallel}^2 B_y + i k_{\parallel} (S_{\perp}^{\parallel})_x^{(1)+(2)}, \end{aligned} \right\} \quad (6.2)$$

and the components of the non-gyrotropic (FLR) heat flux  $S_{\perp}^{\parallel}$  read

$$\left. \begin{aligned} (S_{\perp}^{\parallel})_x^{(1)} &= -\beta_{\parallel}(1 - a_p) i k_{\parallel} B_y; & (S_{\perp}^{\parallel})_x^{(2)} &= +i \omega (S_{\perp}^{\parallel})_y^{(1)} - \beta_{\parallel} i k_{\parallel} \Pi_{yz}^{(1)}; \\ (S_{\perp}^{\parallel})_y^{(1)} &= +\beta_{\parallel}(1 - a_p) i k_{\parallel} B_x; & (S_{\perp}^{\parallel})_y^{(2)} &= -i \omega (S_{\perp}^{\parallel})_x^{(1)} + \beta_{\parallel} i k_{\parallel} \Pi_{xz}^{(1)}. \end{aligned} \right\} \quad (6.3)$$

When the entire  $\Pi$  is neglected, yields the Hall-CGL solution

$$\omega = \pm \frac{k_{\parallel}^2}{2} + k_{\parallel} \sqrt{v_{A\parallel}^2 + \frac{k_{\parallel}^2}{4}}. \quad (6.4)$$

Neglecting the  $\Pi^{(2)}$  contributions yields the Hall-CGL-FLR1 solution

$$\omega = \pm \frac{k_{\parallel}^2}{2}(1 + v_b) + k_{\parallel} \sqrt{v_{A\parallel}^2 + \frac{k_{\parallel}^2}{4}(1 - v_b)^2}, \quad (6.5)$$

and neglecting the heat flux  $S_{\perp}^{\parallel}$  yields the Hall-CGL-FLR2 solution

$$\omega = \frac{1}{1 + v_b k_{\parallel}^2} \left[ \pm \frac{k_{\parallel}^2}{2}(1 + v_b(1 + k_{\parallel}^2)) + k_{\parallel} \sqrt{\left(v_{A\parallel}^2 - \frac{\beta_{\parallel}}{2} k_{\parallel}^2\right)(1 + v_b k_{\parallel}^2) + \frac{k_{\parallel}^2}{4}(1 + v_b(1 + k_{\parallel}^2))^2} \right]. \quad (6.6)$$

Finally, the full model yields Hall-CGL-FLR3 solution

$$\omega = \frac{1}{1 + v_b k_{\parallel}^2} \left\{ \pm \frac{k_{\parallel}^2}{2} \left[ 1 + v_b(1 + k_{\parallel}^2) + k_{\parallel}^2 \beta_{\parallel}^2 \left( \frac{3}{2} - a_p \right) \right] + k_{\parallel} \sqrt{\left[ v_{A\parallel}^2 - \frac{\beta_{\parallel}}{2} k_{\parallel}^2(1 + \beta_{\parallel}(1 - a_p)) + \frac{\beta_{\parallel}^2}{2}(a_p - 2)k_{\parallel}^4 \right] (1 + v_b k_{\parallel}^2) + \frac{k_{\parallel}^2}{4} \left[ 1 + v_b(1 + k_{\parallel}^2) + k_{\parallel}^2 \beta_{\parallel}^2 \left( \frac{3}{2} - a_p \right) \right]^2} \right\}. \quad (6.7)$$

The quantities  $v_b$  and  $v_{A\parallel}^2$  used in the solutions are defined as

$$v_b = \beta_{\parallel} \left( 1 - \frac{a_p}{2} \right); \quad v_{A\parallel}^2 = 1 + \frac{\beta_{\parallel}}{2}(a_p - 1). \quad (6.8)$$

Importantly, the solutions are written here for  $\omega$  (and not  $\omega^2$ ), and all the models of course yield 4 solutions, the other two solutions are obtained by substituting  $\omega$  with  $-\omega$ . For a reader who just jumped to this section, and is confused on how the solutions were split, see §4.2 where the Hall-CGL model is discussed with final equation (4.36). The solutions can be written in various forms, and one possibility is to keep  $|k_{\parallel}|$  in front of the square roots.

The (parallel and oblique) firehose instability was studied in detail by Hunana & Zank (2017), who focused on the Hall-CGL and Hall-CGL-FLR1 models. One of the conclusions reached in that paper was that the main reason for the relatively large discrepancy between the usual fluid models and the kinetic description is the appearance of a huge ‘bump’ in the imaginary phase speed, when close to the firehose threshold. The situation is demonstrated in figure 7, where kinetic solutions obtained by the WHAMP code (blue solid lines) are compared to solutions of the Hall-CGL-FLR2 model (*a*, blue dashed lines) and the Hall-CGL-FLR3 fluid model (*b*, black dashed lines). The plasma  $\beta_{\parallel} = 4$ , so the long-wavelength firehose threshold, that we call the ‘hard threshold’, is at  $a_p = 0.5$ . The temperature anisotropy is varied from  $a_p = 0$  to  $a_p = 0.501$ . For solutions with the Hall-CGL and Hall-CGL-FLR1 models see figures 3 and 4 of Hunana & Zank (2017). In the WHAMP code the value of  $a_p = 0$  cannot be used, and  $a_p = 10^{-4}$  was chosen instead. Also, since here we concentrate on proton dynamics, the influence of electrons in the WHAMP code was eliminated by making the electrons cold (with  $T_e/T_p = 10^{-8}$ ).

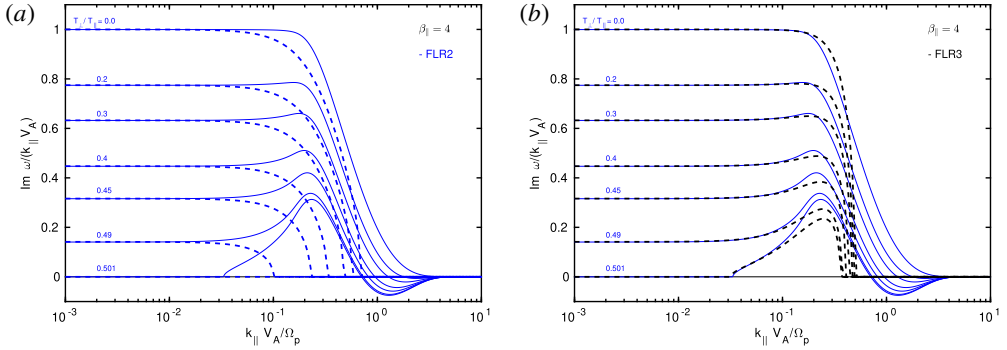


FIGURE 7. Imaginary phase speed (growth rate normalized to the wavenumber) of the parallel firehose instability. The value of  $\beta_{\parallel} = 4$ , and the temperature anisotropy is varied so that the whistler mode is in the firehose-unstable regime. Solid lines (blue) are kinetic solutions, obtained with the WHAMP code. Dashed lines are fluid solutions. (a) FLR2 solutions (blue), (b) FLR3 solutions (black). It is shown that in contrast to the FLR2 model, the FLR3 model reproduces the large ‘bump’ when close to the long-wavelength ‘hard’ firehose threshold  $T_{\perp}/T_{\parallel} = 0.5$ .

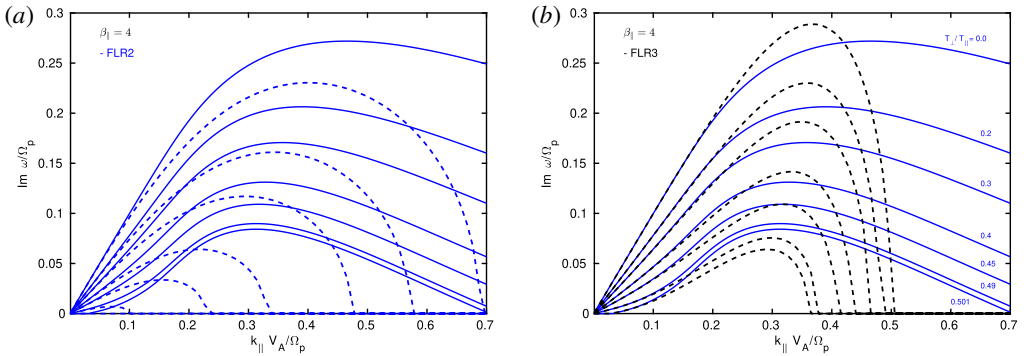


FIGURE 8. Same parameters as in figure 7, but the growth rate is plotted, and a linear scale is used for the  $x$ -axis. Notice the excellent precision of the FLR3 model for small wavenumbers up to  $kd_i = 0.1$ – $0.2$ . In comparison to kinetic theory, the fluid solutions are stabilized much more ‘rapidly’. Nevertheless, the value of the maximum growth rate, and the wavenumber where the maximum growth rate is achieved, is surprisingly close to kinetic theory. This is an excellent result for a fluid model, which does not contain collisionless ion-cyclotron damping.

Very surprisingly, figure 7 shows that the large ‘bump’ is reproduced by the FLR3 model, and the precision is quite good! The situation is further analysed in figure 8, where instead of the imaginary phase speed, the growth rate is plotted. The same conclusion is obtained. For the temperature anisotropy  $a_p = 0.501$ , all simpler fluid models (Hall-CGL, FLR1, FLR2) are fully stable for all ranges of wavenumbers. In contrast, the FLR3 model still develops a strong firehose instability. Therefore, it is the non-gyrotropic heat flux  $S_{\perp}^{\parallel}$  of the FLR3 model that is responsible for the ‘bump’ (for the case of strictly parallel propagation). The appearance of the ‘bump’ can be understood analytically, by evaluating the fluid dispersion relations exactly at the ‘hard’ firehose threshold  $a_p = 1 - 2/\beta_{\parallel}$ , where the quantities  $v_{A\parallel}^2 = 0$  and  $v_b = 1 + \beta_{\parallel}/2$ . At

the hard firehose threshold, the expression under the square root of the FLR2 solution (6.6) can be factorized as

$$k_{\parallel}^2 \left[ 1 - \frac{\beta_{\parallel}}{4} + \frac{k_{\parallel}^2}{2} \left( 1 + \frac{\beta_{\parallel}}{2} \right) \right]^2 + k_{\parallel}^2 \frac{\beta_{\parallel}}{2}, \quad (6.9)$$

implying that the FLR2 model is always stable for all values of  $k_{\parallel}$  and  $\beta_{\parallel}$ . Such a factorization cannot be achieved for the FLR3 model, and the solution (6.7) still becomes unstable at some range of wavenumbers, where

$$\begin{aligned} & \left[ v_{A\parallel}^2 - \frac{\beta_{\parallel}}{2} k_{\parallel}^2 (1 + \beta_{\parallel} (1 - a_p)) + \frac{\beta_{\parallel}^2}{2} (a_p - 2) k_{\parallel}^4 \right] (1 + v_b k_{\parallel}^2) \\ & + \frac{k_{\parallel}^2}{4} \left[ 1 + v_b (1 + k_{\parallel}^2) + k_{\parallel}^2 \beta_{\parallel}^2 \left( \frac{3}{2} - a_p \right) \right]^2 < 0, \end{aligned} \quad (6.10)$$

which is the (strictly parallel) firehose instability criterion of the FLR3 model. We note that the model can become unstable also for the temperature anisotropy  $a_p > 1$ . This can be perhaps considered as some remnant of the ion-cyclotron anisotropy instability, but we did not study the situation further since the instability should not be reproduced correctly. The parallel firehose instability for high plasma beta value  $\beta_{\parallel} = 100$  is shown in figure 9. The left figure is from Hunana & Zank (2017), and it shows solutions of the Hall-CGL model (blue dashed lines) and of the Hall-CGL-FLR1 model (green dashed lines). The kinetic solutions are blue solid lines. It is shown that the FLR corrections are crucial for the correct stabilization of the firehose instability. Figure 9(b) shows refinement with FLR2 (blue dashed lines) and FLR3 tensors (black dashed lines). Obviously, for high plasma beta values, the maximum growth rate of the (strictly) parallel firehose instability is captured very precisely by the Hall-CGL-FLR3 model. The maximum growth rate (for parallel propagation) can be easily found analytically only for the Hall-CGL and Hall-CGL-FLR1 models. For models with the FLR2 and FLR3 tensors it is easier to find the maximum growth rate numerically. For example, let us consider the Hall-CGL model. Assuming the firehose-unstable regime, the frequency of the Hall-CGL solution (6.4) can be split into  $\omega = \omega_r + i\omega_i$ , where the imaginary part  $\omega_i = k_{\parallel} \sqrt{-v_{A\parallel}^2 - k_{\parallel}^2/4}$ . Then, calculating  $\partial\omega_i/\partial k_{\parallel} = 0$ , yields the wavenumber  $k_{\parallel\max}$  and the maximum growth rate  $\gamma_{\max}$  in the following form

$$k_{\parallel\max}^2 = -2v_{A\parallel}^2; \quad \gamma_{\max} = -v_{A\parallel}^2 = - \left( 1 + \frac{\beta_{\parallel}}{2} (a_p - 1) \right). \quad (6.11)$$

Similarly, the Hall-CGL-FLR1 model yields

$$k_{\parallel\max}^2 = -\frac{2v_{A\parallel}^2}{(1 - v_b)^2}; \quad \gamma_{\max} = \frac{v_{A\parallel}^2}{v_{A\parallel}^2 - \frac{\beta_{\parallel}}{2}} = \frac{1 + \frac{\beta_{\parallel}}{2} (a_p - 1)}{1 + \frac{\beta_{\parallel}}{2} (a_p - 2)}. \quad (6.12)$$

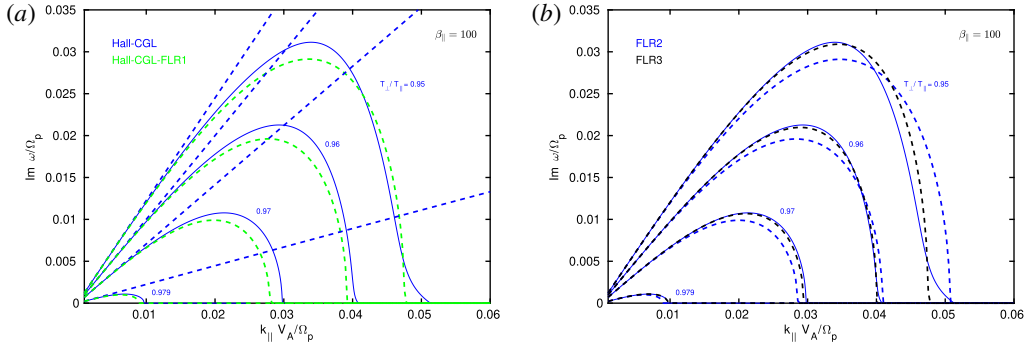


FIGURE 9. Growth rate of the parallel firehose instability for  $\beta_{\parallel} = 100$ . Kinetic solutions are solid blue lines. The temperature anisotropy is varied as  $a_p = 0.95; 0.96; 0.97; 0.979$  (the hard firehose threshold is at  $a_p = 0.98$ ). Four different fluid models are plotted, and all fluid solutions have dashed lines. (a) Hall-CGL (blue), Hall-CGL-FLR1 (green). (b) Hall-CGL-FLR2 (blue), Hall-CGL-FLR3 (black).

### 6.2. Oblique propagation

Here, we briefly investigate the parallel and oblique firehose instability for oblique propagation directions. Contour plots of the growth rate in the  $k - \theta$  plane are shown in figure 10. The plasma  $\beta_{\parallel} = 4$  and the temperature anisotropy  $a_p = 0.49$ , so all the fluid models are in the firehose-unstable regime. We compare four different fluid models, Hall-CGL (a), and enhancement with FLR1 (b), FLR2 (c) and FLR3 (d). The FLR2 solution shows an additional instability at higher wavenumbers that we did not study further. Note the large differences between the four solutions. Importantly, the FLR3 solution shows large enhancement of the growth rate. We use the WHAMP code for kinetic calculations, and we do not provide a contour plot for the kinetic theory. Instead, we plot the growth rate for several propagation angles, so that the comparison with the contour plots can be made easily. The figure (e) is the FLR3 model, and (f) is the kinetic result. It is noted that, similarly to other fluid models with higher-order FLR corrections, the FLR3 model can develop secondary instabilities at scales below the proton gyroscale.

### 6.3. Hellinger's contours for the Hall-CGL-FLR3 model

Here, we prescribe a fixed value for the maximum growth rate,  $\gamma_{\max} = 10^{-3}; 10^{-2}; 10^{-1}$ , and plot solutions in the  $\beta_{\parallel} - a_p$  plane, which is shown in figure 11. (a,c) Show the parallel firehose instability, and (b,d) show the oblique firehose instability. Panels (a,b) are plotted with the usual logarithmic scales, and (c,d) with linear scales. Additionally, only solutions with  $\beta_{\parallel} < 6$  are shown at the bottom panels. Solutions for the Hall-CGL-FLR3 fluid model are black dashed lines, and kinetic solutions are solid blue lines (a,c) and solid red lines (b,d). The kinetic solutions were provided to us by P. Hellinger (private communication) and are identical to solutions shown in figure 1 of Hellinger *et al.* (2006). It is shown that the  $\gamma_{\max} = 10^{-3}; 10^{-2}$  contours clearly lie below the hard firehose threshold, i.e. the firehose instability in the FLR3 fluid model develops at some range of wavenumbers, even if the model is stable in the long-wavelength limit. Importantly, the kinetic contours of Hellinger *et al.* (2006) were not calculated for cold electrons, but for isotropic electrons with  $\beta_e = 1$ . This does not matter for the parallel propagation, since the isotropic electron pressure

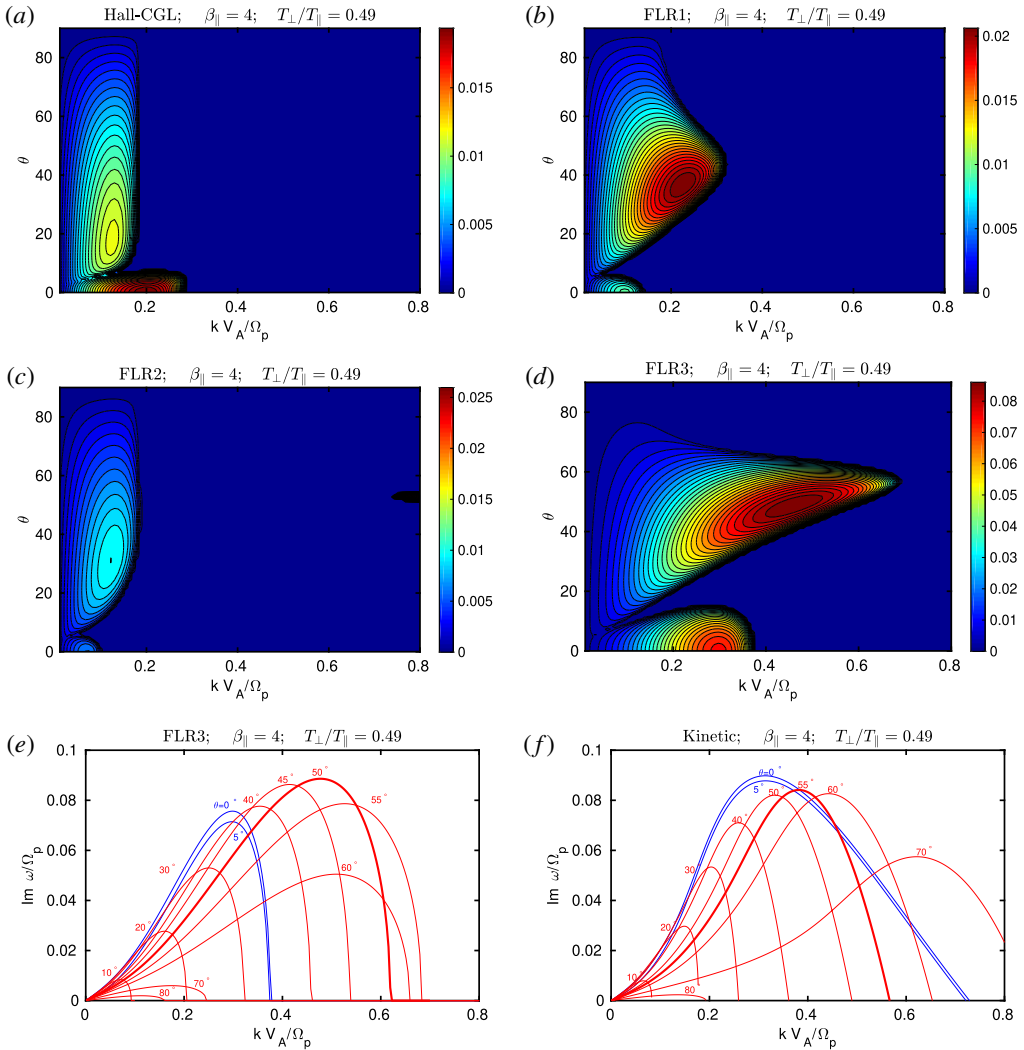


FIGURE 10. The growth rate plotted in the  $k-\theta$  plane, with fixed  $\beta_{\parallel} = 4$  and  $a_p = 0.49$  (the hard firehose threshold is at  $a_p = 0.5$ ), showing the parallel and oblique firehose instability. Four different fluid models are plotted. (a) Hall-CGL; (b) Hall-CGL-FLR1; (c) Hall-CGL-FLR2, (d) Hall-CGL-FLR3. We do not provide a contour plot for kinetic theory. Nevertheless, the bottom figures show solutions for several propagation angles. (e) Hall-CGL-FLR3 model, (f) kinetic.

does not influence the dispersion relations for the parallel propagating whistler and ion-cyclotron modes. Only the effect of electron inertia will enter, however, the effect should be negligible at the scales considered here. Unfortunately, the solution for the parallel firehose instability cannot be further improved by currently developed fluid models. The oblique case is different, since even isotropic electron pressure will influence the dynamics. The contours for the oblique firehose instability should be therefore recalculated by using a proper proton–electron two fluid model. The oblique case could be further improved, by considering FLR contributions from the

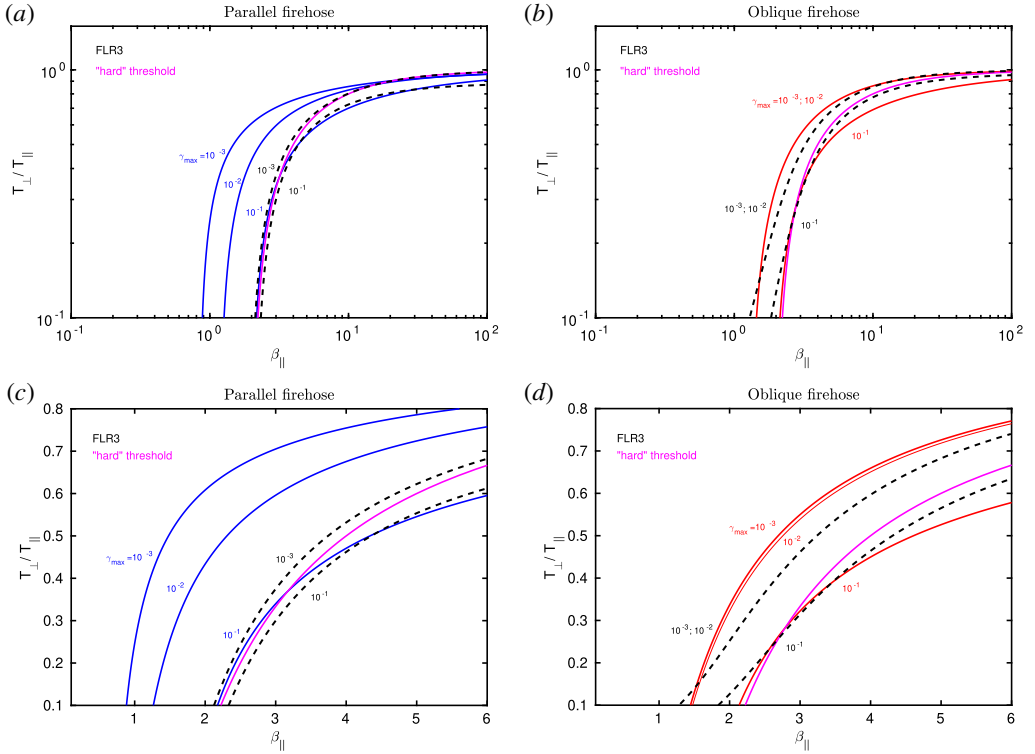


FIGURE 11. (a,b) Solutions for the parallel (a,c) and oblique (b,d) firehose instability in the  $\beta_{\parallel} - a_p$  plane for a prescribed maximum growth rate  $\gamma_{\max} = 10^{-3}; 10^{-2}; 10^{-1}$ . Solid blue and red lines are kinetic solutions from Hellinger *et al.* (2006). Black dashed lines are solutions of the Hall-CGL-FLR3 model. The magenta line is the ‘hard’ (long-wavelength limit) firehose threshold. (c,d) Same as top panels, but with linear scales for both axes, and only showing results for  $\beta_{\parallel} < 6$ .

non-gyrotropic heat flux tensor  $\sigma$  that were neglected here (contributions for parallel propagation are zero). For the oblique case, one can also use higher-order fluid models with the gyrotropic heat flux fluctuations  $q_{\parallel}, q_{\perp}$  (such as the CGL2 model discussed later), or even Landau fluid models. The main reason for the relatively large discrepancies found in the contours for the parallel firehose instability is demonstrated in figure 12.

## 7. Heat flux tensor equation

Multiply the Vlasov equation by  $mc_i c_j c_k$  and integrate over the velocity space. Naturally, the other possibilities are to multiply by  $mc_i c_j v_k$ ,  $mc_i v_j v_k$  or  $mv_i v_j v_k$ , although none of them is more revealing and we will use the first choice. It is convenient to define the symmetric operator  $^S$  that acts on a tensor of third rank according to

$$A_{ijk}^S = A_{ijk} + A_{jki} + A_{kij}, \quad (7.1)$$



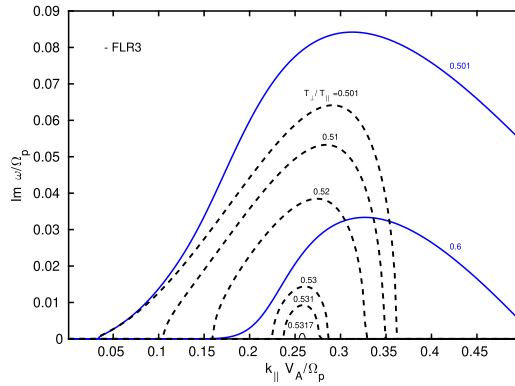


FIGURE 12. Parallel firehose instability,  $\beta_{\parallel} = 4$ . It is shown that the Hall-CGL-FLR3 model (black dashed lines) indeed develops the firehose instability even for values of  $a_p > 0.5$ , i.e. when there is no instability in the long-wavelength limit. Nevertheless, by increasing the value of  $a_p$  beyond 0.5, the growth rate in the FLR3 model quickly falls off, and the instability disappears for  $a_p > 0.5317$ . In contrast, kinetic theory (blue solid lines) develops a strong firehose instability even at  $a_p = 0.6$ .

i.e. the operator represents all possible cyclic permutations. For the first term we will need the following identities

$$\frac{\partial}{\partial t}(c_i c_j c_k) = -\frac{\partial u_i}{\partial t} c_j c_k - c_i \frac{\partial u_j}{\partial t} c_k - c_i c_j \frac{\partial u_k}{\partial t}; \quad (7.2)$$

$$\begin{aligned} m \int f \frac{\partial}{\partial t}(c_i c_j c_k) d^3 v &= -\frac{\partial u_i}{\partial t} \underbrace{m \int c_j c_k f d^3 v}_{p_{jk}} - \frac{\partial u_j}{\partial t} \underbrace{m \int c_i c_k f d^3 v}_{p_{ik}} - \frac{\partial u_k}{\partial t} \underbrace{m \int c_i c_j f d^3 v}_{p_{ij}} \\ &= -\frac{\partial u_i}{\partial t} p_{jk} - \frac{\partial u_j}{\partial t} p_{ik} - \frac{\partial u_k}{\partial t} p_{ij}, \end{aligned} \quad (7.3)$$

and the entire first term of the integrated Vlasov equation calculates as

$$\begin{aligned} \textcircled{1} &= m \int c_i c_j c_k \frac{\partial f}{\partial t} d^3 v = \frac{\partial}{\partial t} \left( \underbrace{m \int c_i c_j c_k f d^3 v}_{q_{ijk}} \right) - m \int f \frac{\partial}{\partial t}(c_i c_j c_k) d^3 v \\ &= \frac{\partial}{\partial t} q_{ijk} + \frac{\partial u_i}{\partial t} p_{jk} + \frac{\partial u_j}{\partial t} p_{ik} + \frac{\partial u_k}{\partial t} p_{ij} = \frac{\partial}{\partial t} q_{ijk} + \left[ \frac{\partial \mathbf{u}}{\partial t} \mathbf{p} \right]_{ijk}^s. \end{aligned} \quad (7.4)$$

For the second term we will need identities

$$m \int \mathbf{c} \mathbf{c} \mathbf{c} v f d^3 v = \mathbf{r} + \mathbf{q} \mathbf{u}; \quad (7.5)$$

$$\begin{aligned} m \int c_i c_j c_k v_l f d^3 v &= m \int c_i c_j c_k (v_l - u_l + u_l) f d^3 v \\ &= \underbrace{m \int c_i c_j c_k c_l f d^3 v}_{r_{ijkl}} + \underbrace{u_l m \int c_i c_j c_k f d^3 v}_{q_{ijk} u_l} = r_{ijkl} + q_{ijk} u_l; \end{aligned} \quad (7.6)$$

$$\partial_l(c_i c_j c_k) = -c_j c_k \partial_l u_i - c_i c_k \partial_l u_j - c_i c_j \partial_l u_k; \quad (7.7)$$

$$\begin{aligned} m \int f v_l \partial_l (c_i c_j c_k) d^3 v &= -(\partial_l u_i) m \underbrace{\int c_j c_k v_l f d^3 v}_{q_{jkl} + p_{jk} u_l} \\ &\quad - (\partial_l u_j) m \underbrace{\int c_i c_k v_l f d^3 v}_{q_{ikl} + p_{ik} u_l} - (\partial_l u_k) m \underbrace{\int c_i c_j v_l f d^3 v}_{q_{ijl} + p_{ij} u_l} \\ &= -(q_{jkl} + p_{jk} u_l) \partial_l u_i - (q_{ikl} + p_{ik} u_l) \partial_l u_j - (q_{ijl} + p_{ij} u_l) \partial_l u_k, \end{aligned} \quad (7.8)$$

and the second term calculates as

$$\begin{aligned} \textcircled{2} &= m \int c_i c_j c_k v_l \partial_l f d^3 v = \partial_l \left( m \int c_i c_j c_k v_l f d^3 v \right) - m \int f v_l \partial_l (c_i c_j c_k) d^3 v \\ &= \partial_l (r_{lij} + u_l q_{ijk}) + (q_{jkl} + p_{jk} u_l) \partial_l u_i + (q_{ikl} + p_{ik} u_l) \partial_l u_j + (q_{ijl} + p_{ij} u_l) \partial_l u_k \\ &= [\nabla \cdot (\mathbf{r} + \mathbf{u} \mathbf{q})]_{ijk} + [(\mathbf{q} + \mathbf{p} \mathbf{u}) \cdot \nabla \mathbf{u}]_{jki} + [(\mathbf{q} + \mathbf{p} \mathbf{u}) \cdot \nabla \mathbf{u}]_{ikj} + [(\mathbf{q} + \mathbf{p} \mathbf{u}) \cdot \nabla \mathbf{u}]_{ijk} \\ &= [\nabla \cdot (\mathbf{r} + \mathbf{u} \mathbf{q})]_{ijk} + [(\mathbf{q} + \mathbf{p} \mathbf{u}) \cdot \nabla \mathbf{u}]_{ijk}^s. \end{aligned} \quad (7.9)$$

Note again that the divergence of a tensor operates through its first component  $\partial_l(r_{lij} + u_l q_{ijk}) = [\nabla \cdot (\mathbf{r} + \mathbf{u} \mathbf{q})]_{ijk}$ . Similarly, if the operator acts on a tensor from the right-hand side, the most natural way is to define that it operates through the last component, i.e.  $[\mathbf{q} \cdot \nabla]_{ij} = q_{ijl} \partial_l$ . In the expressions above  $[\mathbf{q} \cdot \nabla \mathbf{u}]_{ijk} = q_{ijl} \partial_l u_k$  and  $[\mathbf{p} \mathbf{u} \cdot \nabla \mathbf{u}]_{ijk} = (\mathbf{p} \mathbf{u})_{ijl} \partial_l u_k = p_{ij} u_l \partial_l u_k$ . The last term can also be rewritten as  $p_{ij} (\mathbf{u} \cdot \nabla \mathbf{u})_k$ , where the expression  $(\mathbf{u} \cdot \nabla \mathbf{u})_k = u_l \partial_l u_k$  is familiar from MHD. For the third term we will need identities

$$\frac{\partial c_i}{\partial v_l} = \delta_{il}; \quad (7.10)$$

$$\frac{\partial}{\partial v_l} (c_i c_j c_k) = \delta_{il} c_j c_k + c_i \delta_{jl} c_k + c_i c_j \delta_{kl}, \quad (7.11)$$

and the entire third term calculates as

$$\begin{aligned} \textcircled{3} &= q \int c_i c_j c_k E_l \frac{\partial f}{\partial v_l} d^3 v = q E_l \underbrace{\int \frac{\partial}{\partial v_l} (c_i c_j c_k f) d^3 v}_{\rightarrow 0} - q E_l \int \frac{\partial}{\partial v_l} (c_i c_j c_k) f d^3 v \\ &= -q E_l \int (\delta_{il} c_j c_k + \delta_{jl} c_i c_k + \delta_{kl} c_i c_j) f d^3 v = -\frac{q}{m} (E_i p_{jk} + E_j p_{ik} + E_k p_{ij}) \\ &= -\frac{q}{m} [\mathbf{E} \mathbf{p}]_{ijk}^s. \end{aligned} \quad (7.12)$$

For the fourth term we will need identities

$$\frac{\partial}{\partial v_l} (\mathbf{v} \times \mathbf{B})_l = 0; \quad (7.13)$$

$$\frac{\partial}{\partial v_l} [c_i c_j c_k (\mathbf{v} \times \mathbf{B})_l] = c_j c_k (\mathbf{v} \times \mathbf{B})_i + c_i c_k (\mathbf{v} \times \mathbf{B})_j + c_i c_j (\mathbf{v} \times \mathbf{B})_k; \quad (7.14)$$

$$\int c_j (\mathbf{v} \times \mathbf{B})_i f d^3v = -\frac{1}{m} (\mathbf{B} \times \mathbf{p})_{ij}; \quad (7.15)$$

$$\int c_j c_k (\mathbf{v} \times \mathbf{B})_i f d^3v = -\frac{1}{m} [\mathbf{B} \times (\mathbf{q} + \mathbf{up})]_{ijk}, \quad (7.16)$$

and the entire fourth term calculates as

$$\begin{aligned} \textcircled{4} &= \frac{q}{c} \int c_i c_j c_k (\mathbf{v} \times \mathbf{B})_l \frac{\partial f}{\partial v_l} d^3v \\ &= \frac{q}{c} \underbrace{\int \frac{\partial}{\partial v_l} [c_i c_j c_k (\mathbf{v} \times \mathbf{B})_l f] d^3v}_{\rightarrow 0} - \frac{q}{c} \int f \frac{\partial}{\partial v_l} [c_i c_j c_k (\mathbf{v} \times \mathbf{B})_l] d^3v \\ &= -\frac{q}{c} \left\{ \int f c_j c_k (\mathbf{v} \times \mathbf{B})_i d^3v + \int f c_k c_i (\mathbf{v} \times \mathbf{B})_j d^3v + \int f c_i c_j (\mathbf{v} \times \mathbf{B})_k d^3v \right\} \\ &= \frac{q}{mc} \{ [\mathbf{B} \times (\mathbf{q} + \mathbf{up})]_{ijk} + [\mathbf{B} \times (\mathbf{q} + \mathbf{up})]_{jki} + [\mathbf{B} \times (\mathbf{q} + \mathbf{up})]_{kij} \} \\ &\equiv \frac{q}{mc} [\mathbf{B} \times (\mathbf{q} + \mathbf{up})]_{ijk}^S. \end{aligned} \quad (7.17)$$

Combining all the results together  $\textcircled{1} + \textcircled{2} + \textcircled{3} + \textcircled{4} = 0$ , the entire heat flux tensor equation obtained by direct integration of the Vlasov equation reads

$$\frac{\partial \mathbf{q}}{\partial t} + \nabla \cdot (\mathbf{r} + \mathbf{uq}) + \left[ \frac{\partial \mathbf{u}}{\partial t} \mathbf{p} + (\mathbf{q} + \mathbf{pu}) \cdot \nabla \mathbf{u} - \frac{q}{m} \mathbf{E} \mathbf{p} + \frac{q}{mc} \mathbf{B} \times (\mathbf{q} + \mathbf{up}) \right]^S = 0. \quad (7.18)$$

Now we will need to use 3 different momentum equations that will cancel various terms. We will also need identity

$$[(\mathbf{u} \times \mathbf{B})\mathbf{p}]_{ijk} = -[\mathbf{B} \times (\mathbf{up})]_{ijk}. \quad (7.19)$$

We need to multiply the momentum equation for  $\partial u_i / \partial t$  by  $p_{jk}$ , the equation for  $\partial u_j / \partial t$  by  $p_{ki}$  and the equation for  $\partial u_k / \partial t$  by  $p_{ij}$ . All the three momentum equations that we want to subtract from the heat flux equation can be written together as

$$\left[ \left( \frac{\partial \mathbf{u}}{\partial t} + \mathbf{u} \cdot \nabla \mathbf{u} + \frac{1}{mn} \nabla \cdot \mathbf{p} - \frac{q}{m} \mathbf{E} - \frac{q}{mc} \mathbf{u} \times \mathbf{B} \right) \mathbf{p} \right]_{ijk}^S = 0, \quad (7.20)$$

and because of identity (7.19), this is equivalent to

$$\left[ \left( \frac{\partial \mathbf{u}}{\partial t} + \mathbf{u} \cdot \nabla \mathbf{u} + \frac{1}{mn} \nabla \cdot \mathbf{p} - \frac{q}{m} \mathbf{E} \right) \mathbf{p} + \frac{q}{mc} [\mathbf{B} \times (\mathbf{up})] \right]_{ijk}^S = 0. \quad (7.21)$$

Note that, because of the symmetric operator, it does not matter if the tensor  $\mathbf{p}$  is applied to the momentum equation from the left or right, since all the expressions are symmetric in this regard, i.e. for example  $[(\nabla \cdot \mathbf{p})\mathbf{p}]^S = [\mathbf{p}(\nabla \cdot \mathbf{p})]^S$ . By subtracting (7.21) from (7.18), the final heat flux tensor equation reads

$$\frac{\partial}{\partial t} \mathbf{q} + \nabla \cdot (\mathbf{r} + \mathbf{uq}) + \left[ \mathbf{q} \cdot \nabla \mathbf{u} + \frac{q}{mc} \mathbf{B} \times \mathbf{q} - \frac{1}{\rho} \mathbf{p}(\nabla \cdot \mathbf{p}) \right]^S = 0. \quad (7.22)$$

By defining the cyclotron frequency vector  $\boldsymbol{\Omega} = q\mathbf{B}/mc$ , this equation identifies with equation (A 5) in Chust & Belmont (2006). By using the scalar cyclotron frequency defined with respect to  $|\mathbf{B}|$  as  $\Omega = q|\mathbf{B}|/mc$ , the heat flux tensor equation is written explicitly in the index notation

$$\begin{aligned} \frac{\partial}{\partial t} q_{ijk} + \partial_l(r_{lij} + u_l q_{ijk}) + q_{ijl} \partial_l u_k + q_{jkl} \partial_l u_i + q_{kil} \partial_l u_j + \Omega \hat{b}_l (\epsilon_{ilm} q_{mjk} + \epsilon_{jlm} q_{mki} + \epsilon_{klm} q_{mij}) \\ - \frac{1}{mn} (p_{ij} \partial_l p_{lk} + p_{jk} \partial_l p_{li} + p_{ki} \partial_l p_{lj}) = 0, \end{aligned} \quad (7.23)$$

which is equivalent to (4) in Goswami *et al.* (2005) with the term

$$\begin{aligned} (\hat{\mathbf{b}} \times \mathbf{q})_{ijk}^S &= \epsilon_{ilm} \hat{b}_l q_{mjk} + \epsilon_{jlm} \hat{b}_l q_{mki} + \epsilon_{klm} \hat{b}_l q_{mij} \\ &= -\hat{b}_l (\epsilon_{iml} q_{jkm} + \epsilon_{jml} q_{ikm} + \epsilon_{klm} q_{ijm}). \end{aligned} \quad (7.24)$$

### 7.1. Heat flux tensor decomposition

Considering only the gyrotropic part of the heat flux tensor ( $3 \times 3 \times 3$ ) cube, the heat flux is decomposed into the scalar parallel and perpendicular heat flux components  $q_{\parallel}$ ,  $q_{\perp}$  according to

$$\mathbf{q}^g = q_{\parallel} \hat{\mathbf{b}} \hat{\mathbf{b}} \hat{\mathbf{b}} + q_{\perp} [(\mathbf{I} - \hat{\mathbf{b}} \hat{\mathbf{b}}) \hat{\mathbf{b}}]^S, \quad (7.25)$$

and which in the index notation reads

$$q_{ijk}^g = q_{\parallel} \hat{b}_i \hat{b}_j \hat{b}_k + q_{\perp} (\delta_{ij} \hat{b}_k + \delta_{jk} \hat{b}_i + \delta_{ki} \hat{b}_j - 3 \hat{b}_i \hat{b}_j \hat{b}_k). \quad (7.26)$$

This part of the pressure tensor represents only the gyrotropic part, and the full heat flux decomposition can be written as  $\mathbf{q} = \mathbf{q}^g + \mathbf{q}^{ng}$ . In the heat flux tensor equation, the term  $\mathbf{B} \times \mathbf{q}$  is proportional to the cyclotron frequency  $\Omega = qB_0/mc$  and the equation rewrites as

$$\frac{\partial}{\partial t} \mathbf{q} + \nabla \cdot (\mathbf{r} + \mathbf{u} \mathbf{q}) + \left[ \mathbf{q} \cdot \nabla \mathbf{u} + \Omega \frac{|\mathbf{B}|}{B_0} \hat{\mathbf{b}} \times \mathbf{q} - \frac{1}{\rho} \mathbf{p} (\nabla \cdot \mathbf{p}) \right]^S = 0. \quad (7.27)$$

The situation is now similar to the previously studied pressure tensor. At long spatial scales (low frequencies  $\omega$ ), this term will dominate and the gyrotropic contribution must be equal to zero

$$[\hat{\mathbf{b}} \times \mathbf{q}^g]^S = 0. \quad (7.28)$$

At first look, it is not that obvious that the decomposition (7.25) satisfies this equation. It is however possible to verify that indeed

$$\begin{aligned} (\hat{\mathbf{b}} \times \mathbf{q}^g)_{ijk} &= \epsilon_{irs} \hat{b}_r q_{sjk}^g = \epsilon_{irs} \hat{b}_r [q_{\parallel} \hat{b}_s \hat{b}_j \hat{b}_k + q_{\perp} (\delta_{sj} \hat{b}_k + \delta_{jk} \hat{b}_s + \delta_{ks} \hat{b}_j - 3 \hat{b}_s \hat{b}_j \hat{b}_k)] \\ &= q_{\parallel} \hat{b}_j \hat{b}_k \underbrace{\epsilon_{irs} \hat{b}_r \hat{b}_s}_{=0} + q_{\perp} (\epsilon_{irj} \hat{b}_r \hat{b}_k + \delta_{jk} \underbrace{\epsilon_{irs} \hat{b}_r \hat{b}_s}_{=0} + \epsilon_{irk} \hat{b}_r \hat{b}_j - 3 \hat{b}_j \hat{b}_k \underbrace{\epsilon_{irs} \hat{b}_r \hat{b}_s}_{=0}) \\ &= q_{\perp} \hat{b}_r (\epsilon_{irj} \hat{b}_k + \epsilon_{irk} \hat{b}_j); \end{aligned} \quad (7.29)$$

$$(\hat{\mathbf{b}} \times \mathbf{q}^g)_{jki} = q_{\perp} \hat{b}_r (\epsilon_{jrk} \hat{b}_i + \epsilon_{jri} \hat{b}_k); \quad (7.30)$$

$$(\hat{\mathbf{b}} \times \mathbf{q}^g)_{kij} = q_{\perp} \hat{b}_r (\epsilon_{kri} \hat{b}_j + \epsilon_{krj} \hat{b}_i). \quad (7.31)$$

Putting all terms together

$$\begin{aligned} [\hat{\mathbf{b}} \times \mathbf{q}^g]_{ijk}^S &= [\hat{\mathbf{b}} \times \mathbf{q}^g]_{ijk} + [\hat{\mathbf{b}} \times \mathbf{q}^g]_{jki} + [\hat{\mathbf{b}} \times \mathbf{q}^g]_{kij} \\ &= q_{\perp} \hat{b}_r [\hat{b}_i (\underbrace{\epsilon_{jrk} + \epsilon_{krj}}_{=0}) + \hat{b}_j (\underbrace{\epsilon_{irk} + \epsilon_{kri}}_{=0}) + \hat{b}_k (\underbrace{\epsilon_{irj} + \epsilon_{jri}}_{=0})] = 0, \end{aligned} \quad (7.32)$$

and we see that all three parts of the symmetric operator are required to make this term equal to zero. Therefore, the exact heat flux tensor equation reads

$$\frac{\partial}{\partial t} \mathbf{q} + \nabla \cdot (\mathbf{r} + \mathbf{u} \mathbf{q}) + \left[ \mathbf{q} \cdot \nabla \mathbf{u} + \Omega \frac{|\mathbf{B}|}{B_0} \hat{\mathbf{b}} \times \mathbf{q}^{\text{ng}} - \frac{1}{\rho} \mathbf{p} (\nabla \cdot \mathbf{p}) \right]^S = 0. \quad (7.33)$$

Since this should be an introductory text, we do not want to be bothered right now with the complicated algebra of the non-gyrotropic heat flux  $\mathbf{q}^{\text{ng}}$ . For the clarity of the presented material, here we separate the non-gyrotropic heat flux into a separate term  $\mathbf{Q}^{\text{ng}}$ , and write the heat flux tensor equation in the following form

$$\frac{\partial}{\partial t} \mathbf{q}^g + \nabla \cdot (\mathbf{r} + \mathbf{u} \mathbf{q}^g) + \left[ \mathbf{q}^g \cdot \nabla \mathbf{u} - \frac{1}{\rho} \mathbf{p} (\nabla \cdot \mathbf{p}) \right]^S + \mathbf{Q}^{\text{ng}} = 0, \quad (7.34)$$

where

$$\mathbf{Q}^{\text{ng}} = \frac{\partial}{\partial t} \mathbf{q}^{\text{ng}} + \nabla \cdot (\mathbf{u} \mathbf{q}^{\text{ng}}) + \left[ \mathbf{q}^{\text{ng}} \cdot \nabla \mathbf{u} + \Omega \frac{|\mathbf{B}|}{B_0} \hat{\mathbf{b}} \times \mathbf{q}^{\text{ng}} \right]^S. \quad (7.35)$$

We will address the non-gyrotropic heat flux contributions in appendix D.

In a similar fashion to the pressure decomposition (2.35), we are going to frequently apply double contractions with  $\hat{\mathbf{b}}\hat{\mathbf{b}}$  and  $(\mathbf{I} - \hat{\mathbf{b}}\hat{\mathbf{b}})/2$ . Applying these operators to the heat flux tensor  $q_{ijk}$ , must yield quantities that are vectors. It is therefore logical to define parallel and perpendicular heat flux vectors

$$\mathbf{S}^{\parallel} \equiv \mathbf{q} : \hat{\mathbf{b}}\hat{\mathbf{b}}; \quad \mathbf{S}^{\perp} \equiv \mathbf{q} : (\mathbf{I} - \hat{\mathbf{b}}\hat{\mathbf{b}})/2. \quad (7.36)$$

The scalar parallel and perpendicular heat flux components  $q_{\parallel}, q_{\perp}$  (which are the only parts that are gyrotropic) are obtained by further projecting these heat flux vectors along the magnetic field lines, i.e. by performing  $\hat{\mathbf{b}} \cdot$ , so that

$$q_{\parallel} = (\mathbf{q} : \hat{\mathbf{b}}\hat{\mathbf{b}}) \cdot \hat{\mathbf{b}}; \quad q_{\perp} = (\mathbf{q} : (\mathbf{I} - \hat{\mathbf{b}}\hat{\mathbf{b}})/2) \cdot \hat{\mathbf{b}}. \quad (7.37)$$

Briefly considering only the gyrotropic heat flux (7.26) on the right-hand side, it is easy to verify that the parallel decomposition indeed works

$$\begin{aligned} (\mathbf{q}^g : \hat{\mathbf{b}}\hat{\mathbf{b}})_k &= q_{kij}^g \hat{b}_i \hat{b}_j = q_{kij}^g \hat{b}_i \hat{b}_j = [q_{\parallel} \hat{b}_i \hat{b}_j \hat{b}_k + q_{\perp} (\delta_{ij} \hat{b}_k + \delta_{jk} \hat{b}_i + \delta_{ki} \hat{b}_j - 3 \hat{b}_i \hat{b}_j \hat{b}_k)] \hat{b}_i \hat{b}_j \\ &= q_{\parallel} \hat{b}_k + q_{\perp} (\hat{b}_k + \hat{b}_k + \hat{b}_k - 3 \hat{b}_k) = q_{\parallel} \hat{b}_k; \end{aligned} \quad (7.38)$$

$$\mathbf{q}^g : \hat{\mathbf{b}}\hat{\mathbf{b}} = q_{\parallel} \hat{\mathbf{b}}; \quad (7.39)$$

$$(\mathbf{q}^g : \hat{\mathbf{b}}\hat{\mathbf{b}}) \cdot \hat{\mathbf{b}} = q_{\parallel}. \quad (7.40)$$

For the perpendicular decomposition, it is useful to specifically calculate components

$$q_{ik}^g = q_{\parallel} \hat{b}_k + q_{\perp} (\underbrace{\delta_{ii}}_{=3} \hat{b}_k + \hat{b}_k + \hat{b}_k - 3 \hat{b}_k) = q_{\parallel} \hat{b}_k + 2 q_{\perp} \hat{b}_k, \quad (7.41)$$

so that

$$\text{Tr } \mathbf{q}^g = \mathbf{q}^g : \mathbf{I} = q_{\parallel} \hat{\mathbf{b}} + 2q_{\perp} \hat{\mathbf{b}}; \quad (7.42)$$

$$\mathbf{q}^g : (\mathbf{I} - \hat{\mathbf{b}}\hat{\mathbf{b}})/2 = (\mathbf{q}^g : \mathbf{I} - \mathbf{q}^g : \hat{\mathbf{b}}\hat{\mathbf{b}})/2 = (q_{\parallel} \hat{\mathbf{b}} + 2q_{\perp} \hat{\mathbf{b}} - q_{\parallel} \hat{\mathbf{b}})/2 = q_{\perp} \hat{\mathbf{b}}; \quad (7.43)$$

$$(\mathbf{q}^g : (\mathbf{I} - \hat{\mathbf{b}}\hat{\mathbf{b}})/2) \cdot \hat{\mathbf{b}} = q_{\perp}, \quad (7.44)$$

which verifies that the perpendicular decomposition (7.37) is satisfied for  $\mathbf{q}^g$ . Now we apply the decomposition (7.37) directly to the definition of the entire heat flux tensor (2.6), which yields

$$\begin{aligned} (\mathbf{q} : \hat{\mathbf{b}}\hat{\mathbf{b}})_k &= \left( m \int c c c f \, d^3 v : \hat{\mathbf{b}}\hat{\mathbf{b}} \right)_k = m \int c_i c_j c_k f \, d^3 v \hat{b}_i \hat{b}_j \\ &= m \int (v_i - u_{\parallel} \hat{b}_i) \hat{b}_i (v_j - u_{\parallel} \hat{b}_j) \hat{b}_j (v_k - u_{\parallel} \hat{b}_k) f \, d^3 v \\ &= m \int (v_{\parallel} - u_{\parallel})^2 (v_k - u_{\parallel}) f \, d^3 v \equiv (\mathbf{S}^{\parallel})_k; \end{aligned} \quad (7.45)$$

$$(\mathbf{q} : \hat{\mathbf{b}}\hat{\mathbf{b}}) \cdot \hat{\mathbf{b}} = m \int (v_{\parallel} - u_{\parallel})^3 f \, d^3 v \equiv q_{\parallel}; \quad (7.46)$$

and the perpendicular decomposition

$$(\mathbf{q} : \mathbf{I}/2)_k = \frac{m}{2} \int c_i c_j c_k f \, d^3 v \delta_{ij} = \frac{m}{2} \int |c|^2 c_k f \, d^3 v = \frac{m}{2} \int |\mathbf{v} - \mathbf{u}|^2 (v_k - u_{\parallel}) f \, d^3 v; \quad (7.47)$$

$$\begin{aligned} (\mathbf{q} : (\mathbf{I} - \hat{\mathbf{b}}\hat{\mathbf{b}})/2)_k &= \frac{m}{2} \int (|\mathbf{v} - \mathbf{u}|^2 - (v_{\parallel} - u_{\parallel})^2) (v_k - u_{\parallel}) f \, d^3 v \\ &= \frac{m}{2} \int |\mathbf{v}_{\perp} - \mathbf{u}_{\perp}|^2 (v_k - u_{\parallel}) f \, d^3 v \equiv (\mathbf{S}^{\perp})_k; \end{aligned} \quad (7.48)$$

$$(\mathbf{q} : (\mathbf{I} - \hat{\mathbf{b}}\hat{\mathbf{b}})/2) \cdot \hat{\mathbf{b}} = \frac{m}{2} \int |\mathbf{v}_{\perp} - \mathbf{u}_{\perp}|^2 (v_{\parallel} - u_{\parallel}) f \, d^3 v \equiv q_{\perp}. \quad (7.49)$$

The decomposition (7.37) obviously works for the entire heat flux  $\mathbf{q}$ , as well as for the gyrotropic part  $\mathbf{q}^g$ . Similarly to the pressure decomposition, this further yields the required properties that the non-gyrotropic heat flux  $\mathbf{q}^{\text{ng}}$  must satisfy. By using (7.37), the entire heat flux decomposition reads

$$\mathbf{q} = q_{\parallel} \hat{\mathbf{b}}\hat{\mathbf{b}}\hat{\mathbf{b}} + q_{\perp} [(\mathbf{I} - \hat{\mathbf{b}}\hat{\mathbf{b}}) \hat{\mathbf{b}}]^S + \mathbf{q}^{\text{ng}}; \quad (7.50)$$

$$\mathbf{q} = [(\mathbf{q} : \hat{\mathbf{b}}\hat{\mathbf{b}}) \cdot \hat{\mathbf{b}}] \hat{\mathbf{b}}\hat{\mathbf{b}} + [(\mathbf{q} : (\mathbf{I} - \hat{\mathbf{b}}\hat{\mathbf{b}})/2) \cdot \hat{\mathbf{b}}] [(\mathbf{I} - \hat{\mathbf{b}}\hat{\mathbf{b}}) \hat{\mathbf{b}}]^S + \mathbf{q}^{\text{ng}}. \quad (7.51)$$

Obviously, it would be useful to introduce a triple-contraction operator and instead of  $(\mathbf{q} : \hat{\mathbf{b}}\hat{\mathbf{b}}) \cdot \hat{\mathbf{b}}$  to write something like  $\mathbf{q} : \hat{\mathbf{b}}\hat{\mathbf{b}}\hat{\mathbf{b}}$ , but we do not want to introduce new notations. Also, an alternative and perhaps prettier expression is to move the  $\cdot \hat{\mathbf{b}}$  to the left-hand side, as  $(\mathbf{q} : \hat{\mathbf{b}}\hat{\mathbf{b}}) \cdot \hat{\mathbf{b}} = \hat{\mathbf{b}} \cdot \mathbf{q} : \hat{\mathbf{b}}\hat{\mathbf{b}}$ , but we will keep the first choice. Applying  $: \hat{\mathbf{b}}\hat{\mathbf{b}}$  and  $\cdot \hat{\mathbf{b}}$  to the above equation yields

$$(\mathbf{q} : \hat{\mathbf{b}}\hat{\mathbf{b}}) \cdot \hat{\mathbf{b}} = (\mathbf{q} : \hat{\mathbf{b}}\hat{\mathbf{b}}) \cdot \hat{\mathbf{b}} + (\mathbf{q}^{\text{ng}} : \hat{\mathbf{b}}\hat{\mathbf{b}}) \cdot \hat{\mathbf{b}}, \quad (7.52)$$

implying the first requirement for the non-gyrotropic heat flux

$$(\mathbf{q}^{\text{ng}} : \hat{\mathbf{b}}\hat{\mathbf{b}}) \cdot \hat{\mathbf{b}} = 0. \quad (7.53)$$

Similarly, the second requirement is obtained by applying  $:(\mathbf{I} - \hat{\mathbf{b}}\hat{\mathbf{b}})$  and  $\cdot \hat{\mathbf{b}}$ , yielding

$$(\mathbf{q}^{\text{ng}} : (\mathbf{I} - \hat{\mathbf{b}}\hat{\mathbf{b}})) \cdot \hat{\mathbf{b}} = 0. \quad (7.54)$$

By using the first requirement, the second requirement simplifies to

$$(\mathbf{q}^{\text{ng}} : \mathbf{I}) \cdot \hat{\mathbf{b}} = \text{Tr } \mathbf{q}^{\text{ng}} \cdot \hat{\mathbf{b}} = 0, \quad (7.55)$$

where obviously the trace and  $\cdot \hat{\mathbf{b}}$  operators commute. The two requirements in the index notation read

$$q_{ijk}^{\text{ng}} \hat{b}_i \hat{b}_j \hat{b}_k = 0; \quad q_{iik}^{\text{ng}} \hat{b}_k = 0. \quad (7.56)$$

Instead of decomposing  $\mathbf{q} = \mathbf{q}^{\text{g}} + \mathbf{q}^{\text{ng}}$ , an alternative and very useful decomposition of the entire heat flux tensor reads

$$\mathbf{q} = \mathbf{S} + \boldsymbol{\sigma}, \quad (7.57)$$

with the requirement  $\boldsymbol{\sigma} : \hat{\mathbf{b}}\hat{\mathbf{b}} = 0$  and  $\boldsymbol{\sigma} : (\mathbf{I} - \hat{\mathbf{b}}\hat{\mathbf{b}}) = 0$  (or equivalently  $\boldsymbol{\sigma} : \mathbf{I} = 0$ ). The heat flux vectors (7.36) therefore satisfy  $\mathbf{S}^{\parallel} = \mathbf{S} : \hat{\mathbf{b}}\hat{\mathbf{b}}$  and  $\mathbf{S}^{\perp} = \mathbf{S} : (\mathbf{I} - \hat{\mathbf{b}}\hat{\mathbf{b}})/2$ . The heat flux vectors contain both gyrotropic and non-gyrotropic contributions. Since the gyrotropic contributions are obtained by projecting these vectors along the magnetic field lines  $q_{\parallel} = \mathbf{S}^{\parallel} \cdot \hat{\mathbf{b}}$ ,  $q_{\perp} = \mathbf{S}^{\perp} \cdot \hat{\mathbf{b}}$ , it is useful to introduce the following decomposition

$$\mathbf{S}^{\parallel} \equiv q_{\parallel} \hat{\mathbf{b}} + \mathbf{S}_{\perp}^{\parallel}; \quad \mathbf{S}^{\perp} \equiv q_{\perp} \hat{\mathbf{b}} + \mathbf{S}_{\perp}^{\perp}. \quad (7.58)$$

The vectors  $\mathbf{S}_{\perp}^{\parallel}$ ,  $\mathbf{S}_{\perp}^{\perp}$  are referred to as the non-gyrotropic heat flux vectors, and their algebra is addressed in appendix D. Here, we only state that the  $\mathbf{q}^{\text{ng}}$  can be decomposed into vectors  $\mathbf{S}_{\perp}^{\parallel}$ ,  $\mathbf{S}_{\perp}^{\perp}$  and tensor  $\boldsymbol{\sigma}$  according to

$$\mathbf{q}^{\text{ng}} = [\mathbf{S}_{\perp}^{\parallel} \hat{\mathbf{b}}\hat{\mathbf{b}}]^{\text{S}} + \frac{1}{2}[\mathbf{S}_{\perp}^{\perp}(\mathbf{I} - \hat{\mathbf{b}}\hat{\mathbf{b}})]^{\text{S}} + \boldsymbol{\sigma}. \quad (7.59)$$

The entire heat flux tensor  $\boldsymbol{\sigma}$  is of course non-gyrotropic.

Now we need to verify the heat flux contributions (2.72), (2.73) that we used in the pressure equations (2.74), (2.75). To calculate the heat flux contributions to the pressure equations, we will need

$$\begin{aligned} \text{Tr}(\nabla \cdot \mathbf{q}^{\text{g}}) &= \delta_{ij}(\nabla \cdot \mathbf{q}^{\text{g}})_{ij} = \delta_{ij} \partial_k q_{kij}^{\text{g}} = \partial_k q_{iik}^{\text{g}} = \partial_k (q_{\parallel} \hat{b}_k + 2q_{\perp} \hat{b}_k) \\ &= \nabla \cdot (q_{\parallel} \hat{\mathbf{b}}) + 2\nabla \cdot (q_{\perp} \hat{\mathbf{b}}); \end{aligned} \quad (7.60)$$

$$q_{ijk}^{\text{g}} \hat{b}_i \hat{b}_j = q_{\parallel} \hat{b}_k; \quad (7.61)$$

$$\begin{aligned} q_{ijk}^{\text{g}} \partial_k (\hat{b}_i \hat{b}_j) &= [q_{\parallel} \hat{b}_i \hat{b}_j \hat{b}_k + q_{\perp} (\delta_{ij} \hat{b}_k + \delta_{jk} \hat{b}_i + \delta_{ki} \hat{b}_j - 3\hat{b}_i \hat{b}_j \hat{b}_k)] \partial_k (\hat{b}_i \hat{b}_j) \\ &= q_{\parallel} \hat{b}_k \underbrace{\hat{b}_i \hat{b}_j \partial_k (\hat{b}_i \hat{b}_j)}_{=0} + q_{\perp} \hat{b}_k \underbrace{\partial_k (\hat{b}_i \hat{b}_j)}_{=0} \end{aligned}$$

$$\begin{aligned}
& + q_{\perp} \hat{b}_i \partial_j (\hat{b}_i \hat{b}_j) + q_{\perp} \hat{b}_j \partial_i (\hat{b}_i \hat{b}_j) - 3 q_{\perp} \hat{b}_k \underbrace{\hat{b}_i \hat{b}_j \partial_k (\hat{b}_i \hat{b}_j)}_{=0} \\
& = q_{\perp} \partial_j \hat{b}_j + q_{\perp} \hat{b}_j \underbrace{\hat{b}_i \partial_j \hat{b}_i}_{=0} + q_{\perp} \partial_i \hat{b}_i + q_{\perp} \hat{b}_i \underbrace{\hat{b}_j \partial_i \hat{b}_j}_{=0} \\
& = 2 q_{\perp} \nabla \cdot \hat{\mathbf{b}},
\end{aligned} \tag{7.62}$$

and the contributions are

$$\begin{aligned}
\hat{\mathbf{b}} \cdot (\nabla \cdot \mathbf{q}^g) \cdot \hat{\mathbf{b}} &= \hat{b}_i (\partial_k q_{kij}^g) \hat{b}_j = \partial_k (q_{ijk}^g \hat{b}_i \hat{b}_j) - q_{ijk}^g \partial_k (\hat{b}_i \hat{b}_j) = \partial_k (q_{\parallel} \hat{b}_k) - 2 q_{\perp} \nabla \cdot \hat{\mathbf{b}} \\
&= \nabla \cdot (q_{\parallel} \hat{\mathbf{b}}) - 2 q_{\perp} \nabla \cdot \hat{\mathbf{b}};
\end{aligned} \tag{7.63}$$

$$\frac{1}{2} [\text{Tr} \nabla \cdot \mathbf{q}^g - \hat{\mathbf{b}} \cdot (\nabla \cdot \mathbf{q}^g) \cdot \hat{\mathbf{b}}] = \nabla \cdot (q_{\perp} \hat{\mathbf{b}}) + q_{\perp} \nabla \cdot \hat{\mathbf{b}}. \tag{7.64}$$

The left-hand side of the above equation can be also written naturally as  $(\nabla \cdot \mathbf{q}^g) : (\mathbf{I} - \hat{\mathbf{b}}\hat{\mathbf{b}})/2$ , which is consistent with an alternative derivation of the perpendicular pressure equation, where instead of doing trace and subtracting the parallel pressure equation, one can directly perform  $:(\mathbf{I} - \hat{\mathbf{b}}\hat{\mathbf{b}})/2$ .

## 7.2. Parallel heat flux equation

Now we apply the decomposition (7.37) to the heat flux tensor equation (7.34) step by step. We consider only gyrotropic heat flux components. Starting with the equation for the parallel heat flux, the first term calculates as

$$\begin{aligned}
\left( \frac{\partial \mathbf{q}^g}{\partial t} : \hat{\mathbf{b}}\hat{\mathbf{b}} \right) \cdot \hat{\mathbf{b}} &= \frac{\partial}{\partial t} [q_{\parallel} \hat{b}_i \hat{b}_j \hat{b}_k + q_{\perp} (\delta_{ij} \hat{b}_k + \delta_{jk} \hat{b}_i + \delta_{ki} \hat{b}_j - 3 \hat{b}_i \hat{b}_j \hat{b}_k)] \hat{b}_i \hat{b}_j \hat{b}_k \\
&= \frac{\partial q_{\parallel}}{\partial t} \underbrace{\hat{b}_i \hat{b}_j \hat{b}_k \hat{b}_i \hat{b}_j \hat{b}_k}_{=1} + q_{\parallel} \underbrace{\hat{b}_i \hat{b}_j \hat{b}_k \frac{\partial}{\partial t} (\hat{b}_i \hat{b}_j \hat{b}_k)}_{=0} + \frac{\partial q_{\perp}}{\partial t} \underbrace{(1 + 1 + 1 - 3)}_{=0} \\
&\quad + q_{\perp} \left[ \underbrace{\hat{b}_k \frac{\partial}{\partial t} \hat{b}_k}_{=0} + \underbrace{\hat{b}_i \frac{\partial}{\partial t} \hat{b}_i}_{=0} + \underbrace{\hat{b}_j \frac{\partial}{\partial t} \hat{b}_j}_{=0} - 3 \hat{b}_i \hat{b}_j \hat{b}_k \underbrace{\frac{\partial}{\partial t} (\hat{b}_i \hat{b}_j \hat{b}_k)}_{=0} \right] \\
&= \frac{\partial q_{\parallel}}{\partial t};
\end{aligned} \tag{7.65}$$

the second term calculates as

$$\begin{aligned}
(\nabla \cdot (\mathbf{u} \mathbf{q}^g) : \hat{\mathbf{b}}\hat{\mathbf{b}}) \cdot \hat{\mathbf{b}} &= \partial_l (u_l q_{ijk}^g) \hat{b}_i \hat{b}_j \hat{b}_k = (\partial_l u_l) \underbrace{q_{ijk}^g \hat{b}_i \hat{b}_j \hat{b}_k}_{=q_{\parallel}} + u_l \underbrace{(\partial_l q_{ijk}^g) \hat{b}_i \hat{b}_j \hat{b}_k}_{=\partial_l q_{\parallel}} \\
&= q_{\parallel} \nabla \cdot \mathbf{u} + \mathbf{u} \cdot \nabla q_{\parallel} = \nabla \cdot (q_{\parallel} \mathbf{u});
\end{aligned} \tag{7.66}$$

$$((\nabla \cdot \mathbf{r}) : \hat{\mathbf{b}}\hat{\mathbf{b}}) \cdot \hat{\mathbf{b}} = (\partial_l r_{ljk}) \hat{b}_i \hat{b}_j \hat{b}_k = \text{unchanged}. \tag{7.67}$$

Note that in all of the expressions here, the operator  $\cdot \hat{\mathbf{b}}$  can be naturally moved to the left as  $\hat{\mathbf{b}} \cdot$ . The third term calculates as



$$\begin{aligned}
((\mathbf{q}^g \cdot \nabla \mathbf{u}) : \hat{\mathbf{b}}\hat{\mathbf{b}}) \cdot \hat{\mathbf{b}} &= (\mathbf{q}^g \cdot \nabla \mathbf{u})_{ijk} \hat{b}_i \hat{b}_j \hat{b}_k = q_{ijl}^g (\partial_l u_k) \hat{b}_i \hat{b}_j \hat{b}_k \\
&= (\partial_l u_k) [q_{\parallel} \hat{b}_i \hat{b}_j \hat{b}_l + q_{\perp} (\delta_{ij} \hat{b}_l + \delta_{jl} \hat{b}_i + \delta_{li} \hat{b}_j - 3 \hat{b}_i \hat{b}_j \hat{b}_l)] \hat{b}_i \hat{b}_j \hat{b}_k; \\
&= (\partial_l u_k) [q_{\parallel} \hat{b}_i \hat{b}_k + q_{\perp} \underbrace{(\hat{b}_l \hat{b}_k + \hat{b}_l \hat{b}_k + \hat{b}_l \hat{b}_k - 3 \hat{b}_l \hat{b}_k)}_{=0}] \\
&= q_{\parallel} \hat{b}_l (\partial_l u_k) \hat{b}_k = q_{\parallel} \hat{\mathbf{b}} \cdot \nabla \mathbf{u} \cdot \hat{\mathbf{b}};
\end{aligned} \tag{7.68}$$

$$\begin{aligned}
[\mathbf{q}^g \cdot \nabla \mathbf{u}]_{ijk}^s \hat{b}_i \hat{b}_j \hat{b}_k &= [q_{ijl}^g \partial_l u_k + q_{jkl}^g \partial_l u_i + q_{kil}^g \partial_l u_j] \hat{b}_i \hat{b}_j \hat{b}_k \\
&= \underbrace{q_{ijl}^g \hat{b}_i \hat{b}_j (\partial_l u_k) \hat{b}_k}_{=q_{\parallel} \hat{b}_l} + \underbrace{q_{jkl}^g \hat{b}_j \hat{b}_k (\partial_l u_i) \hat{b}_i}_{=q_{\parallel} \hat{b}_l} + \underbrace{q_{kil}^g \hat{b}_k \hat{b}_i (\partial_l u_j) \hat{b}_j}_{=q_{\parallel} \hat{b}_l} \\
&= 3q_{\parallel} \hat{\mathbf{b}} \cdot \nabla \mathbf{u} \cdot \hat{\mathbf{b}}.
\end{aligned} \tag{7.69}$$

For the final fourth term we will need

$$\begin{aligned}
(\nabla \cdot \mathbf{p}) \cdot \hat{\mathbf{b}} &= \hat{b}_k \partial_l p_{lk} = \hat{b}_k \partial_l [(p_{\parallel} - p_{\perp}) \hat{b}_l \hat{b}_k + p_{\perp} \delta_{lk} + \Pi_{lk}] \\
&= \hat{b}_k [\hat{b}_l \hat{b}_k \partial_l (p_{\parallel} - p_{\perp}) + (p_{\parallel} - p_{\perp}) \partial_l (\hat{b}_l \hat{b}_k) + \partial_k p_{\perp} + \partial_l \Pi_{lk}] \\
&= [\hat{b}_l \partial_l (p_{\parallel} - p_{\perp}) + (p_{\parallel} - p_{\perp}) \underbrace{\hat{b}_k \partial_l (\hat{b}_l \hat{b}_k)}_{=\nabla \cdot \hat{\mathbf{b}}} + \hat{b}_k \partial_k p_{\perp} + (\partial_l \Pi_{lk}) \hat{b}_k] \\
&= \hat{\mathbf{b}} \cdot \nabla p_{\parallel} + (p_{\parallel} - p_{\perp}) \nabla \cdot \hat{\mathbf{b}} + (\nabla \cdot \Pi) \cdot \hat{\mathbf{b}},
\end{aligned} \tag{7.70}$$

and the fourth term calculates as

$$\begin{aligned}
[\mathbf{p}(\nabla \cdot \mathbf{p})]_{ijk} \hat{b}_i \hat{b}_j \hat{b}_k &= p_{ij} (\partial_l p_{lk}) \hat{b}_i \hat{b}_j \hat{b}_k = \underbrace{p_{ij} \hat{b}_i \hat{b}_j}_{=p_{\parallel}} \hat{b}_k \partial_l p_{lk} = p_{\parallel} \hat{b}_k \partial_l p_{lk} \\
&= p_{\parallel} \hat{\mathbf{b}} \cdot \nabla p_{\parallel} + p_{\parallel} (p_{\parallel} - p_{\perp}) \nabla \cdot \hat{\mathbf{b}} + p_{\parallel} (\nabla \cdot \Pi) \cdot \hat{\mathbf{b}};
\end{aligned} \tag{7.71}$$

$$\begin{aligned}
[\mathbf{p}(\nabla \cdot \mathbf{p})]_{ijk}^s \hat{b}_i \hat{b}_j \hat{b}_k &= [p_{ij} \partial_l p_{lk} + p_{jk} \partial_l p_{li} + p_{ki} \partial_l p_{lj}] \hat{b}_i \hat{b}_j \hat{b}_k \\
&= \underbrace{p_{ij} \hat{b}_i \hat{b}_j}_{=p_{\parallel}} \hat{b}_k \partial_l p_{lk} + \underbrace{p_{jk} \hat{b}_j \hat{b}_k}_{=p_{\parallel}} \hat{b}_i \partial_l p_{li} + \underbrace{p_{ki} \hat{b}_k \hat{b}_i}_{=p_{\parallel}} \hat{b}_j \partial_l p_{lj} = 3p_{\parallel} \hat{b}_k \partial_l p_{lk};
\end{aligned} \tag{7.72}$$

$$-\frac{1}{\rho} [\mathbf{p}(\nabla \cdot \mathbf{p})]_{ijk}^s \hat{b}_i \hat{b}_j \hat{b}_k = -3 \frac{p_{\parallel}}{\rho} \hat{\mathbf{b}} \cdot \nabla p_{\parallel} - 3 \frac{p_{\parallel}}{\rho} (p_{\parallel} - p_{\perp}) \nabla \cdot \hat{\mathbf{b}} - 3 \frac{p_{\parallel}}{\rho} (\nabla \cdot \Pi) \cdot \hat{\mathbf{b}}. \tag{7.73}$$

Combining all the terms yields the equation for the scalar parallel heat flux

$$\begin{aligned}
\frac{\partial q_{\parallel}}{\partial t} + \nabla \cdot (q_{\parallel} \mathbf{u}) + \hat{\mathbf{b}} \cdot (\nabla \cdot \mathbf{r}) : \hat{\mathbf{b}}\hat{\mathbf{b}} + 3q_{\parallel} \hat{\mathbf{b}} \cdot \nabla \mathbf{u} \cdot \hat{\mathbf{b}} - 3 \frac{p_{\parallel}}{\rho} \hat{\mathbf{b}} \cdot \nabla p_{\parallel} - 3 \frac{p_{\parallel}}{\rho} (p_{\parallel} - p_{\perp}) \nabla \cdot \hat{\mathbf{b}} \\
- 3 \frac{p_{\parallel}}{\rho} (\nabla \cdot \Pi) \cdot \hat{\mathbf{b}} + Q_{\parallel}^{\text{ng}} = 0,
\end{aligned} \tag{7.74}$$

where  $Q_{\parallel}^{\text{ng}} \equiv (\mathbf{Q}^{\text{ng}} : \hat{\mathbf{b}}\hat{\mathbf{b}}) \cdot \hat{\mathbf{b}}$ . The heat flux equations contain the fourth-order moment  $r_{ijkl}$ , which we will consider in the next section and at this stage it is unspecified.

## 7.3. Perpendicular heat flux equation

We will apply the trace operator to (7.34) and also perform  $\hat{\mathbf{b}}$  step by step. In the index notation, we are basically multiplying the entire equation by  $\delta_{ij}\hat{b}_k$ . The first term calculates

$$\text{Tr} \frac{\partial}{\partial t} \mathbf{q}^g = \frac{\partial}{\partial t} \text{Tr} \mathbf{q}^g = \frac{\partial}{\partial t} (q_{\parallel} \hat{\mathbf{b}} + 2q_{\perp} \hat{\mathbf{b}}) = \hat{\mathbf{b}} \frac{\partial}{\partial t} (q_{\parallel} + 2q_{\perp}) + (q_{\parallel} + 2q_{\perp}) \frac{\partial \hat{\mathbf{b}}}{\partial t}; \quad (7.75)$$

$$\left( \text{Tr} \frac{\partial}{\partial t} \mathbf{q}^g \right) \cdot \hat{\mathbf{b}} = \underbrace{\hat{\mathbf{b}} \cdot \hat{\mathbf{b}}}_{=1} \frac{\partial}{\partial t} (q_{\parallel} + 2q_{\perp}) + (q_{\parallel} + 2q_{\perp}) \underbrace{\frac{\partial \hat{\mathbf{b}}}{\partial t} \cdot \hat{\mathbf{b}}}_{=0} = \frac{\partial}{\partial t} (q_{\parallel} + 2q_{\perp}). \quad (7.76)$$

The second term  $\nabla \cdot (\mathbf{r} + \mathbf{u} \mathbf{q}^g)$  calculates as

$$[\text{Tr} \nabla \cdot (\mathbf{u} \mathbf{q}^g)]_k = \delta_{ij} \partial_l (u_l q_{ijk}^g) = \partial_l (u_l q_{lik}^g); \quad (7.77)$$

$$\begin{aligned} [\text{Tr} \nabla \cdot (\mathbf{u} \mathbf{q}^g)] \cdot \hat{\mathbf{b}} &= \partial_l (u_l q_{lik}^g) \hat{b}_k = \partial_l (u_l q_{lik}^g \hat{b}_k) - u_l \underbrace{q_{lik}^g \partial_l \hat{b}_k}_{=0} = \partial_l [u_l (q_{\parallel} + 2q_{\perp})] \\ &= (q_{\parallel} + 2q_{\perp}) \nabla \cdot \mathbf{u} + \mathbf{u} \cdot \nabla (q_{\parallel} + 2q_{\perp}); \end{aligned} \quad (7.78)$$

$$[\text{Tr} \nabla \cdot \mathbf{r}]_k = \delta_{ij} \partial_l r_{lijk} = \partial_l r_{lik}; \quad (7.79)$$

$$[\text{Tr} \nabla \cdot \mathbf{r}] \cdot \hat{\mathbf{b}} = (\partial_l r_{lik}) \hat{b}_k = \text{unchanged}. \quad (7.80)$$

The components  $r_{lik} = r_{ikl}$  are nothing else but the trace of the moment  $\mathbf{r}$

$$[\text{Tr} \mathbf{r}]_{kl} = \delta_{ij} r_{ijkl} = r_{ikl}. \quad (7.81)$$

The third term in (7.34) calculates as

$$[\mathbf{q}^g \cdot \nabla \mathbf{u}]_{ijk}^S = q_{ijl}^g \partial_l u_k + q_{jkl}^g \partial_l u_i + q_{kil}^g \partial_l u_j; \quad (7.82)$$

$$\text{Tr} [\mathbf{q}^g \cdot \nabla \mathbf{u}]_{ijk}^S = q_{iil}^g \partial_l u_k + q_{ikl}^g \partial_l u_i + \underbrace{q_{kil}^g}_{=q_{ikl}^g} \partial_l u_i = q_{iil}^g \partial_l u_k + 2q_{ikl}^g \partial_l u_i \quad (7.83)$$

$$\begin{aligned} \text{Tr} [\mathbf{q}^g \cdot \nabla \mathbf{u}]_{ijk}^S \hat{b}_k &= q_{iil}^g (\partial_l u_k) \hat{b}_k + 2q_{ikl}^g \hat{b}_k \partial_l u_i \\ &= (q_{\parallel} + 2q_{\perp}) \hat{b}_l (\partial_l u_k) \hat{b}_k + 2[q_{\parallel} \hat{b}_i \hat{b}_l + q_{\perp} (\delta_{il} - \hat{b}_i \hat{b}_l)] \partial_l u_i \\ &= (q_{\parallel} + 2q_{\perp}) \hat{b}_l (\partial_l u_k) \hat{b}_k + 2(q_{\parallel} - q_{\perp}) \hat{b}_l (\partial_l u_i) \hat{b}_i + 2q_{\perp} \partial_i u_i \\ &= 3q_{\parallel} \hat{\mathbf{b}} \cdot \nabla \mathbf{u} \cdot \hat{\mathbf{b}} + 2q_{\perp} \nabla \cdot \mathbf{u}, \end{aligned} \quad (7.84)$$

where we have used that  $q_{iil}^g = (q_{\parallel} + 2q_{\perp}) \hat{b}_l$  and  $q_{ikl}^g \hat{b}_k = q_{\parallel} \hat{b}_i \hat{b}_l + q_{\perp} (\delta_{il} - \hat{b}_i \hat{b}_l)$ . The fourth term in (7.34) calculates as

$$[\mathbf{p}(\nabla \cdot \mathbf{p})]_{ijk}^S = p_{ij} \partial_l p_{lk} + p_{jk} \partial_l p_{li} + p_{ki} \partial_l p_{lj}; \quad (7.85)$$

$$\text{Tr} [\mathbf{p}(\nabla \cdot \mathbf{p})]_{ijk}^S = p_{ii} \partial_l p_{lk} + p_{ik} \partial_l p_{li} + \underbrace{p_{ki}}_{=p_{ik}} \partial_l p_{li} = p_{ii} \partial_l p_{lk} + 2p_{ik} \partial_l p_{li}; \quad (7.86)$$

$$\text{Tr} [\mathbf{p}(\nabla \cdot \mathbf{p})]_{ijk}^S \hat{b}_k = p_{ii} \hat{b}_k \partial_l p_{lk} + 2p_{ik} \hat{b}_k \partial_l p_{li}, \quad (7.87)$$

and we have to calculate each term separately. Since  $p_{ii} = p_{\parallel} + 2p_{\perp}$  and by using the already calculated equation (7.70) for  $\hat{b}_k \partial_l p_{lk}$  and identities  $p_{ik} \hat{b}_k = p_{\parallel} \hat{b}_i + \Pi_{ik} \hat{b}_k$  and  $\Pi_{ik} \hat{b}_i \hat{b}_k = 0$  one obtains

$$p_{ii} \hat{b}_k \partial_l p_{lk} = (p_{\parallel} + 2p_{\perp})[\hat{\mathbf{b}} \cdot \nabla p_{\parallel} + (p_{\parallel} - p_{\perp})\nabla \cdot \hat{\mathbf{b}} + (\nabla \cdot \Pi) \cdot \hat{\mathbf{b}}]; \quad (7.88)$$

$$\begin{aligned} p_{ik} \hat{b}_k \partial_l p_{li} &= (p_{\parallel} \hat{b}_i + \Pi_{ik} \hat{b}_k)[\partial_i p_{\perp} + (p_{\parallel} - p_{\perp})(\hat{b}_i \partial_l \hat{b}_l + \hat{b}_l \partial_l \hat{b}_i) + \hat{b}_i \hat{b}_l \partial_l (p_{\parallel} - p_{\perp}) + \partial_l \Pi_{li}] \\ &= p_{\parallel}[\hat{\mathbf{b}} \cdot \nabla p_{\parallel} + (p_{\parallel} - p_{\perp})\nabla \cdot \hat{\mathbf{b}} + (\nabla \cdot \Pi) \cdot \hat{\mathbf{b}}] \\ &\quad + \Pi_{ik} \hat{b}_k [\partial_i p_{\perp} + (p_{\parallel} - p_{\perp})\hat{b}_l \partial_l \hat{b}_i + \partial_l \Pi_{li}] \\ &= p_{\parallel}[\hat{\mathbf{b}} \cdot \nabla p_{\parallel} + (p_{\parallel} - p_{\perp})\nabla \cdot \hat{\mathbf{b}} + (\nabla \cdot \Pi) \cdot \hat{\mathbf{b}}] \\ &\quad + [\nabla p_{\perp} + (p_{\parallel} - p_{\perp})\hat{\mathbf{b}} \cdot \nabla \hat{\mathbf{b}} + \nabla \cdot \Pi] \cdot \Pi \cdot \hat{\mathbf{b}}; \end{aligned} \quad (7.89)$$

and the final result of (7.87) is

$$\begin{aligned} \text{Tr}[p(\nabla \cdot \mathbf{p})]_{ijk}^S \hat{b}_k &= (3p_{\parallel} + 2p_{\perp})[\hat{\mathbf{b}} \cdot \nabla p_{\parallel} + (p_{\parallel} - p_{\perp})\nabla \cdot \hat{\mathbf{b}} + (\nabla \cdot \Pi) \cdot \hat{\mathbf{b}}] \\ &\quad + 2[\nabla p_{\perp} + (p_{\parallel} - p_{\perp})\hat{\mathbf{b}} \cdot \nabla \hat{\mathbf{b}} + \nabla \cdot \Pi] \cdot \Pi \cdot \hat{\mathbf{b}}. \end{aligned} \quad (7.90)$$

Collecting all the terms, one obtains

$$\begin{aligned} \frac{\partial}{\partial t}(q_{\parallel} + 2q_{\perp}) + (q_{\parallel} + 2q_{\perp})\nabla \cdot \mathbf{u} + \mathbf{u} \cdot \nabla(q_{\parallel} + 2q_{\perp}) + (\text{Tr} \nabla \cdot \mathbf{r}) \cdot \hat{\mathbf{b}} + 3q_{\parallel} \hat{\mathbf{b}} \cdot \nabla \mathbf{u} \cdot \hat{\mathbf{b}} \\ + 2q_{\perp} \nabla \cdot \mathbf{u} - \frac{1}{\rho}(3p_{\parallel} + 2p_{\perp})[\hat{\mathbf{b}} \cdot \nabla p_{\parallel} + (p_{\parallel} - p_{\perp})\nabla \cdot \hat{\mathbf{b}} + (\nabla \cdot \Pi) \cdot \hat{\mathbf{b}}] \\ - \frac{2}{\rho}[\nabla p_{\perp} + (p_{\parallel} - p_{\perp})\hat{\mathbf{b}} \cdot \nabla \hat{\mathbf{b}} + \nabla \cdot \Pi] \cdot \Pi \cdot \hat{\mathbf{b}} + (Q_{\parallel}^{\text{ng}} + 2Q_{\perp}^{\text{ng}}) = 0, \end{aligned} \quad (7.91)$$

and subtracting the parallel heat flux equation (7.74) and dividing by two yields the perpendicular heat flux equation

$$\begin{aligned} \frac{\partial}{\partial t} q_{\perp} + \mathbf{u} \cdot \nabla q_{\perp} + 2q_{\perp} \nabla \cdot \mathbf{u} + \frac{1}{2}[(\text{Tr} \nabla \cdot \mathbf{r}) \cdot \hat{\mathbf{b}} - \hat{\mathbf{b}} \cdot (\nabla \cdot \mathbf{r}) : \hat{\mathbf{b}} \hat{\mathbf{b}}] \\ - \frac{p_{\perp}}{\rho}[\hat{\mathbf{b}} \cdot \nabla p_{\parallel} + (p_{\parallel} - p_{\perp})\nabla \cdot \hat{\mathbf{b}} + (\nabla \cdot \Pi) \cdot \hat{\mathbf{b}}] \\ - \frac{1}{\rho}[\nabla p_{\perp} + (p_{\parallel} - p_{\perp})\hat{\mathbf{b}} \cdot \nabla \hat{\mathbf{b}} + \nabla \cdot \Pi] \cdot \Pi \cdot \hat{\mathbf{b}} + Q_{\perp}^{\text{ng}} = 0, \end{aligned} \quad (7.92)$$

where  $Q_{\perp}^{\text{ng}} \equiv (Q^{\text{ng}} : (\mathbf{I} - \hat{\mathbf{b}} \hat{\mathbf{b}})/2) \cdot \hat{\mathbf{b}}$ . The term containing  $\mathbf{r}$  in the above equation can be rewritten to many possible forms, for example

$$\begin{aligned} \frac{1}{2}[(\text{Tr} \nabla \cdot \mathbf{r}) \cdot \hat{\mathbf{b}} - \hat{\mathbf{b}} \cdot (\nabla \cdot \mathbf{r}) : \hat{\mathbf{b}} \hat{\mathbf{b}}] &= \frac{1}{2}[(\partial_l r_{lik}) \hat{b}_k - (\partial_l r_{lik}) \hat{b}_i \hat{b}_j \hat{b}_k] \\ &= \frac{1}{2}[(\text{Tr} \nabla \cdot \mathbf{r}) - (\nabla \cdot \mathbf{r}) : \hat{\mathbf{b}} \hat{\mathbf{b}}] \cdot \hat{\mathbf{b}} \\ &= \hat{\mathbf{b}} \cdot (\nabla \cdot \mathbf{r}) : (\mathbf{I} - \hat{\mathbf{b}} \hat{\mathbf{b}})/2. \end{aligned} \quad (7.93)$$

The last expression is of course consistent with an alternative way to derive the perpendicular heat flux equation, where instead of doing trace of the heat flux tensor equation and subtracting the scalar parallel heat flux equation, one can directly apply operators  $:(\mathbf{I} - \hat{\mathbf{b}} \hat{\mathbf{b}})/2$  and  $\cdot \hat{\mathbf{b}}$ .

## 7.4. Scalar heat flux equations continued

The parallel and perpendicular heat flux equations (7.74), (7.92) contain expressions for the fourth-order moment  $\mathbf{r}$ , and even though we will consider this moment in detail in the next section, here we want to finish the derivation of the scalar heat flux equations, and we write down the required expressions. Similarly to the pressure tensor and the heat flux tensor, the fourth-order moment can be decomposed into its gyrotropic and non-gyrotropic parts,  $\mathbf{r} = \mathbf{r}^g + \mathbf{r}^{ng}$ . The gyrotropic part  $\mathbf{r}^g$  is decomposed according to (8.10), and it can be shown (see later in the text), that direct calculation yields

$$\hat{\mathbf{b}} \cdot (\nabla \cdot \mathbf{r}^g) : \hat{\mathbf{b}}\hat{\mathbf{b}} = \nabla \cdot (r_{\parallel\parallel} \hat{\mathbf{b}}) - 3r_{\parallel\perp} \nabla \cdot \hat{\mathbf{b}}; \quad (7.94)$$

$$\frac{1}{2}[(\text{Tr} \nabla \cdot \mathbf{r}^g) \cdot \hat{\mathbf{b}} - \hat{\mathbf{b}} \cdot (\nabla \cdot \mathbf{r}^g) : \hat{\mathbf{b}}\hat{\mathbf{b}}] = \nabla \cdot (r_{\parallel\perp} \hat{\mathbf{b}}) + (r_{\parallel\perp} - r_{\perp\perp}) \nabla \cdot \hat{\mathbf{b}}, \quad (7.95)$$

which yields the parallel heat flux equation

$$\begin{aligned} \frac{\partial q_{\parallel}}{\partial t} + \nabla \cdot (q_{\parallel} \mathbf{u}) + \nabla \cdot (r_{\parallel\parallel} \hat{\mathbf{b}}) - 3r_{\parallel\perp} \nabla \cdot \hat{\mathbf{b}} + 3q_{\parallel} \hat{\mathbf{b}} \cdot \nabla \mathbf{u} \cdot \hat{\mathbf{b}} - 3 \frac{p_{\parallel}}{\rho} \hat{\mathbf{b}} \cdot \nabla p_{\parallel} \\ - 3 \frac{p_{\parallel}}{\rho} (p_{\parallel} - p_{\perp}) \nabla \cdot \hat{\mathbf{b}} + \hat{\mathbf{b}} \cdot (\nabla \cdot \mathbf{r}^{ng}) : \hat{\mathbf{b}}\hat{\mathbf{b}} - 3 \frac{p_{\parallel}}{\rho} (\nabla \cdot \Pi) \cdot \hat{\mathbf{b}} + Q_{\parallel}^{ng} = 0, \end{aligned} \quad (7.96)$$

and the perpendicular heat flux equation

$$\begin{aligned} \frac{\partial q_{\perp}}{\partial t} + \mathbf{u} \cdot \nabla q_{\perp} + 2q_{\perp} \nabla \cdot \mathbf{u} + \nabla \cdot (r_{\parallel\perp} \hat{\mathbf{b}}) + (r_{\parallel\perp} - r_{\perp\perp}) \nabla \cdot \hat{\mathbf{b}} \\ - \frac{p_{\perp}}{\rho} [\hat{\mathbf{b}} \cdot \nabla p_{\parallel} + (p_{\parallel} - p_{\perp}) \nabla \cdot \hat{\mathbf{b}}] + \frac{1}{2}[(\text{Tr} \nabla \cdot \mathbf{r}^{ng}) \cdot \hat{\mathbf{b}} - \hat{\mathbf{b}} \cdot (\nabla \cdot \mathbf{r}^{ng}) : \hat{\mathbf{b}}\hat{\mathbf{b}}] \\ - \frac{p_{\perp}}{\rho} (\nabla \cdot \Pi) \cdot \hat{\mathbf{b}} - \frac{1}{\rho} [\nabla p_{\perp} + (p_{\parallel} - p_{\perp}) \hat{\mathbf{b}} \cdot \nabla \hat{\mathbf{b}} + \nabla \cdot \Pi] \cdot \Pi \cdot \hat{\mathbf{b}} + Q_{\perp}^{ng} = 0. \end{aligned} \quad (7.97)$$

Exact nonlinear expressions for  $Q_{\parallel}^{ng}$  and  $Q_{\perp}^{ng}$  that represent contributions from the non-gyrotropic heat flux  $\mathbf{q}^{ng}$  are calculated in appendix D, see (D 110), (D 120). The scalar heat flux equations (7.96) and (7.97) are completely general at this stage, since no distribution function has been prescribed yet, and no simplification has been introduced. These equations are exact.

Nevertheless, equations (7.96), (7.97) are very complicated, and at this stage, it is beneficial to simplify them. One possibility is to cancel all the non-gyrotropic contributions  $\Pi$ ,  $\mathbf{q}^{ng}$  and  $\mathbf{r}^{ng}$ , and we will study such fluid models later. Another possibility is to keep only those non-gyrotropic terms that have some non-zero contribution at the linear level. This eliminates the last term in the second line of (7.97) that is proportional to  $[\dots] \cdot \Pi \cdot \hat{\mathbf{b}}$ . Also, by following the derivations in appendix D, it is easy to show that the terms  $Q_{\parallel}^{ng}$ ,  $Q_{\perp}^{ng}$  (that represent the non-gyrotropic heat flux  $\mathbf{q}^{ng}$ ), do not contribute at the linear level. Therefore, the heat flux equations simplify

$$\begin{aligned} \frac{\partial q_{\parallel}}{\partial t} + \nabla \cdot (q_{\parallel} \mathbf{u}) + \nabla \cdot (r_{\parallel\parallel} \hat{\mathbf{b}}) - 3r_{\parallel\perp} \nabla \cdot \hat{\mathbf{b}} + 3q_{\parallel} \hat{\mathbf{b}} \cdot \nabla \mathbf{u} \cdot \hat{\mathbf{b}} - 3 \frac{p_{\parallel}}{\rho} \hat{\mathbf{b}} \cdot \nabla p_{\parallel} \\ - 3 \frac{p_{\parallel}}{\rho} (p_{\parallel} - p_{\perp}) \nabla \cdot \hat{\mathbf{b}} + \hat{\mathbf{b}} \cdot (\nabla \cdot \mathbf{r}^{ng}) : \hat{\mathbf{b}}\hat{\mathbf{b}} - 3 \frac{p_{\parallel}}{\rho} (\nabla \cdot \Pi) \cdot \hat{\mathbf{b}} = 0; \end{aligned} \quad (7.98)$$

$$\begin{aligned}
& \frac{\partial q_{\perp}}{\partial t} + \mathbf{u} \cdot \nabla q_{\perp} + 2q_{\perp} \nabla \cdot \mathbf{u} + \nabla \cdot (r_{\parallel\perp} \hat{\mathbf{b}}) + (r_{\parallel\perp} - r_{\perp\perp}) \nabla \cdot \hat{\mathbf{b}} \\
& - \frac{p_{\perp}}{\rho} [\hat{\mathbf{b}} \cdot \nabla p_{\parallel} + (p_{\parallel} - p_{\perp}) \nabla \cdot \hat{\mathbf{b}}] + \frac{1}{2} [(\text{Tr} \nabla \cdot \mathbf{r}^{\text{ng}}) \cdot \hat{\mathbf{b}} - \hat{\mathbf{b}} \cdot (\nabla \cdot \mathbf{r}^{\text{ng}}) : \hat{\mathbf{b}} \hat{\mathbf{b}}] \\
& - \frac{p_{\perp}}{\rho} (\nabla \cdot \Pi) \cdot \hat{\mathbf{b}} = 0.
\end{aligned} \tag{7.99}$$

Still, no specific distribution function was assumed. However, to evaluate the non-gyrotropic  $\mathbf{r}^{\text{ng}}$ , and correctly evaluate possible cancellations with terms containing  $\Pi$ , we need to use  $\mathbf{r}^{\text{ng}}$  decomposition (8.58). Importantly, the decomposition (8.58) is valid only for perturbations around a bi-Maxwellian distribution function. By keeping only terms that have non-zero contribution at the linear level, see equations (8.63), (8.65) later in the text, the  $\mathbf{r}^{\text{ng}}$  terms are evaluated as

$$\hat{\mathbf{b}} \cdot (\nabla \cdot \mathbf{r}^{\text{ng}}) : \hat{\mathbf{b}} \hat{\mathbf{b}} = \frac{3p_{\parallel}}{\rho} (\nabla \cdot \Pi) \cdot \hat{\mathbf{b}}; \tag{7.100}$$

$$\frac{1}{2} [(\text{Tr} \nabla \cdot \mathbf{r}^{\text{ng}}) \cdot \hat{\mathbf{b}} - \hat{\mathbf{b}} \cdot (\nabla \cdot \mathbf{r}^{\text{ng}}) : \hat{\mathbf{b}} \hat{\mathbf{b}}] = \frac{2p_{\perp}}{\rho} (\nabla \cdot \Pi) \cdot \hat{\mathbf{b}}. \tag{7.101}$$

Importantly, both non-gyrotropic contributions in the parallel heat flux equation (7.98) completely cancel out! The cancellation demonstrates the importance of keeping the non-gyrotropic  $\mathbf{r}^{\text{ng}}$ , if one wants to keep the non-gyrotropic  $\Pi$ . Also, there is a partial cancellation in the perpendicular heat flux equation (7.99). The heat flux equations therefore read

$$\begin{aligned}
& \frac{\partial q_{\parallel}}{\partial t} + \nabla \cdot (q_{\parallel} \mathbf{u}) + \nabla \cdot (r_{\parallel\parallel} \hat{\mathbf{b}}) - 3r_{\parallel\perp} \nabla \cdot \hat{\mathbf{b}} + 3q_{\parallel} \hat{\mathbf{b}} \cdot \nabla \mathbf{u} \cdot \hat{\mathbf{b}} - 3 \frac{p_{\parallel}}{\rho} \hat{\mathbf{b}} \cdot \nabla p_{\parallel} \\
& - 3 \frac{p_{\parallel}}{\rho} (p_{\parallel} - p_{\perp}) \nabla \cdot \hat{\mathbf{b}} = 0;
\end{aligned} \tag{7.102}$$

$$\begin{aligned}
& \frac{\partial q_{\perp}}{\partial t} + \mathbf{u} \cdot \nabla q_{\perp} + 2q_{\perp} \nabla \cdot \mathbf{u} + \nabla \cdot (r_{\parallel\perp} \hat{\mathbf{b}}) + (r_{\parallel\perp} - r_{\perp\perp}) \nabla \cdot \hat{\mathbf{b}} \\
& - \frac{p_{\perp}}{\rho} [\hat{\mathbf{b}} \cdot \nabla p_{\parallel} + (p_{\parallel} - p_{\perp}) \nabla \cdot \hat{\mathbf{b}}] + \frac{p_{\perp}}{\rho} (\nabla \cdot \Pi) \cdot \hat{\mathbf{b}} = 0.
\end{aligned} \tag{7.103}$$

The last term can be evaluated with respect to the mean magnetic field, and assuming that  $\hat{\mathbf{b}}_0$  is in the  $z$ -direction, the term is equal to  $(p_{\perp}/\rho)(\partial_x \Pi_{xz} + \partial_y \Pi_{yz})$ , since  $\Pi_{zz} = 0$ . Essentially, the term should be written in a fully linearized form  $(p_{\perp}^{(0)}/\rho_0)(\nabla \cdot \Pi) \cdot \hat{\mathbf{b}}_0$ , since other non-gyrotropic nonlinear terms were neglected. It is important to emphasize that it is because of this one term, that the equations (7.102), (7.103) are valid only for perturbations around a bi-Maxwellian distribution function. If the term is neglected, i.e. if one neglects from the beginning the non-gyrotropic contributions in (7.98), (7.99), the heat flux equations are valid for perturbations around any distribution function. We will consider a bi-kappa distribution function later.

### 7.5. Heat flux equations with ‘normal’ closure

Detailed calculations with the fourth-order moment will be presented in the next section. Here we want to finish the derivation, and by using the bi-Maxwellian ‘normal’ fluid closure

$$r_{\parallel\parallel} = \frac{3p_{\parallel}^2}{\rho}; \quad r_{\parallel\perp} = \frac{p_{\parallel} p_{\perp}}{\rho}; \quad r_{\perp\perp} = \frac{2p_{\perp}^2}{\rho}, \tag{7.104}$$

terms entering the heat flux equations directly calculate as

$$\nabla \cdot (r_{\parallel\parallel} \hat{\mathbf{b}}) - 3r_{\parallel\perp} \nabla \cdot \hat{\mathbf{b}} = 3 \frac{p_{\parallel}^2}{\rho} \nabla \cdot \hat{\mathbf{b}} + 3p_{\parallel} \hat{\mathbf{b}} \cdot \nabla \left( \frac{p_{\parallel}}{\rho} \right) + 3 \frac{p_{\parallel}}{\rho} \hat{\mathbf{b}} \cdot \nabla p_{\parallel} - 3 \frac{p_{\parallel} p_{\perp}}{\rho} \nabla \cdot \hat{\mathbf{b}}; \quad (7.105)$$

$$\nabla \cdot (r_{\parallel\perp} \hat{\mathbf{b}}) + (r_{\parallel\perp} - r_{\perp\perp}) \nabla \cdot \hat{\mathbf{b}} = \frac{p_{\perp}}{\rho} \hat{\mathbf{b}} \cdot \nabla p_{\parallel} + p_{\parallel} \hat{\mathbf{b}} \cdot \nabla \left( \frac{p_{\perp}}{\rho} \right) + 2 \frac{p_{\perp}}{\rho} (p_{\parallel} - p_{\perp}) \nabla \cdot \hat{\mathbf{b}}. \quad (7.106)$$

The use of these expressions in (7.102), (7.103) cancels various terms, and the scalar heat flux equations read

$$\frac{\partial q_{\parallel}}{\partial t} + \nabla \cdot (q_{\parallel} \mathbf{u}) + 3p_{\parallel} \hat{\mathbf{b}} \cdot \nabla \left( \frac{p_{\parallel}}{\rho} \right) + 3q_{\parallel} \hat{\mathbf{b}} \cdot \nabla \mathbf{u} \cdot \hat{\mathbf{b}} = 0; \quad (7.107)$$

$$\frac{\partial q_{\perp}}{\partial t} + \mathbf{u} \cdot \nabla q_{\perp} + 2q_{\perp} \nabla \cdot \mathbf{u} + p_{\parallel} \hat{\mathbf{b}} \cdot \nabla \left( \frac{p_{\perp}}{\rho} \right) + \frac{p_{\perp}}{\rho} (p_{\parallel} - p_{\perp}) \nabla \cdot \hat{\mathbf{b}} + \frac{p_{\perp}}{\rho} (\nabla \cdot \Pi) \cdot \hat{\mathbf{b}}_0 = 0. \quad (7.108)$$

## 8. Fourth-order fluid moment

The algebra of the fourth-order fluid moment  $\mathbf{r} = m \int \mathbf{c} \mathbf{c} \mathbf{c} \mathbf{c} f d^3 v$  can be quite intimidating at first, since the moment is a 4-D cube ( $3 \times 3 \times 3 \times 3$ ). We are not going to derive the time evolution equation for this moment step by step, nevertheless, later in the text we derive the evolution equation of the  $n$ th-order fluid moment  $\mathbf{X}^{(n)}$ , see (12.16), and therefore, for  $n = 4$ , the equation reads

$$\frac{\partial}{\partial t} \mathbf{r} + \nabla \cdot (\mathbf{X}^{(5)} + \mathbf{u} \mathbf{r}) + \left[ \mathbf{r} \cdot \nabla \mathbf{u} + \frac{q}{mc} \mathbf{B} \times \mathbf{r} - \frac{1}{\rho} (\nabla \cdot \mathbf{p}) \mathbf{q} \right]^S = 0. \quad (8.1)$$

The symmetric operator ‘ $S$ ’ is here defined as

$$A_{ijkl}^S = A_{ijkl} + A_{jkli} + A_{klij} + A_{lijk}. \quad (8.2)$$

Here, we are not that interested in the evolution equation for  $\mathbf{r}$ , we just want to clarify its decomposition, that we need in the heat flux equations. The definition of gyrotropy means that the integral has to be evaluated only over combinations of  $(v_{\parallel} - u_{\perp}) \equiv c_{\parallel}$  and  $|\mathbf{v}_{\perp} - \mathbf{u}_{\perp}|^2 \equiv c_{\perp}^2$ . For the fourth moment  $\mathbf{r}$ , there are obviously only 3 possibilities:  $c_{\parallel}^4$ ,  $c_{\parallel}^2 c_{\perp}^2$  and  $(c_{\perp}^2)^2$ , and these gyrotropic components will be called  $r_{\parallel\parallel}$ ,  $r_{\parallel\perp}$  and  $r_{\perp\perp}$ .

We have already seen that the double contractions with  $\hat{\mathbf{b}}\hat{\mathbf{b}}$  and  $(\mathbf{I} - \hat{\mathbf{b}}\hat{\mathbf{b}})/2$ , were very useful operators to extract the gyrotropic components for the lower-order moments,  $\mathbf{p}$  and  $\mathbf{q}$ . To obtain any scalar quantity from the fourth moment, we obviously need to apply two double contractions. There are 3 possibilities: we can apply  $\hat{\mathbf{b}}\hat{\mathbf{b}}$  twice, we can apply  $\hat{\mathbf{b}}\hat{\mathbf{b}}$  and  $(\mathbf{I} - \hat{\mathbf{b}}\hat{\mathbf{b}})/2$  or we can apply  $(\mathbf{I} - \hat{\mathbf{b}}\hat{\mathbf{b}})/2$  twice. Not surprisingly, these double contractions with  $\mathbf{r}$  indeed extract the 3 possible gyrotropic parts, as it is easy to verify that

$$\begin{aligned} (\mathbf{r} : \hat{\mathbf{b}}\hat{\mathbf{b}}) : \hat{\mathbf{b}}\hat{\mathbf{b}} &= r_{ijkl} \hat{b}_i \hat{b}_j \hat{b}_k \hat{b}_l = m \int c_i c_j c_k c_l f d^3 v \hat{b}_i \hat{b}_j \hat{b}_k \hat{b}_l = m \int c_i \hat{b}_i c_j \hat{b}_j c_k \hat{b}_k c_l \hat{b}_l f d^3 v \\ &= m \int c_{\parallel} c_{\parallel} c_{\parallel} c_{\parallel} f d^3 v = m \int (v_{\parallel} - u_{\parallel})^4 f d^3 v \equiv r_{\parallel\parallel}; \end{aligned} \quad (8.3)$$

$$\begin{aligned}
(\mathbf{r} : \hat{\mathbf{b}}\hat{\mathbf{b}}) : (\mathbf{I} - \hat{\mathbf{b}}\hat{\mathbf{b}})/2 &= \frac{m}{2} \int c_i c_j c_k c_l \hat{b}_i \hat{b}_j (\delta_{kl} - \hat{b}_k \hat{b}_l) f \, d^3 v = \frac{m}{2} \int c_{\parallel}^2 c_k c_l (\delta_{kl} - \hat{b}_k \hat{b}_l) f \, d^3 v \\
&= \frac{m}{2} \int c_{\parallel}^2 (|\mathbf{c}|^2 - c_{\parallel}^2) f \, d^3 v = \frac{m}{2} \int c_{\parallel}^2 c_{\perp}^2 f \, d^3 v \\
&= \frac{m}{2} \int (v_{\parallel} - u_{\parallel})^2 |\mathbf{v}_{\perp} - \mathbf{u}_{\perp}|^2 f \, d^3 v \equiv r_{\parallel\perp};
\end{aligned} \tag{8.4}$$

$$\begin{aligned}
(\mathbf{r} : (\mathbf{I} - \hat{\mathbf{b}}\hat{\mathbf{b}})/2) : (\mathbf{I} - \hat{\mathbf{b}}\hat{\mathbf{b}})/2 &= \frac{m}{4} \int c_i c_j c_k c_l (\delta_{ij} - \hat{b}_i \hat{b}_j) (\delta_{kl} - \hat{b}_k \hat{b}_l) f \, d^3 v = \frac{m}{4} \int c_{\perp}^2 c_{\perp}^2 f \, d^3 v \\
&= \frac{m}{4} \int |\mathbf{v}_{\perp} - \mathbf{u}_{\perp}|^4 f \, d^3 v \equiv r_{\perp\perp}.
\end{aligned} \tag{8.5}$$

We are now ready to guess how to write the decomposition of the gyrotropic fourth-order moment. Motivated by the previous decompositions, it obviously has to be something in the form of

$$r^g = r_{\parallel\parallel} \hat{\mathbf{b}}\hat{\mathbf{b}}\hat{\mathbf{b}}\hat{\mathbf{b}} + r_{\parallel\perp} [\hat{\mathbf{b}}\hat{\mathbf{b}}(\mathbf{I} - \hat{\mathbf{b}}\hat{\mathbf{b}})]^{\text{Sym}} + r_{\perp\perp} [(\mathbf{I} - \hat{\mathbf{b}}\hat{\mathbf{b}})(\mathbf{I} - \hat{\mathbf{b}}\hat{\mathbf{b}})]^{\text{Sym}}, \tag{8.6}$$

where we still did not determine how the symmetric operator acts here. Importantly, the symmetric operator ‘Sym’ is not equivalent to the symmetric operator ‘S’ that cycles all the indices around, equation (8.2). The fluid hierarchy obviously needs two symmetric operators, one unique ‘S’ that is used to derive the evolution equation of a given fluid moment  $\mathbf{X}^{(n)}$ , and one non-unique ‘Sym’ that is used for the decomposition of that fluid moment. The determination of how ‘Sym’ acts here is not that obvious. Nevertheless, one can consider in how many ways one can extract the gyrotropic components from  $r_{ijkl}$ . For  $r_{\parallel\parallel}$ , one does two double contractions with  $(\hat{\mathbf{b}}\hat{\mathbf{b}})$ . The possible choices are  $(\hat{\mathbf{b}}\hat{\mathbf{b}})_{ij}(\hat{\mathbf{b}}\hat{\mathbf{b}})_{kl}$ ;  $(\hat{\mathbf{b}}\hat{\mathbf{b}})_{ik}(\hat{\mathbf{b}}\hat{\mathbf{b}})_{jl}$  and  $(\hat{\mathbf{b}}\hat{\mathbf{b}})_{il}(\hat{\mathbf{b}}\hat{\mathbf{b}})_{jk}$ , however, all of these choices are equivalent. To obtain the  $r_{\parallel\perp}$ , we perform double contractions with  $(\hat{\mathbf{b}}\hat{\mathbf{b}})$  and  $(\mathbf{I} - \hat{\mathbf{b}}\hat{\mathbf{b}})/2$ , where the last operator contains a function  $\delta_{ij}$ . How many different delta functions we can obtain from 4 indices  $i, j, k, l$ ? There are  $\binom{4}{2} = 4!/2!(4-2)! = 6$  different possibilities:

$$\delta_{ij}; \quad \delta_{ik}; \quad \delta_{il}; \quad \delta_{jk}; \quad \delta_{jl}; \quad \delta_{kl}, \tag{8.7}$$

and all of them are non-equivalent. The symmetric operator acting in the second term therefore has 6 components

$$\begin{aligned}
[\hat{\mathbf{b}}\hat{\mathbf{b}}(\mathbf{I} - \hat{\mathbf{b}}\hat{\mathbf{b}})]_{ijkl}^{\text{Sym}} &= (\mathbf{I} - \hat{\mathbf{b}}\hat{\mathbf{b}})_{ij} \hat{b}_k \hat{b}_l + (\mathbf{I} - \hat{\mathbf{b}}\hat{\mathbf{b}})_{ik} \hat{b}_j \hat{b}_l + (\mathbf{I} - \hat{\mathbf{b}}\hat{\mathbf{b}})_{il} \hat{b}_j \hat{b}_k + (\mathbf{I} - \hat{\mathbf{b}}\hat{\mathbf{b}})_{jk} \hat{b}_i \hat{b}_l \\
&\quad + (\mathbf{I} - \hat{\mathbf{b}}\hat{\mathbf{b}})_{jl} \hat{b}_i \hat{b}_k + (\mathbf{I} - \hat{\mathbf{b}}\hat{\mathbf{b}})_{kl} \hat{b}_i \hat{b}_j \\
&= \delta_{ij} \hat{b}_k \hat{b}_l + \delta_{ik} \hat{b}_j \hat{b}_l + \delta_{il} \hat{b}_j \hat{b}_k + \delta_{jk} \hat{b}_i \hat{b}_l + \delta_{jl} \hat{b}_i \hat{b}_k + \delta_{kl} \hat{b}_i \hat{b}_j - 6 \hat{b}_i \hat{b}_j \hat{b}_k \hat{b}_l.
\end{aligned} \tag{8.8}$$

Lastly, to obtain the  $r_{\perp\perp}$  component, one applies two double contractions with  $(\mathbf{I} - \hat{\mathbf{b}}\hat{\mathbf{b}})/2$  that results in combinations of  $\delta_{ij}\delta_{kl}$ . How many possibilities do we have? Obviously, there are 6 possibilities for the first delta function, which, by pairing with the other delta function in a way that indices are not repeated, yields together 6 possibilities  $\delta_{ij}\delta_{kl}$ ;  $\delta_{ik}\delta_{jl}$ ;  $\delta_{il}\delta_{jk}$ ;  $\delta_{jk}\delta_{il}$ ;  $\delta_{jl}\delta_{ik}$ ;  $\delta_{kl}\delta_{ij}$ . However, 3 possibilities have an equivalent pair,  $\delta_{ij}\delta_{kl} = \delta_{kl}\delta_{ij}$ ;  $\delta_{ik}\delta_{jl} = \delta_{jl}\delta_{ik}$  and  $\delta_{il}\delta_{jk} = \delta_{jk}\delta_{il}$  and there are only 3

non-equivalent combinations. The symmetric operator acting on the last term in (8.6) can be determined to be

$$\begin{aligned}
 & [(I - \hat{\mathbf{b}}\hat{\mathbf{b}})(I - \hat{\mathbf{b}}\hat{\mathbf{b}})]_{ijkl}^{\text{Sym}} \\
 &= \frac{1}{2} [(I - \hat{\mathbf{b}}\hat{\mathbf{b}})_{ij}(I - \hat{\mathbf{b}}\hat{\mathbf{b}})_{kl} + (I - \hat{\mathbf{b}}\hat{\mathbf{b}})_{ik}(I - \hat{\mathbf{b}}\hat{\mathbf{b}})_{jl} + (I - \hat{\mathbf{b}}\hat{\mathbf{b}})_{il}(I - \hat{\mathbf{b}}\hat{\mathbf{b}})_{jk}] \\
 &= \frac{1}{2} \left[ \delta_{ij}\delta_{kl} - \delta_{kl}\hat{b}_i\hat{b}_j - \delta_{ij}\hat{b}_k\hat{b}_l + \delta_{ik}\delta_{jl} - \delta_{jl}\hat{b}_i\hat{b}_k - \delta_{ik}\hat{b}_j\hat{b}_l \right. \\
 &\quad \left. + \delta_{il}\delta_{jk} - \delta_{jk}\hat{b}_i\hat{b}_l - \delta_{il}\hat{b}_j\hat{b}_k + 3\hat{b}_i\hat{b}_j\hat{b}_k\hat{b}_l \right]. \tag{8.9}
 \end{aligned}$$

The factor 1/2 is actually not that obvious, and one needs to verify that the decomposition indeed satisfies (8.5). The entire fourth-order gyrotropic moment  $r^g$  is decomposed in the index notation according to

$$\begin{aligned}
 r_{ijkl}^g &= r_{\parallel} \hat{b}_i\hat{b}_j\hat{b}_k\hat{b}_l + r_{\perp\perp} [\delta_{ij}\hat{b}_k\hat{b}_l + \delta_{ik}\hat{b}_j\hat{b}_l + \delta_{il}\hat{b}_j\hat{b}_k + \delta_{jk}\hat{b}_i\hat{b}_l + \delta_{jl}\hat{b}_i\hat{b}_k + \delta_{kl}\hat{b}_i\hat{b}_j - 6\hat{b}_i\hat{b}_j\hat{b}_k\hat{b}_l] \\
 &\quad + \frac{r_{\perp\perp}}{2} \left[ \delta_{ij}\delta_{kl} + \delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk} - \delta_{ij}\hat{b}_k\hat{b}_l - \delta_{ik}\hat{b}_j\hat{b}_l - \delta_{il}\hat{b}_j\hat{b}_k - \delta_{jk}\hat{b}_i\hat{b}_l \right. \\
 &\quad \left. - \delta_{jl}\hat{b}_i\hat{b}_k - \delta_{kl}\hat{b}_i\hat{b}_j + 3\hat{b}_i\hat{b}_j\hat{b}_k\hat{b}_l \right]. \tag{8.10}
 \end{aligned}$$

It might be tempting to rearrange this decomposition to a more compact form

$$\begin{aligned}
 r_{ijkl}^g &= \left( r_{\parallel} - 6r_{\perp\perp} + \frac{3}{2}r_{\perp\perp} \right) \hat{b}_i\hat{b}_j\hat{b}_k\hat{b}_l \\
 &\quad + \left( r_{\perp\perp} - \frac{r_{\perp\perp}}{2} \right) [\delta_{ij}\hat{b}_k\hat{b}_l + \delta_{ik}\hat{b}_j\hat{b}_l + \delta_{il}\hat{b}_j\hat{b}_k + \delta_{jk}\hat{b}_i\hat{b}_l + \delta_{jl}\hat{b}_i\hat{b}_k + \delta_{kl}\hat{b}_i\hat{b}_j] \\
 &\quad + \frac{r_{\perp\perp}}{2} [\delta_{ij}\delta_{kl} + \delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk}], \tag{8.11}
 \end{aligned}$$

nevertheless, in actual calculations we find the form (8.10) to be more useful. It is important to verify that the decomposition (8.10) really works. A straightforward calculation yields

$$r_{ijkl}^g \hat{b}_i = r_{\parallel} \hat{b}_j\hat{b}_k\hat{b}_l + r_{\perp\perp} [\delta_{jk}\hat{b}_l + \delta_{jl}\hat{b}_k + \delta_{kl}\hat{b}_j - 3\hat{b}_j\hat{b}_k\hat{b}_l]; \tag{8.12}$$

$$r_{ijkl}^g \hat{b}_i\hat{b}_j = r_{\parallel} \hat{b}_k\hat{b}_l + r_{\perp\perp} [\delta_{kl} - \hat{b}_k\hat{b}_l]; \tag{8.13}$$

$$r_{ijkl}^g \hat{b}_i\hat{b}_j\hat{b}_k = r_{\parallel} \hat{b}_l; \tag{8.14}$$

$$r_{ijkl}^g \hat{b}_i\hat{b}_j\hat{b}_k\hat{b}_l = r_{\parallel}, \tag{8.15}$$

which is consistent with (8.3). The  $r_{\perp\perp}$  component calculates as

$$r_{ijkl}^g \hat{b}_i\hat{b}_j\delta_{kl} = r_{\parallel} + 2r_{\perp\perp}; \tag{8.16}$$

$$r_{ijkl}^g \hat{b}_i\hat{b}_j(\delta_{kl} - \hat{b}_k\hat{b}_l)/2 = r_{ijkl}^g \hat{b}_i\hat{b}_j\delta_{kl}/2 - r_{ijkl}^g \hat{b}_i\hat{b}_j\hat{b}_k\hat{b}_l/2 = r_{\parallel}/2 + r_{\perp\perp} - r_{\parallel}/2 = r_{\perp\perp}, \tag{8.17}$$

which is consistent with (8.4). The  $r_{\perp\perp}$  component calculates as

$$r_{ijkl}^g \delta_{ij} = r_{\parallel}^g = r_{\parallel} \hat{b}_k\hat{b}_l + r_{\perp\perp} (\delta_{kl} + \hat{b}_k\hat{b}_l) + 2r_{\perp\perp} (\delta_{kl} - \hat{b}_k\hat{b}_l); \tag{8.18}$$

$$r_{ijkl}^g \delta_{ij}\delta_{kl} = r_{\parallel}^g = r_{\parallel} + 4r_{\perp\perp} + 4r_{\perp\perp}; \tag{8.19}$$



$$r_{ijkl}^g \delta_{ij} \hat{b}_k \hat{b}_l = r_{\parallel\parallel} + 2r_{\parallel\perp}; \quad (8.20)$$

$$r_{ijkl}^g \delta_{ij} (\delta_{kl} - \hat{b}_k \hat{b}_l) = 2r_{\parallel\perp} + 4r_{\perp\perp}; \quad (8.21)$$

$$\begin{aligned} r_{ijkl}^g (\delta_{ij} - \hat{b}_i \hat{b}_j) (\delta_{kl} - \hat{b}_k \hat{b}_l) / 4 &= r_{ijkl}^g \delta_{ij} (\delta_{kl} - \hat{b}_k \hat{b}_l) / 4 - r_{ijkl}^g \hat{b}_i \hat{b}_j (\delta_{kl} - \hat{b}_k \hat{b}_l) / 4 \\ &= r_{\parallel\parallel} / 2 + r_{\perp\perp} - r_{\parallel\perp} / 2 = r_{\perp\perp}, \end{aligned} \quad (8.22)$$

which is consistent with (8.5). The decomposition (8.10) indeed works.

Now we need to verify expressions (7.94), (7.95) that were used in the scalar heat flux equations. The first expression calculates as

$$\begin{aligned} (\nabla \cdot \mathbf{r}^g)_{ijk} \hat{b}_i \hat{b}_j \hat{b}_k &= (\partial_l r_{ijkl}^g) \hat{b}_i \hat{b}_j \hat{b}_k = \partial_l (r_{ijkl}^g \hat{b}_i \hat{b}_j \hat{b}_k) - r_{ijkl}^g \partial_l (\hat{b}_i \hat{b}_j \hat{b}_k) \\ &= \partial_l (r_{\parallel\parallel} \hat{b}_l) - r_{ijkl}^g \partial_l (\hat{b}_i \hat{b}_j \hat{b}_k) \\ &= \nabla \cdot (r_{\parallel\parallel} \hat{\mathbf{b}}) - r_{ijkl}^g ((\partial_l \hat{b}_i) \hat{b}_j \hat{b}_k + \hat{b}_i (\partial_l \hat{b}_j) \hat{b}_k + \hat{b}_i \hat{b}_j (\partial_l \hat{b}_k)); \end{aligned} \quad (8.23)$$

$$r_{ijkl}^g \hat{b}_i \hat{b}_j \hat{b}_k = r_{\parallel\parallel} \hat{b}_i \hat{b}_l + r_{\parallel\perp} [\delta_{il} - \hat{b}_i \hat{b}_l]; \quad (8.24)$$

$$r_{ijkl}^g \hat{b}_i \hat{b}_j \hat{b}_k \partial_l \hat{b}_i = r_{\parallel\perp} \nabla \cdot \hat{\mathbf{b}}; \quad (8.25)$$

$$(\nabla \cdot \mathbf{r}^g)_{ijk} \hat{b}_i \hat{b}_j \hat{b}_k = \nabla \cdot (r_{\parallel\parallel} \hat{\mathbf{b}}) - 3r_{\parallel\perp} \nabla \cdot \hat{\mathbf{b}}, \quad (8.26)$$

and the second expression calculates similarly as

$$(\nabla \cdot \mathbf{r}^g)_{ijk} \delta_{ij} \hat{b}_k = (\partial_l r_{iikl}^g) \hat{b}_k = \partial_l (r_{iikl}^g \hat{b}_k) - r_{iikl}^g \partial_l \hat{b}_k; \quad (8.27)$$

$$r_{iikl}^g \hat{b}_k = r_{\parallel\parallel} \hat{b}_l + 2r_{\parallel\perp} \hat{b}_l; \quad (8.28)$$

$$r_{iikl}^g \partial_l \hat{b}_k = (r_{\parallel\perp} + 2r_{\perp\perp}) \nabla \cdot \hat{\mathbf{b}}; \quad (8.29)$$

$$(\nabla \cdot \mathbf{r}^g)_{ijk} \delta_{ij} \hat{b}_k = \nabla \cdot (r_{\parallel\parallel} \hat{\mathbf{b}}) + 2\nabla \cdot (r_{\parallel\perp} \hat{\mathbf{b}}) - (r_{\parallel\perp} + 2r_{\perp\perp}) \nabla \cdot \hat{\mathbf{b}}; \quad (8.30)$$

$$(\nabla \cdot \mathbf{r}^g)_{ijk} (\delta_{ij} - \hat{b}_i \hat{b}_j) \hat{b}_k / 2 = \nabla \cdot (r_{\parallel\perp} \hat{\mathbf{b}}) + (r_{\parallel\perp} - r_{\perp\perp}) \nabla \cdot \hat{\mathbf{b}}, \quad (8.31)$$

which verifies (7.94), (7.95).

Furthermore, by exploring the  $\partial \mathbf{r} / \partial t$  equation (8.1), at frequencies that are much smaller than the gyrofrequency, the gyrotropic part of  $\mathbf{r}$  should satisfy

$$(\hat{\mathbf{b}} \times \mathbf{r}^g)^S = 0, \quad (8.32)$$

in the same way that the gyrotropic parts of  $\mathbf{p}$  and  $\mathbf{q}$  satisfied this requirement. By using the gyrotropic (8.10), it is easy to calculate for example that

$$\begin{aligned} (\hat{\mathbf{b}} \times \mathbf{r}^g)_{ijkl} &= \epsilon_{irs} \hat{b}_r r_{sjkl}^g = r_{\parallel\perp} (\epsilon_{irj} \hat{b}_r \hat{b}_k \hat{b}_l + \epsilon_{irk} \hat{b}_r \hat{b}_j \hat{b}_l + \epsilon_{irl} \hat{b}_r \hat{b}_j \hat{b}_k) \\ &\quad + \frac{r_{\perp\perp}}{2} (\epsilon_{irj} \hat{b}_r \delta_{kl} + \epsilon_{irk} \hat{b}_r \delta_{jl} + \epsilon_{irl} \hat{b}_r \delta_{jk} - \epsilon_{irj} \hat{b}_r \hat{b}_k \hat{b}_l - \epsilon_{irk} \hat{b}_r \hat{b}_j \hat{b}_l - \epsilon_{irl} \hat{b}_r \hat{b}_j \hat{b}_k), \end{aligned} \quad (8.33)$$

and by adding together all the 4 representations of the ‘S’ operator, one can indeed verify that all the terms cancel, yielding (8.32). By decomposing the entire fourth-order moment into its gyrotropic and non-gyrotropic parts

$$\mathbf{r} = \mathbf{r}^g + \mathbf{r}^{\text{ng}}, \quad (8.34)$$

the evolution equation (8.1) can therefore be rewritten as

$$\frac{\partial}{\partial t} \mathbf{r} + \nabla \cdot (\mathbf{X}^{(5)} + \mathbf{ur}) + \left[ \mathbf{r} \cdot \nabla \mathbf{u} + \Omega \frac{|\mathbf{B}|}{B_0} \hat{\mathbf{b}} \times \mathbf{r}^{\text{ng}} - \frac{1}{\rho} (\nabla \cdot \mathbf{p}) \mathbf{q} \right]^S = 0. \quad (8.35)$$

8.1. Non-gyrotropic  $\mathbf{r}^{\text{ng}}$ 

Basic properties of the non-gyrotropic tensor  $\mathbf{r}^{\text{ng}}$  can be easily determined with a similar procedure as those we used for the non-gyrotropic pressure  $\Pi$  and the non-gyrotropic heat flux  $\mathbf{q}^{\text{ng}}$ . The decomposition of the entire fourth-order moment is

$$\mathbf{r} = r_{\parallel\parallel} \hat{\mathbf{b}}\hat{\mathbf{b}}\hat{\mathbf{b}}\hat{\mathbf{b}} + r_{\parallel\perp} [\hat{\mathbf{b}}\hat{\mathbf{b}}(\mathbf{I} - \hat{\mathbf{b}}\hat{\mathbf{b}})]^{\text{Sym}} + r_{\perp\perp} [(\mathbf{I} - \hat{\mathbf{b}}\hat{\mathbf{b}})(\mathbf{I} - \hat{\mathbf{b}}\hat{\mathbf{b}})]^{\text{Sym}} + \mathbf{r}^{\text{ng}}, \quad (8.36)$$

and the meaning of the ‘Sym’ operators was specified by (8.8), (8.9). By using definitions (8.3), (8.4), (8.5) for  $r_{\parallel\parallel}$ ,  $r_{\parallel\perp}$ ,  $r_{\perp\perp}$  in the expression above, the full decomposition reads

$$\begin{aligned} \mathbf{r} = & [(\mathbf{r} : \hat{\mathbf{b}}\hat{\mathbf{b}}) : (\hat{\mathbf{b}}\hat{\mathbf{b}})] \hat{\mathbf{b}}\hat{\mathbf{b}}\hat{\mathbf{b}}\hat{\mathbf{b}} + [(\mathbf{r} : \hat{\mathbf{b}}\hat{\mathbf{b}}) : (\mathbf{I} - \hat{\mathbf{b}}\hat{\mathbf{b}})/2] [\hat{\mathbf{b}}\hat{\mathbf{b}}(\mathbf{I} - \hat{\mathbf{b}}\hat{\mathbf{b}})]^{\text{Sym}} \\ & + [(\mathbf{r} : (\mathbf{I} - \hat{\mathbf{b}}\hat{\mathbf{b}})/2) : (\mathbf{I} - \hat{\mathbf{b}}\hat{\mathbf{b}})/2] [(\mathbf{I} - \hat{\mathbf{b}}\hat{\mathbf{b}})(\mathbf{I} - \hat{\mathbf{b}}\hat{\mathbf{b}})]^{\text{Sym}} + \mathbf{r}^{\text{ng}}. \end{aligned} \quad (8.37)$$

Now, applying  $\hat{\mathbf{b}}\hat{\mathbf{b}}$  twice to both sides of the equation yields  $(\mathbf{r} : \hat{\mathbf{b}}\hat{\mathbf{b}}) : (\hat{\mathbf{b}}\hat{\mathbf{b}}) = (\mathbf{r} : \hat{\mathbf{b}}\hat{\mathbf{b}}) : (\hat{\mathbf{b}}\hat{\mathbf{b}}) + (\mathbf{r}^{\text{ng}} : \hat{\mathbf{b}}\hat{\mathbf{b}}) : (\hat{\mathbf{b}}\hat{\mathbf{b}})$ , implying  $(\mathbf{r}^{\text{ng}} : \hat{\mathbf{b}}\hat{\mathbf{b}}) : (\hat{\mathbf{b}}\hat{\mathbf{b}}) = 0$ , which can be also rewritten as

$$\hat{\mathbf{b}}\hat{\mathbf{b}} : \mathbf{r}^{\text{ng}} : \hat{\mathbf{b}}\hat{\mathbf{b}} = 0. \quad (8.38)$$

Similarly, the other two properties are derived by either applying  $\hat{\mathbf{b}}$  and  $(\mathbf{I} - \hat{\mathbf{b}}\hat{\mathbf{b}})$ , or twice  $(\mathbf{I} - \hat{\mathbf{b}}\hat{\mathbf{b}})$ , yielding

$$(\mathbf{I} - \hat{\mathbf{b}}\hat{\mathbf{b}}) : \mathbf{r}^{\text{ng}} : \hat{\mathbf{b}}\hat{\mathbf{b}} = 0; \quad (\mathbf{I} - \hat{\mathbf{b}}\hat{\mathbf{b}}) : \mathbf{r}^{\text{ng}} : (\mathbf{I} - \hat{\mathbf{b}}\hat{\mathbf{b}}) = 0, \quad (8.39)$$

which, by using the property (8.38), are further reduced to

$$\mathbf{I} : \mathbf{r}^{\text{ng}} : \hat{\mathbf{b}}\hat{\mathbf{b}} = 0; \quad \mathbf{I} : \mathbf{r}^{\text{ng}} : \mathbf{I} = 0. \quad (8.40)$$

The last two properties can be also written as  $\text{Tr} \mathbf{r}^{\text{ng}} : \hat{\mathbf{b}}\hat{\mathbf{b}} = 0$ ,  $\text{Tr} \text{Tr} \mathbf{r}^{\text{ng}} = 0$ . Finally, writing all 3 properties in the index notation

$$r_{ijkl}^{\text{ng}} \hat{b}_i \hat{b}_j \hat{b}_k \hat{b}_l = 0; \quad r_{iikl}^{\text{ng}} \hat{b}_k \hat{b}_l = 0; \quad r_{iikk}^{\text{ng}} = 0. \quad (8.41)$$

One would assume that the non-gyrotropic  $\mathbf{r}^{\text{ng}}$  can be evaluated by rewriting (8.35) as

$$(\hat{\mathbf{b}} \times \mathbf{r}^{\text{ng}})^S = -\frac{B_0}{\Omega |\mathbf{B}|} \left[ \frac{\partial}{\partial t} \mathbf{r} + \nabla \cdot (\mathbf{X}^{(5)} + \mathbf{u}\mathbf{r}) + \left( \mathbf{r} \cdot \nabla \mathbf{u} - \frac{1}{\rho} (\nabla \cdot \mathbf{p}) \mathbf{q} \right)^S \right], \quad (8.42)$$

and then expanding the right-hand side similarly to what we did for the non-gyrotropic  $\Pi$  (and also the non-gyrotropic heat flux vectors in appendix D). However, the equation does not seem to yield anything useful. Instead, one needs to consider a specific example of a bi-Maxwellian distribution function.

## 8.2. Bi-Maxwellian distribution

Here we consider a special case of a bi-Maxwellian distribution function

$$f_0 = n \left( \frac{m}{2\pi} \right)^{3/2} \frac{1}{T_{\parallel}^{1/2} T_{\perp}} \exp \left[ -m \frac{(v_{\parallel} - u_{\parallel})^2}{2T_{\parallel}} - m \frac{|\mathbf{v}_{\perp} - \mathbf{u}_{\perp}|^2}{2T_{\perp}} \right]. \quad (8.43)$$

It is useful to use the fluctuating velocity  $\mathbf{c} = \mathbf{v} - \mathbf{u}$ , where  $c^2 = c_{\parallel}^2 + c_{\perp}^2$  and  $c_{\perp}^2 = c_x^2 + c_y^2$ . For brevity of the calculations we introduce

$$\alpha_{\parallel} \equiv \frac{m}{2T_{\parallel}} = \frac{1}{v_{\text{th}\parallel}^2}; \quad \alpha_{\perp} \equiv \frac{m}{2T_{\perp}} = \frac{1}{v_{\text{th}\perp}^2}, \quad (8.44)$$

where here the thermal speeds are meant to be spatially dependent, i.e. they are written with  $T$  and not with  $T^{(0)}$ . The bi-Maxwellian distribution therefore reads

$$f_0 = n \sqrt{\frac{\alpha_{\parallel}}{\pi}} \frac{\alpha_{\perp}}{\pi} e^{-\alpha_{\parallel} c_{\parallel}^2} e^{-\alpha_{\perp} c_{\perp}^2}. \quad (8.45)$$

It is useful to remind ourselves of the following one-dimensional integrals

$$\int_{-\infty}^{\infty} e^{-\alpha x^2} dx = \sqrt{\frac{\pi}{\alpha}}; \quad \int_{-\infty}^{\infty} x^2 e^{-\alpha x^2} dx = \frac{1}{2} \sqrt{\frac{\pi}{\alpha^3}}; \quad \int_{-\infty}^{\infty} x^4 e^{-\alpha x^2} dx = \frac{3}{4} \sqrt{\frac{\pi}{\alpha^5}}, \quad (8.46)$$

and that the general integral with  $x^n$ , where  $n$  is an integer, reads

$$\int_{-\infty}^{\infty} x^n e^{-\alpha x^2} dx = \frac{(n-1)!!}{2^{n/2}} \sqrt{\frac{\pi}{\alpha^{n+1}}}; \quad n=0, 2, 4, \dots \quad (n=\text{even}); \quad (8.47)$$

$$= 0; \quad n=1, 3, 5, \dots \quad (n=\text{odd}). \quad (8.48)$$

The double factorial  $(n-1)!! = 1 \cdot 3 \cdot 5 \cdots (n-1)$ , with  $(-1)!! = 1$ . The identity (8.47) is easily obtained by performing differentiation  $\partial/\partial\alpha$  of the first result in (8.46), and this technique is useful in the calculation of the fluid moments.

To become more familiar with the bi-Maxwellian distribution, it is useful to verify whether integrals over  $f_0$  indeed yield the expected fluid moments. Since the macroscopic velocity  $\mathbf{u}$  is independent of  $\mathbf{v}$ , changing from variable  $\mathbf{v}$  to  $\mathbf{c}$  just yields  $d^3v = d^3c$  (similarly to the substitution  $y = x + \text{const.}$  that yields  $dy = dx$ ). We use the notations  $d^3c = dc_{\parallel} d^2c_{\perp}$  and  $d^2c_{\perp} = dc_x dc_y$ . Also, because the integrals are evaluated from  $-\infty$  to  $\infty$ , substitution from  $\mathbf{v}$  to  $\mathbf{c}$  does not change these bounds. Integration in velocity space therefore yields

$$\int_{-\infty}^{\infty} e^{-\alpha_{\parallel} c_{\parallel}^2} dc_{\parallel} = \sqrt{\frac{\pi}{\alpha_{\parallel}}}; \quad \int_{-\infty}^{\infty} e^{-\alpha_{\perp} c_{\perp}^2} d^2c_{\perp} = \int e^{-\alpha_{\perp}(c_x^2 + c_y^2)} dc_x dc_y = \frac{\pi}{\alpha_{\perp}}. \quad (8.49)$$

For simplicity we have integrated in Cartesian coordinates (instead of perhaps the more elegant cylindrical coordinates), which has an additional benefit in that we can stop writing the integral bounds, since the integrals are always from  $-\infty$  to  $\infty$ . By using the bi-Maxwellian  $f_0$  it is very easy to verify that

$$n = \int f_0 d^3v; \quad n\mathbf{u} = \int \mathbf{v} f_0 d^3v; \quad p_{\parallel} = m \int c_{\parallel}^2 f_0 d^3v; \quad p_{\perp} = \frac{m}{2} \int c_{\perp}^2 f_0 d^3v. \quad (8.50)$$

Continuing with the heat flux components yields

$$q_{\parallel} = m \int c_{\parallel}^3 f_0 d^3v = mn \underbrace{\sqrt{\frac{\alpha_{\parallel}}{\pi}} \int c_{\parallel}^3 e^{-\alpha_{\parallel} c_{\parallel}^2} dc_{\parallel}}_{=0} \frac{\alpha_{\perp}}{\pi} \int e^{-\alpha_{\perp} c_{\perp}^2} d^2c_{\perp} = 0; \quad (8.51)$$

$$q_{\perp} = \frac{m}{2} \int c_{\parallel} c_{\perp}^2 f_0 d^3v = \frac{m}{2} n \underbrace{\sqrt{\frac{\alpha_{\parallel}}{\pi}} \int c_{\parallel} e^{-\alpha_{\parallel} c_{\parallel}^2} dc_{\parallel}}_{=0} \frac{\alpha_{\perp}}{\pi} \int c_{\perp}^2 e^{-\alpha_{\perp} c_{\perp}^2} d^2c_{\perp} = 0. \quad (8.52)$$

It is important to emphasize that the zero heat flux values were obtained by assuming some prescribed equilibrium  $f_0$ , here bi-Maxwellian. A crucial principle used in constructing fluid models with higher-order moments is that the specific  $f_0$  is assumed only for the last retained moment. Indeed, if one closes the hierarchy by prescribing  $q_{\parallel} = 0$ ,  $q_{\perp} = 0$ , the CGL fluid model is obtained. However, to go higher in the fluid hierarchy, evolution equations for  $q_{\parallel}$  and  $q_{\perp}$  cannot be eliminated, even if the hierarchy will be eventually closed, for example at the fourth-order moment level, by assuming a bi-Maxwellian  $f_0$ . Or in another words, the zero heat flux values were obtained for a distribution function that is strictly  $f_0$ , however, fluctuations/perturbations around  $f_0$  are still allowed. Therefore, to obtain the next available model after the CGL, the heat flux equations must be retained, and a closure is performed on the fourth-order moment. Only this last moment is calculated by assuming a specific distribution function, that is strictly  $f_0$ . For a bi-Maxwellian, the gyrotropic fourth-order moments are calculated according to

$$\begin{aligned} r_{\parallel\parallel} &\equiv m \int c_{\parallel}^4 f_0 d^3v = mn \underbrace{\sqrt{\frac{\alpha_{\parallel}}{\pi}} \int c_{\parallel}^4 e^{-\alpha_{\parallel} c_{\parallel}^2} dc_{\parallel}}_{=3/(4\alpha_{\parallel}^2)} \underbrace{\frac{\alpha_{\perp}}{\pi} \int e^{-\alpha_{\perp} c_{\perp}^2} d^2c_{\perp}}_{=1} \\ &= \frac{3}{4} \frac{mn}{\alpha_{\parallel}^2} = 3 \frac{n^2 T_{\parallel}^2}{nm} = 3 \frac{p_{\parallel}^2}{\rho}; \end{aligned} \quad (8.53)$$

$$\begin{aligned} r_{\parallel\perp} &\equiv \frac{m}{2} \int c_{\parallel}^2 c_{\perp}^2 f_0 d^3v = \frac{m}{2} n \underbrace{\sqrt{\frac{\alpha_{\parallel}}{\pi}} \int c_{\parallel}^2 e^{-\alpha_{\parallel} c_{\parallel}^2} dc_{\parallel}}_{=1/(2\alpha_{\parallel})} \underbrace{\frac{\alpha_{\perp}}{\pi} \int c_{\perp}^2 e^{-\alpha_{\perp} c_{\perp}^2} d^2c_{\perp}}_{=1/\alpha_{\perp}} \\ &= \frac{mn}{4\alpha_{\parallel}\alpha_{\perp}} = \frac{n^2 T_{\parallel} T_{\perp}}{mn} = \frac{p_{\parallel} p_{\perp}}{\rho}; \end{aligned} \quad (8.54)$$

$$\begin{aligned} r_{\perp\perp} &\equiv \frac{m}{4} \int c_{\perp}^4 f_0 d^3v = \frac{m}{4} n \underbrace{\sqrt{\frac{\alpha_{\parallel}}{\pi}} \int e^{-\alpha_{\parallel} c_{\parallel}^2} dc_{\parallel}}_{=1} \underbrace{\frac{\alpha_{\perp}}{\pi} \int c_{\perp}^4 e^{-\alpha_{\perp} c_{\perp}^2} d^2c_{\perp}}_{=2/\alpha_{\perp}^2} \\ &= \frac{mn}{2\alpha_{\perp}^2} = 2 \frac{n^2 T_{\perp}^2}{mn} = 2 \frac{p_{\perp}^2}{\rho}. \end{aligned} \quad (8.55)$$

Therefore, for a bi-Maxwellian, the following closure can be constructed

$$r_{\parallel\parallel} = \frac{3p_{\parallel}^2}{\rho}; \quad r_{\parallel\perp} = \frac{p_{\parallel} p_{\perp}}{\rho}; \quad r_{\perp\perp} = \frac{2p_{\perp}^2}{\rho}, \quad (8.56)$$

and this closure is known as the ‘normal’ closure, a name suggested by Chust & Belmont (2006).

8.3. Bi-Maxwellian non-gyrotropic  $\mathbf{r}^{\text{ng}}$  contributions

Let us consider a specific case of a bi-Maxwellian distribution function. Then, an expansion procedure analogous to the one developed by Grad (1949) for rarefied gases can be used, where the distribution function is expanded in a series of Hermite polynomials. By keeping only the first few terms, one can express the fourth-order moment through the second-order (pressure) moments. The procedure is written down in Oraevskii *et al.* (1968), and the entire fourth-order moment can be decomposed in the following way: see (2.31) in Oraevskii *et al.* (1968) (where a small typo on the left-hand side is present, where instead of  $q_{\alpha\beta\gamma\epsilon}$ , there should be  $q_{\alpha\beta\gamma\delta}$ ), equation (24) in Goswami *et al.* (2005), equation (9) in Passot & Sulem (2007)

$$\rho r_{ijkl} = p_{ij}^g p_{kl}^g + p_{ik}^g p_{jl}^g + p_{il}^g p_{jk}^g + \rho r_{ijkl}^{\text{ng}}, \quad (8.57)$$

where the non-gyrotropic contributions read

$$\rho r_{ijkl}^{\text{ng}} = p_{ij}^g \Pi_{kl} + p_{ik}^g \Pi_{jl} + p_{il}^g \Pi_{jk} + \Pi_{ij} p_{kl}^g + \Pi_{ik} p_{jl}^g + \Pi_{il} p_{jk}^g, \quad (8.58)$$

and where terms  $\Pi_{ij} \Pi_{lk} + \Pi_{ik} \Pi_{jl} + \Pi_{il} \Pi_{jk}$  were neglected. By using  $p_{ij}^g = p_{\parallel} \hat{b}_i \hat{b}_j + p_{\perp} (\delta_{ij} - \hat{b}_i \hat{b}_j)$ , it can be indeed shown by direct straightforward calculation that the gyrotropic part

$$\begin{aligned} \rho r_{ijkl}^g &= p_{ij}^g p_{kl}^g + p_{ik}^g p_{jl}^g + p_{il}^g p_{jk}^g = 3p_{\parallel}^2 \hat{b}_i \hat{b}_j \hat{b}_k \hat{b}_l \\ &+ p_{\parallel} p_{\perp} [\delta_{ij} \hat{b}_k \hat{b}_l + \delta_{ik} \hat{b}_j \hat{b}_l + \delta_{il} \hat{b}_j \hat{b}_k + \delta_{jk} \hat{b}_i \hat{b}_l + \delta_{jl} \hat{b}_i \hat{b}_k + \delta_{kl} \hat{b}_i \hat{b}_j - 6\hat{b}_i \hat{b}_j \hat{b}_k \hat{b}_l] \\ &+ p_{\perp}^2 \left[ \delta_{ij} \delta_{kl} + \delta_{ik} \delta_{jl} + \delta_{il} \delta_{jk} - \delta_{ij} \hat{b}_k \hat{b}_l - \delta_{ik} \hat{b}_j \hat{b}_l - \delta_{il} \hat{b}_j \hat{b}_k - \delta_{jk} \hat{b}_i \hat{b}_l \right. \\ &\left. - \delta_{jl} \hat{b}_i \hat{b}_k - \delta_{kl} \hat{b}_i \hat{b}_j + 3\hat{b}_i \hat{b}_j \hat{b}_k \hat{b}_l \right]. \end{aligned} \quad (8.59)$$

This agrees with the decomposition (8.10) valid for a general distribution function, after one specifies that for a bi-Maxwellian distribution  $r_{\parallel\parallel} = 3p_{\parallel}^2/\rho$ ,  $r_{\parallel\perp} = p_{\parallel} p_{\perp}/\rho$  and  $r_{\perp\perp} = 2p_{\perp}^2/\rho$  (last term in (8.10) contains  $r_{\perp\perp}/2$ ).

Now it is possible to calculate the non-gyrotropic contributions  $\mathbf{r}^{\text{ng}}$ , by using (8.58). It is useful to pre-calculate the trace of (8.58) that reads

$$\rho r_{iikl} = (p_{\parallel} + 6p_{\perp}) \Pi_{kl} + 2(p_{\parallel} - p_{\perp}) \hat{b}_i (\hat{b}_k \Pi_{il} + \hat{b}_l \Pi_{ik}). \quad (8.60)$$

The two terms that enter the parallel and perpendicular heat flux equations (7.96), (7.97) calculate as

$$\hat{\mathbf{b}} \cdot (\nabla \cdot \mathbf{r}^{\text{ng}}) : \hat{\mathbf{b}} \hat{\mathbf{b}} = \frac{3p_{\parallel}}{\rho} (\nabla \cdot \Pi) \cdot \hat{\mathbf{b}} + \frac{3p_{\parallel}}{\rho} (\hat{\mathbf{b}} \cdot \nabla \Pi) : \hat{\mathbf{b}} \hat{\mathbf{b}} + 3 \left( \nabla \frac{p_{\parallel}}{\rho} \right) \cdot \Pi \cdot \hat{\mathbf{b}}; \quad (8.61)$$

$$\begin{aligned} \mathbf{I} : (\nabla \cdot \mathbf{r}^{\text{ng}}) \cdot \hat{\mathbf{b}} &= \left( \nabla \frac{3p_{\parallel} + 4p_{\perp}}{\rho} \right) \cdot \Pi \cdot \hat{\mathbf{b}} + \frac{(3p_{\parallel} + 4p_{\perp})}{\rho} (\nabla \cdot \Pi) \cdot \hat{\mathbf{b}} \\ &+ \frac{2}{\rho} (p_{\parallel} - p_{\perp}) [\Pi : (\nabla \hat{\mathbf{b}}) + (\hat{\mathbf{b}} \cdot \nabla \hat{\mathbf{b}}) \cdot \Pi \cdot \hat{\mathbf{b}} + \hat{\mathbf{b}} \cdot (\nabla \Pi) : \hat{\mathbf{b}} \hat{\mathbf{b}}]. \end{aligned} \quad (8.62)$$

At this stage, we need to further simplify, and we keep only those terms that contribute at the linear level, yielding

$$\hat{\mathbf{b}} \cdot (\nabla \cdot \mathbf{r}^{\text{ng}}) : \hat{\mathbf{b}} \hat{\mathbf{b}} = \frac{3p_{\parallel}}{\rho} (\nabla \cdot \Pi) \cdot \hat{\mathbf{b}}; \quad (8.63)$$

$$I : (\nabla \cdot \mathbf{r}^{\text{ng}}) \cdot \hat{\mathbf{b}} = \frac{(3p_{\parallel} + 4p_{\perp})}{\rho} (\nabla \cdot \Pi) \cdot \hat{\mathbf{b}}, \quad (8.64)$$

so the term entering the  $q_{\perp}$  equation (7.97) reads

$$\frac{1}{2} (I - \hat{\mathbf{b}}\hat{\mathbf{b}}) : (\nabla \cdot \mathbf{r}^{\text{ng}}) \cdot \hat{\mathbf{b}} = \frac{2p_{\perp}}{\rho} (\nabla \cdot \Pi) \cdot \hat{\mathbf{b}}. \quad (8.65)$$

Note that  $(\hat{\mathbf{b}} \cdot \nabla \Pi) : \hat{\mathbf{b}}\hat{\mathbf{b}} = -\Pi : (\hat{\mathbf{b}}(\hat{\mathbf{b}} \cdot \nabla)\hat{\mathbf{b}})^S$ , and the term does not contribute at the linear level. Of course, many nonlinear terms were neglected, and technically, the only fully consistent procedure is to explicitly write those terms at the linear level, in the form

$$\hat{\mathbf{b}} \cdot (\nabla \cdot \mathbf{r}^{\text{ng}}) : \hat{\mathbf{b}}\hat{\mathbf{b}} = \frac{3p_{\parallel}^{(0)}}{\rho_0} (\nabla \cdot \Pi) \cdot \hat{\mathbf{b}}_0 = \frac{3p_{\parallel}^{(0)}}{\rho_0} (\partial_x \Pi_{xz} + \partial_y \Pi_{yz}); \quad (8.66)$$

$$\frac{1}{2} (I - \hat{\mathbf{b}}\hat{\mathbf{b}}) : (\nabla \cdot \mathbf{r}^{\text{ng}}) \cdot \hat{\mathbf{b}} = \frac{2p_{\perp}^{(0)}}{\rho_0} (\nabla \cdot \Pi) \cdot \hat{\mathbf{b}}_0 = \frac{2p_{\perp}^{(0)}}{\rho_0} (\partial_x \Pi_{xz} + \partial_y \Pi_{yz}). \quad (8.67)$$

The results show that, if the  $\Pi$  contributions are kept in the heat flux equations, the  $\mathbf{r}^{\text{ng}}$  contributions cannot just be straightforwardly neglected, since the  $\Pi$  and  $\mathbf{r}^{\text{ng}}$  contributions partially cancel out.

## 9. Bi-Maxwellian fluid model, second-order CGL (CGL2)

In the previous section we have seen that for a bi-Maxwellian distribution function, one can close the fluid hierarchy by prescribing the ‘normal’ closure

$$r_{\parallel\parallel} = \frac{3p_{\parallel}^2}{\rho}; \quad r_{\perp\perp} = \frac{p_{\parallel}p_{\perp}}{\rho}; \quad r_{\perp\perp} = \frac{2p_{\perp}^2}{\rho}. \quad (9.1)$$

It is useful to summarize the fully nonlinear model that we want to consider here. We consider only proton species, and we make the electrons massless and cold. We simplify the pressure equations, by neglecting the FLR stress forces (which enter only nonlinearly), and we keep only those non-gyrotropic contributions in the pressure and heat flux equations that have a non-zero linear contributions. The nonlinear model reads

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{u}) = 0; \quad (9.2)$$

$$\frac{\partial \mathbf{u}}{\partial t} + \mathbf{u} \cdot \nabla \mathbf{u} + \frac{1}{\rho} \nabla \cdot [p_{\parallel} \hat{\mathbf{b}}\hat{\mathbf{b}} + p_{\perp} (I - \hat{\mathbf{b}}\hat{\mathbf{b}}) + \Pi] - \frac{1}{4\pi\rho} (\nabla \times \mathbf{B}) \times \mathbf{B} = 0; \quad (9.3)$$

$$\frac{\partial \mathbf{B}}{\partial t} = \nabla \times (\mathbf{u} \times \mathbf{B}) - \frac{1}{\Omega_p} \frac{B_0}{4\pi} \nabla \times \left[ \frac{1}{\rho} (\nabla \times \mathbf{B}) \times \mathbf{B} \right]; \quad (9.4)$$

$$\frac{\partial p_{\parallel}}{\partial t} + \mathbf{u} \cdot \nabla p_{\parallel} + p_{\parallel} \nabla \cdot \mathbf{u} + 2p_{\parallel} \hat{\mathbf{b}} \cdot \nabla \mathbf{u} \cdot \hat{\mathbf{b}} + \nabla \cdot (q_{\parallel} \hat{\mathbf{b}}) - 2q_{\perp} \nabla \cdot \hat{\mathbf{b}} + \nabla \cdot \mathbf{S}_{\perp}^{\parallel} = 0; \quad (9.5)$$

$$\frac{\partial p_{\perp}}{\partial t} + \mathbf{u} \cdot \nabla p_{\perp} + 2p_{\perp} \nabla \cdot \mathbf{u} - p_{\perp} \hat{\mathbf{b}} \cdot \nabla \mathbf{u} \cdot \hat{\mathbf{b}} + \nabla \cdot (q_{\perp} \hat{\mathbf{b}}) + q_{\perp} \nabla \cdot \hat{\mathbf{b}} + \nabla \cdot \mathbf{S}_{\perp}^{\perp} = 0; \quad (9.6)$$

$$\frac{\partial q_{\parallel}}{\partial t} + \mathbf{u} \cdot \nabla q_{\parallel} + q_{\parallel} \nabla \cdot \mathbf{u} + 3p_{\parallel} \hat{\mathbf{b}} \cdot \nabla \left( \frac{p_{\parallel}}{\rho} \right) + 3q_{\parallel} \hat{\mathbf{b}} \cdot \nabla \mathbf{u} \cdot \hat{\mathbf{b}} = 0; \quad (9.7)$$

$$\frac{\partial q_{\perp}}{\partial t} + \mathbf{u} \cdot \nabla q_{\perp} + 2q_{\perp} \nabla \cdot \mathbf{u} + p_{\parallel} \hat{\mathbf{b}} \cdot \nabla \left( \frac{p_{\perp}}{\rho} \right) + \frac{p_{\perp}}{\rho} (p_{\parallel} - p_{\perp}) \nabla \cdot \hat{\mathbf{b}} + \frac{p_{\perp}}{\rho} (\nabla \cdot \boldsymbol{\Pi}) \cdot \hat{\mathbf{b}}_0 = 0. \quad (9.8)$$

To develop a vocabulary, let us first consider the above fluid model with all the non-gyrotropic fluctuations neglected, i.e. when all the terms with  $\boldsymbol{\Pi}$ ,  $\mathbf{S}_{\perp}^{\parallel}$ ,  $\mathbf{S}_{\perp}^{\perp}$  are neglected. Additionally, let us neglect the Hall term in the induction equation. Such a fluid model is non-dispersive, and represents a generalization of the non-dispersive CGL model. Even though such a fluid model was not explicitly considered by Chew, Goldberger and Low, the pressure equations that include the gyrotropic heat flux contributions  $q_{\parallel}$ ,  $q_{\perp}$ , were indeed given by Chew *et al.* (1956). As discussed previously, the authors just used a different notation with  $q_n = q_{\parallel} - 3q_{\perp}$  and  $q_s = q_{\perp}$ , but their pressure equations contain the gyrotropic heat flux. The authors did not consider evolution equations for the gyrotropic heat fluxes, equations (9.7), (9.8). Nevertheless, the name CGL is so well recognized by the community, i.e. the name CGL is essentially recognized as the collisionless MHD, that we suggest calling the non-dispersive fluid model that used the above gyrotropic heat flux equations, ‘the second-order CGL’, abbreviated as ‘CGL2’. The name CGL2 has a nice advantage in that the abbreviations for many models considered previously can be easily and naturally generalized.

The CGL2 fluid model does not contain any dispersive effects and it is length-scale invariant, similarly to the CGL and MHD models. If the Hall term in the induction equation is considered, this yields the ‘Hall-CGL2’ model. Considering also the first-order FLR corrections to the pressure tensor (FLR1) or the second-order corrections (FLR2), yields the ‘Hall-CGL2-FLR1’ and the ‘Hall-CGL2-FLR2’ fluid models. Finally, considering the non-gyrotropic heat flux fluctuations  $\mathbf{S}_{\perp}^{\parallel}$ ,  $\mathbf{S}_{\perp}^{\perp}$ , yields the Hall-CGL2-FLR3 model. Fluid models with the Hall term neglected can be easily abbreviated as ‘CGL2-FLR1’, ‘CGL2-FLR2’ etc. Even though, if the dispersive effects of the FLR corrections are considered, there is really no reason to neglect the simple Hall term, and the use of such fluid models is discouraged. Obviously, the name ‘CGL2’ is extremely beneficial and the very natural for the classification of fluid models, and we will use this name henceforth.

It is important to correctly normalize the heat flux, and the normalization is according to  $\tilde{q}_{\parallel,\perp} = q_{\parallel,\perp} / (p_{\parallel}^{(0)} V_A)$ , so both the parallel and perpendicular heat fluxes are normalized in the same way (similarly, both  $p_{\parallel}$  and  $p_{\perp}$  are normalized with respect to  $p_{\parallel}^{(0)}$ ). By dropping the tilde, the scalar pressure equations (9.5), (9.6) remain unchanged, and the normalized nonlinear heat flux equations read

$$\frac{\partial q_{\parallel}}{\partial t} + \mathbf{u} \cdot \nabla q_{\parallel} + q_{\parallel} \nabla \cdot \mathbf{u} + 3 \frac{\beta_{\parallel}}{2} p_{\parallel} \hat{\mathbf{b}} \cdot \nabla \left( \frac{p_{\parallel}}{\rho} \right) + 3 q_{\parallel} \hat{\mathbf{b}} \cdot \nabla \mathbf{u} \cdot \hat{\mathbf{b}} = 0; \quad (9.9)$$

$$\begin{aligned} \frac{\partial q_{\perp}}{\partial t} + \mathbf{u} \cdot \nabla q_{\perp} + 2q_{\perp} \nabla \cdot \mathbf{u} + \frac{\beta_{\parallel}}{2} p_{\parallel} \hat{\mathbf{b}} \cdot \nabla \left( \frac{p_{\perp}}{\rho} \right) \\ + \frac{\beta_{\parallel}}{2} \frac{p_{\perp}}{\rho} (p_{\parallel} - p_{\perp}) \nabla \cdot \hat{\mathbf{b}} + \frac{\beta_{\perp}}{2} \frac{p_{\perp}}{\rho} (\nabla \cdot \boldsymbol{\Pi}) \cdot \hat{\mathbf{b}}_0 = 0. \end{aligned} \quad (9.10)$$

Now we need to linearize the system, where one prescribes  $q_{\parallel}^{(0)} = 0$  and  $q_{\perp}^{(0)} = 0$ . Once normalized, linearized, transformed to Fourier space and written in the  $x$ - $z$  plane, the model reads (dropping the tilde everywhere)

$$-\omega \rho + k_{\perp} u_x + k_{\parallel} u_z = 0; \quad (9.11)$$

$$-\omega u_x + \frac{\beta_{\parallel}}{2} k_{\perp} p_{\perp} - v_{A\parallel}^2 k_{\parallel} B_x + k_{\perp} B_z + \frac{\beta_{\parallel}}{2} (k_{\perp} \Pi_{xx} + k_{\parallel} \Pi_{xz}) = 0; \quad (9.12)$$

$$-\omega u_y - v_{A\parallel}^2 k_{\parallel} B_y + \frac{\beta_{\parallel}}{2} (k_{\perp} \Pi_{xy} + k_{\parallel} \Pi_{yz}) = 0; \quad (9.13)$$

$$-\omega u_z + \frac{\beta_{\parallel}}{2} k_{\parallel} p_{\parallel} + \frac{\beta_{\parallel}}{2} (1 - a_p) k_{\perp} B_x + \frac{\beta_{\parallel}}{2} k_{\perp} \Pi_{xz} = 0; \quad (9.14)$$

$$-\omega B_x - k_{\parallel} u_x - i k_{\parallel}^2 B_y = 0; \quad (9.15)$$

$$-\omega B_y - k_{\parallel} u_y - i k_{\parallel} k_{\perp} B_z + i k_{\parallel}^2 B_x = 0; \quad (9.16)$$

$$-\omega B_z + k_{\perp} u_x + i k_{\parallel} k_{\perp} B_y = 0; \quad (9.17)$$

$$-\omega p_{\parallel} + k_{\perp} u_x + 3 k_{\parallel} u_z + k_{\parallel} q_{\parallel} + k_{\perp} S_{\perp}^{\parallel} = 0; \quad (9.18)$$

$$-\omega p_{\perp} + 2 a_p k_{\perp} u_x + a_p k_{\parallel} u_z + k_{\parallel} q_{\perp} + k_{\perp} S_{\perp}^{\perp} = 0; \quad (9.19)$$

$$-\omega q_{\parallel} + \frac{3}{2} \beta_{\parallel} k_{\parallel} (p_{\parallel} - \rho) = 0; \quad (9.20)$$

$$-\omega q_{\perp} + \frac{\beta_{\parallel}}{2} k_{\parallel} (p_{\perp} - a_p \rho) + \frac{\beta_{\parallel}}{2} a_p (1 - a_p) k_{\perp} B_x + \frac{\beta_{\parallel}}{2} a_p k_{\perp} \Pi_{xz} = 0. \quad (9.21)$$

As previously,  $v_{A\parallel}^2 = 1 + (\beta_{\parallel}/2)(a_p - 1)$ .

### 9.1. CGL2 dispersion relation

Neglecting all the dispersive effects, prescribing  $\boldsymbol{\Pi} = 0$ ,  $S_{\perp}^{\parallel} = 0$ ,  $S_{\perp}^{\perp} = 0$ , and eliminating the Hall term in the induction equation (i.e. terms proportional to  $ik^2$  in the above induction equation), yields the CGL2 model. By exploring the system, it is obvious that the Alfvén mode separates from the system in the  $u_y$ ,  $B_y$  components. The dispersion relation of the Alfvén mode in the (non-dispersive) CGL2 model is therefore the same as in the CGL model, and reads (the normalization tildes are dropped)

$$\omega = \pm k_{\parallel} v_{A\parallel} = \pm k \cos(\theta) \sqrt{1 + \frac{\beta_{\parallel}}{2} (a_p - 1)}. \quad (9.22)$$

The ‘hard’ threshold of the oblique firehose instability is therefore the same as in the CGL model, and correctly reproduced. The entire dispersion relation for the CGL2 fluid model can be written in the following form

$$(\omega^2 - v_{A\parallel}^2 k_{\parallel}^2)(\omega^8 - A_6 \omega^6 + A_4 \omega^4 - A_2 \omega^2 + A_0) = 0; \quad (9.23)$$

$$A_6 = k_{\parallel}^2 (v_{A\parallel}^2 + \frac{7}{2} \beta_{\parallel}) + k_{\perp}^2 (1 + a_p \beta); \quad (9.24)$$

$$A_4 = k_{\parallel}^2 \beta_{\parallel} \left\{ k_{\parallel}^2 \left( \frac{7}{2} v_{A\parallel}^2 + \frac{9}{4} \beta_{\parallel} \right) + k_{\perp}^2 \frac{7}{2} \left( 1 + a_p \beta_{\parallel} - \frac{1}{7} a_p^2 \beta_{\parallel} \right) \right\}; \quad (9.25)$$

$$A_2 = k_{\parallel}^4 \beta_{\parallel}^2 \left\{ k_{\parallel}^2 \left( \frac{9}{4} v_{A\parallel}^2 + \frac{3}{8} \beta_{\parallel} \right) + k_{\perp}^2 \frac{9}{4} \left( 1 + a_p \beta_{\parallel} - \frac{5}{9} a_p^2 \beta_{\parallel} \right) \right\}; \quad (9.26)$$

$$A_0 = \frac{3}{8} k_{\parallel}^6 \beta_{\parallel}^3 \{ k_{\parallel}^2 v_{A\parallel}^2 + k_{\perp}^2 (1 + a_p \beta_{\parallel} - a_p^2 \beta_{\parallel}) \}. \quad (9.27)$$

The CGL2 fluid model therefore contains 5 forward and 5 backward propagating modes. The oblique Alfvén mode is separated as stated above, and is incompressible. The remaining 4 forward and 4 backward waves are generally compressible and coupled together through the eighth-order polynomial in  $\omega$  (fourth-order polynomial in  $\omega^2$ ). Let us ignore for a moment the distinction between the forward and backward modes, since these will always be symmetric in  $\omega$ . The CGL2 model contains 5 waves/modes. The oblique Alfvén mode is separated, and the remaining 4



compressible modes are in general strongly coupled. In contrast to the simpler CGL (and MHD) model, that contains only 2 compressible modes (slow and fast), we therefore have 2 ‘new’ modes, that do not have an analogy in the usual CGL and MHD descriptions. There is no standardized vocabulary on how the modes should be named, perhaps the expression ‘higher-order modes’ is the most appropriate. For the highly oblique propagation, considering the mirror instability, sometimes the expression ‘mirror modes’ is used. Nevertheless, since none of the 4 compressible modes match the slow and fast CGL dispersion relations, the distinction between modes becomes blurry. The reader has to get used to the fact that, by considering higher-order fluid moments, we perhaps came closer to the kinetic theory, which admits an infinite number of modes that are difficult to classify, unless a specific situation is considered. The difference between the CGL2 model and the kinetic description is that, unless a firehose or mirror instability threshold is reached, the fluid modes are un-damped.

For strictly parallel propagation ( $k_{\perp} = 0$ ), in addition to the parallel Alfvén mode (9.22)  $\omega = \pm k_{\parallel} v_{A\parallel}$  that was already separated, the solution of the eighth-order polynomial contains another Alfvén mode  $\omega = \pm k_{\parallel} v_{A\parallel}$ , which was of course expected and is equivalent to the CGL model, since the components  $u_x, u_y, B_x, B_y$  de-couple for parallel propagation. The remaining solutions are

$$\omega = \pm k_{\parallel} \sqrt{\beta_{\parallel} \left( \frac{3}{2} + \sqrt{\frac{3}{2}} \right)} = \pm 1.65 k_{\parallel} \sqrt{\beta_{\parallel}}; \quad (9.28)$$

$$\omega = \pm k_{\parallel} \sqrt{\beta_{\parallel} \left( \frac{3}{2} - \sqrt{\frac{3}{2}} \right)} = \pm 0.52 k_{\parallel} \sqrt{\beta_{\parallel}}; \quad (9.29)$$

$$\omega = \pm k_{\parallel} \sqrt{\frac{\beta_{\parallel}}{2}} = \pm 0.71 k_{\parallel} \sqrt{\beta_{\parallel}}. \quad (9.30)$$

Solutions (9.28) and (9.29) can be obtained by considering a pure 1-D geometry (where  $p_{\perp}$  and  $q_{\perp}$  disappear since the kinetic velocity  $v_{\perp}$  disappears) and considering only fluctuations in the quantities  $\rho, u_z, p_{\parallel}, q_{\parallel}$ . The  $q_{\parallel}$  heat flux fluctuations in the CGL2 fluid model are therefore responsible for ‘splitting’ of the CGL ion-acoustic mode  $\omega = \pm k_{\parallel} \sqrt{3\beta_{\parallel}/2}$  into two modes (9.28) and (9.29).

Considering solutions for strictly perpendicular propagation ( $k_{\parallel} = 0$ ), the only non-zero mode is the fast mode

$$\omega = \pm k_{\perp} \sqrt{1 + a_p \beta_{\parallel}}, \quad (9.31)$$

and the dispersion relation is unchanged from the CGL model (for perpendicular propagation the  $q_{\parallel}, q_{\perp}$  contributions to the equations for  $p_{\parallel}, p_{\perp}$  naturally vanish and the CGL2 model is equivalent to the CGL model).

## 9.2. Mirror instability

A very interesting direction of propagation for this fluid model is the highly oblique limit,  $k_{\perp} \gg k_{\parallel}$ , since the CGL2 fluid model contains the correct mirror instability threshold. The correct mirror threshold can be already seen in the quantity  $A_0$ , where, in contrast to the usual CGL model, the factor  $1/6$  disappears in the last term  $\sim a_p^2 \beta_{\parallel}$ .

In the highly oblique limit, the fast mode (9.31) can be separated from the coupled eighth degree polynomial (actually fourth degree polynomial in  $\omega^2$ ) in the following

way. Importantly, one cannot prescribe a completely perpendicular propagation  $k_{\parallel} = 0$ , and a highly oblique limit  $k_{\perp} \gg k_{\parallel}$ , or  $k_{\parallel} \rightarrow 0$ , has to be considered. The polynomial coefficients can be approximated as

$$A_6 = k_{\perp}^2 (1 + a_p \beta); \quad (9.32)$$

$$A_4 = k_{\parallel}^2 \beta_{\parallel} k_{\perp}^2 \frac{7}{2} (1 + a_p \beta_{\parallel} - \frac{1}{7} a_p^2 \beta_{\parallel}); \quad (9.33)$$

$$A_2 = k_{\parallel}^4 \beta_{\parallel}^2 k_{\perp}^2 \frac{9}{4} (1 + a_p \beta_{\parallel} - \frac{5}{9} a_p^2 \beta_{\parallel}); \quad (9.34)$$

$$A_0 = \frac{3}{8} k_{\parallel}^6 \beta_{\parallel}^3 k_{\perp}^2 (1 + a_p \beta_{\parallel} - a_p^2 \beta_{\parallel}). \quad (9.35)$$

An alternative approximation is to use  $k_{\perp} \rightarrow k$  in the above expressions. The coefficient  $A_6$  does not contain any  $k_{\parallel}$  and is very large compared to  $A_4, A_2, A_0$  that all contain  $k_{\parallel}$ . The solutions will inevitably be one fast mode, and 3 very slow modes. In this specific case, the fast mode can be quickly separated from the eighth-order polynomial in (9.23) with a neat trick

$$\left. \begin{aligned} (\omega^8 - A_6 \omega^6 + A_4 \omega^4 - A_2 \omega^2 + A_0) &= 0; \\ (\omega^2 - A_6)(-A_6 \omega^6 + A_4 \omega^4 - A_2 \omega^2 + A_0) &= 0. \end{aligned} \right\} \quad (9.36)$$

The trick can be verified by multiplying both brackets term by term, dividing by  $-A_6$  and observing that  $(A_4, A_2, A_0)/A_6$  is negligible. The fast mode is always obtained correctly in this way (when the  $A_6$  is much larger than all the other coefficients), however, the dispersion relation for the 3 slow modes is correct only in this specific case when the system is non-dispersive, and all 3 slow modes have constant phase speeds. A proper limit and expansion should be always checked.

Nevertheless, in the highly oblique limit, the fast mode  $\omega = \pm k_{\perp} \sqrt{1 + a_p \beta_{\parallel}}$  is separated, and the rest of the dispersion is of sixth order in  $\omega$  and contains 3 ‘slow’ modes. By cancelling out the  $k_{\perp}^2$  that is common in all of the remaining coefficients (or  $k^2$ ), an obvious substitution is offered that reads  $\bar{\omega} = \omega/(k_{\parallel} \sqrt{\beta_{\parallel}})$ , and that yields a dispersion equation of third order in  $\bar{\omega}^2$  in the form

$$\begin{aligned} -(1 + a_p \beta_{\parallel})(\bar{\omega}^2)^3 + \frac{7}{2} (1 + a_p \beta_{\parallel} - \frac{1}{7} a_p^2 \beta_{\parallel}) (\bar{\omega}^2)^2 - \frac{9}{4} (1 + a_p \beta_{\parallel} - \frac{5}{9} a_p^2 \beta_{\parallel}) (\bar{\omega}^2) \\ + \frac{3}{8} (1 + a_p \beta_{\parallel} - a_p^2 \beta_{\parallel}) = 0. \end{aligned} \quad (9.37)$$

Now, at exactly the mirror threshold  $a_p^2 \beta_{\parallel} = 1 + a_p \beta_{\parallel}$ , the last term disappears and we have one solution  $\bar{\omega}^2 = 0$ , and two solutions

$$\bar{\omega}^2 = (3 \pm \sqrt{5})/2, \quad (9.38)$$

that are both positive. Slightly below or beyond the mirror threshold, we can prescribe  $a_p^2 \beta_{\parallel} = (1 + a_p \beta_{\parallel})(1 + \epsilon)$ , where  $\epsilon$  is a small parameter, meaning that for  $\epsilon < 0$ , the system is slightly below the expected mirror threshold, i.e. stable. For  $\epsilon > 0$ , the system is slightly beyond the expected mirror threshold, i.e. unstable. The polynomial transforms to

$$(\bar{\omega}^2)^3 - \left(3 - \frac{\epsilon}{2}\right) (\bar{\omega}^2)^2 + \left(1 - \frac{5}{4}\epsilon\right) (\bar{\omega}^2) + \frac{3}{8}\epsilon = 0. \quad (9.39)$$

It is possible to show that this polynomial has a discriminant<sup>12</sup>  $\Delta > 0$ , implying the existence of 3 distinct real roots. Now, when  $\epsilon$  is small, the two solutions (9.38) will remain almost the same and change only slightly, importantly, the solutions will remain positive. Solutions of a cubic polynomial  $(x - x_1)(x - x_2)(x - x_3) = 0$  form the last coefficient in that polynomial,  $-x_1x_2x_3$ , here equal to  $\frac{3}{8}\epsilon$ , and here  $x_1$  and  $x_2$  are positive. Therefore,  $\epsilon < 0$  (slightly below the threshold) implies  $x_3 > 0$ ; and  $\epsilon > 0$  (slightly beyond the threshold) implies  $x_3 < 0$ . The negative solution (in  $(\bar{\omega}^2)$ ) represents the mirror instability, which finishes the analytic proof that the (highly oblique) mirror threshold in CGL2 model is indeed  $a_p^2\beta_{\parallel} = 1 + a_p\beta_{\parallel}$ . The CGL2 fluid model therefore contains the same mirror instability threshold, as is found in kinetic theory.

The conclusion can be double checked by numerically exploring solutions of (9.39) for several values of  $\epsilon$ , and it is indeed possible to conclude that, for  $\epsilon < 0$  (slightly below the threshold), the signs of the solutions are  $\bar{\omega}^2 = +, +, +$ ; and that, for  $\epsilon > 0$  (slightly beyond the mirror threshold), the signs of the solutions are  $\bar{\omega}^2 = +, +, -$ . The solution with the minus sign represents the mirror instability. It is noted that the Landau fluid description allows for the construction of a fluid model with quasi-static heat fluxes, that also correctly reproduces the ‘hard’ mirror instability threshold (Sulem & Passot 2012). Considering Landau fluid models, the mirror instability was numerically investigated in detail for example by Passot & Sulem (2007), Passot *et al.* (2012) and Sulem & Passot (2015). As a suggestion for possible future work, it would be beneficial to numerically calculate ‘Hellinger’s’ contours (see § 6.3) with the prescribed maximum growth rate for the mirror instability. If one uses Landau fluid models with sufficiently precise FLR corrections, the results should match the kinetic contours very precisely.

### 9.3. Hall-CGL2 dispersion relation

Considering the Hall term, the dispersion relation of the Hall-CGL2 fluid model reads

$$\begin{aligned} &(\omega^2 - v_{A\parallel}^2 k_{\parallel}^2)(\omega^8 - A_6\omega^6 + A_4\omega^4 - A_2\omega^2 + A_0) \\ &= k^2 k_{\parallel}^2 \omega^2 \left[ \omega^6 - \omega^4 \beta_{\parallel} \left( \frac{7}{2} k_{\parallel}^2 + a_p k_{\perp}^2 \right) + \omega^2 \beta_{\parallel}^2 k_{\parallel}^2 \left( \frac{9}{4} k_{\parallel}^2 + 3a_p k_{\perp}^2 \right) - k_{\parallel}^4 \beta_{\parallel}^3 \left( \frac{3}{8} k_{\parallel}^2 + a_p k_{\perp}^2 \right) \right], \end{aligned} \quad (9.40)$$

where the left-hand side represents the CGL2 dispersion relation (9.23) and the right-hand side is the Hall-term contribution. For strictly parallel propagation, the solutions are the usual Hall-CGL whistler and ion-cyclotron waves  $\omega = \pm k_{\parallel}^2/2 + k_{\parallel} \sqrt{v_{A\parallel}^2 + k_{\parallel}^2/4}$ ,  $\omega = \pm k_{\parallel}^2/2 - k_{\parallel} \sqrt{v_{A\parallel}^2 + k_{\parallel}^2/4}$ , that are accompanied by the CGL2 solutions (9.28), (9.30) and (9.29). For strictly perpendicular propagation, the solution is (9.31) since the Hall term vanishes. The dispersion relation of the Hall-CGL-FLR1 model was already quite large to write down, and we do not provide the final dispersion relation for the Hall-CGL2-FLR1 model, instead we recommend working with the full system of linearized equations and solving it numerically or using analytic software such as Maple.

<sup>12</sup>Discriminant of a cubic polynomial  $ax^3 + bx^2 + cx + d = 0$  is calculated according to  $\Delta = 18abcd - 4b^3d + b^2c^2 - 4ac^3 - 27a^2d^2$ .

## 9.4. ‘Static’ closure – generalized isothermal closure

There exists even simpler fluid closure that recovers the correct mirror threshold, called the ‘static’ or ‘quasi-static’ closure (Constantinescu 2002; Chust & Belmont 2006; Passot *et al.* 2006). By using heat flux equations (9.7), (9.8) and considering the quasi-static regime with  $\partial q_{\parallel}/\partial t = 0$ ,  $\partial q_{\perp}/\partial t = 0$ ,  $\mathbf{u} = 0$  yields

$$\hat{\mathbf{b}} \cdot \nabla \left( \frac{p_{\parallel}}{\rho} \right) = 0; \quad p_{\parallel} \hat{\mathbf{b}} \cdot \nabla \left( \frac{p_{\perp}}{\rho} \right) + \frac{p_{\perp}}{\rho} (p_{\parallel} - p_{\perp}) \nabla \cdot \hat{\mathbf{b}} = 0, \quad (9.41)$$

where the FLR pressure tensor  $\Pi$  was neglected in the second equation. It is useful to define  $\partial_{\parallel} \equiv \hat{\mathbf{b}} \cdot \nabla$ , which represents the gradient along the magnetic field lines, and to use the identity

$$\nabla \cdot \hat{\mathbf{b}} = -\frac{1}{|\mathbf{B}|} \hat{\mathbf{b}} \cdot \nabla |\mathbf{B}|, \quad (9.42)$$

further yielding

$$\partial_{\parallel} T_{\parallel} = 0; \quad \frac{\partial_{\parallel} T_{\perp}}{T_{\perp}} = \left( 1 - \frac{T_{\perp}}{T_{\parallel}} \right) \frac{\partial_{\parallel} |\mathbf{B}|}{|\mathbf{B}|}. \quad (9.43)$$

Equation (9.43) describes the evolution of the parallel and perpendicular temperature fluctuations along the magnetic field lines, and it can be verified that the solution is

$$T_{\parallel} = T_{\parallel}^{(0)}; \quad T_{\perp} = T_{\perp}^{(0)} \frac{\frac{|\mathbf{B}|}{B_0}}{1 - a_p + a_p \frac{|\mathbf{B}|}{B_0}}, \quad (9.44)$$

where  $a_p = T_{\perp}^{(0)}/T_{\parallel}^{(0)}$ . Solution (9.44) is referred to as the ‘static’ closure and it is, for example, equivalent to (19), (20) of Passot *et al.* (2006), and (2) of Constantinescu (2002). For isotropic mean temperatures  $a_p = 1$ , equation (9.44) yields  $T_{\perp} = T_{\perp}^{(0)} = T_{\parallel}^{(0)}$  and both parallel and perpendicular temperatures are isothermal. A similar result is obtained for general  $a_p$  with  $|\mathbf{B}| = B_0$ , which yields  $T_{\perp} = T_{\perp}^{(0)}$ . The closure (9.44) can be therefore viewed as a generalization of the isothermal closure in the presence of temperature anisotropy and variations of magnetic field strength. The closure is used directly in the momentum equation (9.3), and the time-dependent pressure equations (9.5), (9.6) are disregarded. Therefore, even though the closure was derived by considering fourth-order moments, the resulting fluid model is actually simpler than the CGL2 and CGL models. As discussed for example by Passot *et al.* (2006), the closure prescribes a limitation for the minimal value of  $|\mathbf{B}|/B_0 > 1 - 1/a_p$  that is required to prevent temperature singularity. Or in other words, by separating  $|\mathbf{B}| = B_0 + B^{(1)}$  the requirement reads  $B^{(1)} > -1/a_p$ , meaning that the fluctuating  $B^{(1)}$  cannot be too negative. Mirror instability is often associated with nonlinear structures in the form of magnetic holes and humps, and the requirement implies that the magnetic holes cannot be too deep. To easily analyse the dispersion relations, it is useful to write the closure at the linear level, in the following form

$$p_{\parallel} = nT_{\parallel}^{(0)}; \quad p_{\perp} \stackrel{\text{lin}}{=} nT_{\perp}^{(0)} + p_{\perp}^{(0)}(1 - a_p) \frac{B_z^{(1)}}{B_0}. \quad (9.45)$$

The result can be obtained by linearizing (9.44) or (9.43), and at the linear level  $\partial_i |\mathbf{B}| \stackrel{\text{lin}}{=} \partial_i B_z$ . The normalized closure reads (tildes are dropped)  $p_{\parallel} = n$  and  $p_{\perp} = a_p n +$

$a_p(1 - a_p)B_z$ . Considering the simplest non-dispersive model (by neglecting the Hall term and  $\Pi$ ), the dispersion relation of this fluid model reads

$$\left. \begin{aligned} (\omega^2 - k_{\parallel}^2 v_{A\parallel}^2)(\omega^4 - A_2 \omega^2 + A_0) &= 0; \\ A_2 &= k_{\parallel}^2 \left( v_{A\parallel}^2 + \frac{\beta_{\parallel}}{2} \right) + k_{\perp}^2 \left( 1 + a_p \beta_{\parallel} - a_p^2 \frac{\beta_{\parallel}}{2} \right); \\ A_0 &= k_{\parallel}^2 \frac{\beta_{\parallel}}{2} [k_{\parallel}^2 v_{A\parallel}^2 + k_{\perp}^2 (1 + a_p \beta_{\parallel} - a_p^2 \beta_{\parallel})]. \end{aligned} \right\} \quad (9.46)$$

In comparison with the dispersion relation (3.193) of the adiabatic CGL model, the ‘static’ closure eliminates the erroneous 1/6 factor in the  $A_0$  coefficient and yields the correct mirror threshold. The erroneous 1/6 factor in the CGL model can be therefore interpreted as a result of inadequacy of adiabatic closures in the very slow-dynamics context, such as the mirror instability. For completeness, solutions for parallel propagation ( $k_{\perp} = 0$ ) are  $\omega = \pm v_{A\parallel} k_{\parallel}$  and  $\omega = \pm \sqrt{(\beta_{\parallel}/2)} k_{\parallel}$ . For perpendicular propagation ( $k_{\parallel} = 0$ ) the solution is  $\omega = \pm \sqrt{1 + a_p \beta_{\parallel} - a_p^2 (\beta_{\parallel}/2)} k_{\perp}$ . Obviously, somewhere beyond the mirror threshold, the perpendicular fast mode experiences unphysical instability, which can be interpreted as a result of inadequacy of ‘static’ (isothermal) closures in the fast-dynamics context.

To clearly demonstrate that variations of magnetic field in the ‘static’ closures (9.45) are indeed crucial in recovering the mirror instability, let us neglect them and quickly consider strictly isothermal closure  $p_{\parallel} = nT_{\parallel}^{(0)}$ ;  $p_{\perp} = nT_{\perp}^{(0)}$ , which yields the following dispersion relation

$$\left. \begin{aligned} (\omega^2 - k_{\parallel}^2 v_{A\parallel}^2)(\omega^4 - A_2 \omega^2 + A_0) &= 0; \\ A_2 &= k_{\parallel}^2 \left( v_{A\parallel}^2 + \frac{\beta_{\parallel}}{2} \right) + k_{\perp}^2 \left( 1 + a_p \frac{\beta_{\parallel}}{2} \right); \\ A_0 &= k_{\parallel}^2 \frac{\beta_{\parallel}}{2} \left[ k_{\parallel}^2 v_{A\parallel}^2 + k_{\perp}^2 \left( 1 + a_p \frac{\beta_{\parallel}}{2} - a_p^2 \frac{\beta_{\parallel}}{2} \right) \right]. \end{aligned} \right\} \quad (9.47)$$

As can be seen from the  $A_0$  coefficient, this model does not recover the mirror instability. For parallel propagation the solutions are the same as for the ‘static’ closure, nevertheless, the perpendicular fast mode is now always stable with  $\omega = \pm \sqrt{1 + a_p (\beta_{\parallel}/2)} k_{\perp}$ . As a double check, after prescribing  $a_p = 1$ , the dispersion relation (9.47) is equal to (9.46), since the  $B_z$  contributions in the ‘static’ closure disappear. Also, we considered models with polytropic indices  $\gamma_{\parallel}, \gamma_{\perp}$  in § 3.11, and the dispersion relation (9.47) is consistent with (3.171), after prescribing  $\gamma_{\parallel} = 1$ ,  $\gamma_{\perp} = 1$ .

## 10. Bi-kappa fluid model (BiKappa)

### 10.1. Bi-kappa distribution function

Often in the solar wind and in space plasma physics, distribution functions with elongated high-energy ‘suprathermal’ tails are observed, that cannot be modelled with the bi-Maxwellian distribution. A slightly more general distribution function is the bi-kappa distribution, that contains one free parameter  $\kappa$ , see for example Summers & Thorne (1992), Pierrard & Lazar (2010), Livadiotis & McComas (2013)

and references therein. The  $\kappa$  parameter can be different in the parallel and transverse directions (see e.g. Basu (2009), dos Santos, Ziebell & Gaelzer (2015)). Here, we consider only one free  $\kappa$  parameter. The distribution has the following form

$$f_0 = n \frac{\Gamma(\kappa + 1)}{\pi^{3/2} \kappa^{3/2} \Gamma(\kappa - \frac{1}{2})} \frac{1}{\theta_{\parallel} \theta_{\perp}^2} \left[ 1 + \frac{(v_{\parallel} - u_{\parallel})^2}{\kappa \theta_{\parallel}^2} + \frac{|\mathbf{v}_{\perp} - \mathbf{u}_{\perp}|^2}{\kappa \theta_{\perp}^2} \right]^{-\kappa-1}, \quad (10.1)$$

where the generalized thermal speeds

$$\theta_{\parallel} = \sqrt{1 - \frac{3}{2\kappa}} \sqrt{\frac{2T_{\parallel}}{m}}; \quad \theta_{\perp} = \sqrt{1 - \frac{3}{2\kappa}} \sqrt{\frac{2T_{\perp}}{m}}, \quad (10.2)$$

and  $\Gamma$  is the usual gamma function. In the limit  $\kappa \rightarrow \infty$ , the generalized thermal speeds are the usual thermal speeds and the distribution function is bi-Maxwellian, since

$$\lim_{x \rightarrow \infty} \left(1 + \frac{a}{x}\right)^{-x} = e^{-a}; \quad \lim_{x \rightarrow \infty} \left(1 + \frac{a}{x}\right)^{-x-1} = e^{-a}. \quad (10.3)$$

Why the generalized thermal speeds are defined this way will become clear from later calculations of the moments, and the choice of the power law index  $-(\kappa + 1)$ , instead of perhaps the more logical  $-\kappa$ , is attributed to purely historical reasons. Calculating second-order (pressure) integrals with this distribution requires for convergence  $\kappa > 3/2$ , and this value is also the minimum value for the generalized thermal speeds (10.2) to have real values. Sometimes, in the definition of the isotropic kappa distribution, a value of  $\Gamma(3/2) = \sqrt{\pi}/2$  is used and  $\pi^{3/2} = 2\pi\sqrt{\pi}/2 = 2\pi\Gamma(3/2)$ . Again, for brevity of the calculations, it is useful to introduce  $\alpha_{\parallel} = 1/(\kappa\theta_{\parallel}^2)$ ,  $\alpha_{\perp} = 1/(\kappa\theta_{\perp}^2)$ , and rewrite the bi-kappa distribution (10.1) in the following form

$$f_0 = n \frac{\Gamma(\kappa + 1)}{\Gamma(\kappa - \frac{1}{2})} \sqrt{\frac{\alpha_{\parallel}}{\pi}} \frac{\alpha_{\perp}}{\pi} [1 + \alpha_{\parallel} c_{\parallel}^2 + \alpha_{\perp} c_{\perp}^2]^{-\kappa-1}. \quad (10.4)$$

The important integral is

$$\int_{-\infty}^{\infty} (1 + x^2)^{-a} dx = \sqrt{\pi} \frac{\Gamma(a - \frac{1}{2})}{\Gamma(a)}, \quad a > \frac{1}{2}. \quad (10.5)$$

For the critical value of  $a = 1/2$ , the integral  $\int 1/\sqrt{1+x^2} dx = \text{arcsinh}(x)$ , and since  $\lim_{x \rightarrow \pm\infty} \text{arcsinh}(x) = \pm\infty$ , the integral diverges. For an isotropic kappa distribution  $f(c^2)$ , the three-dimensional integrals can be conveniently evaluated by transformation to spherical coordinates

$$\int_{-\infty}^{\infty} f(c^2) d^3c = 4\pi \int_0^{\infty} c^2 f(c^2) dc. \quad (10.6)$$

In this case, the important integral is

$$\int_0^{\infty} x^2 (1 + x^2)^{-a} dx = \frac{\sqrt{\pi}}{4} \frac{\Gamma(a - \frac{3}{2})}{\Gamma(a)}; \quad a > \frac{3}{2}, \quad (10.7)$$

and by a direct substitution

$$\int_0^{\infty} x^2 (1 + \alpha x^2)^{-(\kappa+1)} dx = \frac{\sqrt{\pi}}{4\alpha^{3/2}} \frac{\Gamma(\kappa - \frac{1}{2})}{\Gamma(\kappa + 1)}; \quad \kappa > \frac{1}{2}. \quad (10.8)$$

For an isotropic kappa distribution therefore

$$\int_{-\infty}^{\infty} (1 + \alpha c^2)^{-(\kappa+1)} d^3c = 4\pi \int_0^{\infty} c^2 (1 + \alpha c^2)^{-(\kappa+1)} dc = \left(\frac{\pi}{\alpha}\right)^{3/2} \frac{\Gamma(\kappa - \frac{1}{2})}{\Gamma(\kappa + 1)}, \quad (10.9)$$

which explains the normalization factors. For an anisotropic bi-kappa distributions  $f(c_{\parallel}^2, c_{\perp}^2)$ , the integration is slightly more complicated because, in contrast to a bi-Maxwellian distribution, the integration cannot be fully separated into two independent integrations over parallel and perpendicular velocity components, and has to be performed successively. It is possible to use a cylindrical coordinate system and integrate

$$\int_{-\infty}^{\infty} f(c_{\parallel}^2, c_{\perp}^2) d^3c = 2\pi \int_{c_{\perp}=0}^{\infty} \int_{c_{\parallel}=-\infty}^{\infty} f(c_{\parallel}^2, c_{\perp}^2) c_{\perp} dc_{\parallel} dc_{\perp}. \quad (10.10)$$

However, here we keep the old fashioned Cartesian coordinates, and at the expense of a slightly bit more algebra (i.e. one more integration), we will integrate over  $d^3c = dc_{\parallel} d^2c_{\perp} = dc_{\parallel} dc_x dc_y$ . A small added benefit is that we do not have to follow whether the integrals are  $\int_{-\infty}^{\infty}$  or  $\int_0^{\infty}$ , and we can stop writing the boundaries.

The most important integral is a slight generalization of (10.5) in the form

$$\int_{-\infty}^{\infty} (1 + b + \alpha x^2)^{-a} dx = \sqrt{\frac{\pi}{\alpha}} \frac{\Gamma(a - \frac{1}{2})}{\Gamma(a)} (1 + b)^{-(a-\frac{1}{2})}; \quad a > \frac{1}{2}, \quad (10.11)$$

where for simplicity  $b, \alpha$  are assumed to be positive constants. So one can calculate

$$\int_{-\infty}^{\infty} (1 + \alpha_x x^2 + \alpha_y y^2 + \alpha_z z^2)^{-a} dx = \sqrt{\frac{\pi}{\alpha_x}} \frac{\Gamma(a - \frac{1}{2})}{\Gamma(a)} (1 + \alpha_y y^2 + \alpha_z z^2)^{-(a-1/2)}, \quad (10.12)$$

where, importantly, the ‘power law’ changed from  $-a$  to  $-(a - 1/2)$ . Performing a successive three-dimensional integral therefore yields

$$\begin{aligned} & \int_{-\infty}^{\infty} (1 + \alpha_x x^2 + \alpha_y y^2 + \alpha_z z^2)^{-a} dx dy dz \\ &= \sqrt{\frac{\pi}{\alpha_x}} \frac{\Gamma(a - \frac{1}{2})}{\Gamma(a)} \int (1 + \alpha_y y^2 + \alpha_z z^2)^{-(a-1/2)} dy dz \\ &= \sqrt{\frac{\pi}{\alpha_x}} \sqrt{\frac{\pi}{\alpha_y}} \frac{\Gamma(a - \frac{1}{2})}{\Gamma(a)} \frac{\Gamma(a - 1)}{\Gamma(a - \frac{1}{2})} \int (1 + \alpha_z z^2)^{-(a-1)} dz \\ &= \sqrt{\frac{\pi}{\alpha_x}} \sqrt{\frac{\pi}{\alpha_y}} \sqrt{\frac{\pi}{\alpha_z}} \frac{\Gamma(a - \frac{1}{2})}{\Gamma(a)} \frac{\Gamma(a - 1)}{\Gamma(a - \frac{1}{2})} \frac{\Gamma(a - \frac{3}{2})}{\Gamma(a - 1)} \\ &= \sqrt{\frac{\pi}{\alpha_x}} \sqrt{\frac{\pi}{\alpha_y}} \sqrt{\frac{\pi}{\alpha_z}} \frac{\Gamma(a - \frac{3}{2})}{\Gamma(a)}; \quad a > \frac{3}{2}, \end{aligned} \quad (10.13)$$

and integrating the bi-kappa distribution over the velocity space

$$\int_{-\infty}^{\infty} (1 + \alpha_{\parallel} c_{\parallel}^2 + \alpha_{\perp} c_{\perp}^2)^{-(\kappa+1)} dc_{\parallel} d^2c_{\perp} = \sqrt{\frac{\pi}{\alpha_{\parallel}}} \frac{\pi}{\alpha_{\perp}} \frac{\Gamma(\kappa - \frac{1}{2})}{\Gamma(\kappa + 1)}; \quad \kappa > \frac{1}{2}, \quad (10.14)$$

which verifies that the normalization constants in the bi-kappa definition (10.4) are indeed correct, and that  $\int f_0 d^3c = n$ . For later calculations, it is useful to write down two partial integrals when the integration is done over  $dc_{\parallel}$  and  $d^2c_{\perp}$  that read

$$\int_{-\infty}^{\infty} (1 + \alpha_{\parallel} c_{\parallel}^2)^{-a} dc_{\parallel} = \sqrt{\frac{\pi}{\alpha_{\parallel}}} \frac{\Gamma(a - \frac{1}{2})}{\Gamma(a)}; \quad (10.15)$$

$$\int_{-\infty}^{\infty} (1 + \alpha_{\parallel} c_{\parallel}^2 + \alpha_{\perp} c_{\perp}^2)^{-(\kappa+1)} dc_{\parallel} = \sqrt{\frac{\pi}{\alpha_{\parallel}}} \frac{\Gamma(\kappa + \frac{1}{2})}{\Gamma(\kappa + 1)} (1 + \alpha_{\perp} c_{\perp}^2)^{-(\kappa+1/2)}; \quad (10.16)$$

$$\int f_0 dc_{\parallel} = n \frac{\Gamma(\kappa + \frac{1}{2})}{\Gamma(\kappa - \frac{1}{2})} \frac{\alpha_{\perp}}{\pi} (1 + \alpha_{\perp} c_{\perp}^2)^{-(\kappa+1/2)}, \quad (10.17)$$

and

$$\int_{-\infty}^{\infty} (1 + \alpha_{\perp} c_{\perp}^2)^{-a} d^2c_{\perp} = \frac{\pi}{\alpha_{\perp}} \frac{\Gamma(a - 1)}{\Gamma(a)}; \quad (10.18)$$

$$\int (1 + \alpha_{\parallel} c_{\parallel}^2 + \alpha_{\perp} c_{\perp}^2)^{-(\kappa+1)} d^2c_{\perp} = \frac{\pi}{\alpha_{\perp}} \frac{\Gamma(\kappa)}{\Gamma(\kappa + 1)} (1 + \alpha_{\parallel} c_{\parallel}^2)^{-\kappa}; \quad (10.19)$$

$$\int f_0 d^2c_{\perp} = n \frac{\Gamma(\kappa)}{\Gamma(\kappa - \frac{1}{2})} \sqrt{\frac{\alpha_{\parallel}}{\pi}} (1 + \alpha_{\parallel} c_{\parallel}^2)^{-\kappa}. \quad (10.20)$$

Continuing with the calculation of velocity moments, it is easy to show that  $\int c f_0 d^3c = 0$ , which verifies  $n\mathbf{u} = \int \mathbf{v} f_0 d^3v$ . Continuing with the higher-order moments, it is possible to calculate those from table integrals (a great help for double checking is an analytic software like Maple or Mathematica) or similarly for a bi-Maxwellian distribution to use a trick with the differentiation with respect to  $\alpha$  as

$$\begin{aligned} \int_{-\infty}^{\infty} x^2 (1 + \alpha x^2)^{-(\kappa+1)} dx &= -\frac{1}{\kappa} \frac{\partial}{\partial \alpha} \int_{-\infty}^{\infty} (1 + \alpha x^2)^{-\kappa} dx = -\frac{1}{\kappa} \sqrt{\pi} \frac{\Gamma(\kappa - \frac{1}{2})}{\Gamma(\kappa)} \frac{\partial}{\partial \alpha} \alpha^{-1/2} \\ &= \frac{\sqrt{\pi}}{2\alpha^{3/2}} \frac{\Gamma(\kappa - \frac{1}{2})}{\Gamma(\kappa + 1)}, \end{aligned} \quad (10.21)$$

where we have also used a property of the gamma function  $x\Gamma(x) = \Gamma(x+1)$ , and the last integral requires  $\kappa > \frac{1}{2}$ . A slightly more general integral is

$$\int_{-\infty}^{\infty} x^2 (1 + b + \alpha x^2)^{-a} dx = \frac{\sqrt{\pi}}{2\alpha^{3/2}} \frac{\Gamma(a - \frac{3}{2})}{\Gamma(a)} (1 + b)^{-(a-3/2)}, \quad (10.22)$$

which can be used to evaluate the following integrals

$$\int_{-\infty}^{\infty} c_{\parallel}^2 (1 + \alpha_{\parallel} c_{\parallel}^2)^{-a} dc_{\parallel} = \frac{\sqrt{\pi}}{2\alpha_{\parallel}^{3/2}} \frac{\Gamma(a - \frac{3}{2})}{\Gamma(a)}; \quad (10.23)$$



$$\begin{aligned}
\int_{-\infty}^{\infty} c_x^2 (1 + \alpha_{\perp} c_{\perp}^2)^{-a} d^2 c_{\perp} &= \int_{-\infty}^{\infty} c_x^2 (1 + \alpha_{\perp} c_x^2 + \alpha_{\perp} c_y^2)^{-a} dc_x dc_y \\
&= \frac{\sqrt{\pi}}{2\alpha_{\perp}^{3/2}} \frac{\Gamma(a - \frac{3}{2})}{\Gamma(a)} \int_{-\infty}^{\infty} (1 + \alpha_{\perp} c_y^2)^{-(a-3/2)} dc_y \\
&= \frac{\sqrt{\pi}}{2\alpha_{\perp}^{3/2}} \frac{\Gamma(a - \frac{3}{2})}{\Gamma(a)} \sqrt{\frac{\pi}{\alpha_{\perp}}} \frac{\Gamma(a-2)}{\Gamma(a - \frac{3}{2})} = \frac{\pi}{2\alpha_{\perp}^2} \frac{\Gamma(a-2)}{\Gamma(a)};
\end{aligned} \tag{10.24}$$

$$\int_{-\infty}^{\infty} c_{\perp}^2 (1 + \alpha_{\perp} c_{\perp}^2)^{-a} d^2 c_{\perp} = \frac{\pi}{\alpha_{\perp}^2} \frac{\Gamma(a-2)}{\Gamma(a)}. \tag{10.25}$$

Since we like to double check everything, the above integral can be also calculated as

$$\begin{aligned}
\int_{-\infty}^{\infty} c_{\perp}^2 (1 + \alpha_{\perp} c_{\perp}^2)^{-a} d^2 c_{\perp} &= -\frac{1}{a-1} \frac{\partial}{\partial \alpha_{\perp}} \int_{-\infty}^{\infty} (1 + \alpha_{\perp} c_{\perp}^2)^{-(a-1)} d^2 c_{\perp} \\
&= -\frac{1}{a-1} \frac{\Gamma(a-2)}{\Gamma(a-1)} \frac{\partial}{\partial \alpha_{\perp}} \frac{\pi}{\alpha_{\perp}} \\
&= \frac{\pi}{\alpha_{\perp}^2} \frac{\Gamma(a-2)}{\Gamma(a)}.
\end{aligned} \tag{10.26}$$

For the calculation of the parallel pressure, we first integrate over  $d^2 c_{\perp}$  so we can use result (10.20) and then over  $dc_{\parallel}$

$$\begin{aligned}
p_{\parallel} \equiv m \int c_{\parallel}^2 f_0 d^2 c_{\perp} dc_{\parallel} &= mn \frac{\Gamma(\kappa)}{\Gamma(\kappa - \frac{1}{2})} \sqrt{\frac{\alpha_{\parallel}}{\pi}} \int c_{\parallel}^2 (1 + \alpha_{\parallel} c_{\parallel}^2)^{-\kappa} dc_{\parallel} \\
&= mn \frac{\Gamma(\kappa)}{\Gamma(\kappa - \frac{1}{2})} \sqrt{\frac{\alpha_{\parallel}}{\pi}} \frac{\sqrt{\pi}}{2\alpha_{\parallel}^{3/2}} \frac{\Gamma(\kappa - \frac{3}{2})}{\Gamma(\kappa)} \\
&= mn \frac{\Gamma(\kappa - \frac{3}{2})}{\Gamma(\kappa - \frac{1}{2})} \frac{1}{2\alpha_{\parallel}} = mn \frac{\Gamma(\kappa - \frac{3}{2})}{\Gamma(\kappa - \frac{1}{2})} \frac{1}{2} \left( \kappa - \frac{3}{2} \right) \frac{2T_{\parallel}}{m} = nT_{\parallel},
\end{aligned} \tag{10.27}$$

where we have used  $(\kappa - \frac{3}{2})\Gamma(\kappa - \frac{3}{2}) = \Gamma(\kappa - \frac{1}{2})$ . Calculating the perpendicular pressure, we here first integrate over  $dc_{\parallel}$  so we can use result (10.17) and then over  $d^2 c_{\perp}$

$$\begin{aligned}
p_{\perp} \equiv \frac{m}{2} \int c_{\perp}^2 f_0 dc_{\parallel} d^2 c_{\perp} &= \frac{m}{2} n \frac{\Gamma(\kappa + \frac{1}{2})}{\Gamma(\kappa - \frac{1}{2})} \frac{\alpha_{\perp}}{\pi} \int c_{\perp}^2 (1 + \alpha_{\perp} c_{\perp}^2)^{-(\kappa+1/2)} d^2 c_{\perp} \\
&= \frac{m}{2} n \frac{\Gamma(\kappa + \frac{1}{2})}{\Gamma(\kappa - \frac{1}{2})} \frac{\alpha_{\perp}}{\pi} \frac{\pi}{\alpha_{\perp}^2} \frac{\Gamma(\kappa - \frac{3}{2})}{\Gamma(\kappa + \frac{1}{2})} \\
&= mn \frac{\Gamma(\kappa - \frac{3}{2})}{\Gamma(\kappa - \frac{1}{2})} \frac{1}{2\alpha_{\perp}} = mn \frac{\Gamma(\kappa - \frac{3}{2})}{\Gamma(\kappa - \frac{1}{2})} \frac{1}{2} \left( \kappa - \frac{3}{2} \right) \frac{2T_{\perp}}{m} = nT_{\perp}.
\end{aligned} \tag{10.28}$$

The last two calculations clarify the reasoning behind the definition of the bi-kappa generalized thermal speeds  $\theta_{\parallel}$ ,  $\theta_{\perp}$  (here rewritten with the notation  $\alpha_{\parallel}$  and  $\alpha_{\perp}$ ), where one starts with a bi-kappa distribution with unspecified  $\alpha_{\parallel}$  and  $\alpha_{\perp}$ , and calculating the parallel and perpendicular pressure integrals, yields the required forms  $\alpha_{\parallel}^{-1} = (\kappa - \frac{3}{2})(2T_{\parallel}/m)$ ;  $\alpha_{\perp}^{-1} = (\kappa - \frac{3}{2})(2T_{\perp}/m)$ . Then, the split to  $\alpha_{\parallel}^{-1} = \kappa\theta_{\parallel}^2$ ;  $\alpha_{\perp}^{-1} = \kappa\theta_{\perp}^2$  is dictated by the requirement that for  $\kappa \rightarrow \infty$ , the generalized thermal speeds converge to the usual thermal speeds, and which also yields that in this limit the distribution is bi-Maxwellian.

The bi-kappa distribution was specified with perhaps a somewhat obscure power law  $-(\kappa + 1)$  instead of the more logical  $-\kappa$ , which is attributed to purely historical reasons. If one wants to define an anisotropic power-law distribution, a first guess would be

$$\tilde{f}_0 \sim n[1 + \alpha_{\parallel}c_{\parallel}^2 + \alpha_{\perp}c_{\perp}^2]^{-\kappa}, \quad (10.29)$$

which then yields normalization constants

$$\tilde{f}_0 = n \frac{\Gamma(\kappa)}{\Gamma(\kappa - \frac{3}{2})} \sqrt{\frac{\alpha_{\parallel}}{\pi}} \frac{\alpha_{\perp}}{\pi} [1 + \alpha_{\parallel}c_{\parallel}^2 + \alpha_{\perp}c_{\perp}^2]^{-\kappa}. \quad (10.30)$$

Then, calculating the parallel and perpendicular pressures yields the requirement  $\alpha_{\parallel,\perp}^{-1} = (\kappa - \frac{5}{2})(2T_{\parallel,\perp}/m)$ . Further writing  $\alpha_{\parallel,\perp}^{-1} = \kappa\theta_{\parallel,\perp}^2$  yields a distribution function

$$\tilde{f}_0 = n \frac{\Gamma(\kappa)}{\pi^{3/2} \kappa^{3/2} \Gamma(\kappa - \frac{3}{2})} \frac{1}{\theta_{\parallel}\theta_{\perp}^2} \left[ 1 + \frac{c_{\parallel}^2}{\kappa\theta_{\parallel}^2} + \frac{c_{\perp}^2}{\kappa\theta_{\perp}^2} \right]^{-\kappa}, \quad (10.31)$$

with generalized thermal speeds

$$\theta_{\parallel} = \sqrt{1 - \frac{5}{2\kappa}} \sqrt{\frac{2T_{\parallel}}{m}}; \quad \theta_{\perp} = \sqrt{1 - \frac{5}{2\kappa}} \sqrt{\frac{2T_{\perp}}{m}}. \quad (10.32)$$

The convergence of the pressure integrals requires  $\kappa > \frac{5}{2}$ , which is also the limiting case for the thermal speeds to be real. In the limit  $\kappa \rightarrow \infty$  the generalized thermal speeds converge to the usual thermal speeds and the distribution converges to the bi-Maxwellian. To double check that the distribution (10.31) is really correct, one can of course substitute  $\kappa \rightarrow \kappa + 1$ , which fully recovers the original bi-kappa distribution (10.1), since  $\kappa\theta^2 \rightarrow (\kappa + 1)(1 - 5/2(\kappa + 1))(2T/m) = \kappa(1 - 3/2\kappa)(2T/m)$ . From now on, let us continue calculations with the original bi-kappa distribution (10.1), that uses the  $-(\kappa + 1)$  power-law index.

## 10.2. Bi-kappa fluid closure

Now that we are familiar with the bi-kappa distribution, we are ready to calculate higher-order moments. The heat fluxes are zero and easy to calculate, since both  $q_{\parallel}$ ,  $q_{\perp}$  are anti-symmetric in  $c_{\parallel}$ , and by a direct calculation

$$\begin{aligned} & \int_{-\infty}^{\infty} c_{\parallel}^3 (1 + \alpha_{\parallel}c_{\parallel}^2 + \alpha_{\perp}c_{\perp}^2)^{-a} dc_{\parallel} d^2c_{\perp} \\ &= -\frac{1}{a-1} \frac{\partial}{\partial \alpha_{\parallel}} \underbrace{\int_{-\infty}^{\infty} c_{\parallel} (1 + \alpha_{\parallel}c_{\parallel}^2 + \alpha_{\perp}c_{\perp}^2)^{-(a-1)} dc_{\parallel} d^2c_{\perp}}_{=0} = 0, \end{aligned} \quad (10.33)$$

implying

$$q_{\parallel} \equiv m \int c_{\parallel}^3 f_0 d^3 c = 0; \quad q_{\perp} \equiv \frac{m}{2} \int c_{\parallel} c_{\perp}^2 f_0 d^3 c = 0. \quad (10.34)$$

As discussed previously for the bi-Maxwellian distribution, the heat fluxes must be kept in the general form, since we want to consider closure performed at the fourth-order moment level. To calculate the fourth-order fluid moments, we will need the following integrals

$$\int_{-\infty}^{\infty} c_{\parallel}^4 (1 + \alpha_{\parallel} c_{\parallel}^2)^{-a} dc_{\parallel} = \frac{3}{4} \frac{\sqrt{\pi}}{\alpha_{\parallel}^{5/2}} \frac{\Gamma(a - \frac{5}{2})}{\Gamma(a)}; \quad a > \frac{5}{2}; \quad (10.35)$$

$$\int_{-\infty}^{\infty} c_{\parallel}^2 (1 + \alpha_{\parallel} c_{\parallel}^2 + \alpha_{\perp} c_{\perp}^2)^{-a} dc_{\parallel} = \frac{\sqrt{\pi}}{2\alpha_{\parallel}^{3/2}} \frac{\Gamma(a - \frac{3}{2})}{\Gamma(a)} (1 + \alpha_{\perp} c_{\perp}^2)^{-(a-3/2)}; \quad a > \frac{3}{2}; \quad (10.36)$$

$$\int_{-\infty}^{\infty} c_{\parallel}^2 c_{\perp}^2 (1 + \alpha_{\parallel} c_{\parallel}^2 + \alpha_{\perp} c_{\perp}^2)^{-a} dc_{\parallel} d^2 c_{\perp} = \frac{\sqrt{\pi}}{2\alpha_{\parallel}^{3/2}} \frac{\pi}{\alpha_{\perp}^2} \frac{\Gamma(a - \frac{7}{2})}{\Gamma(a)}; \quad a > \frac{7}{2}; \quad (10.37)$$

$$\int_{-\infty}^{\infty} c_{\perp}^4 (1 + \alpha_{\perp} c_{\perp}^2)^{-a} d^2 c_{\perp} = \frac{2\pi}{\alpha_{\perp}^3} \frac{\Gamma(a - 3)}{\Gamma(a)}; \quad a > 3. \quad (10.38)$$

The  $r_{\parallel\parallel}$  moment then calculates as

$$\begin{aligned} r_{\parallel\parallel} &\equiv m \int c_{\parallel}^4 f_0 d^2 c_{\perp} dc_{\parallel} = mn \frac{\Gamma(\kappa)}{\Gamma(\kappa - \frac{1}{2})} \sqrt{\frac{\alpha_{\parallel}}{\pi}} \int c_{\parallel}^4 (1 + \alpha_{\parallel} c_{\parallel}^2)^{-\kappa} dc_{\parallel} \\ &= mn \frac{\Gamma(\kappa)}{\Gamma(\kappa - \frac{1}{2})} \sqrt{\frac{\alpha_{\parallel}}{\pi}} \frac{3}{4} \frac{\sqrt{\pi}}{\alpha_{\parallel}^{5/2}} \frac{\Gamma(\kappa - \frac{5}{2})}{\Gamma(\kappa)} \\ &= mn \frac{3}{4} \frac{\Gamma(\kappa - \frac{5}{2})}{\Gamma(\kappa - \frac{1}{2})} \frac{1}{\alpha_{\parallel}^2} = mn \frac{3}{4} \frac{\Gamma(\kappa - \frac{5}{2})}{\Gamma(\kappa - \frac{1}{2})} (\kappa - \frac{3}{2})^2 \frac{4T_{\parallel}^2}{m^2} = 3 \frac{p_{\parallel}^2}{\rho} \frac{(\kappa - 3/2)}{(\kappa - 5/2)}; \quad \kappa > \frac{5}{2}, \end{aligned} \quad (10.39)$$

where in the last step we used that  $\Gamma(x)(x+1)^2/\Gamma(x+2) = (x+1)/x$  with  $x = \kappa - 5/2$ . The  $r_{\perp\perp}$  moment calculates as

$$\begin{aligned} r_{\perp\perp} &\equiv \frac{m}{2} \int c_{\parallel}^2 c_{\perp}^2 f_0 dc_{\parallel} d^2 c_{\perp} = \frac{m}{2} n \frac{\Gamma(\kappa + 1)}{\Gamma(\kappa - \frac{1}{2})} \sqrt{\frac{\alpha_{\parallel}}{\pi}} \frac{\alpha_{\perp}}{\pi} \frac{\sqrt{\pi}}{2\alpha_{\parallel}^{3/2}} \frac{\pi}{\alpha_{\perp}^2} \frac{\Gamma(\kappa - \frac{5}{2})}{\Gamma(\kappa + 1)} \\ &= \frac{1}{4} mn \frac{\Gamma(\kappa - \frac{5}{2})}{\Gamma(\kappa - \frac{1}{2})} \frac{1}{\alpha_{\parallel} \alpha_{\perp}} \\ &= \frac{1}{4} mn \frac{\Gamma(\kappa - \frac{5}{2})}{\Gamma(\kappa - \frac{1}{2})} \left( \kappa - \frac{3}{2} \right)^2 \frac{4T_{\parallel} T_{\perp}}{m^2} = \frac{p_{\parallel} p_{\perp}}{\rho} \frac{(\kappa - 3/2)}{(\kappa - 5/2)}; \quad \kappa > \frac{5}{2}, \end{aligned} \quad (10.40)$$

and finally the  $r_{\perp\perp}$  moment calculates as

$$\begin{aligned} r_{\perp\perp} &\equiv \frac{m}{4} \int c_{\perp}^4 f_0 dc_{\parallel} d^2 c_{\perp} = \frac{m}{4} n \frac{\Gamma(\kappa + \frac{1}{2})}{\Gamma(\kappa - \frac{1}{2})} \frac{\alpha_{\perp}}{\pi} \int c_{\perp}^4 (1 + \alpha_{\perp} c_{\perp}^2)^{-(\kappa+1/2)} d^2 c_{\perp} \\ &= \frac{m}{4} n \frac{\Gamma(\kappa + \frac{1}{2})}{\Gamma(\kappa - \frac{1}{2})} \frac{\alpha_{\perp}}{\pi} \frac{2\pi}{\alpha_{\perp}^3} \frac{\Gamma(\kappa - \frac{5}{2})}{\Gamma(\kappa + \frac{1}{2})} = \frac{1}{2} mn \frac{\Gamma(\kappa - \frac{5}{2})}{\Gamma(\kappa - \frac{1}{2})} \frac{1}{\alpha_{\perp}^2} = 2 \frac{p_{\perp}^2}{\rho} \frac{(\kappa - 3/2)}{(\kappa - 5/2)}; \quad \kappa > \frac{5}{2}. \end{aligned} \quad (10.41)$$

To summarize, the following closure can be constructed

$$r_{\parallel\parallel} = \alpha_\kappa \frac{3p_\parallel^2}{\rho}; \quad r_{\parallel\perp} = \alpha_\kappa \frac{p_\parallel p_\perp}{\rho}; \quad r_{\perp\perp} = \alpha_\kappa \frac{2p_\perp^2}{\rho}, \quad (10.42)$$

where for the bi-kappa distribution (10.1), the  $\alpha_\kappa$  coefficient reads

$$\alpha_\kappa = \frac{\kappa - \frac{3}{2}}{\kappa - \frac{5}{2}}; \quad \kappa > \frac{5}{2}. \quad (10.43)$$

It is noteworthy that all three fourth-order moments are just multiplied by the same constant  $\alpha_\kappa$ . Obviously, in the limit  $\kappa \rightarrow \infty$ ,  $\alpha_\kappa \rightarrow 1$ , and the ‘normal’ bi-Maxwellian closure is obtained.

As suggested for example by Chust & Belmont (2006), a very large class of distribution functions can be accounted for by considering closures

$$r_{\parallel\parallel} = \alpha_{\parallel\parallel} \frac{3p_\parallel^2}{\rho}; \quad r_{\parallel\perp} = \alpha_{\parallel\perp} \frac{p_\parallel p_\perp}{\rho}; \quad r_{\perp\perp} = \alpha_{\perp\perp} \frac{2p_\perp^2}{\rho}, \quad (10.44)$$

where the 3 proportionality constants  $\alpha_{\parallel\parallel}$ ,  $\alpha_{\parallel\perp}$ ,  $\alpha_{\perp\perp}$  have to be determined, once a specific distribution function is prescribed. Such an unspecified closure can account for a very wide class of distribution functions. Also, a much more complicated distribution functions can be modelled by considering multi-fluid models. For example, one could consider a plasma consisting of 3 different electron species, that would describe the core, halo and the tail/strahl parts of the electron distribution function. Further information about the core, halo and strahl components of the electron distribution function can be found for example in Vocks *et al.* (2005), Vocks, Mann & Rausche (2008), Pierrard, Lazar & Schlickeiser (2011) and references therein.

Here, we consider closure (10.42), (10.43). Nevertheless, in the following calculations one can keep the value of  $\alpha_\kappa$  unspecified. Therefore, the results are actually valid for a much larger class of fluid models, and not just for a bi-kappa distribution. Let us completely neglect the non-gyrotropic  $\Pi$  and  $\mathbf{r}^{\text{ng}}$  in the heat flux equations (7.98) and (7.99). We need to calculate

$$\nabla \cdot (r_{\parallel\parallel} \hat{\mathbf{b}}) - 3r_{\parallel\perp} \nabla \cdot \hat{\mathbf{b}} = \alpha_\kappa \left[ 3 \frac{p_\parallel^2}{\rho} \nabla \cdot \hat{\mathbf{b}} + 3p_\parallel \hat{\mathbf{b}} \cdot \nabla \left( \frac{p_\parallel}{\rho} \right) + 3 \frac{p_\parallel}{\rho} \hat{\mathbf{b}} \cdot \nabla p_\parallel - 3 \frac{p_\parallel p_\perp}{\rho} \nabla \cdot \hat{\mathbf{b}} \right]; \quad (10.45)$$

$$\nabla \cdot (r_{\parallel\perp} \hat{\mathbf{b}}) + (r_{\parallel\perp} - r_{\perp\perp}) \nabla \cdot \hat{\mathbf{b}} = \alpha_\kappa \left[ \frac{p_\perp}{\rho} \hat{\mathbf{b}} \cdot \nabla p_\parallel + p_\parallel \hat{\mathbf{b}} \cdot \nabla \left( \frac{p_\perp}{\rho} \right) + 2 \frac{p_\perp}{\rho} (p_\parallel - p_\perp) \nabla \cdot \hat{\mathbf{b}} \right], \quad (10.46)$$

and the nonlinear heat flux equations for the bi-kappa fluid model read

$$\begin{aligned} \frac{\partial q_\parallel}{\partial t} + \mathbf{u} \cdot \nabla q_\parallel + q_\parallel \nabla \cdot \mathbf{u} + \alpha_\kappa 3p_\parallel \hat{\mathbf{b}} \cdot \nabla \left( \frac{p_\parallel}{\rho} \right) + 3q_\parallel \hat{\mathbf{b}} \cdot \nabla \mathbf{u} \cdot \hat{\mathbf{b}} \\ + (\alpha_\kappa - 1) \frac{3p_\parallel}{\rho} [(p_\parallel - p_\perp) \nabla \cdot \hat{\mathbf{b}} + \hat{\mathbf{b}} \cdot \nabla p_\parallel] = 0; \end{aligned} \quad (10.47)$$

$$\begin{aligned} \frac{\partial q_{\perp}}{\partial t} + \mathbf{u} \cdot \nabla q_{\perp} + 2q_{\perp} \nabla \cdot \mathbf{u} + \alpha_{\kappa} p_{\parallel} \hat{\mathbf{b}} \cdot \nabla \left( \frac{p_{\perp}}{\rho} \right) + (2\alpha_{\kappa} - 1) \frac{p_{\perp}}{\rho} (p_{\parallel} - p_{\perp}) \nabla \cdot \hat{\mathbf{b}} \\ + (\alpha_{\kappa} - 1) \frac{p_{\perp}}{\rho} \hat{\mathbf{b}} \cdot \nabla p_{\parallel} = 0. \end{aligned} \quad (10.48)$$

Obviously, for  $\alpha_{\kappa} = 1$  (i.e. for  $\kappa \rightarrow \infty$ ), the bi-Maxwellian heat flux equations are obtained. The model is now closed, and it is accompanied by (9.2)–(9.6). The model is nonlinear and numerical simulations for a chosen (and fixed) value of  $\kappa$  can be considered. Even though the  $\kappa$  coefficient is fixed, similarly to bi-Maxwellian nonlinear simulations, the temperature anisotropy is free to develop, and complicated plasma heating processes can be studied (we recommend adding the FLR pressure stress forces to the scalar pressure equations). We abbreviate the non-dispersive fluid model that uses the above heat flux equations as ‘BiKappa’. If the Hall term is considered in the induction equation, the ‘Hall-BiKappa’ fluid model is obtained. Similarly, considering FLR1 and FLR2 corrections to the pressure tensor, yields the ‘Hall-BiKappa-FLR1’ and ‘Hall-BiKappa-FLR2’ fluid models. The heat flux equations are normalized, linearized, written in the  $x$ – $z$  plane and Fourier transformed according to

$$-\omega q_{\parallel} + \alpha_{\kappa} \frac{3}{2} \beta_{\parallel} k_{\parallel} (p_{\parallel} - \rho) + (\alpha_{\kappa} - 1) \frac{3}{2} \beta_{\parallel} [(1 - a_p) k_{\perp} B_x + k_{\parallel} p_{\parallel}] = 0; \quad (10.49)$$

$$-\omega q_{\perp} + \alpha_{\kappa} \frac{\beta_{\parallel}}{2} k_{\parallel} (p_{\perp} - a_p \rho) + (2\alpha_{\kappa} - 1) \frac{\beta_{\parallel}}{2} a_p (1 - a_p) k_{\perp} B_x + (\alpha_{\kappa} - 1) \frac{\beta_{\parallel}}{2} a_p k_{\parallel} p_{\parallel} = 0, \quad (10.50)$$

where we have used the same plasma beta definition as always  $\beta_{\parallel}/2 = p_{\parallel}^{(0)}/(V_A^2 \rho_0)$ . The heat flux equations are accompanied by the linearized system (9.11)–(9.19).

To consider FLR3 corrections, one needs to derive non-gyrotropic heat flux vectors  $\mathbf{S}_{\perp}^{\parallel}$ ,  $\mathbf{S}_{\perp}^{\perp}$  for the bi-kappa distribution. For the first-order vectors the derivation is straightforward, and by following the calculations in appendix D, it is easy to show that the nonlinear

$$\begin{aligned} \mathbf{S}_{\perp}^{\parallel(1)} = \frac{B_0}{\Omega |\mathbf{B}|} \hat{\mathbf{b}} \times \left[ \alpha_{\kappa} p_{\perp} \nabla \left( \frac{p_{\parallel}}{\rho} \right) + (\alpha_{\kappa} - 1) \frac{p_{\parallel}}{\rho} \nabla p_{\perp} \right. \\ \left. + (3\alpha_{\kappa} - 1) \frac{p_{\parallel}}{\rho} (p_{\parallel} - p_{\perp}) \hat{\mathbf{b}} \cdot \nabla \hat{\mathbf{b}} + 2q_{\parallel} \hat{\mathbf{b}} \cdot \nabla \mathbf{u} + 2q_{\perp} \hat{\mathbf{b}} \times \boldsymbol{\omega} \right]; \end{aligned} \quad (10.51)$$

$$\begin{aligned} \mathbf{S}_{\perp}^{\perp(1)} = \frac{B_0}{\Omega |\mathbf{B}|} \hat{\mathbf{b}} \times \left[ 2\alpha_{\kappa} p_{\perp} \nabla \left( \frac{p_{\perp}}{\rho} \right) \right. \\ \left. + 2(\alpha_{\kappa} - 1) \frac{p_{\perp}}{\rho} \nabla p_{\perp} + 2(\alpha_{\kappa} - 1) \frac{p_{\perp}}{\rho} (p_{\parallel} - p_{\perp}) \hat{\mathbf{b}} \cdot \nabla \hat{\mathbf{b}} + 4q_{\perp} \hat{\mathbf{b}} \cdot \nabla \mathbf{u} \right]. \end{aligned} \quad (10.52)$$

Here, we neglected perturbations  $\tilde{r}$ , since we will not consider Landau fluid models for bi-kappa distribution. For  $\alpha_{\kappa} = 1$ , the results are of course equivalent to the bi-Maxwellian expressions (D 37), (D 60). The second-order heat flux vectors are not addressed here, since we would have to derive  $\mathbf{r}^{\text{ng}}$  for the bi-kappa distribution.

## 10.3. BiKappa dispersion relation

The BiKappa dispersion relation can be written in the following form

$$(\omega^2 - v_{A\parallel}^2 k_{\parallel}^2)(\omega^8 - A_6 \omega^6 + A_4 \omega^4 - A_2 \omega^2 + A_0) = 0; \quad (10.53)$$

$$A_6 = k_{\parallel}^2 (v_{A\parallel}^2 + \alpha_{\kappa} \frac{7}{2} \beta_{\parallel}) + k_{\perp}^2 (1 + a_p \beta); \quad (10.54)$$

$$A_4 = \alpha_{\kappa} k_{\parallel}^2 \beta_{\parallel} \left\{ k_{\parallel}^2 \left( \frac{7}{2} v_{A\parallel}^2 + \frac{(3 + 6\alpha_{\kappa})}{4} \beta_{\parallel} \right) + k_{\perp}^2 \frac{7}{2} \left( 1 + a_p \beta_{\parallel} - \frac{1}{7} a_p^2 \beta_{\parallel} \right) \right\}; \quad (10.55)$$

$$A_2 = \alpha_{\kappa} k_{\parallel}^4 \beta_{\parallel}^2 \left\{ k_{\parallel}^2 \left( \frac{(3 + 6\alpha_{\kappa})}{4} v_{A\parallel}^2 + \alpha_{\kappa} \frac{3}{8} \beta_{\parallel} \right) + k_{\perp}^2 \frac{9}{4} \left( \frac{(1 + 2\alpha_{\kappa})}{3} (1 + a_p \beta_{\parallel}) - \frac{(1 + 9\alpha_{\kappa})}{18} a_p^2 \beta_{\parallel} \right) \right\}; \quad (10.56)$$

$$A_0 = \alpha_{\kappa}^2 \frac{3}{8} k_{\parallel}^6 \beta_{\parallel}^3 \{ k_{\parallel}^2 v_{A\parallel}^2 + k_{\perp}^2 (1 + a_p \beta_{\parallel} - a_p^2 \beta_{\parallel}) \}, \quad (10.57)$$

and for  $\alpha_{\kappa} = 1$  it is of course equivalent to the CGL2 dispersion relation (9.23). The oblique Alfvén mode is not affected, and is identical to (9.22), implying that the long-wavelength threshold of the oblique firehose instability is identical as well.

For strictly parallel propagation ( $k_{\perp} = 0$ ), the solutions of the remaining eighth-order polynomial are

$$\omega = \pm k_{\parallel} v_{A\parallel}; \quad (10.58)$$

$$\omega = \pm k_{\parallel} \sqrt{\beta_{\parallel} \left( \frac{3}{2} \alpha_{\kappa} + \frac{1}{2} \sqrt{3\alpha_{\kappa}(3\alpha_{\kappa} - 1)} \right)}; \quad (10.59)$$

$$\omega = \pm k_{\parallel} \sqrt{\alpha_{\kappa} \frac{\beta_{\parallel}}{2}}; \quad (10.60)$$

$$\omega = \pm k_{\parallel} \sqrt{\beta_{\parallel} \left( \frac{3}{2} \alpha_{\kappa} - \frac{1}{2} \sqrt{3\alpha_{\kappa}(3\alpha_{\kappa} - 1)} \right)}. \quad (10.61)$$

Note that in the last solution (10.61), the expression under the square root remains always positive, regardless of how large  $\alpha_{\kappa}$  is (i.e. when the  $\kappa$ -index approaches 5/2). For strictly perpendicular propagation ( $k_{\parallel} = 0$ ), the fast mode is not affected, and is equivalent to (9.31).

## 10.4. Mirror instability

By considering the highly oblique limit  $k_{\perp} \gg k_{\parallel}$ , and by assuming that  $\alpha_{\kappa}$  is not enormously large (i.e. that  $\kappa$  is not very close to 5/2) so that  $\alpha_{\kappa} k_{\parallel}$  can be neglected with respect to  $k_{\perp}$ , the polynomial coefficients can be approximated as

$$A_6 = k_{\perp}^2 (1 + a_p \beta_{\parallel}); \quad (10.62)$$

$$A_4 = \alpha_{\kappa} k_{\parallel}^2 \beta_{\parallel} k_{\perp}^2 \frac{7}{2} \left( 1 + a_p \beta_{\parallel} - \frac{1}{7} a_p^2 \beta_{\parallel} \right); \quad (10.63)$$

$$A_2 = \alpha_{\kappa} k_{\parallel}^4 \beta_{\parallel}^2 k_{\perp}^2 \frac{9}{4} \left( \frac{(1 + 2\alpha_{\kappa})}{3} (1 + a_p \beta_{\parallel}) - \frac{(1 + 9\alpha_{\kappa})}{18} a_p^2 \beta_{\parallel} \right); \quad (10.64)$$

$$A_0 = \alpha_{\kappa}^2 \frac{3}{8} k_{\parallel}^6 \beta_{\parallel}^3 k_{\perp}^2 (1 + a_p \beta_{\parallel} - a_p^2 \beta_{\parallel}). \quad (10.65)$$

The fast mode can be quickly separated with the trick (9.36), and by using the substitution  $\bar{\omega} = \omega/(k_{\parallel}\sqrt{\beta_{\parallel}})$ , the resulting polynomial reads

$$\begin{aligned} & -(1 + a_p\beta_{\parallel})\bar{\omega}^6 + \alpha_{\kappa}\frac{7}{2}\left(1 + a_p\beta_{\parallel} - \frac{1}{7}a_p^2\beta_{\parallel}\right)\bar{\omega}^4 \\ & - \alpha_{\kappa}\frac{9}{4}\left(\frac{(1 + 2\alpha_{\kappa})}{3}(1 + a_p\beta_{\parallel}) - \frac{(1 + 9\alpha_{\kappa})}{18}a_p^2\beta_{\parallel}\right)\bar{\omega}^2 \\ & + \alpha_{\kappa}^2\frac{3}{8}(1 + a_p\beta_{\parallel} - a_p^2\beta_{\parallel}) = 0. \end{aligned} \quad (10.66)$$

At exactly the (assumed) mirror threshold  $a_p^2\beta_{\parallel} = 1 + a_p\beta_{\parallel}$ , the coefficient  $A_0$  disappears, so we have one solution  $\bar{\omega}^2 = 0$ ; from the remaining polynomial the  $1 + a_p\beta_{\parallel}$  can be factored out and the solutions are

$$\bar{\omega}^2 = \frac{3}{2}\alpha_{\kappa} \pm \frac{1}{4}\sqrt{10\alpha_{\kappa}(3\alpha_{\kappa} - 1)}, \quad (10.67)$$

both always positive. Slightly below ( $\epsilon < 0$ ) or beyond ( $\epsilon > 0$ ) the mirror threshold, one can prescribe  $a_p^2\beta_{\parallel} = (1 + a_p\beta_{\parallel})(1 + \epsilon)$ , where  $\epsilon$  is a small parameter and the polynomial transforms to

$$(\bar{\omega}^2)^3 - \alpha_{\kappa}\left(3 - \frac{\epsilon}{2}\right)(\bar{\omega}^2)^2 + \frac{\alpha_{\kappa}}{8}(5 + 3\alpha_{\kappa} - \epsilon(1 + 9\alpha_{\kappa}))(\bar{\omega}^2) + \frac{3}{8}\alpha_{\kappa}^2\epsilon = 0, \quad (10.68)$$

and which for  $\alpha_{\kappa} = 1$  is equivalent to the bi-Maxwellian polynomial (9.39). As previously, for small  $\epsilon$  the solutions (10.67) remain almost the same, and positive. Furthermore, since multiplying all 3 solutions together must yield  $-\frac{3}{8}\alpha_{\kappa}^2\epsilon$ , this implies that below ( $\epsilon < 0$ ) the threshold the third solution is positive, and beyond ( $\epsilon > 0$ ) the threshold the third solution is negative, which analytically proves that the (highly oblique) mirror instability threshold is  $a_p^2\beta_{\parallel} = 1 + a_p\beta_{\parallel}$ , regardless of  $\alpha_{\kappa}$ . The long-wavelength limit or ‘hard’ mirror threshold for the bi-kappa distribution is therefore the same as for the bi-Maxwellian distribution.

### Conclusion

We investigated solutions of the (non-dispersive) BiKappa fluid model. The calculations were done with closure (10.42) for a general  $\alpha_{\kappa}$  coefficient, i.e. without using the specific value (10.43) for the bi-kappa distribution. The calculations are therefore more general and valid for any distribution function that can be closed as (10.42). We conclude that all 3 long-wavelength ‘hard’ thresholds – for the parallel and oblique firehose instability and the highly oblique mirror instability – are not affected by the value of  $\alpha_{\kappa}$ .

### 10.5. Hall-BiKappa dispersion relation

The Hall-BiKappa dispersion relation can be written in the following form

$$\begin{aligned} & (\omega^2 - v_{A\parallel}^2 k_{\parallel}^2)(\omega^8 - A_6\omega^6 + A_4\omega^4 - A_2\omega^2 + A_0) \\ & = k^2 k_{\parallel}^2 \omega^2 \left[ \omega^6 - \omega^4 \beta_{\parallel} \left( \frac{7}{2}\alpha_{\kappa} k_{\parallel}^2 + a_p k_{\perp}^2 \right) + \omega^2 \alpha_{\kappa} \beta_{\parallel}^2 k_{\parallel}^2 \left( \frac{3}{4}(1 + 2\alpha_{\kappa})k_{\parallel}^2 + 3a_p k_{\perp}^2 \right) \right. \\ & \quad \left. - \alpha_{\kappa} k_{\parallel}^4 \beta_{\parallel}^3 \left( \frac{3}{8}\alpha_{\kappa} k_{\parallel}^2 + \frac{1}{8}(5 + 3\alpha_{\kappa})a_p k_{\perp}^2 \right) \right], \end{aligned} \quad (10.69)$$

where we kept the BiKappa dispersion relation (10.53)–(10.57) on the left-hand side, and the Hall term contributions to the right-hand side. For the bi-Maxwellian value  $\alpha_k = 1$ , the result is equivalent to the Hall-CGL2 dispersion relation (9.40). For strictly parallel propagation, the solutions are the usual dispersive ion-cyclotron and whistler waves (that are not affected by the  $\alpha_k$  coefficient) and the non-dispersive solutions (10.59)–(10.61) of the BiKappa model. For strictly perpendicular propagation, the Hall term naturally disappears and the fast mode solution is (9.31).

## 11. Collisionless damping in fluid models – Landau fluid models

In the previous sections, we have calculated the fluid hierarchy up to the fourth-order fluid moment  $\mathbf{r}$ . We have seen that, unless we crossed some instability threshold, the dispersion relations yielded frequencies  $\omega$  that were purely real, and there was no collisionless damping present. Remarkably, collisionless damping, i.e. Landau damping, will be absent even if we continue to develop the fluid hierarchy to higher-order moments. In fact, we will see in the next section that, by considering parallel 1-D propagation (where all the fluctuations are along  $\mathbf{B}_0$ ), the fluid models closed by a bi-Maxwellian fluid closure at higher-order moments than  $\mathbf{r}$  are always unstable. Such fluid models are physically ill posed. To go higher in the fluid hierarchy, one needs to incorporate the Landau damping phenomenon into the fluid framework, and consider Landau fluid closures.

Nevertheless, to capture the Landau damping, it is not necessary to go beyond the fourth-order moment and closures performed at the fourth moment level describe the Landau damping with very good accuracy. Even though the entire Part 2 of this text is dedicated to Landau fluid closures, it is beneficial to briefly show why the Landau damping was omitted in the previous calculations. Let us consider the simplest example of 1-D electrostatic propagation. The collisionless Vlasov equation reads

$$\frac{\partial f}{\partial t} + v_z \partial_z f + \frac{q}{m} E_z \frac{\partial f}{\partial v_z} = 0. \quad (11.1)$$

Direct integration of the Vlasov equation (over  $v_z$ ) for a general distribution function  $f$ , of course yields the usual density equation,  $\partial n / \partial t + \partial_z (n u_z) = 0$ . The same result is obtained if one considers a specific  $f_0$ , such as for example 1-D Maxwellian  $f_0 = n \sqrt{(\alpha/\pi)} e^{-\alpha(v_z - u_z)^2}$ , 1-D kappa distribution, etc. The last term of the Vlasov equation  $\sim E_z$  vanishes by the integration, since  $\int_{-\infty}^{\infty} (\partial f / \partial v_z) dv_z = 0$ . Now let us do the calculation differently, by expanding the distribution function  $f = f_0 + f^{(1)}$ , where  $f_0$  does not have any time or spatial dependence, so  $\partial f_0 / \partial t = 0$ ,  $\partial_z f_0 = 0$ . We can consider the Maxwellian  $f_0 = n_0 \sqrt{(\alpha/\pi)} e^{-\alpha v_z^2}$  where  $\alpha = m/(2T^{(0)}) = 1/v_{th}^2$ . The electric field is also expanded  $E_z = E_z^{(0)} + E_z^{(1)}$ , and one assumes that there is no large-scale electric field, so  $E_z^{(0)} = 0$ . By neglecting the term with two small quantities  $E_z^{(1)} \partial f^{(1)} / \partial v_z$ , the Vlasov equation reads

$$\frac{\partial f^{(1)}}{\partial t} + v_z \partial_z f^{(1)} + \frac{q}{m} E_z^{(1)} \frac{\partial f_0}{\partial v_z} = 0, \quad (11.2)$$

and transformation to Fourier space yields

$$-i(\omega - k_z v_z) f^{(1)} + \frac{q}{m} E_z^{(1)} \frac{\partial f_0}{\partial v_z} = 0. \quad (11.3)$$



The equation (11.3) is actually not that much different from (11.1), the Vlasov equation is just fully linearized and written in Fourier space. However, in kinetic calculations, the (11.3) is not directly integrated in this form. The crucial technique that reveals the presence of Landau damping in the Vlasov equation is that the equation (11.3) is first divided by  $(\omega - k_z v_z)$ , and an expression for  $f^{(1)}$  is obtained. Only then is the equation integrated over the velocity space, yielding

$$f^{(1)} = -\frac{q}{m} i E_z^{(1)} \frac{\partial f_0}{\partial v_z} \frac{1}{\omega - k_z v_z}; \quad \Rightarrow \quad n^{(1)} = \int_{-\infty}^{\infty} f^{(1)} dv_z = -\frac{q}{m} i E_z^{(1)} \int_{-\infty}^{\infty} \frac{\partial f_0}{\partial v_z} \frac{1}{\omega - k_z v_z} dv_z. \quad (11.4)$$

The integral contains a ‘singularity’ in the denominator, and it is not obvious how to calculate this integral. Such singularities are called wave–particle resonances, and can have a more general form  $\omega - k_z v_z \pm l\Omega$ , where  $\Omega$  is the cyclotron frequency of the considered species, and  $l$  is an integer. The resonance for  $l=0$  is called the Landau resonance, and the other resonances for  $l \neq 0$  are called cyclotron resonances. It is the presence of these resonances in integrals such as (11.4), that yields collisionless damping mechanisms, the Landau damping and the cyclotron damping. In Part 2 of our guide, we will see that the integral actually cannot be ‘calculated’, i.e. the integral cannot be expressed through elementary functions, even if  $f_0$  is a simple Maxwellian. The integral is just expressed through the famous plasma dispersion function  $Z(\zeta)$ , where  $\zeta = \omega/(|k_z|v_{th})$ , since the plasma dispersion function was specifically developed to describe this integral. Skipping many technical (but very important) details, such as the definition of  $Z(\zeta)$ , for a Maxwellian  $f_0$  the density integral is simply expressed as

$$n^{(1)} = -\frac{q}{m} \frac{i E_z^{(1)}}{k_z} (1 + \zeta Z(\zeta)). \quad (11.5)$$

One can define function  $R(\zeta) = 1 + \zeta Z(\zeta)$ , that is called the plasma response function. Now, for example, considering a proton–electron plasma with  $T_e^{(0)} = T_p^{(0)}$  at wavelengths that are much longer than the Debye length, yields the following dispersion relation and numerical solution

$$R(\zeta_p) + R(\zeta_e) = 0; \quad \Rightarrow \quad \zeta_p = \frac{\omega}{|k_z|v_{thp}} = \pm 1.457 - 0.627i. \quad (11.6)$$

The negative imaginary part represents the Landau damping, and the effect comes from the Landau resonance in the integral (11.4). Obviously, integrating the Vlasov equation directly in the form (11.1) or (2.1) neglects the wave–particle resonances in velocity space, which clarifies why traditional fluid models do not contain collisionless damping mechanisms. It is  $f^{(1)}$  – the fluctuations/perturbations around  $f_0$  – where the wave–particle resonances are written down explicitly, that yields the collisionless damping mechanisms in kinetic theory. To introduce collisionless damping into fluid models, we have no other choice, and we have to calculate the hierarchy of moments again, this time by integrating over the ‘kinetic’  $f^{(1)}$ . We find it useful to introduce the following vocabulary:

- (i) Fluid moments = integrals over a general (unspecified) distribution function  $f$ , or over a specific  $f_0$ . Evolution equations for fluid moments are obtained by direct integration of the Vlasov equation written in a form that does not explicitly contain wave–particle resonances. The resulting expressions can be nonlinear or linear and contain only fluid variables, i.e. expressions do not contain  $Z(\zeta)$ .

- (ii) Kinetic moments = integrals over  $f^{(1)}$  (perturbations  $f^{(1)} = f - f_0$ , where a specific equilibrium  $f_0$  is assumed), where the wave-particle resonances in velocity space are written down explicitly. The resulting expressions are linear and contain  $Z(\zeta)$ , which is not a fluid variable.

Importantly, we want to keep the entire nonlinear fluid hierarchy discussed so far, including the nonlinear heat flux equations. Therefore, considering bi-Maxwellian distribution  $f_0$ , instead of the ‘normal’ closure, the fourth-order moment is decomposed in the following form

$$r_{\parallel\parallel} = \frac{3p_{\parallel}^2}{\rho} + \tilde{r}_{\parallel\parallel}; \quad r_{\parallel\perp} = \frac{p_{\parallel}p_{\perp}}{\rho} + \tilde{r}_{\parallel\perp}; \quad r_{\perp\perp} = \frac{2p_{\perp}^2}{\rho} + \tilde{r}_{\perp\perp}. \quad (11.7)$$

The first terms represent the ‘normal’ closure, i.e. when an exact bi-Maxwellian  $f_0$  is prescribed. The tilde components represent fluctuations/perturbations around  $f_0$ , and a good notation would also be  $\delta r$ . The decomposition (11.7) yields the following form of the nonlinear heat flux equations

$$\frac{\partial q_{\parallel}}{\partial t} + \nabla \cdot (q_{\parallel} \mathbf{u}) + \nabla \cdot (\tilde{r}_{\parallel\parallel} \hat{\mathbf{b}}) - 3\tilde{r}_{\parallel\perp} \nabla \cdot \hat{\mathbf{b}} + 3p_{\parallel} \hat{\mathbf{b}} \cdot \nabla \left( \frac{p_{\parallel}}{\rho} \right) + 3q_{\parallel} \hat{\mathbf{b}} \cdot \nabla \mathbf{u} \cdot \hat{\mathbf{b}} = 0; \quad (11.8)$$

$$\begin{aligned} \frac{\partial q_{\perp}}{\partial t} + \nabla \cdot (q_{\perp} \mathbf{u}) + q_{\perp} \nabla \cdot \mathbf{u} + \nabla \cdot (\tilde{r}_{\parallel\perp} \hat{\mathbf{b}}) + (\tilde{r}_{\parallel\perp} - \tilde{r}_{\perp\perp}) \nabla \cdot \hat{\mathbf{b}} + p_{\parallel} \hat{\mathbf{b}} \cdot \nabla \left( \frac{p_{\perp}}{\rho} \right) \\ + \frac{p_{\perp}}{\rho} (p_{\parallel} - p_{\perp}) \nabla \cdot \hat{\mathbf{b}} + \frac{p_{\perp}}{\rho} (\nabla \cdot \Pi) \cdot \hat{\mathbf{b}}_0 = 0. \end{aligned} \quad (11.9)$$

Similarly to decomposition (11.7), ‘kinetic’ corrections with wave-particle resonances can also be considered for the non-gyrotropic  $\mathbf{r}^{\text{ng}}$ , which is not addressed in this guide. Neglecting those contributions, the heat flux equations are equivalent for example to equations (14), (15) of Passot & Sulem (2007), and equations (9), (10) of Passot *et al.* (2012).

In Part 2 of this guide, we will calculate the ‘kinetic moments’ for  $\tilde{r}_{\parallel\parallel}$ ,  $\tilde{r}_{\parallel\perp}$ ,  $\tilde{r}_{\perp\perp}$ , by integrating over  $f^{(1)}$ . In the 3-D electromagnetic geometry, we will use a more sophisticated  $f^{(1)}$  in the gyrotropic limit, which describes both Landau damping and its magnetic analogue, the transit-time damping. The expressions for the entire ‘kinetic’ hierarchy will contain  $Z(\zeta)$ , or actually  $R(\zeta)$ , which is not a fluid variable. Nevertheless, we will see that by introducing the concept of the Padé approximation to  $R(\zeta)$ , that sometimes a rare possibility arises, when the ‘kinetic’ moment  $\tilde{r}$  can be expressed through lower-order ‘kinetic’ moments. Such a closure, where the last retained kinetic moment (in this case  $\tilde{r}$ ) is expressed through lower-order moments, in a way such that the kinetic  $Z(\zeta)$  function is completely eliminated and that the closure is valid for all  $\zeta$  values, will be called a Landau fluid closure. For the impatient reader, we will reveal that one of the possibilities is going to be (written in Fourier space)

$$\tilde{r}_{\parallel\parallel} = \frac{32 - 9\pi}{2(3\pi - 8)} n_0 v_{\text{th}}^2 T_{\parallel} - \frac{2\sqrt{\pi}}{3\pi - 8} v_{\text{th}} \frac{ik_{\parallel}}{|k_{\parallel}|} q_{\parallel}; \quad (11.10)$$

$$\tilde{r}_{\parallel\perp} = -\frac{\sqrt{\pi}}{2} v_{\text{th}} \frac{ik_{\parallel}}{|k_{\parallel}|} q_{\perp}; \quad (11.11)$$

$$\tilde{r}_{\perp\perp} = 0. \quad (11.12)$$

The first closure was obtained in the slab geometry by Hammett & Perkins (1990) and the remaining closures were obtained in the 3-D geometry by Snyder *et al.* (1997) (note that thermal speeds in those papers are defined without the factor of 2). In Part 2 of our guide, we are going to map all the Landau fluid closures that can be constructed at the fourth-moment level, and we will identify the most precise closures. A summary of these findings can be found in Hunana *et al.* (2018).

## 12. Evolution equation for $n$ th-order fluid moment

Now we are proficient with calculating the fluid moments for general  $f$ , and before we start with kinetic calculations, it is useful to derive the time-dependent evolution equation for a general  $n$ th-order fluid moment. Before going through this part, we highly recommend going through the parts of the paper where the pressure tensor and heat flux tensor equations are obtained. Then the algebra and the notation used here will look hopefully logical and natural. We define the  $n$ th-order fluid moment  $\mathbf{X}^{(n)}$  according to

$$X_{ij\dots n}^{(n)} = m \int c_i c_j \dots c_n f \, d^3 v. \quad (12.1)$$

The symmetric operator is defined such that it cycles around through all possibilities according to

$$[X^{(n)}]_{ijk\dots n}^S = X_{ijk\dots n}^{(n)} + X_{jik\dots ni}^{(n)} + X_{k\dots nij}^{(n)} + \dots + X_{nijk\dots n-1}^{(n)}, \quad (12.2)$$

and it is compatible with the previously used symmetric operators for a second-order tensor  $[X^{(2)}]_{ij}^S = X_{ij}^{(2)} + X_{ji}^{(2)}$  and for a third-order tensor  $[X^{(3)}]_{ijk}^S = X_{ijk}^{(3)} + X_{jki}^{(3)} + X_{kij}^{(3)}$ . We multiply the Vlasov equation by the mass  $m$  and by the velocities  $c_i c_j \dots c_n$  and integrate over  $d^3 v$ . To integrate the first term of the Vlasov equation we will need

$$\frac{\partial}{\partial t}(c_i c_j \dots c_n) = - \left( \frac{\partial u_i}{\partial t} c_j \dots c_n \right) - \left( c_i \frac{\partial u_j}{\partial t} \dots c_n \right) - \dots - \left( c_i c_j \dots \frac{\partial u_n}{\partial t} \right); \quad (12.3)$$

and the first term of the Vlasov equation calculates as

$$\begin{aligned} \textcircled{1} &= m \int c_i c_j \dots c_n \frac{\partial f}{\partial t} \, d^3 v \\ &= \frac{\partial}{\partial t} \left( \underbrace{m \int c_i c_j \dots c_n f \, d^3 v}_{X_{ij\dots n}^{(n)}} \right) - m \int f \frac{\partial}{\partial t} (c_i c_j \dots c_n) \, d^3 v = \frac{\partial}{\partial t} X_{ij\dots n}^{(n)} \\ &\quad + \frac{\partial u_i}{\partial t} \underbrace{m \int f c_j c_k \dots c_n \, d^3 v}_{=X_{jk\dots n}^{(n-1)}} + \frac{\partial u_j}{\partial t} \underbrace{m \int c_i c_k \dots c_n \, d^3 v}_{=X_{ik\dots n}^{(n-1)}} + \dots \\ &\quad + \frac{\partial u_n}{\partial t} \underbrace{m \int f c_i c_j c_k \dots c_{n-1} \, d^3 v}_{=X_{ij\dots n-1}^{(n-1)}} \\ &= \left\{ \frac{\partial}{\partial t} \mathbf{X}^{(n)} + \left[ \frac{\partial \mathbf{u}}{\partial t} \mathbf{X}^{(n-1)} \right]^S \right\}_{ij\dots n}. \end{aligned} \quad (12.4)$$

For the second term we will need

$$\begin{aligned}
 m \int c_i c_j c_k \dots c_n v_l f \, d^3 v &= m \int c_i c_j c_k \dots c_n \underbrace{(v_l - u_l + u_l)}_{=c_l} f \, d^3 v \\
 &= m \underbrace{\int c_i c_j c_k \dots c_n c_l f \, d^3 v}_{=X_{ij\dots nl}^{(n+1)}} + u_l m \underbrace{\int c_i c_j c_k \dots c_n f \, d^3 v}_{=X_{ij\dots n}^{(n)}} \\
 &= X_{ij\dots nl}^{(n+1)} + X_{ij\dots n}^{(n)} u_l; \tag{12.5}
 \end{aligned}$$

$$\partial_l(c_i c_j \dots c_n) = -((\partial_l u_i) c_j \dots c_n) - (c_i (\partial_l u_j) \dots c_n) - \dots - (c_i c_j \dots (\partial_l u_n)); \tag{12.6}$$

$$\begin{aligned}
 m \int v_l f \partial_l(c_i c_j c_k \dots c_n) \, d^3 v &= m \int v_l f [ -((\partial_l u_i) c_j \dots c_n) - (c_i (\partial_l u_j) \dots c_n) - \dots - (c_i c_j \dots (\partial_l u_n)) ] \, d^3 v \\
 &= -(\partial_l u_i) m \int f v_l c_j \dots c_n \, d^3 v - (\partial_l u_j) m \\
 &\quad \times \int f v_l c_i \dots c_n \, d^3 v (\partial_l u_n) - \dots - (\partial_l u_n) m \int f v_l c_i c_j \dots c_{n-1} \, d^3 v \\
 &= -(\partial_l u_i) (X_{jk\dots nl}^{(n)} + X_{jk\dots n}^{(n-1)} u_l) - (\partial_l u_j) (X_{ik\dots nl}^{(n)} + X_{ik\dots n}^{(n-1)} u_l) - \dots \\
 &\quad - (\partial_l u_n) (X_{ij\dots (n-1)l}^{(n)} + X_{ij\dots n-1}^{(n-1)} u_l) \\
 &= -(X_{jk\dots nl}^{(n)} + X_{jk\dots n}^{(n-1)} u_l) (\partial_l u_i) - (X_{ik\dots nl}^{(n)} + X_{ik\dots n}^{(n-1)} u_l) (\partial_l u_j) - \dots \\
 &\quad - (X_{ij\dots (n-1)l}^{(n)} + X_{ij\dots n-1}^{(n-1)} u_l) (\partial_l u_n) \\
 &= -\{(\mathbf{X}^{(n)} + \mathbf{X}^{(n-1)} \mathbf{u}) \cdot \nabla \mathbf{u}\}_{ij\dots n}^S. \tag{12.7}
 \end{aligned}$$

The second term calculates as

$$\begin{aligned}
 \textcircled{2} &= m \int c_i c_j c_k \dots c_n \underbrace{\mathbf{v} \cdot \nabla f}_{=v_l \partial_l} \, d^3 v \\
 &= \partial_l \left( m \int c_i c_j c_k \dots c_n v_l f \, d^3 v \right) - m \int v_l f \partial_l(c_i c_j c_k \dots c_n) \, d^3 v \\
 &= \partial_l (X_{ij\dots nl}^{(n+1)} + u_l X_{ij\dots n}^{(n)}) + \{(\mathbf{X}^{(n)} + \mathbf{X}^{(n-1)} \mathbf{u}) \cdot \nabla \mathbf{u}\}_{ij\dots n}^S \\
 &= \{\nabla \cdot (\mathbf{X}^{(n+1)} + \mathbf{u} \mathbf{X}^{(n)})\}_{ij\dots n} + \{(\mathbf{X}^{(n)} + \mathbf{X}^{(n-1)} \mathbf{u}) \cdot \nabla \mathbf{u}\}_{ij\dots n}^S. \tag{12.8}
 \end{aligned}$$

The third term calculates as

$$\begin{aligned}
 \textcircled{3} &= q \int c_i c_j \dots c_n E_l \frac{\partial f}{\partial v_l} \, d^3 v = q E_l \underbrace{\int \frac{\partial}{\partial v_l} (c_i c_j \dots c_n f) \, d^3 v}_{\rightarrow 0} - q E_l \int f \frac{\partial}{\partial v_l} (c_i c_j \dots c_n) \, d^3 v \\
 &= -q E_l \int f (\delta_{il} c_j \dots c_n + c_i \delta_{jl} \dots c_n + \dots + c_i c_j \dots c_{n-1} \delta_{nl}) \, d^3 v \\
 &= -q E_i \int f c_j \dots c_n \, d^3 v - q E_j \int f c_i \dots c_n \, d^3 v - \dots - q E_n \int f c_i c_j \dots c_{n-1} \, d^3 v
 \end{aligned}$$

$$\begin{aligned}
&= -\frac{q}{m} E_i X_{jk\dots n}^{(n-1)} - \frac{q}{m} E_j \underbrace{X_{ik\dots n}^{(n-1)}}_{=X_{k\dots ni}^{(n-1)}} - \dots - \frac{q}{m} E_n X_{ij\dots n-1}^{(n-1)} \\
&= -\frac{q}{m} [\mathbf{E} \mathbf{X}^{(n-1)}]_{ij\dots n}^S.
\end{aligned} \tag{12.9}$$

For the fourth term we will need

$$\begin{aligned}
\frac{\partial}{\partial v_l} [c_i c_j \dots c_n (\mathbf{v} \times \mathbf{B})_l] &= \delta_{il} c_j \dots c_n (\mathbf{v} \times \mathbf{B})_l + c_i \delta_{jl} \dots c_n (\mathbf{v} \times \mathbf{B})_l + \dots \\
&\quad + c_i c_j \dots c_{n-1} \delta_{nl} (\mathbf{v} \times \mathbf{B})_l + c_i c_j \dots c_n \underbrace{\frac{\partial}{\partial v_l} (\mathbf{v} \times \mathbf{B})_l}_{=0} \\
&= c_j \dots c_n (\mathbf{v} \times \mathbf{B})_i + c_i \dots c_n (\mathbf{v} \times \mathbf{B})_j + \dots \\
&\quad + c_i c_j \dots c_{n-1} (\mathbf{v} \times \mathbf{B})_n
\end{aligned} \tag{12.10}$$

$$\begin{aligned}
\int c_j c_k \dots c_n (\mathbf{v} \times \mathbf{B})_l f \, d^3 v &= \int c_j c_k \dots c_n \epsilon_{irs} v_r B_s f \, d^3 v = \epsilon_{irs} B_s \int c_j c_k \dots c_n v_r f \, d^3 v \\
&= \frac{1}{m} \underbrace{\epsilon_{irs}}_{=-\epsilon_{isr}} B_s (X_{jk\dots nr}^{(n)} + X_{jk\dots n}^{(n-1)} u_r) = -\frac{1}{m} \epsilon_{isr} B_s (X_{rjk\dots n}^{(n)} + u_r X_{jk\dots n}^{(n-1)}) \\
&= -\frac{1}{m} [\mathbf{B} \times (\mathbf{X}^{(n)} + \mathbf{u} \mathbf{X}^{(n-1)})]_{ijk\dots n}.
\end{aligned} \tag{12.11}$$

The fourth term calculates as

$$\begin{aligned}
\textcircled{4} &= \frac{q}{c} \int c_i c_j \dots c_n (\mathbf{v} \times \mathbf{B})_l \frac{\partial f}{\partial v_l} \, d^3 v \\
&= \frac{q}{c} \underbrace{\int \frac{\partial}{\partial v_l} [c_i c_j \dots c_n (\mathbf{v} \times \mathbf{B})_l f] \, d^3 v}_{\rightarrow 0} - \frac{q}{c} \int f \frac{\partial}{\partial v_l} [c_i c_j \dots c_n (\mathbf{v} \times \mathbf{B})_l] \, d^3 v \\
&= -\frac{q}{c} \int f (c_j \dots c_n (\mathbf{v} \times \mathbf{B})_i + c_i \dots c_n (\mathbf{v} \times \mathbf{B})_j + \dots + c_i c_j \dots c_{n-1} (\mathbf{v} \times \mathbf{B})_n) \, d^3 v \\
&= +\frac{q}{mc} [\mathbf{B} \times (\mathbf{X}^{(n)} + \mathbf{u} \mathbf{X}^{(n-1)})]_{ijk\dots n}^S
\end{aligned} \tag{12.12}$$

Combining all the results together  $\textcircled{1} + \textcircled{2} + \textcircled{3} + \textcircled{4} = 0$ , the evolution equation for  $n$ th-order fluid moment  $\mathbf{X}_{ij\dots n}^{(n)}$  reads (written in the vector notation without indices)

$$\begin{aligned}
\frac{\partial}{\partial t} \mathbf{X}^{(n)} + \nabla \cdot (\mathbf{X}^{(n+1)} + \mathbf{u} \mathbf{X}^{(n)}) + \left[ \frac{\partial \mathbf{u}}{\partial t} \mathbf{X}^{(n-1)} + (\mathbf{X}^{(n)} + \mathbf{X}^{(n-1)} \mathbf{u}) \cdot \nabla \mathbf{u} - \frac{q}{m} \mathbf{E} \mathbf{X}^{(n-1)} \right. \\
\left. + \frac{q}{mc} \mathbf{B} \times (\mathbf{X}^{(n)} + \mathbf{u} \mathbf{X}^{(n-1)}) \right]^S = 0.
\end{aligned} \tag{12.13}$$

The final result should have only the time derivative of  $\mathbf{X}^{(n)}$  and, similarly to the heat flux tensor, we need to use the momentum equation to eliminate  $\partial_t \mathbf{u}$ . We subtract a total of  $n$  momentum equations combined with  $\mathbf{X}^{(n)}$  in the form

$$\left[ \left( \frac{\partial \mathbf{u}}{\partial t} + \mathbf{u} \cdot \nabla \mathbf{u} + \frac{1}{\rho} \nabla \cdot \mathbf{p} - \frac{q}{m} \mathbf{E} - \frac{q}{mc} \mathbf{u} \times \mathbf{B} \right) \mathbf{X}^{(n-1)} \right]^S = 0. \tag{12.14}$$

Again, because of the symmetric operator, it does not matter if  $\mathbf{X}^{(n)}$  is applied to the momentum equation from the left or right. The subtraction also conveniently eliminates the electric field, however it introduces  $\nabla \cdot \mathbf{p}$ . We also need an identity

$$\begin{aligned} [(\mathbf{u} \times \mathbf{B})\mathbf{X}^{(n-1)}]_{ijk\dots n} &= (\mathbf{u} \times \mathbf{B})_i X_{jk\dots n}^{(n-1)} = -(\mathbf{B} \times \mathbf{u})_i X_{jk\dots n}^{(n-1)} = -[(\mathbf{B} \times \mathbf{u})\mathbf{X}^{(n-1)}]_{ijk\dots n} \\ &= -[\mathbf{B} \times (\mathbf{u}\mathbf{X}^{(n-1)})]_{ijk\dots n}, \end{aligned} \quad (12.15)$$

where in the last step we emphasized that it does not matter if the  $\mathbf{B} \times$  applies only to  $\mathbf{u}$  or to the entire tensor  $\mathbf{u}\mathbf{X}^{(n-1)}$ , since the vector product operates only through the first component of the tensor anyway. The final evolution equation reads

$$\boxed{\frac{\partial}{\partial t}\mathbf{X}^{(n)} + \nabla \cdot (\mathbf{X}^{(n+1)} + \mathbf{u}\mathbf{X}^{(n)}) + \left[ \mathbf{X}^{(n)} \cdot \nabla \mathbf{u} + \frac{q}{mc}\mathbf{B} \times \mathbf{X}^{(n)} - \frac{1}{\rho}(\nabla \cdot \mathbf{p})\mathbf{X}^{(n-1)} \right]^S} = 0, \quad (12.16)$$

and it is valid for  $n \geq 2$ . Evaluation at  $m=3$  recovers the heat flux tensor equation (7.22) and evaluation at  $m=2$  recovers the pressure tensor equation (2.28). Note that  $\mathbf{X}^{(4)} = \mathbf{r}$ ,  $\mathbf{X}^{(3)} = \mathbf{q}$ ,  $\mathbf{X}^{(2)} = \mathbf{p}$ , however, and very importantly, according to our definition (12.1),  $\mathbf{X}^{(1)} = 0$  and it is not equal to the velocity  $\mathbf{u}$ , which is defined according to (2.4). Additionally,  $\mathbf{X}^{(0)} = \rho$ . The momentum equation can still be recovered by evaluating (12.13) at  $n=1$ , and by dividing by  $\rho$  (where naturally the symmetric operator does not have any influence on a vector, and  $u_i^S = u_i$ ). Similar equation to our (12.13) was also obtained for example by Waelbroeck (2010), equation (22), even though we did not verify whether it is consistent with ours since  $\mathbf{X}^{(n)}$  in that work is defined with  $v$  instead of our  $c$ .

We want to always decompose the tensor  $\mathbf{X}^{(n)}$  into its gyrotropic and non-gyrotropic parts, and to have space to write the ‘g’ and ‘ng’, let us move the index  $(n)$  down, so that

$$\mathbf{X}_{(n)} = \mathbf{X}_{(n)}^g + \mathbf{X}_{(n)}^{ng}. \quad (12.17)$$

By generalizing the results we have seen for the gyrotropic  $\mathbf{p}^g, \mathbf{q}^g, \mathbf{r}^g$ , the same result is obtained for the  $n$ th moment, and at low frequencies the gyrotropic part must satisfy

$$[\hat{\mathbf{b}} \times \mathbf{X}_{(n)}^g]^S = 0, \quad (12.18)$$

which can be indeed viewed as a definition of gyrotropy. The general (12.16) therefore rewrites as

$$\frac{\partial}{\partial t}\mathbf{X}_{(n)} + \nabla \cdot (\mathbf{X}_{(n+1)} + \mathbf{u}\mathbf{X}_{(n)}) + \left[ \mathbf{X}_{(n)} \cdot \nabla \mathbf{u} + \Omega \frac{|\mathbf{B}|}{B_0} \hat{\mathbf{b}} \times \mathbf{X}_{(n)}^{ng} - \frac{1}{\rho}(\nabla \cdot \mathbf{p})\mathbf{X}_{(n-1)} \right]^S = 0, \quad (12.19)$$

yielding an implicit equation for the non-gyrotropic part in the form

$$[\hat{\mathbf{b}} \times \mathbf{X}_{(n)}^{ng}]^S = -\frac{B_0}{\Omega |\mathbf{B}|} \left\{ \frac{\partial}{\partial t}\mathbf{X}_{(n)} + \nabla \cdot (\mathbf{X}_{(n+1)} + \mathbf{u}\mathbf{X}_{(n)}) + \left[ \mathbf{X}_{(n)} \cdot \nabla \mathbf{u} - \frac{1}{\rho}(\nabla \cdot \mathbf{p})\mathbf{X}_{(n-1)} \right]^S \right\}. \quad (12.20)$$

Ideally, an ‘inversion procedure’ should be found for the left-hand side, so that an equation for  $\mathbf{X}_{(n)}^{ng}$  can be found, and then by expanding the right-hand side (for

example at the leading order by considering only gyrotropic  $\mathbf{X}^g$ ), a fully nonlinear expressions for  $\mathbf{X}_{(n)}^{ng}$  can be obtained. Nevertheless, we have seen that the inversion procedure was already quite complicated for the matrix  $(\hat{\mathbf{b}} \times \boldsymbol{\Pi})^S$ , see (5.9). For a tensor of third rank, such as  $(\hat{\mathbf{b}} \times \mathbf{q}^{ng})^S$ , the full inversion procedure is actually not addressed in this guide and an advanced reader is referred to (43) of Ramos (2005). The inversion procedure was obtained only for a part of the  $\mathbf{q}^{ng}$ , that can be decomposed into the non-gyrotropic heat flux vectors  $\mathbf{S}_{\perp}^{\parallel}, \mathbf{S}_{\perp}^{\perp}$ , which is done in detail in appendix D. For the rest of  $\mathbf{q}^{ng}$  that consists of heat flux tensor  $\boldsymbol{\sigma}$ , the inversion procedure cannot be avoided. Additionally, we have seen that for fourth-order moment  $\mathbf{r}^{ng}$ , the equation (12.20) does not seem to yield anything useful even at the linear level, and a different procedure was used, where the specific bi-Maxwellian distribution was expanded in Hermite polynomials. Therefore, for higher-order moments, the equation (12.20) is immensely complicated.

It is useful to briefly consider how many scalar gyrotropic moments an  $n$ th-order tensor  $\mathbf{X}_{(n)}$  has. The scalar gyrotropic moments can be constructed only from combinations of  $c_{\parallel}$  and  $c_{\perp}^2$ . It is beneficial to write down the following self-explanatory table

$\mathbf{X}_{(n)}$	Possible combinations of $c_{\parallel}, c_{\perp}^2$	Common names	Number
0	1	$n$	1
1	$c_{\parallel}$	$u_{\parallel}$	1
2	$c_{\parallel}^2; c_{\perp}^2$	$p_{\parallel}; p_{\perp}$	2
3	$c_{\parallel}^3; c_{\parallel}c_{\perp}^2$	$q_{\parallel}; q_{\perp}$	2
4	$c_{\parallel}^4; c_{\parallel}^2c_{\perp}^2; c_{\perp}^4$	$r_{\parallel\parallel}; r_{\parallel\perp}; r_{\perp\perp}$	3
5	$c_{\parallel}^5; c_{\parallel}^3c_{\perp}^2; c_{\parallel}c_{\perp}^4$		3
6	$c_{\parallel}^6; c_{\parallel}^4c_{\perp}^2; c_{\parallel}^2c_{\perp}^4; c_{\perp}^6$		4
7	$c_{\parallel}^7; c_{\parallel}^5c_{\perp}^2; c_{\parallel}^3c_{\perp}^4; c_{\parallel}c_{\perp}^6$		4
8	$c_{\parallel}^8; c_{\parallel}^6c_{\perp}^2; c_{\parallel}^4c_{\perp}^4; c_{\parallel}^2c_{\perp}^6; c_{\perp}^8$		5
9	$c_{\parallel}^9; c_{\parallel}^7c_{\perp}^2; c_{\parallel}^5c_{\perp}^4; c_{\parallel}^3c_{\perp}^6; c_{\parallel}c_{\perp}^8$		5

Obviously, the number of gyrotropic moments for tensor  $\mathbf{X}_{(n)}$  is equal to  $1 + \text{int}[n/2]$ . The function ‘int’ means integer part. In mathematics, the integer function is often abbreviated with square brackets, as  $1 + [n/2]$ . However, this notation might be possibly confusing (we used square brackets in many other places), and we prefer to write explicitly ‘int’. Another possibility would be to use the ‘floor’ function.

### 12.1. Hierarchy of moments in a 1-D geometry (electrostatic)

Let us see if we can use the  $n$ th-order moment equation (12.16) for something useful, since the equation is immensely complicated. Let us consider the special case of strictly parallel propagation along  $B_0$  (in the  $z$ -direction), where all the moments (including the velocity) are along  $B_0$ , i.e. let us consider the electrostatic case. Let us change the index of a moment from  $(n)$  to  $(l)$ , so that the index is not confused with the number density. Let us stop writing the  $\parallel$  subscript on all the variables starting from the pressure, so  $p_{\parallel}, q_{\parallel}, r_{\parallel\parallel}$  are simply  $p, q, r$  and the  $l$ th-order parallel moment is just  $X^{(l)} \equiv m \int c_{\parallel}^l f dv$ . We keep the  $z$ -subscript for  $u_z$  to remind us we are writing 1-D equations. In a 1-D geometry, the nonlinear evolution equation (12.16) simplifies to

$$\frac{\partial}{\partial t} X^{(l)} + \partial_z (X^{(l+1)} + u_z X^{(l)}) + l (\partial_z u_z) X^{(l)} - \frac{l}{\rho} (\partial_z p) X^{(l-1)} = 0; \quad l \geq 2. \quad (12.21)$$

If one wants to be more general, instead of propagation along  $B_0$ , it is possible to consider propagation along magnetic field line  $\hat{\mathbf{b}}$ , and write the same equation with  $u_{\parallel}$  and  $\partial_{\parallel}$ . The equation is accompanied by the density and momentum equations

$$\frac{\partial}{\partial t}\rho + \partial_z(u_z\rho) = 0; \quad (12.22)$$

$$\frac{\partial}{\partial t}u_z + u_z\partial_zu_z + \frac{1}{\rho}\partial_zp - \frac{q_r}{m}E_z = 0, \quad (12.23)$$

where for the charge  $q_r$  it is useful to keep the species index  $r$ , so that it is not confused with the heat flux  $q$ . The last term disappears anyway since in the case considered here  $E_z = 0$ . Evaluating equation (12.21) for  $l = 2, 3, 4, 5, 6 \dots$  yields

$$\frac{\partial}{\partial t}p + \partial_z(q + u_zp) + 2(\partial_zu_z)p = 0; \quad (12.24)$$

$$\frac{\partial}{\partial t}q + \partial_z(r + u_zq) + 3(\partial_zu_z)q - \frac{3}{\rho}(\partial_zp)p = 0; \quad (12.25)$$

$$\frac{\partial}{\partial t}r + \partial_z(X^{(5)} + u_zr) + 4(\partial_zu_z)r - \frac{4}{\rho}(\partial_zp)q = 0; \quad (12.26)$$

$$\frac{\partial}{\partial t}X^{(5)} + \partial_z(X^{(6)} + u_zX^{(5)}) + 5(\partial_zu_z)X^{(5)} - \frac{5}{\rho}(\partial_zp)r = 0; \quad (12.27)$$

$$\frac{\partial}{\partial t}X^{(6)} + \partial_z(X^{(7)} + u_zX^{(6)}) + 6(\partial_zu_z)X^{(6)} - \frac{6}{\rho}(\partial_zp)X^{(5)} = 0, \quad (12.28)$$

and so on. The nonlinear equation (12.21) can be linearized, in the first step by assuming zero mean  $u_{z0}$  value, yielding

$$\frac{\partial}{\partial t}X^{(l)} + \partial_zX^{(l+1)} + (l+1)X_0^{(l)}(\partial_zu_z) - \frac{l}{\rho_0}X_0^{(l-1)}(\partial_zp) = 0, \quad (12.29)$$

and calculating it for  $l = 2 - 8$  and by prescribing  $q_0 = 0$ ,  $X_0^{(5)} = 0$  and in general for  $l$  odd  $X_0^{(l)} = 0$ , yields

$$\frac{\partial}{\partial t}p + \partial_zq + 3p_0(\partial_zu_z) = 0; \quad (12.30)$$

$$\frac{\partial}{\partial t}q + \partial_zr - 3\frac{p_0}{\rho_0}(\partial_zp) = 0; \quad (12.31)$$

$$\frac{\partial}{\partial t}r + \partial_zX^{(5)} + 5r_0(\partial_zu_z) = 0; \quad (12.32)$$

$$\frac{\partial}{\partial t}X^{(5)} + \partial_zX^{(6)} - 5\frac{r_0}{\rho_0}(\partial_zp) = 0; \quad (12.33)$$

$$\frac{\partial}{\partial t}X^{(6)} + \partial_zX^{(7)} + 7X_0^{(6)}(\partial_zu_z) = 0; \quad (12.34)$$

$$\frac{\partial}{\partial t}X^{(7)} + \partial_zX^{(8)} - 7\frac{X_0^{(6)}}{\rho_0}(\partial_zp) = 0; \quad (12.35)$$

$$\frac{\partial}{\partial t}X^{(8)} + \partial_zX^{(9)} + 9X_0^{(8)}(\partial_zu_z) = 0. \quad (12.36)$$



Alternatively, we can split the linearized equations for even and odd  $l$  according to

$$l = \text{even} : \quad \frac{\partial}{\partial t} X^{(l)} + \partial_z X^{(l+1)} + (l+1) X_0^{(l)} (\partial_z u_z) = 0; \quad (12.37)$$

$$l = \text{odd} : \quad \frac{\partial}{\partial t} X^{(l)} + \partial_z X^{(l+1)} - \frac{l}{\rho_0} X_0^{(l-1)} (\partial_z p) = 0. \quad (12.38)$$

When closing the system, one has to be careful how to deal with  $l = \text{even}$  quantities, since we want to separate the ‘deviations’, similarly as we did for  $r = 3p^2/\rho + \tilde{r}$ . For a Maxwellian  $f_0$  and  $l \geq 4$  and even, we therefore write

$$l = \text{even} : \quad X^{(l)} = \frac{l-1}{\rho} p X^{(l-2)} + \tilde{X}^{(l)}, \quad (12.39)$$

or equivalently

$$l = \text{even} : \quad X^{(l)} = (l-1)!! \frac{p^{l/2}}{\rho^{l/2-1}} + \tilde{X}^{(l)}, \quad (12.40)$$

so for example  $r_0 = 3p_0^2/\rho_0$ ,  $X_0^{(6)} = 15p_0^3/\rho_0^2$ ,  $X_0^{(8)} = 105p_0^4/\rho_0^3$ , and linearization of the  $l$ th-order moment yields

$$l = \text{even} : \quad X^{(l)} \stackrel{\text{lin}}{=} X_0^{(l)} \left[ \frac{l}{2} \frac{p}{p_0} - \left( \frac{l}{2} - 1 \right) \frac{\rho}{\rho_0} \right] + \tilde{X}^{(l)}; \quad (12.41)$$

$$\stackrel{\text{lin}}{=} X_0^{(l)} \left[ \frac{l}{2} \frac{T}{T_0} + \frac{\rho}{\rho_0} \right] + \tilde{X}^{(l)}. \quad (12.42)$$

Using (12.41) in the first term of  $l = \text{even}$  equation (12.37), and using the pressure and density equations, yields linear evolution equation for deviations of even moments

$$l = \text{even} : \quad \frac{\partial}{\partial t} \tilde{X}^{(l)} + \partial_z X^{(l+1)} - X_0^{(l)} \frac{l}{2p_0} \partial_z q = 0, \quad (12.43)$$

and using (12.41) in the second term of  $l = \text{odd}$  equation (12.38), and  $T/T_0 = p/p_0 - \rho/\rho_0$  yields linear evolution equation for deviations of odd moments

$$l = \text{odd} : \quad \frac{\partial}{\partial t} X^{(l)} + \partial_z \tilde{X}^{(l+1)} + X_0^{(l-1)} \frac{l(l-1)}{2m} \partial_z T = 0. \quad (12.44)$$

Finishing the calculation with evaluations of  $X_0^{(l)}$  values, the evolution equations of deviations are

$$l = \text{even} : \quad \frac{\partial}{\partial t} \tilde{X}^{(l)} + \partial_z X^{(l+1)} - (l-1)!! \frac{l}{2} \left( \frac{p_0}{\rho_0} \right)^{l/2-1} \partial_z q = 0; \quad (12.45)$$

$$l = \text{odd} : \quad \frac{\partial}{\partial t} X^{(l)} + \partial_z \tilde{X}^{(l+1)} + l!! \frac{(l-1)}{2} n_0 \left( \frac{p_0}{\rho_0} \right)^{(l-1)/2} \partial_z T = 0, \quad (12.46)$$

or by using  $p_0/\rho_0 = v_{\text{th}}^2/2$ , equivalently

$$l = \text{even} : \quad \frac{\partial}{\partial t} \tilde{X}^{(l)} + \partial_z X^{(l+1)} - (l-1)!! \frac{l}{2^{l/2}} v_{\text{th}}^{l-2} \partial_z q = 0; \quad (12.47)$$

$$l = \text{odd} : \quad \frac{\partial}{\partial t} X^{(l)} + \partial_z \tilde{X}^{(l+1)} + l!! \frac{(l-1)}{2^{(l+1)/2}} n_0 v_{\text{th}}^{l-1} \partial_z T = 0, \quad (12.48)$$

Continuing to higher orders, the hierarchy of deviations of Maxwellian moments therefore evaluates as

$$\frac{\partial}{\partial t} q + \partial_z \tilde{r} + \frac{3}{2} n_0 v_{\text{th}}^2 \partial_z T = 0; \quad (12.49)$$

$$\frac{\partial}{\partial t} \tilde{r} + \partial_z X^{(5)} - 3 v_{\text{th}}^2 \partial_z q = 0; \quad (12.50)$$

$$\frac{\partial}{\partial t} X^{(5)} + \partial_z \tilde{X}^{(6)} + \frac{15}{2} n_0 v_{\text{th}}^4 \partial_z T = 0; \quad (12.51)$$

$$\frac{\partial}{\partial t} \tilde{X}^{(6)} + \partial_z X^{(7)} - \frac{45}{4} v_{\text{th}}^4 \partial_z q = 0; \quad (12.52)$$

$$\frac{\partial}{\partial t} X^{(7)} + \partial_z \tilde{X}^{(8)} + \frac{315}{8} n_0 v_{\text{th}}^6 \partial_z T = 0; \quad (12.53)$$

$$\frac{\partial}{\partial t} \tilde{X}^{(8)} + \partial_z X^{(9)} - \frac{105}{2} v_{\text{th}}^6 \partial_z q = 0. \quad (12.54)$$

The equations are written in physical units, and these results will be useful in Part 2, where we will calculate Landau fluid closures for deviations of various moments.

### 12.2. Impossibility of going beyond CGL2 without Landau fluid closures

One can analyse the dispersion relations easily in physical units, but since in this Part 1 we used normalized units with  $\tilde{k} = kV_A/\Omega_p$ ,  $\tilde{\omega} = \omega/\Omega_p$  and  $\beta = v_{\text{th}}^2/V_A^2$  in almost all dispersion relations, let us rewrite the fluid hierarchy to normalized units, so that everything feels more familiar. The  $X^{(5)}$  is normalized with  $p_0 V_A^3$ , the  $X^{(6)}$  with  $p_0 V_A^4$  and  $X^{(n)}$  with  $p_0 V_A^{n-2}$ . By dropping the normalization tilde as usual (but obviously not for even moments such as  $\tilde{r}$ , that should be perhaps called  $\delta r$ ), the equations in Fourier space read

$$-\omega \rho + k u_z = 0; \quad (12.55)$$

$$-\omega u_z + \frac{\beta}{2} k p = 0; \quad (12.56)$$

$$-\omega p + 3k u_z + k q = 0; \quad (12.57)$$

$$-\omega q + k \tilde{r} + \frac{3}{2} \beta k (p - \rho) = 0; \quad (12.58)$$

$$-\omega \tilde{r} + k X^{(5)} - 3 \beta k q = 0; \quad (12.59)$$

$$-\omega X^{(5)} + k \tilde{X}^{(6)} + \frac{15}{2} \beta^2 k (p - \rho) = 0; \quad (12.60)$$

$$-\omega \tilde{X}^{(6)} + k X^{(7)} - \frac{45}{4} \beta^2 k q = 0. \quad (12.61)$$

Now, using just first 3 equations (12.55)–(12.57), and closing the system with a closure  $q = 0$ , yields the CGL model, with solution

$$\text{CGL: } \omega = \pm k \sqrt{\beta \frac{3}{2}}, \quad (12.62)$$

the familiar CGL ion-acoustic mode. Going higher in the fluid hierarchy and using the first 4 equations (12.55)–(12.58), and closing the system with  $\tilde{r} = 0$  (which is

equivalent to prescribing  $r = 3p^2/\rho$ ), yields the CGL2 model. The dispersion relation is  $\omega^4 - 3\beta k^2 \omega^2 + \frac{3}{4}\beta^2 k^4 = 0$ , with solutions

$$\text{CGL2: } \omega = \pm k \sqrt{\beta \left( \frac{3}{2} + \sqrt{\frac{3}{2}} \right)}; \quad (12.63)$$

$$\omega = \pm k \sqrt{\beta \left( \frac{3}{2} - \sqrt{\frac{3}{2}} \right)}, \quad (12.64)$$

the already obtained (9.28), (9.29). The heat flux fluctuations in the CGL2 model therefore did ‘split’ the CGL ion-acoustic mode to two modes. One might assume that this will always be the case when going higher and higher in the fluid hierarchy. With the next example we demonstrate that this does not happen. Let us consider fluctuations in  $\tilde{r}$ , and use the first 5 equations (12.55)–(12.59) with the closure  $X^{(5)} = 0$ . We might call this model CGL3. The dispersion relation reads  $\omega^4 = \frac{15}{4}\beta^2 k^4$ , and the solutions are

$$\text{CGL3: } \omega = \pm k \sqrt{\beta} \left( \frac{15}{4} \right)^{1/4}; \quad \omega = \pm i k \sqrt{\beta} \left( \frac{15}{4} \right)^{1/4}. \quad (12.65)$$

Importantly, the last two modes are imaginary, and one is unstable. Therefore, such a fluid model does not make any physical sense. Why did this happen? For example, expressing everything through pressure fluctuations, and putting the  $\tilde{r}$  fluctuations on the right-hand side yields

$$(\omega^4 - 3\beta k^2 \omega^2 + \frac{3}{4}\beta^2 k^4) p = k^2 \omega^2 \tilde{r}; \quad (12.66)$$

$$= k^3 \omega X^{(5)} + (-3\beta k^2 \omega^2 + \frac{9}{2}\beta^2 k^4) p. \quad (12.67)$$

Therefore, if  $\tilde{r} = 0$ , so when (12.66) is used, the CGL2 model is obtained. However, when fluctuations in  $\tilde{r}$  are considered, so when (12.67) is used, the term  $-3\beta k^2 \omega^2$  cancels on both sides, with the resulting dispersion relation

$$(\omega^4 - \frac{15}{4}\beta^2 k^4) p = k^3 \omega X^{(5)}, \quad (12.68)$$

and by prescribing closure  $X^{(5)} = 0$ , yields the dispersion relation of the CGL3 model.

Now it is easy to go higher in the fluid hierarchy. By considering the first 6 equations (12.55)–(12.60), the dispersion relation reads

$$(\omega^6 - \frac{45}{4}\beta^2 k^4 \omega^2 + \frac{15}{4}\beta^3 k^6) p = k^4 \omega^2 \tilde{X}^{(6)}, \quad (12.69)$$

and prescribing closure  $\tilde{X}^{(6)} = 0$  (which is equivalent to prescribing  $X^{(6)} = 15p^3/\rho^2$ ), yields a model that we can call CGL4, with numerical solutions

$$\text{CGL4: } \frac{\omega}{k\sqrt{\beta}} = \pm 0.58; \quad \frac{\omega}{k\sqrt{\beta}} = \pm 1.75; \quad \frac{\omega}{k\sqrt{\beta}} = \pm 1.87i, \quad (12.70)$$

again, two modes being imaginary, and one unstable.

The variable  $\omega/(k\sqrt{\beta})$ , written here in normalized units, is easily rewritten to physical units (reintroducing tilde for clarity)

$$\frac{\tilde{\omega}}{\tilde{k}\sqrt{\tilde{\beta}}} = \frac{\frac{\omega}{\Omega_p}}{\frac{kV_A}{\Omega_p} \sqrt{\frac{v_{th}^2}{V_A^2}}} = \frac{\omega}{kv_{th}}, \quad (12.71)$$

which is directly related to the ‘kinetic’ variable  $\zeta$ , that we will use in Part 2 of the text, defined as

$$\zeta \equiv \frac{\omega}{|k|v_{\text{th}}}, \quad (12.72)$$

where the absolute value does not make any difference now (since all fluid models here contain  $\zeta^2$ ), and will become important only when Landau fluid closures are considered. The hierarchy of CGL dispersion relations then can be written as

$$\text{CGL} : \quad \zeta^2 - \frac{3}{2} = 0; \quad (12.73)$$

$$\text{CGL2} : \quad \zeta^4 - 3\zeta^2 + \frac{3}{4} = 0; \quad (12.74)$$

$$\text{CGL3} : \quad \zeta^4 - \frac{15}{4} = 0; \quad (12.75)$$

$$\text{CGL4} : \quad \zeta^6 - \frac{45}{4}\zeta^2 + \frac{15}{4} = 0; \quad (12.76)$$

$$\text{CGL5} : \quad \zeta^6 - \frac{105}{8} = 0; \quad (12.77)$$

$$\text{CGL6} : \quad \zeta^8 - \frac{105}{2}\zeta^2 + \frac{315}{16} = 0; \quad (12.78)$$

$$\text{CGL7} : \quad \zeta^8 - \frac{945}{16} = 0. \quad (12.79)$$

All the models beyond CGL2 contain modes that are unstable.

#### Double check

Let us double check our results, to verify that we calculated everything correctly. As stated previously, a closure should be performed only at the last retained fluid moment, and this is especially true when nonlinear numerical simulations are performed.<sup>13</sup> The CGL and CGL2 systems were already discussed at great length. Let us check the CGL3 dispersion relation, and let us work directly in physical units. The system of equations that we have in mind reads

$$\left. \begin{aligned} \frac{\partial}{\partial t} \rho + \rho_0 \partial_z u_z &= 0; \\ \frac{\partial}{\partial t} u_z + \frac{1}{\rho_0} \partial_z p &= 0; \\ \frac{\partial}{\partial t} p + \partial_z q + 3p_0 (\partial_z u_z) &= 0; \\ \frac{\partial}{\partial t} q + \partial_z r - 3 \frac{p_0}{\rho_0} (\partial_z p) &= 0; \\ \frac{\partial}{\partial t} r + \partial_z X^{(5)} + 5r_0 (\partial_z u_z) &= 0, \end{aligned} \right\} \quad (12.80)$$

and the closure is performed by setting  $X^{(5)} = 0$ . Importantly, the fourth-order moment  $r$  is kept undetermined, and it is not separated to its ‘core’ and  $\tilde{r}$ , nor any specific form of a distribution function is prescribed yet (other than the mean values of odd moments  $X_0^{(l)} = 0$ ). Calculating dispersion relation of this system (a matrix multiplied by vector  $(\rho, u_z, p, q, r)$ ) yields result  $\omega^4 - 5k^4 r_0 / \rho_0 = 0$ , which for  $r_0 = 3p_0^2 / \rho_0$  recovers the CGL3 dispersion relation. Our calculations were therefore

<sup>13</sup>An alternative approach that would describe the Landau damping very precisely is to keep the fully nonlinear CGL2 equations, and supplement them with a hierarchy of time-dependent linear equations for higher-order moments, closed with an appropriate Landau fluid closure.

done right. Similarly, double checking the CGL4 dispersion, the system (12.80) is supplemented with (12.33), and a closure is performed with  $X^{(6)} = 15p^3/\rho$ . Again, the  $r$  is left untouched, and dispersion relation of this system recovers the CGL4 result. Checking the CGL5 dispersion relation, the system is supplemented with (12.34), and a closure  $X^{(7)} = 0$ . Again, the  $X^{(6)}$  is kept untouched and generally undetermined. Calculating the dispersion relation yields  $\omega^6 - 7k^6 X_0^{(6)}/\rho_0 = 0$ , and prescribing Maxwellian  $X_0^{(6)} = 15p_0^3/\rho_0^2$  recovers the CGL5 result. Going higher in the hierarchy, and using closure  $X^{(8)} = 105p^4/\rho^3$  recovers the CGL6 results. And finally, keeping all the variables up to  $X^{(8)}$  undetermined and using closure  $X^{(9)} = 0$  yields dispersion relation  $\omega^8 - 9k^8 X_0^{(8)}/\rho_0 = 0$ , which for Maxwellian  $X_0^{(8)} = 105p_0^4/\rho_0^3$  recovers the CGL7 result.

Or, in general, without yet performing a closure, and going higher and higher in the hierarchy step by step, a straightforward calculation yields

$$\text{CGL : } \left( -\omega^2 + 3\frac{p_0}{\rho_0}k^2 \right) p = -k\omega q; \quad (12.81)$$

$$\text{CGL2 : } -\omega^2 p = -k^2 r; \quad (12.82)$$

$$\text{CGL3 : } \left( -\omega^4 + 5\frac{r_0}{\rho_0}k^4 \right) p = -k^3 \omega X^{(5)}; \quad (12.83)$$

$$\text{CGL4 : } -\omega^4 p = -k^4 X^{(6)}; \quad (12.84)$$

$$\text{CGL5 : } \left( -\omega^6 + 7\frac{X_0^{(6)}}{\rho_0}k^6 \right) p = -k^5 \omega X^{(7)}; \quad (12.85)$$

$$\text{CGL6 : } -\omega^6 p = -k^6 X^{(8)}; \quad (12.86)$$

$$\text{CGL7 : } \left( -\omega^8 + 9\frac{X_0^{(8)}}{\rho_0}k^8 \right) p = -k^7 \omega X^{(9)}, \quad (12.87)$$

and for higher orders therefore

$$\text{CGL}l : \quad l = \text{odd}; \quad \left( -\omega^{l+1} + (l+2)\frac{X_0^{(l+1)}}{\rho_0}k^{l+1} \right) p = -k^l \omega X^{(l+2)}; \quad (12.88)$$

$$\text{CGL}l : \quad l = \text{even}; \quad -\omega^l p = -k^l X^{(l+2)}, \quad (12.89)$$

where the validity can be proven by induction and using (12.37), (12.38). CGL $l$  model (or CGL of  $l$ th order) is therefore defined as performing a closure on  $X^{(l+2)}$  moment. In the equations (12.81)–(12.89) we used the word CGL, even though no specific distribution function (nor any closure) was prescribed yet.

The closure in each model is performed by specifying the right-hand side of (12.88), (12.89), which yields dispersion relations

$$\text{CGL}l : \quad l = \text{odd}; \quad \text{closure } X^{(l+2)} = 0; \quad \Rightarrow \quad \left( \frac{\omega}{k} \right)^{l+1} - (l+2)\frac{X_0^{(l+1)}}{\rho_0} = 0; \quad (12.90)$$

$$\begin{aligned} \text{CGL}l : \quad l = \text{even}; \quad \text{closure } X^{(l+2)} &= (l+1)!! \frac{p^{(l+2)/2}}{\rho^{l/2}}; \quad \Rightarrow \\ &\left( \frac{\omega}{k} \right)^{l+2} - \left( \frac{l}{2} + 1 \right) \frac{X_0^{(l+2)}}{\rho_0} \left( \frac{\omega}{k} \right)^2 + \frac{l}{2} \frac{X_0^{(l+2)}}{\rho_0} = 0. \end{aligned} \quad (12.91)$$

The results are written in physical units. Alternatively, by using the  $\zeta$  variable, the results are

$$\text{CGL1: } l = \text{odd}; \Rightarrow \zeta^{l+1} - \frac{(l+2)!!}{2^{(l+1)/2}} = 0; \quad (12.92)$$

$$\text{CGL1: } l = \text{even}; \Rightarrow \zeta^{l+2} - (l+1)!! \frac{l+2}{2^{l/2+1}} \zeta^2 + (l+1)!! \frac{l}{2^{l/2+2}} = 0, \quad (12.93)$$

which for specific  $l$  values recovers results (12.73)–(12.79).

The results (12.90)–(12.91) are written for a closure performed on moment  $X_0^{(l+2)}$ , which is dictated by the CGL1 vocabulary, and which might be perhaps a bit confusing. Alternatively, one can reformulate the results for closures performed on moment  $X_0^{(l)}$  (corresponding name for a Maxwellian is CGL( $l-2$ )) and that read

$$l = \text{odd}; l \geq 3 \quad \text{closure } X^{(l)} = 0; \Rightarrow \left(\frac{\omega}{k}\right)^{l-1} - l \frac{X_0^{(l-1)}}{\rho_0} = 0; \quad (12.94)$$

$$l = \text{even}; l \geq 4 \quad \text{closure } X^{(l)} = (l-1)!! \frac{p^{l/2}}{\rho^{l/2-1}}; \Rightarrow$$

$$\left(\frac{\omega}{k}\right)^l - \frac{l}{2} \left(\frac{\omega}{k}\right)^2 \frac{X_0^{(l)}}{\rho_0} + \left(\frac{l}{2} - 1\right) \frac{X_0^{(l)}}{\rho_0} = 0, \quad (12.95)$$

or rewritten with the  $\zeta$  variable

$$l = \text{odd}; l \geq 3 \quad \text{closure } X^{(l)} = 0; \Rightarrow \zeta^{l-1} - \frac{l!!}{2^{(l-1)/2}} = 0;$$

$$l = \text{even}; l \geq 4 \quad \text{closure } X^{(l)} = (l-1)!! \frac{p^{l/2}}{\rho^{l/2-1}}; \Rightarrow$$

$$\zeta^l - \frac{(l-1)!!}{2^{l/2}} \left( l \zeta^2 - \frac{l}{2} + 1 \right) = 0, \quad (12.96)$$

a result reported in Hunana *et al.* (2018), equation (10). Importantly, solutions of the above dispersion relations for  $l > 4$ , will always yield results with some complex numbers, and some solutions will be unstable.

Therefore, the last physically meaningful fluid model (without considering the Landau fluid closures) is the CGL2 model, and the last closure is the ‘normal’ closure,  $r = 3p^2/\rho$ . To go higher in the fluid hierarchy, we necessarily need to make connection to kinetic theory, and consider Landau fluid closures. These closures are described in Part 2 of our guide, and for direct comparison with the simplest Landau fluid dispersion relations, see Part 2, § 3.12 ‘Parallel ion-acoustic (sound) mode, cold electrons’.

### 13. Conclusions

We offer a brief summary of the major results discussed throughout the text.

- A collisionless plasma is completely described by the kinetic Vlasov equation. Directly integrating the Vlasov equation over velocity space, yields evolution equations for an infinite hierarchy of fluid moments. In addition to the usual

density and momentum equations, it is possible to derive an evolution equation for a general  $n$ th-order fluid moment ( $n \geq 2$ ), see (12.16). Then, evolution equations for the pressure tensor and heat flux tensor can be obtained by simply evaluating (12.16) at  $n=2$  and  $n=3$ .

- Even though the fluid hierarchy is infinite, the reformulation of the kinetic description to a fluid formalism is not necessarily complete, since the usual calculations neglect wave–particle resonances in velocity space. These wave–particle resonances, that are implicitly present in the Vlasov equation, are responsible for collisionless damping mechanisms, such as Landau damping, transit-time damping and cyclotron damping.
- The presence (or absence) of collisionless damping mechanisms in a fluid hierarchy is determined by the type of closure selected to truncate that fluid hierarchy. We differentiate between two classes of closures: a) ‘classical’ (non-Landau fluid) closures, that neglect wave–particle resonances, and b) Landau fluid closures, that account for Landau wave–particle resonances and associated Landau damping and transit-time damping, addressed in Part 2.
- There are currently no known closures that account for cyclotron wave–particle resonances and associated cyclotron damping. Nevertheless, there is *a priori* no reason why such fluid closures can not be found in the future, at least for the simplified case of electromagnetic propagation along the magnetic field (slab geometry).
- Considering ‘classical’ closures, the hierarchy of fluid moments therefore does not contain collisionless damping mechanisms, regardless of the order to which the hierarchy is developed. Or in another words, collisionless damping is beyond all orders in the classical hierarchy of fluid moments.
- In fact, it is impossible to go beyond the fourth-order moment with classical bi-Maxwellian fluid closures  $\mathbf{X}^{(n)} = \mathbf{X}_0^{(n)} + \tilde{\mathbf{X}}^{(n)}$ , where  $\mathbf{X}_0^{(n)}$  represents bi-Maxwellian value of  $\mathbf{X}^{(n)}$  and its perturbation  $\tilde{\mathbf{X}}^{(n)} = 0$ , since all the fluid models contain higher-order modes that are unstable. This is perhaps the most surprising result discussed in Part 1. For a detailed proof, see § 12.2, where a 1-D geometry is considered and a bi-Maxwellian closure is used at the  $n$ th-order fluid moment, which yields dispersion relation (12.96). The same dispersion relation will be valid in a 3-D geometry for the propagation along the magnetic field, where additional de-coupled modes will be present as well. For  $n > 4$ , all the fluid models contain unphysical instabilities. The same result is expected for other distribution functions.
- The classical bi-Maxwellian closure at the fourth-order moment reads  $r_{\parallel\parallel} = 3(p_{\parallel}^2/\rho)$ ;  $r_{\parallel\perp} = p_{\parallel}p_{\perp}/\rho$ ;  $r_{\perp\perp} = 2(p_{\perp}^2/\rho)$  (or equivalently  $\tilde{\mathbf{r}} = 0$ ), and is called the ‘normal’ closure. Therefore, the ‘normal’ closure is the last classical fluid closure ( $\tilde{\mathbf{X}}^{(n)} = 0$ ) and beyond the fourth-order moment, Landau fluid closures ( $\tilde{\mathbf{X}}^{(n)} \neq 0$ ) are required.
- From a collisionless (or weakly collisional) perspective, the long-wavelength low-frequency limit of a distribution function is not necessarily an isotropic Maxwellian, but a general distribution function that is gyrotropic, i.e. isotropic only in its perpendicular velocity components. Fluid moments are therefore typically decomposed into their gyrotropic and non-gyrotropic parts. For example

the pressure tensor  $\mathbf{p} = \mathbf{p}^g + \mathbf{\Pi}$ , the heat flux tensor  $\mathbf{q} = \mathbf{q}^g + \mathbf{q}^{ng}$ , the fourth-order moment  $\mathbf{r} = \mathbf{r}^g + \mathbf{r}^{ng}$ . The non-gyrotropic parts are also referred to as the FLR corrections, since they represent deviations from gyrotropy at small spatial scales, when the Larmor radius (gyroradius) is not infinitely small.

- A general  $n$ th-order fluid moment contains  $1 + \text{int}[n/2]$  scalar gyrotropic moments, where the function ‘int’ means integer part.
- The pressure tensor is second-order fluid moment and contains two gyrotropic moments,  $p_{\parallel}$  and  $p_{\perp}$ . Therefore, any fluid description of collisionless (or weakly collisional) plasmas, requires two separate evolution equations for  $p_{\parallel}$  and  $p_{\perp}$ . Importantly, from a linear perspective the evolution equations remain different even when the distribution function is isotropic, i.e. even when the mean pressure values are equal ( $p_{\parallel}^{(0)} = p_{\perp}^{(0)}$ ), since pressure fluctuations in the directions parallel and perpendicular to the local magnetic field remain anisotropic. We often use proton pressure (temperature) anisotropy coefficient  $a_p = T_{\perp}^{(0)}/T_{\parallel}^{(0)} = p_{\perp}^{(0)}/p_{\parallel}^{(0)}$ .
- The simplest fluid model that describes a collisionless plasma in an adiabatic regime is the CGL description (model) – named after Chew–Goldberger–Law (Chew *et al.* 1956). The model is non-dispersive and obtained by a closure with zero heat flux. The CGL description should be viewed as collisionless magnetohydrodynamics (collisionless MHD), since the usual MHD description with isotropic scalar pressure is highly collisional implicitly.
- The CGL pressure tensor is decomposed with respect to the direction of the local magnetic field lines according to  $\mathbf{p}^g = p_{\parallel} \hat{\mathbf{b}}\hat{\mathbf{b}} + p_{\perp}(\mathbf{I} - \hat{\mathbf{b}}\hat{\mathbf{b}})$ . It is possible to switch to the reference frame of the magnetic field lines, where the pressure tensor is a diagonal matrix. However, it should not be forgotten that in a such a reference frame, imposing an external (mean) magnetic field is not straightforward. For nonlinear numerical simulations of turbulence and plasma heating, the laboratory reference frame is strongly recommended.
- The two pressure equations of the CGL model can be interpreted as conservation laws for the first and second adiabatic invariants, see § 2.5. The first adiabatic invariant is a conservation of the magnetic moment of a particle that is periodically gyrating around a mean magnetic field. The second adiabatic invariant is a conservation of the average parallel momentum of a particle that is completely trapped and periodically bouncing inside of a magnetic bottle.
- Without performing (yet) any kind of closure and leaving the heat flux tensor  $\mathbf{q}$  and FLR pressure tensor  $\mathbf{\Pi}$  unspecified, it is possible to derive rigorously exact evolution equations for  $p_{\parallel}$  and  $p_{\perp}$ , see (2.91)–(2.92).
- Rewriting the pressure equations to an alternative form, equations (2.100)–(2.101), shows that the adiabatic invariants in the CGL model are broken by the inclusion of the Hall term in the induction equation, gyrotropic heat flux, non-gyrotropic FLR corrections to the pressure and heat flux, as well as by coupling of various species together. All of these contributions therefore yield very complicated anisotropic plasma heating processes, that can be studied only by nonlinear numerical simulations.



- The anisotropic plasma heating does not simplify much even in the case of periodic boundary conditions (i.e. when the system can be viewed as completely isolated) and when averaging over the entire spatial volume is performed and expressed as a conservation of total energy, see (2.113)–(2.116). Only when the anisotropic heating is neglected, i.e. when only total plasma heating studied, the conservation of energy significantly simplify, see (2.109)–(2.111).
- By using the concept of polytropic indices, the pressure equations in the ideal CGL model can be interpreted as having  $\gamma_{\parallel} = 3$  and  $\gamma_{\perp} = 2$ . Consequently, the CGL dispersion relation is in general not equal to the MHD dispersion relation, even when  $a_p = 1$ . The exception is the Alfvén mode, which for  $a_p = 1$  propagates with the same phase speed  $\omega/k = V_A \cos \theta$  in both CGL and MHD models, for all the propagation directions  $\theta$ .
- Polytropic indices can be further related to the number of degrees of freedom  $i$  through  $\gamma = (i + 2)/i$ , which for the CGL model yields  $i_{\parallel} = 1$  and  $i_{\perp} = 2$ . The CGL pressure equations can be therefore viewed as being composed of strongly coupled 1-D and 2-D dynamics, whereas the dynamics in the MHD model with  $\gamma = 5/3$  and  $i = 3$  can be viewed as isotropically three-dimensional.
- In the CGL model, the Alfvén mode propagates with the phase speed  $\omega/k = V_A \cos \theta \sqrt{1 + (\beta_{\parallel}/2)(a_p - 1)}$ , which matches the (collisionless) kinetic theory in the long-wavelength limit. Consequently, for  $a_p = 1$ , the Alfvén mode in the MHD model matches the kinetic theory as well. The CGL result is very useful when numerically solving kinetic dispersion relations (for example with the WHAMP solver), since the Alfvén mode can be easily identified at long wavelengths.
- In contrast to MHD, where the ordering of phase speeds is always  $v_s \leq v_A \leq v_f$  (slow, Alfvén, fast), in collisionless plasmas the oblique slow mode can become faster than the oblique Alfvén mode. In general, oblique slow and fast modes in the CGL model do not necessarily match the kinetic theory in the long-wavelength limit. Nevertheless, the effect when  $v_s > v_A$  is present in the CGL model and exists even for  $a_p = 1$ , see § 3.6. The effect is very important for the parallel firehose instability.
- Considering strictly perpendicular propagation ( $\theta = 90^\circ$  or  $k_{\parallel} = 0$ ), the fast mode in the CGL model propagates with the phase speed  $\omega/k = V_A \sqrt{1 + a_p \beta_{\parallel}}$ , which matches the kinetic theory in the long-wavelength limit. The expression can be rewritten as  $(\omega/k)^2 = V_A^2 + v_{\text{th}\perp}^2 = V_A^2 + 2(p_{\perp}^{(0)}/\rho_0)$ , and can be contrasted with the MHD result  $(\omega/k)^2 = V_A^2 + \frac{5}{3}(p^{(0)}/\rho_0)$ . The factor of 2 in the CGL result comes from  $\gamma_{\perp} = 2$ , which shows that the adiabatic assumption of the CGL model is appropriate for the perpendicular fast mode.
- Contributions of the Hall term completely disappear for strictly perpendicular propagation.
- It is more difficult to analytically show where the adiabatic CGL value  $\gamma_{\parallel} = 3$  is appropriate, since for example the parallel propagating ion-acoustic (sound) mode is strongly Landau damped in kinetic theory (unless the electrons are hot), which also modifies its real frequency. The ion-acoustic mode in the CGL model therefore does not match kinetic theory. One needs to consider Landau fluid models with proton and electron species, which is addressed in Part 2.

- Nevertheless, when electrons are hot ( $T_e \gg T_i$ ) and the ion-acoustic mode is Landau damped only weakly, the real phase speed obtained from kinetic theory in the long-wavelength limit reads  $\omega/k = \sqrt{(T_e + 3T_i)/m_i}$ . Thus, in this example the kinetic theory is matched by a fluid model where protons are described adiabatically (with the CGL value  $\gamma_{\parallel} = 3$ ), while electrons are isothermal (with  $\gamma_{\parallel} = 1$ ). Also, a good example for  $\gamma_{\parallel} = 3$  is the real frequency of the Langmuir mode,  $\omega^2 = \omega_{pe}^2 + 3(T_{\parallel e}^{(0)}/m_e)k^2$ , which is addressed in Part 2.
- For anisotropic temperatures ( $a_p \neq 1$ ), some modes can become unstable. A threshold of an instability that is obtained in the long-wavelength (non-dispersive) limit is called a ‘hard’ threshold. The CGL model contains 3 instabilities, the oblique firehose instability, the parallel firehose instability and the mirror instability. The mirror instability is not described correctly.
- The oblique firehose instability is an instability of the oblique Alfvén mode. Since the Alfvén mode is described correctly by the CGL model, its instability threshold  $1 + (\beta_{\parallel}/2)(a_p - 1) < 0$  matches the kinetic theory. The instability requires  $a_p < 1$  and (in the long-wavelength limit)  $\beta_{\parallel} > 2$ . Since the CGL model is non-dispersive, here the instability growth rate only has a simple  $\cos \theta$  dependence. Once dispersive effects are considered, the instability reaches maximum growth rate at some wavenumber and angle  $\theta$  that is oblique, see figure 10.
- The parallel firehose instability is an instability of the whistler mode, which can be shown by considering the Hall-CGL model. However, considering usual (non-causal) analytic solutions (4.32)–(4.35), one arrives at a contradiction that for  $\omega_r > 0$  the whistler mode is unstable, and that for  $\omega_r < 0$  the whistler mode is stable. The contradiction arises, because at the range of wavenumbers where the firehose instability exists, the ion-cyclotron and whistler modes are completely degenerate and distinguishing between them loses sense. The problem is resolved by introducing a small ( $\sim \epsilon$ ) causal dissipation into the momentum equations, which is later removed by the limit  $\lim_{\epsilon \rightarrow 0^+}$ . The procedure yields causal dispersion relations (4.43)–(4.46), where the whistler mode is firehose unstable for both positive and negative  $\omega_r$ , and the ion-cyclotron mode is stable.
- The parallel firehose instability occurs for quasi-parallel (small  $\theta$ ) propagation directions, with the maximum growth rate at  $\theta = 0^\circ$ , see figure 10. In the non-dispersive CGL model it corresponds to the instability of the slow mode. Its threshold obtained for  $\theta = 0^\circ$  is equivalent to the oblique firehose instability (since  $v_s = v_A$ ), and matches the kinetic theory. At first it might sound surprising that the slow mode connects to the whistler mode once dispersive effects are considered. However, considering  $\theta = 0^\circ$ ,  $\beta_{\parallel} > 2$ ,  $a_p < 1$  yields that the fast mode is the ion-acoustic (sound) mode and  $v_s = v_A$ , see figure 1. After dispersive effects are introduced, which split the Alfvén and slow mode to the ion-cyclotron and whistler modes, it is clarified that the instability is associated with the whistler mode. For quasi-parallel propagation directions the situation is more obvious, since  $v_s > v_A$  (see figure 1), and the slow mode clearly becomes the whistler mode.
- In the non-dispersive CGL model, all 3 instabilities are non-propagating, i.e. purely growing with zero real frequency. Once dispersive effects are introduced, the parallel firehose instability becomes propagating.

- The mirror instability requires  $a_p > 1$  and usually develops for highly oblique propagation angles. In the CGL model it corresponds to the instability of the slow mode. Its threshold obtained in the highly oblique (but not completely perpendicular) limit reads  $1 + a_p \beta_{\parallel} - \frac{1}{6} a_p^2 \beta_{\parallel} < 0$ , and with respect to kinetic theory the 1/6 factor is erroneous. The erroneous 1/6 factor shows that the adiabatic CGL closure is not appropriate for the very slow dynamics of the mirror instability.
- The simplest fluid closure that recovers the correct threshold of the mirror instability is the ‘static’ closure, see § 9.4. This closure is derived from the ‘normal’ closure in a simplified quasi-static approximation, and represents a generalization of the isothermal closure (in the presence of temperature anisotropy and variations of magnetic field strength). The ‘static’ closure is very useful since it clarifies that the highly oblique very slow dynamics of the slow mode (and of the associated mirror instability) is better described by a generalized isothermal closure than the adiabatic CGL closure.
- However, the ‘static’ closure modifies the dynamics of the perpendicular fast mode, that now does not match the kinetic theory. Additionally, the fast mode now experiences an unphysical instability, even though the instability is beyond the mirror threshold, and should perhaps not play a role in numerical simulations. The erroneous result shows that the generalized isothermal closure is not appropriate for the relatively fast dynamics of the perpendicular fast mode, where the adiabatic CGL closure is appropriate.
- Heuristically, the correct mirror threshold can be also obtained by keeping the CGL value  $\gamma_{\perp} = 2$  and modifying the parallel polytropic index to  $\gamma_{\parallel} = 1/2$ . Such a closure does not alter the dynamics of the perpendicular fast mode, since the phase speed depends only on  $\gamma_{\perp}$  (and not on  $\gamma_{\parallel}$ ). Thresholds of parallel and oblique firehose instability are not altered either. Nevertheless, the  $\gamma_{\parallel} = 1/2$  is even below the isothermal value  $\gamma_{\parallel} = 1$ , the closure is heuristic, and it is mentioned only as a curiosity.
- Obviously, to correctly capture both the slow dynamics of the highly oblique mirror instability, as well as the fast dynamics of the perpendicular fast mode, neither adiabatic nor (generalized) isothermal closures are sufficient. Considering classical (rigorously derived) closures, one has no other choice but to keep the correct adiabatic CGL values  $\gamma_{\parallel} = 3$ ,  $\gamma_{\perp} = 2$  unaltered, and instead, break the (long-wavelength) adiabaticity by going higher in the fluid hierarchy. One needs to consider fluid models with evolution equations for the heat flux tensor  $\mathbf{q}$ , that has two gyrotropic components,  $q_{\parallel}$  and  $q_{\perp}$ . The system is closed at the fourth-order moment  $\mathbf{r}$ , that has 3 gyrotropic components,  $r_{\parallel\parallel}$ ;  $r_{\parallel\perp}$ ;  $r_{\perp\perp}$ . To perform a closure, a specific distribution function (or a class of distribution functions) has to be assumed.
- Considering a bi-Maxwellian distribution function, the closure  $\tilde{\mathbf{r}} = 0$  is known as the ‘normal’ closure (see above). We call this non-dispersive bi-Maxwellian fluid model the ‘second-order CGL’, abbreviated as CGL2. The abbreviation CGL2 seems beneficial, since the name CGL is associated with collisionless magnetohydrodynamics, now just taken to one higher order. Additionally, by showing that all classical bi-Maxwellian models beyond CGL2 do not make physical sense, fluid models can be easily classified as based on CGL or CGL2 descriptions. For direct comparison of the two systems, see dispersion relations

of the CGL model (3.193) and the CGL2 model (9.23). The Hall term introduces the simplest dispersive effects; see the dispersion relations of the Hall-CGL model (4.81), and the Hall-CGL2 model (9.40).

- By comparing the dispersion relations, it is obvious that the gyrotropic heat flux fluctuations in the CGL2 model modify the general dynamics (the mean heat flux values are zero  $q_{\parallel}^{(0)} = q_{\perp}^{(0)} = 0$ ). One of the exceptions is the perpendicular fast mode, which is not altered by the gyrotropic heat fluxes, and therefore described correctly. Other exceptions are the oblique Alfvén mode (the oblique firehose instability), and once dispersive effects are introduced also the parallel propagating ion-cyclotron and whistler modes (the parallel firehose instability).
- Two additional evolution equations typically create two additional modes. For example, for the case of highly oblique CGL2 propagation, one obtains the Alfvén mode, the fast mode, and 3 ‘slow’ modes. One of these slow modes is responsible for the mirror instability, and is sometimes called the mirror mode. The CGL2 mirror threshold matches the kinetic theory. The CGL2 model is therefore the simplest classical model that correctly captures (in the long-wavelength limit) both the mirror instability threshold, as well as the perpendicular fast mode. Nevertheless, only the ‘hard’ mirror threshold is correctly recovered, and capturing the (long-wavelength) mirror instability growth rate requires Landau damping.
- One can construct closures for other distribution functions than bi-Maxwellians. For example, one can consider the bi-kappa distribution function (10.1) with a free parameter  $\kappa > 3/2$ . The closure then reads  $r_{\parallel\parallel} = 3\alpha_{\kappa}(p_{\parallel}^2/\rho)$ ;  $r_{\parallel\perp} = \alpha_{\kappa}(p_{\parallel}p_{\perp}/\rho)$ ;  $r_{\perp\perp} = 2\alpha_{\kappa}(p_{\perp}^2/\rho)$ , where the coefficient  $\alpha_{\kappa} = (\kappa - \frac{3}{2})/(\kappa - \frac{5}{2})$ . The closure is restricted to  $\kappa > \frac{5}{2}$ , which can be seen from the  $\alpha_{\kappa}$  coefficient, and the same requirement is necessary for the convergence of the velocity integrals. We call this closure and resulting fluid model simply ‘BiKappa’.
- If a closure for the bi-kappa distribution is performed at a higher  $n$ th-order moment (which is presumably not possible with classical closures since it is expected that for  $n > 4$  all models contain unphysical instabilities), the closure will be restricted to  $\kappa > (n + 1)/2$ , which increases with  $n$ . The requirement for the minimum  $\kappa$  value is there, since the power law in the bi-kappa distribution is assumed all the way till infinite velocities, and the requirement is necessary for the convergence of the velocity integrals. Therefore, by going higher and higher in the bi-kappa fluid hierarchy, the minimum value of  $\kappa$  increases, and in the limit  $\kappa \rightarrow \infty$  the required distribution function converges towards bi-Maxwellian.
- The dispersion relation for the BiKappa model is given by (10.53). A general  $\alpha_{\kappa}$  coefficient is used, so the dispersion relation is valid for a much larger class of distribution functions that can be closed with an analogous closure with just one  $\alpha_{\kappa}$  coefficient. The dispersion relation shows that the general dynamics is of course altered by the value of the  $\alpha_{\kappa}$  coefficient. Nevertheless, few exceptions are: the oblique Alfvén mode (oblique firehose threshold), the perpendicular fast mode, the parallel firehose threshold and the mirror threshold. When dispersive effects are considered, and figures similar to figure 10 created, the growth rates for oblique propagation will of course be affected by the  $\alpha_{\kappa}$  value. The Hall-BiKappa dispersion relation is given by (10.69).

- The non-gyrotropic (FLR) corrections  $\Pi$  to the pressure tensor were studied in § 5. The tensor  $\Pi$  is described by the pressure tensor equation implicitly. To prevent introducing several new independent fluid variables for components of the tensor  $\Pi$ , each with its own evolution equation, expansion on temporal and spatial scales is required. Correct evaluation of the FLR tensor with respect to magnetic field lines is quite cumbersome, and at the leading order (in temporal and spatial scales) one obtains (5.43)–(5.45), or equivalently (5.51)–(5.53). The FLR tensor  $\Pi$  is therefore often evaluated in the linear approximation, which is appropriate when the magnetic field lines are not too distorted.
- Even though the FLR pressure tensor  $\Pi$  has a zero trace, its contributions should not be called ‘off diagonal’, since the diagonal components are non-zero. In the collisionless case considered here with the mean field in the  $z$ -direction, only  $\Pi_{zz} = 0$ , and  $\Pi_{xx} = -\Pi_{yy} \neq 0$ . In the collisional case (which was not considered) also  $\Pi_{zz} \neq 0$ .
- For the purpose of comparing various contributions, we differentiate between FLR1, FLR2 and FLR3 corrections, even though the classification can be a bit blurry and there are various possibilities. Perhaps the best definition that we used later is as follows. The classical FLR1 tensor  $\Pi$  contains only velocity gradients, see (5.74). The FLR2 tensor also contains  $\partial\Pi/\partial t$  and the Hall term from the induction equation, see (5.109). It can also contain gyrotropic heat flux contributions, see (5.140). The FLR3 contains non-gyrotropic heat flux vectors  $S_{\perp}^{\parallel}$ ,  $S_{\perp}^{\perp}$ , described in § 5.8, with the detailed algebra presented in appendix D. It can also contain the non-gyrotropic heat flux tensor  $\sigma$ , which we neglected. For complete clarity in analysing the dispersion relations, the entire Hall-CGL-FLR3 model (linearized in the  $x$ – $z$  plane, normalized and Fourier transformed) is written down in § 5.9, see (5.160)–(5.164).
- Considering the strictly perpendicular fast mode, kinetic theory yields the leading-order FLR corrections (in the long-wavelength limit) to the phase speed in the following form  $(\omega/k)^2 = V_A^2(1 - \frac{1}{8}k^2\rho_i^2) + v_{th\perp}^2(1 - \frac{5}{16}k^2\rho_i^2)$ . The kinetic result is reproduced by the CGL-FLR3 model (the Hall contributions are zero). Importantly, both the first- and second-order non-gyrotropic heat flux vectors have to be retained.
- If the second-order non-gyrotropic heat flux contributions are neglected in the FLR3 model, yields solution for the perpendicular fast mode (5.159). In such a model, the correction to the Alfvén speed is captured correctly, however, the correction to the thermal speed is  $+\frac{1}{16}k^2\rho_i^2$  instead of  $-\frac{5}{16}k^2\rho_i^2$ , i.e. the correction has the wrong sign. This is surprising, since the solution of the FLR2 model, equation (5.120), has a correction to the thermal speed  $-\frac{1}{16}k^2\rho_i^2$ , i.e. at least the sign is correct. Finally, the FLR1 model yields solution (5.119), and does not capture any correction to the Alfvén speed, additionally, the correction to the thermal speed has the wrong sign as well.
- We used the Hall-CGL-FLR3 model to investigate the parallel and oblique firehose instability, and compare it with results of the FLR2 and FLR1 models, see figures 7–11. It is shown that the growth rates of the parallel and oblique firehose instability are strongly enhanced by the non-gyrotropic heat flux vectors of the FLR3 model, see figure 10 (see the scales of the colour bars).

- In general, when the maximum growth rates are sufficiently large (let us say  $\omega_i/\Omega \sim 0.1$ ) and not tiny (such as 0.001), the FLR3 model reproduces the kinetic results with unexpectedly good accuracy. See for example bottom panels of figure 10. Not only is the value of the maximum growth rate for the oblique firehose instability (red lines) approximately the same in the FLR3 model and kinetic theory, the maximum growth rate is also reached approximately for the same angle of propagation and the same wavenumber.
- The oblique firehose instability can be (in general) reproduced with better accuracy than the parallel firehose instability. Nevertheless, for very high  $\beta_{\parallel}$  values, the parallel firehose instability is reproduced very accurately, see figure 9.
- Importantly, we show that the non-gyrotropic heat flux vectors in the FLR3 model, partially reproduce the large ‘bump’ in the imaginary phase speed (growth rate normalized to the wavenumber), when the plasma is close to the long-wavelength limit ‘hard’ firehose threshold, see figure 7. The result clearly shows that, similarly to kinetic theory, fluid models can develop the firehose instability at small spatial scales, even when they are stable in the long-wavelength limit. Or in other words, fluid models can become unstable even before the ‘hard’ firehose threshold is reached. This unexpected result is perhaps the second most surprising result discussed in Part 1.

### Acknowledgements

We acknowledge support of the NSF EPSCoR RII-Track-1 Cooperative Agreement no. OIA-1655280 ‘Connecting the Plasma Universe to Plasma Technology in Alabama’, led by G.P.Z. This work was supported by the European Research Council in the frame of the Consolidating grant ERC-2017-CoG771310-PI2FA ‘Partial Ionisation: Two-Fluid Approach’, led by Elena Khomenko. A.T. acknowledges support of the NASA Heliophysics Supporting Research grant no. 80NSSC18K1211. P.H. thanks T. Passot, M. Laurenza, N. Vitas, P. Hellinger and S. P. Gary for many useful discussions. We are also very thankful to two (out of three) anonymous referees whose comments and suggestions had a great impact on this text. Significant effort has been made to eliminate all the misprints from the equations. However, we will amend possible misprints, if found, in a corrigendum.

### Appendix A. Fourier transformations

If we have an equation in real space, we decompose each quantity as a superposition of waves according to

$$f(\mathbf{x}, t) = \frac{1}{(2\pi)^4} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \hat{f}(\mathbf{k}, \omega) e^{i\mathbf{k} \cdot \mathbf{x} - i\omega t} d^3k d\omega. \quad (\text{A } 1)$$

Most of plasma physics books and kinetic papers use  $e^{i\mathbf{k} \cdot \mathbf{x} - i\omega t}$  in the definition (A 1). Decomposing the frequency to real and imaginary parts  $\omega = \omega_r + i\omega_i$ , so  $e^{-i\omega t} = e^{-i\omega_r t} e^{+\omega_i t}$ , implies that a wave with  $\omega_i > 0$  grows (i.e. has a positive growth rate) and a wave with  $\omega_i < 0$  is damped. Of course, an alternative definition of (A 1) with  $e^{-i\mathbf{k} \cdot \mathbf{x} + i\omega t}$  is allowed, one just has to remember that a wave with  $\omega_i > 0$  is damped and a wave with  $\omega_i < 0$  grows. However, in Part 2 of the text, we will calculate the Landau damping and we will see that following all the correct minus signs in the Landau



integral can be very confusing. We therefore recommend using the first decomposition, where waves with  $\omega_i < 0$  are damped. This choice yields that Fourier transformations in the  $x$ - $z$  plane (with  $B_0$  in the  $z$ -direction) are performed according to a shortcut

$$\frac{\partial}{\partial t} \leftrightarrow -i\omega; \quad \partial_z \leftrightarrow ik_{\parallel}; \quad \partial_x \leftrightarrow ik_{\perp}. \quad (\text{A } 2)$$

The transformation (A 1) is technically the inverse/backward Fourier transform  $\mathcal{F}^{-1}\hat{f}(\mathbf{k}, \omega)$ . The forward Fourier transform reads

$$f(\mathbf{k}, \omega) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \hat{f}(\mathbf{x}, t) e^{-i\mathbf{k} \cdot \mathbf{x} + i\omega t} d^3x dt. \quad (\text{A } 3)$$

The location of the normalization constants  $1/(2\pi)$  is an *ad hoc* choice, one just needs to be consistent in using them, especially when calculating convolutions.

## Appendix B. Generalized vector (cross) product

In the collisionless fluid hierarchy one encounters a vector product between vectors and tensors, as for example in the pressure tensor equation (2.26). It is useful to clarify the definition of this generalized vector product. Consider matrix  $\mathbf{A}$ , which is written as a tensor product between vectors  $\mathbf{a}$ ,  $\mathbf{b}$ , so  $\mathbf{A} = \mathbf{a} \otimes \mathbf{b} = \mathbf{ab}$ , where, as everywhere in this text, we omit writing the tensor product  $\otimes$ . Or in the index notation  $A_{ij} = a_i b_j$ . Now consider vector  $\mathbf{c}$ , and how can one define a generalized vector product  $\mathbf{c} \times \mathbf{A}$ , by using the usual vector product between two vectors. The natural definition is  $\mathbf{c} \times \mathbf{A} = \mathbf{c} \times (\mathbf{ab}) \equiv (\mathbf{c} \times \mathbf{a})\mathbf{b}$ , or in the index notation  $(\mathbf{c} \times \mathbf{A})_{ij} = \epsilon_{ikl} c_k a_l b_j = \epsilon_{ikl} c_k A_{lj}$ . Now let us consider how to define the much more confusing  $\mathbf{A} \times \mathbf{c}$ . The natural definition is  $\mathbf{A} \times \mathbf{c} = (\mathbf{ab}) \times \mathbf{c} \equiv \mathbf{a}(\mathbf{b} \times \mathbf{c})$ , and in the index notation  $(\mathbf{A} \times \mathbf{c})_{ij} = a_i \epsilon_{jkl} b_k c_l = \epsilon_{jkl} A_{ik} c_l$ . Now it is straightforward to show that

$$\mathbf{A} \times \mathbf{c} = -(\mathbf{c} \times \mathbf{A}^T)^T. \quad (\text{B } 1)$$

Generalizing the vector product to higher-order tensors is easy and in the fluid hierarchy we always encounter a vector  $\hat{\mathbf{b}}$ , and tensors  $\mathbf{p}$ ,  $\mathbf{q}$ ,  $\mathbf{r}$ , let us therefore write

$$\left. \begin{aligned} (\hat{\mathbf{b}} \times \mathbf{p})_{ij} &= \epsilon_{irs} \hat{b}_r p_{sj}; & (\mathbf{p} \times \hat{\mathbf{b}})_{ij} &= \epsilon_{jrs} p_{ir} \hat{b}_s; \\ (\hat{\mathbf{b}} \times \mathbf{q})_{ijk} &= \epsilon_{irs} \hat{b}_r q_{sjk}; & (\mathbf{q} \times \hat{\mathbf{b}})_{ijk} &= \epsilon_{krs} q_{ijr} \hat{b}_s; \\ (\hat{\mathbf{b}} \times \mathbf{r})_{ijkl} &= \epsilon_{irs} \hat{b}_r r_{sjkl}; & (\mathbf{r} \times \hat{\mathbf{b}})_{ijkl} &= \epsilon_{lrs} r_{ijk} \hat{b}_s. \end{aligned} \right\} \quad (\text{B } 2)$$

The vector product is very useful for decomposing quantities to the directions parallel and perpendicular with respect to  $\hat{\mathbf{b}}$ . For example, the usual decomposition of the velocity  $\mathbf{u}$  reads

$$\mathbf{u} = \underbrace{(\mathbf{u} \cdot \hat{\mathbf{b}})\hat{\mathbf{b}}}_{\mathbf{u}_{\parallel}} + \underbrace{\mathbf{u} \cdot (\mathbf{I} - \hat{\mathbf{b}}\hat{\mathbf{b}})}_{\mathbf{u}_{\perp}}. \quad (\text{B } 3)$$

Nevertheless, it is possible to directly obtain  $\mathbf{u}_{\perp}$ , by applying  $\hat{\mathbf{b}} \times$  twice on  $\mathbf{u}$ , so that

$$\mathbf{u} = (\mathbf{u} \cdot \hat{\mathbf{b}})\hat{\mathbf{b}} - \hat{\mathbf{b}} \times (\hat{\mathbf{b}} \times \mathbf{u}), \quad (\text{B } 4)$$

or  $\mathbf{u}_\perp = \hat{\mathbf{b}} \times \mathbf{u} \times \hat{\mathbf{b}}$ . Similar decompositions can be made for tensors. For example, a general matrix  $\Pi$  can be decomposed according to

$$\Pi = (\Pi \cdot \hat{\mathbf{b}})\hat{\mathbf{b}} + \Pi \cdot (\mathbf{I} - \hat{\mathbf{b}}\hat{\mathbf{b}}), \quad (\text{B } 5)$$

where the perpendicular component can be written as  $\Pi \cdot (\mathbf{I} - \hat{\mathbf{b}}\hat{\mathbf{b}}) = -(\Pi \times \hat{\mathbf{b}}) \times \hat{\mathbf{b}}$ , i.e. where all the  $\hat{\mathbf{b}}$  act on  $\Pi$  from the right. Alternatively, one can write a decomposition when all the  $\hat{\mathbf{b}}$  act on the  $\Pi$  from the left,  $\Pi = \hat{\mathbf{b}}(\hat{\mathbf{b}} \cdot \Pi) + (\mathbf{I} - \hat{\mathbf{b}}\hat{\mathbf{b}}) \cdot \Pi$ , where the perpendicular part  $(\mathbf{I} - \hat{\mathbf{b}}\hat{\mathbf{b}}) \cdot \Pi = -\hat{\mathbf{b}} \times (\hat{\mathbf{b}} \times \Pi)$ .

### Appendix C. MHD dispersion relation

We assume that many people reading this text will have some previous experience with MHD. It is therefore beneficial to obtain the MHD dispersion relation, so that it will be clear how more advanced fluid models are treated in this text. The MHD fluid model reads

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{u}) = 0; \quad (\text{C } 1)$$

$$\frac{\partial \mathbf{u}}{\partial t} + \mathbf{u} \cdot \nabla \mathbf{u} + \frac{1}{\rho} \nabla p - \frac{1}{4\pi\rho} (\nabla \times \mathbf{B}) \times \mathbf{B} = 0; \quad (\text{C } 2)$$

$$\frac{\partial \mathbf{B}}{\partial t} = \nabla \times (\mathbf{u} \times \mathbf{B}); \quad (\text{C } 3)$$

$$\frac{\partial p}{\partial t} + \mathbf{u} \cdot \nabla p + \gamma p \nabla \cdot \mathbf{u} = 0, \quad (\text{C } 4)$$

where  $\gamma = 5/3$ . As in advanced fluid models, it is beneficial to normalize the speed with respect to the Alfvén speed  $V_A$ , and normalize the length with respect to the ion inertial length  $d_i = V_A/\Omega_p$ , i.e. by using normalizations (3.23)–(3.28), which in Fourier space yields  $\tilde{k} = kd_i$ ,  $\tilde{\omega} = \omega/\Omega_p$ . Normalizing the MHD equations and dropping the tilde yields unchanged density, induction and pressure equations, and the momentum equation reads

$$\frac{\partial \mathbf{u}}{\partial t} + \mathbf{u} \cdot \nabla \mathbf{u} + \frac{p^{(0)}}{\rho_0 V_A^2} \frac{1}{\rho} \nabla p - \frac{1}{\rho} (\nabla \times \mathbf{B}) \times \mathbf{B} = 0. \quad (\text{C } 5)$$

One can decide how to rewrite  $p^{(0)}/(\rho_0 V_A^2)$ , either by introducing the usual MHD sound speed  $C_s^2 = \gamma p^{(0)}/\rho_0$ , or by introducing the plasma beta

$$\beta = \frac{v_{\text{th}}^2}{V_A^2} = \frac{2T^{(0)}/m_p}{V_A^2} = \frac{2p^{(0)}/\rho_0}{V_A^2} = \frac{p^{(0)}}{B_0^2/(8\pi)}; \quad \Rightarrow \quad \frac{p^{(0)}}{\rho_0 V_A^2} = \frac{1}{\gamma} \frac{C_s^2}{V_A^2} = \frac{\beta}{2}, \quad (\text{C } 6)$$

where we use the Boltzmann constant  $k_B = 1$ , see the footnote 3 after the definition of thermal speeds (3.30). By specifying the mean magnetic field to be in the  $z$ -direction, the normalized equations are linearized according to

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \mathbf{u} = 0; \quad (\text{C } 7)$$

$$\frac{\partial u_x}{\partial t} + \frac{\beta}{2} \partial_x p + \partial_x B_z - \partial_z B_x = 0; \quad (\text{C } 8)$$



$$\frac{\partial u_y}{\partial t} + \frac{\beta}{2} \partial_y p + \partial_y B_z - \partial_z B_y = 0; \quad (\text{C } 9)$$

$$\frac{\partial u_z}{\partial t} + \frac{\beta}{2} \partial_z p = 0; \quad (\text{C } 10)$$

$$\frac{\partial B_x}{\partial t} = \partial_z u_x; \quad \frac{\partial B_y}{\partial t} = \partial_z u_y; \quad \frac{\partial B_z}{\partial t} = -\partial_x u_x - \partial_y u_y; \quad (\text{C } 11)$$

$$\frac{\partial p}{\partial t} + \gamma \nabla \cdot \mathbf{u} = 0, \quad (\text{C } 12)$$

and without a loss of generality, we consider propagation in the  $x$ - $z$  plane (with  $\partial_y = 0$ ). By considering a wave propagating with wavenumber  $\mathbf{k} = (k \sin \theta, 0, k \cos \theta) = (k_x, 0, k_z) = (k_\perp, 0, k_\parallel)$ , i.e. a wave propagating in the direction that makes an angle  $\theta$  with respect to  $\mathbf{B}_0$ , the MHD system written in Fourier space reads

$$-\omega \rho + (k \sin \theta) u_x + (k \cos \theta) u_z = 0; \quad (\text{C } 13)$$

$$-\omega u_x + \frac{\beta}{2} (k \sin \theta) p + (k \sin \theta) B_z - (k \cos \theta) B_x = 0; \quad (\text{C } 14)$$

$$-\omega u_y - (k \cos \theta) B_y = 0; \quad -\omega u_z + \frac{\beta}{2} (k \cos \theta) p = 0; \quad (\text{C } 15)$$

$$-\omega B_x - (k \cos \theta) u_x = 0; \quad -\omega B_y - (k \cos \theta) u_y = 0; \quad -\omega B_z + (k \sin \theta) u_x = 0; \quad (\text{C } 16)$$

$$-\omega p + \gamma (k \sin \theta) u_x + \gamma (k \cos \theta) u_z = 0. \quad (\text{C } 17)$$

Calculating the determinant of this system yields the MHD dispersion relation in normalized units (bringing back tildes for clarity) that reads

$$(\tilde{\omega}^2 - \tilde{k}^2 \cos^2 \theta) \left( \tilde{\omega}^4 - \left( 1 + \gamma \frac{\beta}{2} \right) \tilde{k}^2 \tilde{\omega}^2 + \gamma \frac{\beta}{2} \tilde{k}^4 \cos^2 \theta \right) = 0, \quad (\text{C } 18)$$

or alternatively by using (C 6) with the sound speed

$$(\tilde{\omega}^2 - \tilde{k}^2 \cos^2 \theta) \left( \tilde{\omega}^4 - \left( 1 + \frac{C_s^2}{V_A^2} \right) \tilde{k}^2 \tilde{\omega}^2 + \frac{C_s^2}{V_A^2} \tilde{k}^4 \cos^2 \theta \right) = 0, \quad (\text{C } 19)$$

or in physical units

$$(\omega^2 - V_A^2 k^2 \cos^2 \theta) (\omega^4 - (V_A^2 + C_s^2) k^2 \omega^2 + V_A^2 C_s^2 k^4 \cos^2 \theta) = 0, \quad (\text{C } 20)$$

yielding the MHD dispersion relation for the Alfvén mode  $\omega_A = \pm V_A k \cos \theta = \pm V_A k_\parallel$  (sometimes called the shear Alfvén mode), and the slow and fast modes (3.110).

In the limit  $\beta \rightarrow 0$  that corresponds to a cold plasma ( $C_s \rightarrow 0$ ) or a plasma with a very strong magnetic field ( $V_A \rightarrow \infty$ ), the slow mode  $\omega_s = \pm C_s k_\parallel$ , and the fast mode  $\omega_f = \pm V_A k$  (sometimes called the compressional Alfvén mode). In the limit  $\beta \rightarrow \infty$  that corresponds to the incompressible MHD ( $C_s \rightarrow \infty$ ) or a plasma with vanishing magnetic field ( $V_A \rightarrow 0$ ), the slow mode  $\omega_s = \pm V_A k_\parallel$ , and the fast mode  $\omega_f = \pm C_s k$ . Note the difference in  $k_\parallel$  and  $k$  in these two limits.

#### Appendix D. Non-gyrotropic heat flux tensor $q^{\text{ng}}$

The heat flux tensor equation (7.33) is rewritten as

$$(\hat{\mathbf{b}} \times \mathbf{q}^{\text{ng}})^S = -\frac{B_0}{\Omega |\mathbf{B}|} \left[ \frac{d}{dt} \mathbf{q} + \nabla \cdot \mathbf{r} + \mathbf{q} \nabla \cdot \mathbf{u} + (\mathbf{q} \cdot \nabla \mathbf{u})^S - \frac{1}{\rho} (\mathbf{p} (\nabla \cdot \mathbf{p}))^S \right]. \quad (\text{D } 1)$$

The ‘inversion’ procedure for the left-hand side of this is equation is very complicated, and will not be addressed here. Nevertheless, the inversion procedure exists, and an interested reader can check equation (43) of Ramos (2005). The leading-order  $\mathbf{q}^{\text{ng}}$  (first order in frequency and wavenumber) can be obtained by making the quantities  $\mathbf{q}, \mathbf{r}, \mathbf{p}$  on the right-hand side gyrotropic, and we want to solve

$$(\hat{\mathbf{b}} \times \mathbf{q}^{\text{ng}})^S = -\frac{B_0}{\Omega|\mathbf{B}|} \left[ \frac{d}{dt} \mathbf{q}^g + \nabla \cdot \mathbf{r}^g + \mathbf{q}^g \nabla \cdot \mathbf{u} + (\mathbf{q}^g \cdot \nabla \mathbf{u})^S - \frac{1}{\rho} (\mathbf{p}^g (\nabla \cdot \mathbf{p}^g))^S \right]. \quad (\text{D } 2)$$

The heat flux tensor decomposition is  $\mathbf{q} = \mathbf{q}^g + \mathbf{q}^{\text{ng}}$ , or alternatively  $\mathbf{q} = \mathbf{S} + \boldsymbol{\sigma}$ , where  $\boldsymbol{\sigma} : \hat{\mathbf{b}}\hat{\mathbf{b}} = 0$ ,  $\boldsymbol{\sigma} : \mathbf{I} = 0$ . Since

$$\underbrace{\mathbf{q} : \hat{\mathbf{b}}\hat{\mathbf{b}}}_{S^\parallel} = \underbrace{\mathbf{q}^g : \hat{\mathbf{b}}\hat{\mathbf{b}}}_{q_\parallel \hat{\mathbf{b}}} + \mathbf{q}^{\text{ng}} : \hat{\mathbf{b}}\hat{\mathbf{b}}; \quad (\text{D } 3)$$

$$\underbrace{\mathbf{q} : (\mathbf{I} - \hat{\mathbf{b}}\hat{\mathbf{b}})/2}_{S^\perp} = \underbrace{\mathbf{q}^g : (\mathbf{I} - \hat{\mathbf{b}}\hat{\mathbf{b}})/2}_{q_\perp \hat{\mathbf{b}}} + \mathbf{q}^{\text{ng}} : (\mathbf{I} - \hat{\mathbf{b}}\hat{\mathbf{b}})/2, \quad (\text{D } 4)$$

one can define perpendicular components of the heat flux vectors

$$\mathbf{S}_\perp^\parallel = \mathbf{q}^{\text{ng}} : \hat{\mathbf{b}}\hat{\mathbf{b}}; \quad (\text{D } 5)$$

$$\mathbf{S}_\perp^\perp = \mathbf{q}^{\text{ng}} : (\mathbf{I} - \hat{\mathbf{b}}\hat{\mathbf{b}})/2. \quad (\text{D } 6)$$

Therefore, to obtain these vectors, we want to apply  $:\hat{\mathbf{b}}\hat{\mathbf{b}}$ , and  $:(\mathbf{I} - \hat{\mathbf{b}}\hat{\mathbf{b}})/2$  to the entire equation (D 2). A very useful expression also is

$$\text{Tr } \mathbf{q}^{\text{ng}} = \mathbf{q}^{\text{ng}} : \mathbf{I} = \mathbf{S}_\perp^\parallel + 2\mathbf{S}_\perp^\perp, \quad (\text{D } 7)$$

and the trace of the entire heat flux tensor reads

$$\text{Tr } \mathbf{q} = q_\parallel \hat{\mathbf{b}} + 2q_\perp \hat{\mathbf{b}} + \mathbf{S}_\perp^\parallel + 2\mathbf{S}_\perp^\perp. \quad (\text{D } 8)$$

#### D.1. Non-gyrotropic heat flux vector $\mathbf{S}_\perp^\parallel$

Applying  $:\hat{\mathbf{b}}\hat{\mathbf{b}}$  to the left-hand side of (D 2) yields

$$\begin{aligned} (\hat{\mathbf{b}} \times \mathbf{q}^{\text{ng}})_{ijk}^S \hat{b}_i \hat{b}_j &= (\epsilon_{irs} \hat{b}_r q_{sjk}^{\text{ng}} + \epsilon_{jrs} \hat{b}_r q_{ski}^{\text{ng}} + \epsilon_{krs} \hat{b}_r q_{sij}^{\text{ng}}) \hat{b}_i \hat{b}_j \\ &= \cancel{\epsilon_{irs} \hat{b}_r \hat{b}_i \hat{b}_j q_{sjk}^{\text{ng}}} + \cancel{\epsilon_{jrs} \hat{b}_r \hat{b}_i \hat{b}_j q_{ski}^{\text{ng}}} + \epsilon_{krs} \hat{b}_r q_{sij}^{\text{ng}} \hat{b}_i \hat{b}_j \\ &= \epsilon_{krs} \hat{b}_r (\mathbf{S}_\perp^\parallel)_s = (\hat{\mathbf{b}} \times \mathbf{S}_\perp^\parallel)_k. \end{aligned} \quad (\text{D } 9)$$

Terms on the right-hand side of (D 2) calculate as

$$\left( \frac{d}{dt} \mathbf{q}^g \right)_{ijk} \hat{b}_i \hat{b}_j = \frac{dq_\parallel}{dt} \hat{b}_k + q_\parallel \frac{d\hat{b}_k}{dt} - 2q_\perp \frac{d\hat{b}_k}{dt}; \quad (\text{D } 10)$$

$$(\nabla \cdot \mathbf{r}^g)_{ijk} \hat{b}_i \hat{b}_j = \hat{b}_k \hat{\mathbf{b}} \cdot \nabla r_{\parallel\parallel} + (r_{\parallel\parallel} - 3r_{\parallel\perp}) (\hat{\mathbf{b}} \cdot \nabla \hat{b}_k + \hat{b}_k \nabla \cdot \hat{\mathbf{b}}) - \hat{b}_k \hat{\mathbf{b}} \cdot \nabla r_{\perp\perp} + \partial_k r_{\perp\perp}; \quad (\text{D } 11)$$

$$q_{ijk}^g \hat{b}_i = q_\parallel \hat{b}_j \hat{b}_k + q_\perp (\delta_{jk} - \hat{b}_j \hat{b}_k); \quad (\text{D } 12)$$

$$q_{ijk}^g \hat{b}_i \hat{b}_j \nabla \cdot \mathbf{u} = q_{\parallel} \hat{b}_k \nabla \cdot \mathbf{u}; \quad (\text{D } 13)$$

$$(q^g \cdot \nabla \mathbf{u})_{ijk}^S \hat{b}_i \hat{b}_j = q_{\parallel} \hat{\mathbf{b}} \cdot \nabla u_k + 2(q_{\parallel} - q_{\perp}) \hat{b}_k \hat{\mathbf{b}} \cdot (\nabla \mathbf{u}) \cdot \hat{\mathbf{b}} + 2q_{\perp} (\partial_k \mathbf{u}) \cdot \hat{\mathbf{b}}. \quad (\text{D } 14)$$

For the last term on the right-hand side of (D 2)

$$(\mathbf{p}(\nabla \cdot \mathbf{p}))_{ijk}^S \hat{b}_i \hat{b}_j = p_{\parallel} (\nabla \cdot \mathbf{p})_k + 2(p_{\parallel} \hat{b}_k + \Pi_{jk} \hat{b}_j) (\nabla \cdot \mathbf{p}) \cdot \hat{\mathbf{b}}; \quad (\text{D } 15)$$

$$(\nabla \cdot \mathbf{p})_k = \hat{b}_k \hat{\mathbf{b}} \cdot \nabla (p_{\parallel} - p_{\perp}) + (p_{\parallel} - p_{\perp}) (\hat{\mathbf{b}} \cdot \nabla \hat{b}_k + \hat{b}_k \nabla \cdot \hat{\mathbf{b}}) + \partial_k p_{\perp} + (\nabla \cdot \Pi)_k; \quad (\text{D } 16)$$

$$(\nabla \cdot \mathbf{p}) \cdot \hat{\mathbf{b}} = \hat{\mathbf{b}} \cdot \nabla p_{\parallel} + (p_{\parallel} - p_{\perp}) \nabla \cdot \hat{\mathbf{b}} + (\nabla \cdot \Pi) \cdot \hat{\mathbf{b}}, \quad (\text{D } 17)$$

and so for the gyrotropic

$$(\mathbf{p}^g(\nabla \cdot \mathbf{p}^g))_{ijk}^S \hat{b}_i \hat{b}_j = p_{\parallel} \hat{b}_k \hat{\mathbf{b}} \cdot \nabla (3p_{\parallel} - p_{\perp}) + p_{\parallel} (p_{\parallel} - p_{\perp}) (\hat{\mathbf{b}} \cdot \nabla \hat{b}_k + 3\hat{b}_k \nabla \cdot \hat{\mathbf{b}}) + p_{\parallel} \partial_k p_{\perp}. \quad (\text{D } 18)$$

Collecting all the results together yields

$$\begin{aligned} (\hat{\mathbf{b}} \times \mathbf{S}_{\perp}^{\parallel})_k = & -\frac{B_0}{\Omega |\mathbf{B}|} \left[ \frac{dq_{\parallel}}{dt} \hat{b}_k + q_{\parallel} \frac{d\hat{b}_k}{dt} - 2q_{\perp} \frac{d\hat{b}_k}{dt} + \hat{b}_k \hat{\mathbf{b}} \cdot \nabla r_{\parallel\parallel} \right. \\ & + (r_{\parallel\parallel} - 3r_{\parallel\perp}) (\hat{\mathbf{b}} \cdot \nabla \hat{b}_k + \hat{b}_k \nabla \cdot \hat{\mathbf{b}}) - \hat{b}_k \hat{\mathbf{b}} \cdot \nabla r_{\parallel\perp} + \partial_k r_{\parallel\perp} \\ & + q_{\parallel} \hat{b}_k \nabla \cdot \mathbf{u} + q_{\parallel} \hat{\mathbf{b}} \cdot \nabla u_k + 2(q_{\parallel} - q_{\perp}) \hat{b}_k \hat{\mathbf{b}} \cdot (\nabla \mathbf{u}) \cdot \hat{\mathbf{b}} + 2q_{\perp} (\partial_k \mathbf{u}) \cdot \hat{\mathbf{b}} \\ & \left. - \frac{p_{\parallel}}{\rho} \hat{b}_k \hat{\mathbf{b}} \cdot \nabla (3p_{\parallel} - p_{\perp}) - \frac{p_{\parallel}}{\rho} (p_{\parallel} - p_{\perp}) (\hat{\mathbf{b}} \cdot \nabla \hat{b}_k + 3\hat{b}_k \nabla \cdot \hat{\mathbf{b}}) - \frac{p_{\parallel}}{\rho} \partial_k p_{\perp} \right]. \end{aligned} \quad (\text{D } 19)$$

The equation appears very complicated, however, it is not overly so, since it is only a vector equation. The equation just describes perpendicular components of  $\mathbf{S}_{\perp}^{\parallel}$ . By multiplying it with  $\hat{b}_k$ , the left-hand side is zero, and on the right-hand side one recovers the gyrotropic parallel heat flux equation (7.102). To understand the equation better, let us simplify for a moment and evaluate it with respect to  $\hat{\mathbf{b}}_0 = (0, 0, 1)$ . The  $x$ -component  $(\hat{\mathbf{b}}_0 \times \mathbf{S}_{\perp}^{\parallel})_x = -(\mathbf{S}_{\perp}^{\parallel})_y$ , so that

$$\begin{aligned} (\mathbf{S}_{\perp}^{\parallel})_y = & \frac{1}{\Omega} \left[ q_{\parallel} \frac{d\hat{b}_x}{dt} - 2q_{\perp} \frac{d\hat{b}_x}{dt} + (r_{\parallel\parallel} - 3r_{\parallel\perp}) \partial_z \hat{b}_x + \partial_x r_{\parallel\perp} \right. \\ & \left. + q_{\parallel} \partial_z u_x + 2q_{\perp} \partial_x u_z - \frac{p_{\parallel}}{\rho} (p_{\parallel} - p_{\perp}) \partial_z \hat{b}_x - \frac{p_{\parallel}}{\rho} \partial_x p_{\perp} \right] \end{aligned} \quad (\text{D } 20)$$

and similarly for the  $y$ -component  $(\hat{\mathbf{b}}_0 \times \mathbf{S}_{\perp}^{\parallel})_y = +(\mathbf{S}_{\perp}^{\parallel})_x$ , yielding

$$\begin{aligned} (\mathbf{S}_{\perp}^{\parallel})_x = & -\frac{1}{\Omega} \left[ q_{\parallel} \frac{d\hat{b}_y}{dt} - 2q_{\perp} \frac{d\hat{b}_y}{dt} + (r_{\parallel\parallel} - 3r_{\parallel\perp}) \partial_z \hat{b}_y + \partial_y r_{\parallel\perp} \right. \\ & \left. + q_{\parallel} \partial_z u_y + 2q_{\perp} \partial_y u_z - \frac{p_{\parallel}}{\rho} (p_{\parallel} - p_{\perp}) \partial_z \hat{b}_y - \frac{p_{\parallel}}{\rho} \partial_y p_{\perp} \right]. \end{aligned} \quad (\text{D } 21)$$

Now, by using the leading-order MHD induction equation  $d\hat{\mathbf{b}}_x/dt = \partial_z u_x$  and  $d\hat{\mathbf{b}}_y/dt = \partial_z u_y$  (which was for example used to get the FLR1 pressure tensor), and further considering specific example of a bi-Maxwellian expressions for the fourth-order moments, i.e.  $r_{\parallel\parallel} = 3p_{\parallel}^2/\rho + \tilde{r}_{\parallel\parallel}$  etc., yields

$$(\mathbf{S}_{\perp}^{\parallel})_y = \frac{1}{\Omega} \left[ p_{\perp} \partial_x \left( \frac{p_{\parallel}}{\rho} \right) + 2 \frac{p_{\parallel}}{\rho} (p_{\parallel} - p_{\perp}) \partial_z \hat{\mathbf{b}}_x + 2q_{\parallel} \partial_z u_x + 2q_{\perp} (\partial_x u_z - \partial_z u_x) \right. \\ \left. + (\tilde{r}_{\parallel\parallel} - 3\tilde{r}_{\parallel\perp}) \partial_z \hat{\mathbf{b}}_x + \partial_x \tilde{r}_{\parallel\perp} \right]; \quad (\text{D } 22)$$

$$(\mathbf{S}_{\perp}^{\parallel})_x = -\frac{1}{\Omega} \left[ p_{\perp} \partial_y \left( \frac{p_{\parallel}}{\rho} \right) + 2 \frac{p_{\parallel}}{\rho} (p_{\parallel} - p_{\perp}) \partial_z \hat{\mathbf{b}}_y + 2q_{\parallel} \partial_z u_y + 2q_{\perp} (\partial_y u_z - \partial_z u_y) \right. \\ \left. + (\tilde{r}_{\parallel\parallel} - 3\tilde{r}_{\parallel\perp}) \partial_z \hat{\mathbf{b}}_y + \partial_y \tilde{r}_{\parallel\perp} \right]. \quad (\text{D } 23)$$

The above equations are very useful if we ever need linearized expressions for  $\mathbf{S}_{\perp}^{\parallel}$ , which is beneficial to do right now, and the partially linearized expressions read

$$(\mathbf{S}_{\perp}^{\parallel})_y \stackrel{\text{lin}}{=} \frac{1}{\Omega} \left[ p_{\perp} \partial_x \left( \frac{p_{\parallel}}{\rho} \right) + 2 \frac{p_{\parallel}}{\rho} (p_{\parallel} - p_{\perp}) \partial_z \hat{\mathbf{b}}_x + \partial_x \tilde{r}_{\parallel\perp} \right]; \quad (\text{D } 24)$$

$$(\mathbf{S}_{\perp}^{\parallel})_x \stackrel{\text{lin}}{=} -\frac{1}{\Omega} \left[ p_{\perp} \partial_y \left( \frac{p_{\parallel}}{\rho} \right) + 2 \frac{p_{\parallel}}{\rho} (p_{\parallel} - p_{\perp}) \partial_z \hat{\mathbf{b}}_y + \partial_y \tilde{r}_{\parallel\perp} \right]. \quad (\text{D } 25)$$

Note that the mean values of perturbations  $\tilde{r}^{(0)} = 0$ . It is also noteworthy to point out that the contributions from the induction equation  $d\hat{\mathbf{b}}/dt$  are completely eliminated by the linearization process, and more elaborate forms of induction equation will not bring additional precision at the linear level. Coming back to the expressions (D 22), (D 23), it is also possible to introduce vorticity  $\boldsymbol{\omega} = \nabla \times \mathbf{u}$ , with components  $\omega_x = (\partial_y u_z - \partial_z u_y)$  and  $\omega_y = -(\partial_x u_z - \partial_z u_x)$ . By noting that  $(\hat{\mathbf{b}}_0 \times \boldsymbol{\omega})_x = -\omega_y = (\partial_x u_z - \partial_z u_x)$  and  $(\hat{\mathbf{b}}_0 \times \boldsymbol{\omega})_y = \omega_x = (\partial_y u_z - \partial_z u_y)$ , both equations can be written together as

$$\mathbf{S}_{\perp}^{\parallel} = \frac{1}{\Omega} \hat{\mathbf{b}}_0 \times \left[ p_{\perp} \nabla \left( \frac{p_{\parallel}}{\rho} \right) + 2 \frac{p_{\parallel}}{\rho} (p_{\parallel} - p_{\perp}) \partial_z \hat{\mathbf{b}} + 2q_{\parallel} \partial_z \mathbf{u} + 2q_{\perp} \hat{\mathbf{b}}_0 \times \boldsymbol{\omega} \right. \\ \left. + (\tilde{r}_{\parallel\parallel} - 3\tilde{r}_{\parallel\perp}) \partial_z \hat{\mathbf{b}} + \nabla \tilde{r}_{\parallel\perp} \right]. \quad (\text{D } 26)$$

Now that we understand that equation (D 19) just describes the perpendicular components of  $\mathbf{S}_{\perp}^{\parallel}$  with respect to  $\hat{\mathbf{b}}$ , since the parallel components are zero (directly from the decomposition of  $\mathbf{S}^{\parallel}$ )

$$\mathbf{S}_{\perp}^{\parallel} \cdot \hat{\mathbf{b}} = 0; \quad \mathbf{S}_{\perp}^{\perp} \cdot \hat{\mathbf{b}} = 0, \quad (\text{D } 27)$$

we actually do not have to evaluate equation (D 19) with respect to  $\hat{\mathbf{b}}_0$ . For a general vector  $\mathbf{a}$ , the following identity holds  $\hat{\mathbf{b}} \times (\hat{\mathbf{b}} \times \mathbf{a}) = \hat{\mathbf{b}}(\hat{\mathbf{b}} \cdot \mathbf{a}) - \mathbf{a}$ . So for a vector  $\mathbf{a}$  which does not have any parallel components to  $\hat{\mathbf{b}}$ , so that  $\hat{\mathbf{b}} \cdot \mathbf{a} = 0$ , the identity is  $\hat{\mathbf{b}} \times (\hat{\mathbf{b}} \times \mathbf{a}) = -\mathbf{a}$ , and

$$\hat{\mathbf{b}} \times (\hat{\mathbf{b}} \times \mathbf{S}_{\perp}^{\parallel}) = -\mathbf{S}_{\perp}^{\parallel}; \quad \hat{\mathbf{b}} \times (\hat{\mathbf{b}} \times \mathbf{S}_{\perp}^{\perp}) = -\mathbf{S}_{\perp}^{\perp}. \quad (\text{D } 28)$$

We can apply  $\hat{\mathbf{b}} \times$  to the entire equation (D 19), and get the fully nonlinear expression for  $S_{\perp}^{\parallel}$ . Let us apply  $\hat{\mathbf{b}} \times$  to each term separately, since we will need these results later, the terms calculate as

$$\hat{\mathbf{b}} \times \left[ \left( \frac{d}{dt} \mathbf{q}^g \right) : \hat{\mathbf{b}} \hat{\mathbf{b}} \right] = \hat{\mathbf{b}} \times \left[ q_{\parallel} \frac{d\hat{\mathbf{b}}}{dt} - 2q_{\perp} \frac{d\hat{\mathbf{b}}}{dt} \right]; \quad (\text{D } 29)$$

$$\hat{\mathbf{b}} \times [(\nabla \cdot \mathbf{r}^g) : \hat{\mathbf{b}} \hat{\mathbf{b}}] = \hat{\mathbf{b}} \times [(r_{\parallel\parallel} - 3r_{\parallel\perp}) \hat{\mathbf{b}} \cdot \nabla \hat{\mathbf{b}} + \nabla r_{\parallel\perp}]; \quad (\text{D } 30)$$

$$\hat{\mathbf{b}} \times [\mathbf{q}^g (\nabla \cdot \mathbf{u}) : \hat{\mathbf{b}} \hat{\mathbf{b}}] = 0; \quad (\text{D } 31)$$

$$\hat{\mathbf{b}} \times [(\mathbf{q}^g \cdot \nabla \mathbf{u})^S : \hat{\mathbf{b}} \hat{\mathbf{b}}] = \hat{\mathbf{b}} \times [q_{\parallel} \hat{\mathbf{b}} \cdot \nabla \mathbf{u} + 2q_{\perp} (\nabla \mathbf{u}) \cdot \hat{\mathbf{b}}]; \quad (\text{D } 32)$$

$$\hat{\mathbf{b}} \times [(\mathbf{p}^g (\nabla \cdot \mathbf{p}^g))^S : \hat{\mathbf{b}} \hat{\mathbf{b}}] = \hat{\mathbf{b}} \times [p_{\parallel} (p_{\parallel} - p_{\perp}) \hat{\mathbf{b}} \cdot \nabla \hat{\mathbf{b}} + p_{\parallel} \nabla p_{\perp}]. \quad (\text{D } 33)$$

To summarize, applying  $\hat{\mathbf{b}} \hat{\mathbf{b}}$  and  $\hat{\mathbf{b}} \times$  to equation (D 2), yields the following nonlinear expression

$$\begin{aligned} S_{\perp}^{\parallel} = & \frac{B_0}{\Omega |\mathbf{B}|} \hat{\mathbf{b}} \times \left[ q_{\parallel} \left( \frac{d\hat{\mathbf{b}}}{dt} + \hat{\mathbf{b}} \cdot \nabla \mathbf{u} \right) + 2q_{\perp} \left( (\nabla \mathbf{u}) \cdot \hat{\mathbf{b}} - \frac{d\hat{\mathbf{b}}}{dt} \right) + (r_{\parallel\parallel} - 3r_{\parallel\perp}) \hat{\mathbf{b}} \cdot \nabla \hat{\mathbf{b}} \right. \\ & \left. - \frac{p_{\parallel}}{\rho} (p_{\parallel} - p_{\perp}) \hat{\mathbf{b}} \cdot \nabla \hat{\mathbf{b}} + \nabla r_{\parallel\perp} - \frac{p_{\parallel}}{\rho} \nabla p_{\perp} \right]. \end{aligned} \quad (\text{D } 34)$$

Now the calculations proceed in a similar way as before. The usual (MHD) induction equation can be written as

$$\frac{d\hat{\mathbf{b}}}{dt} = \hat{\mathbf{b}} \cdot \nabla \mathbf{u} - \hat{\mathbf{b}} \left[ \hat{\mathbf{b}} \cdot (\nabla \mathbf{u}) \cdot \hat{\mathbf{b}} \right]; \quad \Rightarrow \quad \hat{\mathbf{b}} \times \frac{d\hat{\mathbf{b}}}{dt} = \hat{\mathbf{b}} \times (\hat{\mathbf{b}} \cdot \nabla \mathbf{u}), \quad (\text{D } 35)$$

and since  $(\nabla \mathbf{u}) \cdot \hat{\mathbf{b}} - \hat{\mathbf{b}} \cdot \nabla \mathbf{u} = \hat{\mathbf{b}} \times (\nabla \times \mathbf{u}) = \hat{\mathbf{b}} \times \boldsymbol{\omega}$ ,

$$\begin{aligned} S_{\perp}^{\parallel} = & \frac{B_0}{\Omega |\mathbf{B}|} \hat{\mathbf{b}} \times \left[ 2q_{\parallel} \hat{\mathbf{b}} \cdot \nabla \mathbf{u} + 2q_{\perp} \hat{\mathbf{b}} \times \boldsymbol{\omega} + (r_{\parallel\parallel} - 3r_{\parallel\perp}) \hat{\mathbf{b}} \cdot \nabla \hat{\mathbf{b}} \right. \\ & \left. - \frac{p_{\parallel}}{\rho} (p_{\parallel} - p_{\perp}) \hat{\mathbf{b}} \cdot \nabla \hat{\mathbf{b}} + \nabla r_{\parallel\perp} - \frac{p_{\parallel}}{\rho} \nabla p_{\perp} \right]. \end{aligned} \quad (\text{D } 36)$$

Finally, by prescribing bi-Maxwellian perturbations for the fourth-order moment yields

$$\begin{aligned} S_{\perp}^{\parallel} = & \frac{B_0}{\Omega |\mathbf{B}|} \hat{\mathbf{b}} \times \left[ p_{\perp} \nabla \left( \frac{p_{\parallel}}{\rho} \right) + 2 \frac{p_{\parallel}}{\rho} (p_{\parallel} - p_{\perp}) \hat{\mathbf{b}} \cdot \nabla \hat{\mathbf{b}} + 2q_{\parallel} \hat{\mathbf{b}} \cdot \nabla \mathbf{u} + 2q_{\perp} \hat{\mathbf{b}} \times \boldsymbol{\omega} \right. \\ & \left. + (\tilde{r}_{\parallel\parallel} - 3\tilde{r}_{\parallel\perp}) \hat{\mathbf{b}} \cdot \nabla \hat{\mathbf{b}} + \nabla \tilde{r}_{\parallel\perp} \right]. \end{aligned} \quad (\text{D } 37)$$

We can report a sign typo in (3.7) of Sulem & Passot (2015), whose term  $3\tilde{r}_{\parallel\perp}$  has the opposite sign.

## D.2. Conversion of other notations

In the paper of Ramos (2005), the following notation is used

$$\left. \begin{aligned} q_{B\parallel} &= \frac{1}{2}q_{\parallel}; & q_{T\parallel} &= q_{\perp}; & \mathbf{q}_{B\perp} &= \frac{1}{2}\mathbf{S}_{\perp}^{\parallel}; & \mathbf{q}_{T\perp} &= \mathbf{S}_{\perp}^{\perp}; \\ \tilde{r}_{\parallel}^{(0)} &= m\left(\frac{1}{2}\tilde{r}_{\parallel\parallel} + \tilde{r}_{\parallel\perp}\right); & \tilde{r}_{B\perp}^{(0)} &= \frac{m}{2}\tilde{r}_{\parallel\perp}; & \tilde{r}_{\perp}^{(0)} &= m\left(\frac{1}{2}\tilde{r}_{\parallel\perp} + \tilde{r}_{\perp\perp}\right), \end{aligned} \right\} \quad (\text{D } 38)$$

and his equations (54), (57) are indeed equivalent to (D 37).

In the paper by Macmahon (1965), the following notation is used

$$\left. \begin{aligned} \mathbf{q}^{\parallel} &= \frac{1}{2}\mathbf{S}^{\parallel}; & \mathbf{q}^{\perp} &= \mathbf{S}^{\perp}; & q_{\parallel} &= \frac{1}{2}q_{\parallel}; & q_{\parallel}^{\perp} &= q_{\perp}; & \mathbf{q}_{\perp}^{\parallel} &= \frac{1}{2}\mathbf{S}_{\perp}^{\parallel}; & \mathbf{q}_{\perp}^{\perp} &= \mathbf{S}_{\perp}^{\perp}; \\ R_1 &= r_{\parallel\parallel}; & R_2 &= r_{\parallel\perp}; & R_3 &= 2r_{\perp\perp}. \end{aligned} \right\} \quad (\text{D } 39)$$

and his equation (14) is equivalent to (D 36), if in Macmahon (1965) we continue the calculation of

$$\hat{\mathbf{b}} \times (\nabla \cdot \mathbf{p}^g) = \hat{\mathbf{b}} \times [(p_{\parallel} - p_{\perp})\hat{\mathbf{b}} \cdot \nabla \hat{\mathbf{b}} + \nabla p_{\perp}]. \quad (\text{D } 40)$$

D.3. Non-gyrotropic heat flux vector  $\mathbf{S}_{\perp}^{\perp}$ 

Now we want to apply :  $(\mathbf{I} - \hat{\mathbf{b}}\hat{\mathbf{b}})/2$  to equation (D 2). Additionally, we now know that  $\mathbf{S}_{\perp}^{\perp}$  has only perpendicular components, and that afterward we will be also applying  $\hat{\mathbf{b}} \times$ . The left-hand side calculates trivially as

$$\begin{aligned} \delta_{ij}(\hat{\mathbf{b}} \times \mathbf{q}^g)_{ijk}^S &= \delta_{ij}(\epsilon_{irs}\hat{b}_r q_{sjk}^g + \epsilon_{jrs}\hat{b}_r q_{ski}^g + \epsilon_{krs}\hat{b}_r q_{sij}^g) \\ &= \hat{b}_r \epsilon_{irs} q_{sik}^g + \hat{b}_r \epsilon_{jrs} q_{ski}^g + \epsilon_{krs} \hat{b}_r q_{sij}^g = (\hat{\mathbf{b}} \times (\mathbf{q}^g : \mathbf{I}))_k; \end{aligned} \quad (\text{D } 41)$$

$$(\hat{\mathbf{b}} \times \mathbf{q}^g)^S : (\mathbf{I} - \hat{\mathbf{b}}\hat{\mathbf{b}})/2 = \hat{\mathbf{b}} \times \mathbf{S}_{\perp}^{\perp} \quad (\text{D } 42)$$

and applying  $\hat{\mathbf{b}} \times$  therefore yields

$$\hat{\mathbf{b}} \times [(\hat{\mathbf{b}} \times \mathbf{q}^g)^S : (\mathbf{I} - \hat{\mathbf{b}}\hat{\mathbf{b}})/2] = -\mathbf{S}_{\perp}^{\perp}. \quad (\text{D } 43)$$

Let us continue with applying trace to equation (D 2) term by term

$$\delta_{ij} \left( \frac{d}{dt} \mathbf{q}^g \right)_{ijk} = \hat{b}_k \frac{d}{dt} (q_{\parallel} + 2q_{\perp}) + (q_{\parallel} + 2q_{\perp}) \frac{d}{dt} \hat{b}_k; \quad (\text{D } 44)$$

$$\begin{aligned} \delta_{ij}(\nabla \cdot \mathbf{r}^g)_{ijk} &= \hat{b}_k \hat{\mathbf{b}} \cdot \nabla (r_{\parallel\parallel} + r_{\parallel\perp} - 2r_{\perp\perp}) \\ &\quad + (r_{\parallel\parallel} + r_{\parallel\perp} - 2r_{\perp\perp})(\hat{\mathbf{b}} \cdot \nabla \hat{b}_k + \hat{b}_k \nabla \cdot \hat{\mathbf{b}}) + \partial_k r_{\parallel\perp} + 2\partial_k r_{\perp\perp}; \end{aligned} \quad (\text{D } 45)$$

$$\delta_{ij} q_{ijk}^g \nabla \cdot \mathbf{u} = (q_{\parallel} + 2q_{\perp}) \hat{b}_k \nabla \cdot \mathbf{u}; \quad (\text{D } 46)$$

$$\begin{aligned} \delta_{ij}(\mathbf{q}^g \cdot \nabla \mathbf{u})_{ijk}^S &= q_{\parallel} \hat{\mathbf{b}} \cdot \nabla u_k + 2q_{\parallel} \hat{b}_k \hat{\mathbf{b}} \cdot (\nabla \mathbf{u}) \cdot \hat{\mathbf{b}} \\ &\quad + 2q_{\perp} [2\hat{\mathbf{b}} \cdot \nabla u_k + (\partial_k \mathbf{u}) \cdot \hat{\mathbf{b}} + \hat{b}_k \nabla \cdot \mathbf{u} - 3\hat{b}_k \hat{\mathbf{b}} \cdot (\nabla \mathbf{u}) \cdot \hat{\mathbf{b}}]; \end{aligned} \quad (\text{D } 47)$$

$$\begin{aligned} \delta_{ij}(\mathbf{p}^g(\nabla \cdot \mathbf{p}^g))_{ijk}^S &= (3p_{\parallel} + 2p_{\perp})(p_{\parallel} - p_{\perp})\hat{\mathbf{b}}_k \nabla \cdot \hat{\mathbf{b}} + (p_{\parallel} + 4p_{\perp})(p_{\parallel} - p_{\perp})\hat{\mathbf{b}} \cdot \nabla \hat{\mathbf{b}}_k \\ &\quad + (p_{\parallel} + 4p_{\perp})\partial_k p_{\perp} - (p_{\parallel} + 4p_{\perp})\hat{\mathbf{b}}_k \hat{\mathbf{b}} \cdot \nabla p_{\perp} + (3p_{\parallel} + 2p_{\perp})\hat{\mathbf{b}}_k \hat{\mathbf{b}} \cdot \nabla p_{\parallel}, \end{aligned} \quad (\text{D } 48)$$

and by further applying  $\hat{\mathbf{b}} \times$  to these terms yields

$$\hat{\mathbf{b}} \times \text{Tr} \left( \frac{d}{dt} \mathbf{q}^g \right) = \hat{\mathbf{b}} \times \left[ (q_{\parallel} + 2q_{\perp}) \frac{d\hat{\mathbf{b}}}{dt} \right]; \quad (\text{D } 49)$$

$$\hat{\mathbf{b}} \times \text{Tr}(\nabla \cdot \mathbf{r}^g) = \hat{\mathbf{b}} \times [(r_{\parallel\parallel} + r_{\perp\perp} - 2r_{\perp\perp})\hat{\mathbf{b}} \cdot \nabla \hat{\mathbf{b}} + \nabla r_{\parallel\perp} + 2\nabla r_{\perp\perp}]; \quad (\text{D } 50)$$

$$\hat{\mathbf{b}} \times \text{Tr} \mathbf{q}^g \nabla \cdot \mathbf{u} = 0; \quad (\text{D } 51)$$

$$\hat{\mathbf{b}} \times \text{Tr}(\mathbf{q}^g \cdot \nabla \mathbf{u})^S = \hat{\mathbf{b}} \times [q_{\parallel} \hat{\mathbf{b}} \cdot \nabla \mathbf{u} + 2q_{\perp} (2\hat{\mathbf{b}} \cdot \nabla \mathbf{u} + (\nabla \mathbf{u}) \cdot \hat{\mathbf{b}})]; \quad (\text{D } 52)$$

$$\hat{\mathbf{b}} \times \text{Tr}(\mathbf{p}^g(\nabla \cdot \mathbf{p}^g))^S = \hat{\mathbf{b}} \times [(p_{\parallel} + 4p_{\perp})(p_{\parallel} - p_{\perp})\hat{\mathbf{b}} \cdot \nabla \hat{\mathbf{b}} + (p_{\parallel} + 4p_{\perp})\nabla p_{\perp}], \quad (\text{D } 53)$$

and direct subtraction with results (D 29)–(D 33) gives

$$\hat{\mathbf{b}} \times \left[ \left( \frac{d}{dt} \mathbf{q}^g \right) : (\mathbf{I} - \hat{\mathbf{b}}\hat{\mathbf{b}})/2 \right] = \hat{\mathbf{b}} \times \left[ 2q_{\perp} \frac{d\hat{\mathbf{b}}}{dt} \right]; \quad (\text{D } 54)$$

$$\hat{\mathbf{b}} \times [(\nabla \cdot \mathbf{r}^g) : (\mathbf{I} - \hat{\mathbf{b}}\hat{\mathbf{b}})/2] = \hat{\mathbf{b}} \times [(2r_{\parallel\perp} - r_{\perp\perp})\hat{\mathbf{b}} \cdot \nabla \hat{\mathbf{b}} + \nabla r_{\perp\perp}]; \quad (\text{D } 55)$$

$$\hat{\mathbf{b}} \times [(\mathbf{q}^g \cdot \nabla \mathbf{u})^S : (\mathbf{I} - \hat{\mathbf{b}}\hat{\mathbf{b}})/2] = \hat{\mathbf{b}} \times [2q_{\perp} \hat{\mathbf{b}} \cdot \nabla \mathbf{u}]; \quad (\text{D } 56)$$

$$\hat{\mathbf{b}} \times [(\mathbf{p}^g(\nabla \cdot \mathbf{p}^g))^S : (\mathbf{I} - \hat{\mathbf{b}}\hat{\mathbf{b}})/2] = \hat{\mathbf{b}} \times [2p_{\perp}(p_{\parallel} - p_{\perp})\hat{\mathbf{b}} \cdot \nabla \hat{\mathbf{b}} + 2p_{\perp}\nabla p_{\perp}]. \quad (\text{D } 57)$$

To summarize, applying  $:(\mathbf{I} - \hat{\mathbf{b}}\hat{\mathbf{b}})/2$  and  $\hat{\mathbf{b}} \times$  to equation (D 2) yields the following nonlinear expression

$$\begin{aligned} \mathbf{S}_{\perp}^{\perp} &= \frac{B_0}{\Omega |\mathbf{B}|} \hat{\mathbf{b}} \times \left[ 2q_{\perp} \frac{d\hat{\mathbf{b}}}{dt} + (2r_{\parallel\perp} - r_{\perp\perp})\hat{\mathbf{b}} \cdot \nabla \hat{\mathbf{b}} + \nabla r_{\perp\perp} + 2q_{\perp} \hat{\mathbf{b}} \cdot \nabla \mathbf{u} \right. \\ &\quad \left. - 2\frac{p_{\perp}}{\rho}(p_{\parallel} - p_{\perp})\hat{\mathbf{b}} \cdot \nabla \hat{\mathbf{b}} - 2\frac{p_{\perp}}{\rho}\nabla p_{\perp} \right]. \end{aligned} \quad (\text{D } 58)$$

Using the simple MHD induction equation gives

$$\begin{aligned} \mathbf{S}_{\perp}^{\perp} &= \frac{B_0}{\Omega |\mathbf{B}|} \hat{\mathbf{b}} \times \left[ (2r_{\parallel\perp} - r_{\perp\perp})\hat{\mathbf{b}} \cdot \nabla \hat{\mathbf{b}} + \nabla r_{\perp\perp} + 4q_{\perp} \hat{\mathbf{b}} \cdot \nabla \mathbf{u} \right. \\ &\quad \left. - 2\frac{p_{\perp}}{\rho}(p_{\parallel} - p_{\perp})\hat{\mathbf{b}} \cdot \nabla \hat{\mathbf{b}} - 2\frac{p_{\perp}}{\rho}\nabla p_{\perp} \right], \end{aligned} \quad (\text{D } 59)$$

and finally prescribing bi-Maxwellian values for the fourth-order moments yields

$$\mathbf{S}_{\perp}^{\perp} = \frac{B_0}{\Omega |\mathbf{B}|} \hat{\mathbf{b}} \times \left[ 2p_{\perp} \nabla \left( \frac{p_{\perp}}{\rho} \right) + 4q_{\perp} \hat{\mathbf{b}} \cdot \nabla \mathbf{u} + (2\tilde{r}_{\parallel\perp} - \tilde{r}_{\perp\perp})\hat{\mathbf{b}} \cdot \nabla \hat{\mathbf{b}} + \nabla \tilde{r}_{\perp\perp} \right]. \quad (\text{D } 60)$$

Additionally, evaluation from linear kinetic theory in the gyrotropic limit (at long wavelengths) yields  $\tilde{r}_{\perp\perp} = 0$ , as we will see in Part 2. With this requirement, equation (D 60) is equivalent to (3.6) of Sulem & Passot (2015). The result (D 60) is also

equivalent to expressions (55), (58) of Ramos (2005). However, equation (15) of Macmahon (1965) appears to have a typo, where in comparison to our (D 59), his definition of  $\mathbf{T}$  appears to be missing a factor of 2 in front of the  $\nabla \mathbf{u}$  term. Ramos (2005) however states that his results are equivalent to those of Macmahon (1965), so we might not correctly understand the notation in that paper.

Partial linearization of result (D 60) yields that by components

$$(\mathbf{S}_{\perp}^{\perp})_y \stackrel{\text{lin}}{=} \frac{1}{\Omega} \left[ 2p_{\perp} \partial_x \left( \frac{p_{\perp}}{\rho} \right) + \partial_x \tilde{r}_{\perp\perp} \right]; \quad (\text{D } 61)$$

$$(\mathbf{S}_{\perp}^{\perp})_x \stackrel{\text{lin}}{=} -\frac{1}{\Omega} \left[ 2p_{\perp} \partial_y \left( \frac{p_{\perp}}{\rho} \right) + \partial_y \tilde{r}_{\perp\perp} \right], \quad (\text{D } 62)$$

where we have kept the  $\tilde{r}_{\perp\perp}$  contributions, just in case we need them in the future. Note that, similarly to the linearization of the vector  $\mathbf{S}_{\perp}^{\parallel}$ , the induction equation is completely eliminated at the linear level, and more elaborate forms of induction equation will not have additional contributions.

#### D.4. Second-order heat flux vectors

In a previous section we have seen that by applying :  $\hat{\mathbf{b}}\hat{\mathbf{b}}$ , :  $(\mathbf{I} - \hat{\mathbf{b}}\hat{\mathbf{b}})/2$  and  $\hat{\mathbf{b}} \times$  to equation (D 1), the fully nonlinear (but implicit) expressions for the non-gyrotropic heat flux vectors read

$$\mathbf{S}_{\perp}^{\parallel} = \frac{B_0}{\Omega |\mathbf{B}|} \hat{\mathbf{b}} \times \left\{ \hat{\mathbf{b}}\hat{\mathbf{b}} : \left[ \frac{d}{dt} \mathbf{q} + \nabla \cdot \mathbf{r} + \mathbf{q} \nabla \cdot \mathbf{u} + (\mathbf{q} \cdot \nabla \mathbf{u})^S - \frac{1}{\rho} (\mathbf{p}(\nabla \cdot \mathbf{p}))^S \right] \right\}; \quad (\text{D } 63)$$

$$\mathbf{S}_{\perp}^{\perp} = \frac{B_0}{\Omega |\mathbf{B}|} \hat{\mathbf{b}} \times \left\{ (\mathbf{I} - \hat{\mathbf{b}}\hat{\mathbf{b}})/2 : \left[ \frac{d}{dt} \mathbf{q} + \nabla \cdot \mathbf{r} + \mathbf{q} \nabla \cdot \mathbf{u} + (\mathbf{q} \cdot \nabla \mathbf{u})^S - \frac{1}{\rho} (\mathbf{p}(\nabla \cdot \mathbf{p}))^S \right] \right\}. \quad (\text{D } 64)$$

Here, we are interested only in the linear level contributions, which we want to use in the FLR pressure tensor. Therefore, since we know that here we will linearize everything at the end, we can immediately get rid of the two heat flux terms  $\mathbf{q} \nabla \cdot \mathbf{u}$  and  $(\mathbf{q} \cdot \nabla \mathbf{u})^S$ , and also replace the  $d/dt$  by  $\partial/\partial t$ , so partial linearization yields

$$\mathbf{S}_{\perp}^{\parallel} \stackrel{\text{lin}}{=} \frac{1}{\Omega} \hat{\mathbf{b}} \times \left\{ \hat{\mathbf{b}}\hat{\mathbf{b}} : \left[ \frac{\partial}{\partial t} \mathbf{q} + \nabla \cdot \mathbf{r} - \frac{1}{\rho} (\mathbf{p}(\nabla \cdot \mathbf{p}))^S \right] \right\}; \quad (\text{D } 65)$$

$$\mathbf{S}_{\perp}^{\perp} \stackrel{\text{lin}}{=} \frac{1}{\Omega} \hat{\mathbf{b}} \times \left\{ (\mathbf{I} - \hat{\mathbf{b}}\hat{\mathbf{b}})/2 : \left[ \frac{\partial}{\partial t} \mathbf{q} + \nabla \cdot \mathbf{r} - \frac{1}{\rho} (\mathbf{p}(\nabla \cdot \mathbf{p}))^S \right] \right\}. \quad (\text{D } 66)$$

Additionally, the  $\partial \mathbf{q} / \partial t$  term can be easily precalculated as

$$\left( \frac{\partial}{\partial t} \mathbf{q} \right) : \hat{\mathbf{b}}\hat{\mathbf{b}} = \frac{\partial}{\partial t} (q_{\parallel} \hat{\mathbf{b}} + \mathbf{S}_{\perp}^{\parallel}) - \mathbf{q} : \frac{\partial}{\partial t} (\hat{\mathbf{b}}\hat{\mathbf{b}}) \stackrel{\text{lin}}{=} \frac{\partial q_{\parallel}}{\partial t} \hat{\mathbf{b}} + \frac{\partial \mathbf{S}_{\perp}^{\parallel}}{\partial t}; \quad (\text{D } 67)$$

$$\left( \frac{\partial}{\partial t} \mathbf{q} \right) : (\mathbf{I} - \hat{\mathbf{b}}\hat{\mathbf{b}})/2 = \frac{\partial}{\partial t} (q_{\perp} \hat{\mathbf{b}} + \mathbf{S}_{\perp}^{\perp}) + \mathbf{q} : \frac{\partial}{\partial t} (\hat{\mathbf{b}}\hat{\mathbf{b}}/2) \stackrel{\text{lin}}{=} \frac{\partial q_{\perp}}{\partial t} \hat{\mathbf{b}} + \frac{\partial \mathbf{S}_{\perp}^{\perp}}{\partial t}, \quad (\text{D } 68)$$

and by applying  $\hat{\mathbf{b}} \times$ , terms proportional to  $\hat{\mathbf{b}}$  disappear, yielding

$$\mathbf{S}_{\perp}^{\parallel} \stackrel{\text{lin}}{=} \frac{1}{\Omega} \hat{\mathbf{b}} \times \left\{ \frac{\partial}{\partial t} \mathbf{S}_{\perp}^{\parallel} + \hat{\mathbf{b}}\hat{\mathbf{b}} : \left[ \nabla \cdot \mathbf{r} - \frac{1}{\rho} (\mathbf{p}(\nabla \cdot \mathbf{p}))^S \right] \right\}; \quad (\text{D } 69)$$



$$\mathbf{S}_{\perp}^{\perp} \stackrel{\text{lin}}{=} \frac{1}{\Omega} \hat{\mathbf{b}} \times \left\{ \frac{\partial}{\partial t} \mathbf{S}_{\perp}^{\perp} + (\mathbf{I} - \hat{\mathbf{b}}\hat{\mathbf{b}})/2 : \left[ \nabla \cdot \mathbf{r} - \frac{1}{\rho} (\mathbf{p}(\nabla \cdot \mathbf{p}))^S \right] \right\}. \quad (\text{D } 70)$$

The first-order vectors were obtained by keeping only gyrotropic quantities on the right-hand side. Linearizing results (D 37), (D 60) indeed yields that at the linear level

$$\mathbf{S}_{\perp}^{\parallel(1)} \stackrel{\text{lin}}{=} \frac{1}{\Omega} \hat{\mathbf{b}}_0 \times \left[ p_{\perp}^{(0)} \nabla \left( \frac{p_{\parallel}}{\rho} \right) + 2 \frac{p_{\parallel}^{(0)}}{\rho_0} (p_{\parallel}^{(0)} - p_{\perp}^{(0)}) \partial_z \hat{\mathbf{b}} \right]; \quad (\text{D } 71)$$

$$\mathbf{S}_{\perp}^{\perp(1)} \stackrel{\text{lin}}{=} \frac{1}{\Omega} \hat{\mathbf{b}}_0 \times \left[ 2 p_{\perp}^{(0)} \nabla \left( \frac{p_{\perp}}{\rho} \right) \right], \quad (\text{D } 72)$$

where the gradients are meant to be further linearized, and where we also neglected perturbations  $\tilde{\mathbf{r}}$ , since right now we do not want to calculate these quantities from linear kinetic theory. To obtain the second-order heat flux vectors at the linear level, we need to calculate

$$\mathbf{S}_{\perp}^{\parallel(2)} \stackrel{\text{lin}}{=} \frac{1}{\Omega} \hat{\mathbf{b}} \times \left\{ \frac{\partial}{\partial t} \mathbf{S}_{\perp}^{\parallel(1)} + \hat{\mathbf{b}}\hat{\mathbf{b}} : \left[ \nabla \cdot \mathbf{r}^{\text{ng}} - \frac{1}{\rho} (\mathbf{p}^{\text{g}}(\nabla \cdot \mathbf{\Pi}))^S \right] \right\}; \quad (\text{D } 73)$$

$$\mathbf{S}_{\perp}^{\perp(2)} \stackrel{\text{lin}}{=} \frac{1}{\Omega} \hat{\mathbf{b}} \times \left\{ \frac{\partial}{\partial t} \mathbf{S}_{\perp}^{\perp(1)} + (\mathbf{I} - \hat{\mathbf{b}}\hat{\mathbf{b}})/2 : \left[ \nabla \cdot \mathbf{r}^{\text{ng}} - \frac{1}{\rho} (\mathbf{p}^{\text{g}}(\nabla \cdot \mathbf{\Pi}))^S \right] \right\}. \quad (\text{D } 74)$$

To calculate the  $\mathbf{r}^{\text{ng}}$  contributions it is enough to work with the partially linearized (valid only for bi-Maxwellian)

$$\partial_l \mathbf{r}_{ijkl}^{\text{ng}} \stackrel{\text{lin}}{=} \frac{1}{\rho_0} [p_{ij}^{\text{g}} \partial_l \Pi_{kl} + p_{ik}^{\text{g}} \partial_l \Pi_{jl} + p_{il}^{\text{g}} \partial_l \Pi_{jk} + p_{kl}^{\text{g}} \partial_l \Pi_{ij} + p_{jl}^{\text{g}} \partial_l \Pi_{ik} + p_{jk}^{\text{g}} \partial_l \Pi_{il}], \quad (\text{D } 75)$$

which further yields

$$(\partial_l \mathbf{r}_{ijkl}^{\text{ng}}) \hat{b}_i \hat{b}_j \stackrel{\text{lin}}{=} \frac{1}{\rho_0} [p_{\parallel} \partial_l \Pi_{kl} + 2 p_{\parallel} \hat{b}_k (\partial_l \Pi_{il}) \hat{b}_i + 2 p_{\parallel} \hat{b}_l (\partial_l \Pi_{ik}) \hat{b}_i + p_{kl}^{\text{g}} (\partial_l \Pi_{ij}) \hat{b}_i \hat{b}_j]; \quad (\text{D } 76)$$

$$(\partial_l \mathbf{r}_{ijkl}^{\text{ng}}) \delta_{ij} \stackrel{\text{lin}}{=} \frac{1}{\rho_0} [(p_{\parallel} + 6 p_{\perp}) \partial_l \Pi_{kl} + 2(p_{\parallel} - p_{\perp}) \hat{b}_k (\partial_l \Pi_{il}) \hat{b}_i + 2(p_{\parallel} - p_{\perp}) \hat{b}_l (\partial_l \Pi_{ik}) \hat{b}_i], \quad (\text{D } 77)$$

and further evaluation with  $\hat{\mathbf{b}}_0$  (e.g.  $\hat{b}_l \partial_l \rightarrow \partial_z$ , or  $\Pi_{ik} \hat{b}_i \rightarrow \Pi_{zk}$ , but keeping the  $\hat{b}_k$  intact for now) yields

$$[\hat{\mathbf{b}}\hat{\mathbf{b}} : (\nabla \cdot \mathbf{r}^{\text{ng}})]_k \stackrel{\text{lin}}{=} \frac{1}{\rho_0} [p_{\parallel} \partial_l \Pi_{kl} + 2 p_{\parallel} \hat{b}_k \partial_l \Pi_{zl} + 2 p_{\parallel} \partial_z \Pi_{zk}]; \quad (\text{D } 78)$$

$$[\mathbf{I} : (\nabla \cdot \mathbf{r}^{\text{ng}})]_k \stackrel{\text{lin}}{=} \frac{1}{\rho_0} [(p_{\parallel} + 6 p_{\perp}) \partial_l \Pi_{kl} + 2(p_{\parallel} - p_{\perp}) \hat{b}_k \partial_l \Pi_{zl} + 2(p_{\parallel} - p_{\perp}) \partial_z \Pi_{zk}]; \quad (\text{D } 79)$$

$$[(\mathbf{I} - \hat{\mathbf{b}}\hat{\mathbf{b}})/2 : (\nabla \cdot \mathbf{r}^{\text{ng}})]_k \stackrel{\text{lin}}{=} \frac{1}{\rho_0} [3 p_{\perp} \partial_l \Pi_{kl} - p_{\perp} \hat{b}_k \partial_l \Pi_{zl} - p_{\perp} \partial_z \Pi_{zk}]. \quad (\text{D } 80)$$

We purposely kept the  $\hat{b}_k$  for now, since multiplying the results by  $\hat{b}_k$  and further linearizing yields

$$\hat{\mathbf{b}} \cdot [\hat{\mathbf{b}}\hat{\mathbf{b}} : (\nabla \cdot \mathbf{r}^{\text{ng}})] \stackrel{\text{lin}}{=} 3 \frac{p_{\parallel}^{(0)}}{\rho_0} \partial_l \Pi_{zl}; \quad (\text{D } 81)$$

$$\hat{\mathbf{b}} \cdot [(\mathbf{I} - \hat{\mathbf{b}}\hat{\mathbf{b}})/2 : (\nabla \cdot \mathbf{r}^{\text{ng}})] \stackrel{\text{lin}}{=} 2 \frac{P_{\perp}^{(0)}}{\rho_0} \partial_l \Pi_{zl}, \quad (\text{D } 82)$$

which verifies that the contributions (7.100), (7.101) to the gyrotropic heat flux equations were indeed calculated correctly. In contrast, applying  $\hat{\mathbf{b}} \times$  eliminates terms proportional to  $\hat{\mathbf{b}}_k$ , implying

$$\hat{\mathbf{b}} \times [\hat{\mathbf{b}}\hat{\mathbf{b}} : (\nabla \cdot \mathbf{r}^{\text{ng}})]_k \stackrel{\text{lin}}{=} \hat{\mathbf{b}}_0 \times \frac{P_{\parallel}^{(0)}}{\rho_0} [\partial_l \Pi_{kl} + 2\partial_z \Pi_{kz}]; \quad (\text{D } 83)$$

$$\hat{\mathbf{b}} \times [(\mathbf{I} - \hat{\mathbf{b}}\hat{\mathbf{b}})/2 : (\nabla \cdot \mathbf{r}^{\text{ng}})]_k \stackrel{\text{lin}}{=} \hat{\mathbf{b}}_0 \times \frac{P_{\perp}^{(0)}}{\rho_0} [3\partial_l \Pi_{kl} - \partial_z \Pi_{kz}], \quad (\text{D } 84)$$

where we have used perhaps a somewhat strange mixture of vector and index notations, but which should be otherwise clear. For any vector  $\mathbf{a}$ , the notation just means  $\hat{\mathbf{b}}_0 \times \mathbf{a} = (-a_y, a_x, 0)$ . Similarly straightforward calculation yields

$$\hat{\mathbf{b}} \times [\hat{\mathbf{b}}\hat{\mathbf{b}} : (\mathbf{p}^g \nabla \cdot \Pi)]_k \stackrel{\text{lin}}{=} \hat{\mathbf{b}}_0 \times p_{\parallel}^{(0)} \partial_l \Pi_{lk}; \quad (\text{D } 85)$$

$$\hat{\mathbf{b}} \times [\mathbf{I} : (\mathbf{p}^g \nabla \cdot \Pi)]_k \stackrel{\text{lin}}{=} \hat{\mathbf{b}}_0 \times (p_{\parallel}^{(0)} + 4p_{\perp}^{(0)}) \partial_l \Pi_{lk}; \quad (\text{D } 86)$$

$$\hat{\mathbf{b}} \times [(\mathbf{I} - \hat{\mathbf{b}}\hat{\mathbf{b}})/2 : (\mathbf{p}^g \nabla \cdot \Pi)]_k \stackrel{\text{lin}}{=} \hat{\mathbf{b}}_0 \times 2p_{\perp}^{(0)} \partial_l \Pi_{lk}. \quad (\text{D } 87)$$

The final form of the second-order non-gyrotropic heat flux vectors reads

$$(\mathbf{S}_{\perp}^{\parallel})_k^{(2)} \stackrel{\text{lin}}{=} \frac{1}{\Omega} \hat{\mathbf{b}}_0 \times \left[ \frac{\partial}{\partial t} (\mathbf{S}_{\perp}^{\parallel})_k^{(1)} + 2 \frac{P_{\parallel}^{(0)}}{\rho_0} \partial_z \Pi_{kz} \right]; \quad (\text{D } 88)$$

$$(\mathbf{S}_{\perp}^{\perp})_k^{(2)} \stackrel{\text{lin}}{=} \frac{1}{\Omega} \hat{\mathbf{b}}_0 \times \left[ \frac{\partial}{\partial t} (\mathbf{S}_{\perp}^{\perp})_k^{(1)} + \frac{P_{\perp}^{(0)}}{\rho_0} (\partial_l \Pi_{lk} - \partial_z \Pi_{kz}) \right]. \quad (\text{D } 89)$$

Alternatively, one can write  $\partial_l \Pi_{lk} - \partial_z \Pi_{kz} = \partial_x \Pi_{xk} + \partial_y \Pi_{yk}$ . The second-order heat flux vectors contribute to the scalar (gyrotropic) pressure equations, and also to the FLR pressure tensor  $\Pi$ . The non-gyrotropic pressure tensor  $\Pi$ , and the non-gyrotropic heat flux vectors  $\mathbf{S}_{\perp}^{\parallel}$ ,  $\mathbf{S}_{\perp}^{\perp}$  are therefore generally coupled. To make the system easily solvable, terms with  $\Pi$  on the right-hand side of (D 88), (D 89) are approximated by  $\Pi^{(1)}$ . The equations (D 88), (D 89) can be written in a more elegant vector form, see equations (5.153), (5.154), and the result is equivalent to equations (53), (54) of Passot *et al.* (2012).

Instead of splitting of  $\Pi$ ,  $\mathbf{S}_{\perp}^{\parallel}$  and  $\mathbf{S}_{\perp}^{\perp}$  into the first- and second-order components, one might be interested in the future to check the behaviour of the coupled system, i.e. one might be interested in solving the pressure tensor equation, that is coupled to the heat flux vectors, by making the associated variables independent quantities. For example, one might be interested to double check the solutions for the parallel firehose instability. The procedure of how to obtain time-dependent equations for  $\mathbf{S}_{\perp}^{\parallel}$  and  $\mathbf{S}_{\perp}^{\perp}$  can be perhaps a bit blurry from the steps outlined above, so let us write it one more time, only for  $\mathbf{S}_{\perp}^{\parallel}$ . One can work directly with the heat flux tensor equation (7.33), and keeping only terms that contribute at the linear level, the partially linearized heat flux tensor equation reads

$$\frac{\partial}{\partial t} \mathbf{q} + \Omega (\hat{\mathbf{b}} \times \mathbf{q}^{\text{ng}})^S + \nabla \cdot \mathbf{r} - \frac{1}{\rho_0} (\mathbf{p}(\nabla \cdot \mathbf{p}))^S = 0. \quad (\text{D } 90)$$

Applying  $\hat{\mathbf{b}}\hat{\mathbf{b}}$  and  $\hat{\mathbf{b}}\times$ , yields that, at the linear level,

$$\hat{\mathbf{b}}_0 \times \frac{\partial}{\partial t} \mathbf{S}_{\perp}^{\parallel} - \Omega \mathbf{S}_{\perp}^{\parallel} + \hat{\mathbf{b}}_0 \times \left[ p_{\perp}^{(0)} \nabla \left( \frac{p_{\parallel}}{\rho} \right) + 2 \frac{p_{\parallel}^{(0)}}{\rho_0} (p_{\parallel}^{(0)} - p_{\perp}^{(0)}) \partial_z \hat{\mathbf{b}} \right] + \hat{\mathbf{b}}_0 \times 2 \frac{p_{\parallel}^{(0)}}{\rho_0} \partial_z \vec{\Pi}_z = 0, \quad (\text{D } 91)$$

or in the index notation

$$\frac{\partial}{\partial t} (S_{\perp}^{\parallel})_x - \Omega (S_{\perp}^{\parallel})_y + p_{\perp}^{(0)} \partial_x \left( \frac{p_{\parallel}}{\rho} \right) + 2 \frac{p_{\parallel}^{(0)}}{\rho_0} (p_{\parallel}^{(0)} - p_{\perp}^{(0)}) \partial_z \hat{b}_x + 2 \frac{p_{\parallel}^{(0)}}{\rho_0} \partial_z \Pi_{xz} = 0; \quad (\text{D } 92)$$

$$\frac{\partial}{\partial t} (S_{\perp}^{\parallel})_y + \Omega (S_{\perp}^{\parallel})_x + p_{\perp}^{(0)} \partial_y \left( \frac{p_{\parallel}}{\rho} \right) + 2 \frac{p_{\parallel}^{(0)}}{\rho_0} (p_{\parallel}^{(0)} - p_{\perp}^{(0)}) \partial_z \hat{b}_y + 2 \frac{p_{\parallel}^{(0)}}{\rho_0} \partial_z \Pi_{yz} = 0. \quad (\text{D } 93)$$

Importantly, bi-Maxwellian distribution function was assumed. Normalizing the equations, writing them in the  $x$ - $z$  plane and Fourier transforming yields

$$-i\omega (S_{\perp}^{\parallel})_x - (S_{\perp}^{\parallel})_y + \frac{\beta_{\parallel}}{2} a_p i k_{\perp} (p_{\parallel} - \rho) + \beta_{\parallel} (1 - a_p) i k_{\parallel} B_x + \beta_{\parallel} i k_{\parallel} \Pi_{xz} = 0; \quad (\text{D } 94)$$

$$-i\omega (S_{\perp}^{\parallel})_y + (S_{\perp}^{\parallel})_x + \beta_{\parallel} (1 - a_p) i k_{\parallel} B_y + \beta_{\parallel} i k_{\parallel} \Pi_{yz} = 0. \quad (\text{D } 95)$$

Considering parallel propagation ( $k_{\perp} = 0$ ), these equations are coupled with equations for  $\Pi_{xz}$ ,  $\Pi_{yz}$ , equations (5.157), (5.158), and also with equations for  $u_x$ ,  $u_y$ ,  $B_x$ ,  $B_y$ , system (6.1). Now, it should be easy to derive the dispersion relation and, for example, check if the solutions for the parallel firehose instability are improved.

### D.5. Contributions of $\mathbf{q}^{\text{ng}}$ to various equations

Here we calculate contributions of  $\mathbf{q}^{\text{ng}}$  to the FLR pressure tensor, scalar heat flux equation and scalar pressure equations. Contributions to the FLR pressure tensor are calculated at the linear level by neglecting  $\sigma$ . Contributions to scalar heat flux equations and scalar pressure equations are calculated nonlinearly. In these calculations, no specific form of a distribution function is assumed.

#### D.5.1. Contributions to FLR pressure tensor

For contributions to the FLR tensor  $\Pi$ , we need to calculate  $\nabla \cdot \mathbf{q}^{\text{ng}}$ . To calculate it, we will neglect the heat flux tensor  $\sigma$ , and use the decomposition of the heat flux tensor  $\mathbf{S}$ , that can be figured out to be

$$\mathbf{S} = q_{\parallel} \hat{\mathbf{b}}\hat{\mathbf{b}} + q_{\perp} [\hat{\mathbf{b}}(\mathbf{I} - \hat{\mathbf{b}}\hat{\mathbf{b}})]^S + [S_{\perp}^{\parallel} \hat{\mathbf{b}}\hat{\mathbf{b}}]^S + \frac{1}{2} [S_{\perp}^{\perp} (\mathbf{I} - \hat{\mathbf{b}}\hat{\mathbf{b}})]^S, \quad (\text{D } 96)$$

or in the index notation

$$\begin{aligned} S_{ijk} = & q_{\parallel} \hat{b}_i \hat{b}_j \hat{b}_k + q_{\perp} [\hat{b}_i \delta_{jk} + \hat{b}_j \delta_{ik} + \hat{b}_k \delta_{ij} - 3 \hat{b}_i \hat{b}_j \hat{b}_k] \\ & + (S_{\perp}^{\parallel})_i \hat{b}_j \hat{b}_k + (S_{\perp}^{\parallel})_j \hat{b}_i \hat{b}_k + (S_{\perp}^{\parallel})_k \hat{b}_i \hat{b}_j \\ & + \frac{1}{2} [(S_{\perp}^{\perp})_i (\delta_{jk} - \hat{b}_j \hat{b}_k) + (S_{\perp}^{\perp})_j (\delta_{ik} - \hat{b}_i \hat{b}_k) + (S_{\perp}^{\perp})_k (\delta_{ij} - \hat{b}_i \hat{b}_j)]. \end{aligned} \quad (\text{D } 97)$$

Here, we want to calculate only contributions at the linear level, and linearizing  $\nabla \cdot \mathbf{S}$  yields

$$\begin{aligned} \partial_k S_{ijk} \stackrel{\text{lin}}{=} & \hat{b}_i \hat{b}_j \partial_z q_{\parallel} + \hat{b}_i \partial_j q_{\perp} + \hat{b}_j \partial_i q_{\perp} + \delta_{ij} \partial_z q_{\perp} - 3 \hat{b}_i \hat{b}_j \partial_z q_{\perp} \\ & + \hat{b}_j \partial_z (S_{\perp}^{\parallel})_i + \hat{b}_i \partial_z (S_{\perp}^{\parallel})_j + \hat{b}_i \hat{b}_j \partial_k (S_{\perp}^{\parallel})_k \\ & + \frac{1}{2} [\partial_j (S_{\perp}^{\perp})_i - \hat{b}_j \partial_z (S_{\perp}^{\perp})_i + \partial_i (S_{\perp}^{\perp})_j - \hat{b}_i \partial_z (S_{\perp}^{\perp})_j + \delta_{ij} \partial_k (S_{\perp}^{\perp})_k - \hat{b}_i \hat{b}_j \partial_k (S_{\perp}^{\perp})_k], \end{aligned} \quad (\text{D } 98)$$

and straightforward evaluation gives

$$(\nabla \cdot \mathbf{S})_{xx} = \partial_z q_{\perp} + \frac{1}{2} [3 \partial_x (S_{\perp}^{\perp})_x + \partial_y (S_{\perp}^{\perp})_y]; \quad (\text{D } 99)$$

$$(\nabla \cdot \mathbf{S})_{xy} = \frac{1}{2} [\partial_y (S_{\perp}^{\perp})_x + \partial_x (S_{\perp}^{\perp})_y]; \quad (\text{D } 100)$$

$$(\nabla \cdot \mathbf{S})_{xz} = \partial_x q_{\perp} + \partial_z (S_{\perp}^{\parallel})_x; \quad (\text{D } 101)$$

$$(\nabla \cdot \mathbf{S})_{yy} = \partial_z q_{\perp} + \frac{1}{2} [\partial_x (S_{\perp}^{\perp})_x + 3 \partial_y (S_{\perp}^{\perp})_y]; \quad (\text{D } 102)$$

$$(\nabla \cdot \mathbf{S})_{yz} = \partial_y q_{\perp} + \partial_z (S_{\perp}^{\parallel})_y; \quad (\text{D } 103)$$

$$(\nabla \cdot \mathbf{S})_{zz} = \partial_z q_{\parallel} + \partial_x (S_{\perp}^{\parallel})_x + \partial_y (S_{\perp}^{\parallel})_y. \quad (\text{D } 104)$$

#### D.5.2. Contributions to scalar heat flux equations

When we calculated the scalar heat flux equations, we separated the non-gyrotropic heat flux  $\mathbf{q}^{\text{ng}}$  by defining a quantity

$$\mathbf{Q}^{\text{ng}} = \frac{d}{dt} \mathbf{q}^{\text{ng}} + \mathbf{q}^{\text{ng}} \nabla \cdot \mathbf{u} + \left[ \mathbf{q}^{\text{ng}} \cdot \nabla \mathbf{u} + \Omega \frac{|\mathbf{B}|}{B_0} \hat{\mathbf{b}} \times \mathbf{q}^{\text{ng}} \right]^S,$$

and we need to calculate

$$Q_{\parallel}^{\text{ng}} \equiv (\mathbf{Q}^{\text{ng}} : \hat{\mathbf{b}} \hat{\mathbf{b}}) \cdot \hat{\mathbf{b}}; \quad Q_{\perp}^{\text{ng}} \equiv (\mathbf{Q}^{\text{ng}} : (\mathbf{I} - \hat{\mathbf{b}} \hat{\mathbf{b}})/2) \cdot \hat{\mathbf{b}}. \quad (\text{D } 105)$$

Here, we want to obtain exact nonlinear expressions. Direct calculations yield

$$\left( \frac{d}{dt} q_{ijk}^{\text{ng}} \right) \hat{b}_i \hat{b}_j \hat{b}_k = \frac{d}{dt} \underbrace{(q_{ijk}^{\text{ng}} \hat{b}_i \hat{b}_j \hat{b}_k)}_{=0} - q_{ijk}^{\text{ng}} \frac{d}{dt} (\hat{b}_i \hat{b}_j \hat{b}_k) = -3 S_{\perp}^{\parallel} \cdot \frac{d \hat{\mathbf{b}}}{dt}; \quad (\text{D } 106)$$

$$(\mathbf{q}^{\text{ng}} \cdot \nabla \mathbf{u})_{ijk}^S \hat{b}_i \hat{b}_j \hat{b}_k = 3 S_{\perp}^{\parallel} \cdot (\nabla \mathbf{u}) \cdot \hat{\mathbf{b}}; \quad (\text{D } 107)$$

$$(\hat{\mathbf{b}} \times \mathbf{q}^{\text{ng}})_{ijk}^S \hat{b}_i \hat{b}_j \hat{b}_k = 0, \quad (\text{D } 108)$$

and the final fully nonlinear result reads

$$Q_{\parallel}^{\text{ng}} = -3 S_{\perp}^{\parallel} \cdot \frac{d \hat{\mathbf{b}}}{dt} + 3 S_{\perp}^{\parallel} \cdot (\nabla \mathbf{u}) \cdot \hat{\mathbf{b}}. \quad (\text{D } 109)$$

Similar calculations yield a nonlinear expression for  $Q_{\perp}^{\text{ng}}$ , which calculates as

$$\left( \frac{d}{dt} q_{ijk}^{\text{ng}} \right) \delta_{ij} \hat{b}_k = \frac{d}{dt} \underbrace{(q_{ijk}^{\text{ng}} \delta_{ij} \hat{b}_k)}_{=0} - q_{ijk}^{\text{ng}} \delta_{ij} \frac{d}{dt} \hat{b}_k = -(\mathbf{q}^{\text{ng}} : \mathbf{I}) \cdot \frac{d \hat{\mathbf{b}}}{dt} = -(S_{\perp}^{\parallel} + 2 S_{\perp}^{\perp}) \cdot \frac{d \hat{\mathbf{b}}}{dt}; \quad (\text{D } 110)$$

$$\left(\frac{d}{dt}q_{ijk}^{\text{ng}}\right)(\delta_{ij} - \hat{b}_i\hat{b}_j)\hat{b}_k/2 = (S_{\perp}^{\parallel} - S_{\perp}^{\perp}) \cdot \frac{d\hat{\mathbf{b}}}{dt}; \quad (\text{D } 111)$$

$$(\hat{\mathbf{b}} \times \mathbf{q}^{\text{ng}})_{ijk}^S \delta_{ij} \hat{b}_k = 0; \quad (\text{D } 112)$$

$$\begin{aligned} (\mathbf{q}^{\text{ng}} \cdot \nabla \mathbf{u})_{ijk}^S \delta_{ij} \hat{b}_k &= (q_{ijl}^{\text{ng}} \partial_l u_k + q_{jkl}^{\text{ng}} \partial_l u_i + q_{kil}^{\text{ng}} \partial_l u_j) \delta_{ij} \hat{b}_k = (\mathbf{q}^{\text{ng}} : \mathbf{I})_l (\partial_l u_k) \hat{b}_k + 2q_{ilk}^{\text{ng}} \hat{b}_k \partial_l u_i \\ &= (S_{\perp}^{\parallel} + 2S_{\perp}^{\perp}) \cdot (\nabla \mathbf{u}) \cdot \hat{\mathbf{b}} + 2(\mathbf{q}^{\text{ng}} \cdot \hat{\mathbf{b}}) : \nabla \mathbf{u}. \end{aligned} \quad (\text{D } 113)$$

To continue the calculation, we need to use (an exact expression)

$$\mathbf{q}^{\text{ng}} = [S_{\perp}^{\parallel} \hat{\mathbf{b}} \hat{\mathbf{b}}]^S + \frac{1}{2} [S_{\perp}^{\perp} (\mathbf{I} - \hat{\mathbf{b}} \hat{\mathbf{b}})]^S + \boldsymbol{\sigma}, \quad (\text{D } 114)$$

which allows us to calculate

$$q_{ijk}^{\text{ng}} \hat{b}_k = (S_{\perp}^{\parallel})_i \hat{b}_j + (S_{\perp}^{\perp})_j \hat{b}_i + \sigma_{ijk} \hat{b}_k; \quad (\text{D } 115)$$

$$\begin{aligned} q_{ilk}^{\text{ng}} \hat{b}_k \partial_l u_i &= (S_{\perp}^{\parallel})_i \hat{b}_l \partial_l u_i + (S_{\perp}^{\perp})_l (\partial_l u_i) \hat{b}_i + \sigma_{ilk} \hat{b}_k \partial_l u_i \\ &= S_{\perp}^{\parallel} \cdot (\hat{\mathbf{b}} \cdot \nabla \mathbf{u}) + S_{\perp}^{\perp} \cdot (\nabla \mathbf{u}) \cdot \hat{\mathbf{b}} + (\boldsymbol{\sigma} \cdot \hat{\mathbf{b}}) : \nabla \mathbf{u}, \end{aligned} \quad (\text{D } 116)$$

and therefore

$$(\mathbf{q}^{\text{ng}} \cdot \nabla \mathbf{u})_{ijk}^S \delta_{ij} \hat{b}_k = (3S_{\perp}^{\parallel} + 2S_{\perp}^{\perp}) \cdot (\nabla \mathbf{u}) \cdot \hat{\mathbf{b}} + 2S_{\perp}^{\perp} \cdot (\hat{\mathbf{b}} \cdot \nabla \mathbf{u}) + 2(\boldsymbol{\sigma} \cdot \hat{\mathbf{b}}) : \nabla \mathbf{u}; \quad (\text{D } 117)$$

$$(\mathbf{q}^{\text{ng}} \cdot \nabla \mathbf{u})_{ijk}^S (\delta_{ij} - \hat{b}_i \hat{b}_j) \hat{b}_k / 2 = S_{\perp}^{\perp} \cdot (\nabla \mathbf{u}) \cdot \hat{\mathbf{b}} + S_{\perp}^{\parallel} \cdot (\hat{\mathbf{b}} \cdot \nabla \mathbf{u}) + \hat{\mathbf{b}} \cdot \boldsymbol{\sigma} : \nabla \mathbf{u}. \quad (\text{D } 118)$$

The final fully nonlinear result reads

$$\mathcal{Q}_{\perp}^{\text{ng}} = (S_{\perp}^{\parallel} - S_{\perp}^{\perp}) \cdot \frac{d\hat{\mathbf{b}}}{dt} + S_{\perp}^{\perp} \cdot (\nabla \mathbf{u}) \cdot \hat{\mathbf{b}} + (\hat{\mathbf{b}} \cdot \nabla \mathbf{u}) \cdot S_{\perp}^{\parallel} + \hat{\mathbf{b}} \cdot \boldsymbol{\sigma} : \nabla \mathbf{u}. \quad (\text{D } 119)$$

However, it is easy to see that, at the linear level, both  $\mathcal{Q}_{\parallel}^{\text{ng}} \stackrel{\text{lin}}{=} 0$  and  $\mathcal{Q}_{\perp}^{\text{ng}} \stackrel{\text{lin}}{=} 0$ .

### D.5.3. Contributions to scalar pressure equations

The non-gyrotropic heat flux term entering the parallel pressure equation calculates as

$$\begin{aligned} (\nabla \cdot \mathbf{q}^{\text{ng}}) : \hat{\mathbf{b}} \hat{\mathbf{b}} &= (\partial_k q_{kij}^{\text{ng}}) \hat{b}_i \hat{b}_j = \partial_k (q_{kij}^{\text{ng}} \hat{b}_i \hat{b}_j) - q_{kij}^{\text{ng}} \partial_k (\hat{b}_i \hat{b}_j) \\ &= \partial_k (S_{\perp}^{\parallel})_k - 2q_{kij}^{\text{ng}} \hat{b}_i \partial_k \hat{b}_j, \end{aligned} \quad (\text{D } 120)$$

and since

$$q_{ijk}^{\text{ng}} \hat{b}_i = (S_{\perp}^{\parallel})_j \hat{b}_k + (S_{\perp}^{\perp})_k \hat{b}_j + \sigma_{ijk} \hat{b}_i; \quad (\text{D } 121)$$

$$q_{ijk}^{\text{ng}} \hat{b}_i \partial_k \hat{b}_j = (S_{\perp}^{\parallel})_j \hat{b}_k \partial_k \hat{b}_j + (S_{\perp}^{\perp})_k \underbrace{\hat{b}_j \partial_k \hat{b}_j}_{=0} + \sigma_{ijk} \hat{b}_i \partial_k \hat{b}_j = (\hat{\mathbf{b}} \cdot \nabla \hat{\mathbf{b}}) \cdot S_{\perp}^{\parallel} + \hat{\mathbf{b}} \cdot \boldsymbol{\sigma} : \nabla \hat{\mathbf{b}}, \quad (\text{D } 122)$$

which yields the final nonlinear result

$$(\nabla \cdot \mathbf{q}^{\text{ng}}) : \hat{\mathbf{b}} \hat{\mathbf{b}} = \nabla \cdot S_{\perp}^{\parallel} - 2(\hat{\mathbf{b}} \cdot \nabla \hat{\mathbf{b}}) \cdot S_{\perp}^{\parallel} - 2\hat{\mathbf{b}} \cdot \boldsymbol{\sigma} : \nabla \hat{\mathbf{b}}. \quad (\text{D } 123)$$

Similarly, the term entering the perpendicular pressure equation calculates as

$$(\nabla \cdot \mathbf{q}^{\text{ng}}) : \mathbf{I} = (\partial_k q_{kij}^{\text{ng}}) \delta_{ij} = \nabla \cdot (\mathbf{q}^{\text{ng}} : \mathbf{I}) = \nabla \cdot (\mathbf{S}_{\perp}^{\parallel} + 2\mathbf{S}_{\perp}^{\perp}); \quad (\text{D } 124)$$

$$(\nabla \cdot \mathbf{q}^{\text{ng}}) : (\mathbf{I} - \hat{\mathbf{b}}\hat{\mathbf{b}})/2 = \nabla \cdot \mathbf{S}_{\perp}^{\perp} + (\hat{\mathbf{b}} \cdot \nabla \hat{\mathbf{b}}) \cdot \mathbf{S}_{\perp}^{\parallel} + \hat{\mathbf{b}} \cdot \boldsymbol{\sigma} : \nabla \hat{\mathbf{b}}. \quad (\text{D } 125)$$

The nonlinear results (D 124), (D 126) are exact, and here we do not address how to further decompose the tensor  $\boldsymbol{\sigma}$ . Finally, considering contributions at the linear level (assuming that the mean values of the entire heat flux  $\mathbf{q}^{(0)} = 0$ ), it is easy to see that

$$(\nabla \cdot \mathbf{q}^{\text{ng}}) : \hat{\mathbf{b}}\hat{\mathbf{b}} \stackrel{\text{lin}}{=} \nabla \cdot \mathbf{S}_{\perp}^{\parallel} = \partial_x (S_{\perp}^{\parallel})_x + \partial_y (S_{\perp}^{\parallel})_y; \quad (\text{D } 126)$$

$$(\nabla \cdot \mathbf{q}^{\text{ng}}) : (\mathbf{I} - \hat{\mathbf{b}}\hat{\mathbf{b}})/2 \stackrel{\text{lin}}{=} \nabla \cdot \mathbf{S}_{\perp}^{\perp} = \partial_x (S_{\perp}^{\perp})_x + \partial_y (S_{\perp}^{\perp})_y. \quad (\text{D } 127)$$

The heat flux tensor  $\boldsymbol{\sigma}$  therefore does not contribute to the gyrotropic pressure equations at the linear level.

## REFERENCES

- ABRAHAM-SHRAUNER, B. 1967 Propagation of hydromagnetic waves through an anisotropic plasma. *J. Plasma Phys.* **1**, 361–378.
- ABRAHAM-SHRAUNER, B. W. 1973 Small amplitude hydromagnetic waves for a plasma with a generalized polytropic law. *Plasma Phys.* **15**, 375.
- BALE, S. D., KASPER, J. C., HOWES, G. G., QUATAERT, E., SALEM, C. & SUNDKVIST, D. 2009 Magnetic fluctuation power near proton temperature anisotropy instability thresholds in the solar wind. *Phys. Rev. Lett.* **103**, 211101.
- BARAKAT, A. R. & SCHUNK, R. W. 1982 Transport equations for multicomponent anisotropic space plasmas – a review. *Plasma Phys.* **24**, 389–418.
- BASU, B. 2009 Hydromagnetic waves and instabilities in kappa distribution plasma. *Phys. Plasmas* **16** (5), 052106.
- BASU, B. & COPPI, B. 1984 Theory of field-swelling instability in anisotropic plasmas. *Phys. Fluids* **27**, 1187–1193.
- BELMONT, G. & MAZELLE, C. 1992 Polytropic indices in collisionless plasmas – theory and measurements. *J. Geophys. Res.* **97**, 8327–8336.
- BHATNAGAR, P. L., GROSS, E. P. & KROOK, M. 1954 A model for collision processes in gases. I. Small amplitude processes in charged and neutral one-component systems. *Phys. Rev.* **94**, 511–525.
- BIERMANN, L. 1950 Über den Ursprung der Magnetfelder auf Sternen und im interstellaren Raum (miteinem Anhang von A. Schlüter). *Z. Naturforsch. Teil A* **5**, 65.
- BORGOGNO, D., HELLINGER, P., PASSOT, T., SULEM, P. L. & TRÁVNÍČEK, P. M. 2009 Alfvén wave filamentation and dispersive phase mixing in a high-density channel: Landau fluid and hybrid simulations. *Nonlinear Process. Geophys.* **16**, 275–285.
- BRAGINSKII, S. I. 1958 Transport phenomena in a completely ionized two-temperature plasma. *Sov. J. Expl. Theor. Phys.* **6**, 358.
- BRAGINSKII, S. I. 1965 Transport processes in a plasma. *Rev. Plasma Phys.* **1**, 205.
- BRUNO, R. & CARBONE, V. 2013 The solar wind as a turbulence laboratory. *Living Rev. Solar Phys.* **10**, 2.
- CAMPOREALE, E. & BURGESS, D. 2017 Comparison of linear modes in kinetic plasma models. *J. Plasma Phys.* **83**, 535830201.
- CHANDRASEKHAR, S., KAUFMAN, A. N. & WATSON, K. M. 1958 The Stability of the Pinch. *Proc. R. Soc. Lond. Ser. A* **245**, 435–455.
- CHEW, G. F., GOLDBERGER, M. L. & LOW, F. E. 1956 The Boltzmann equation and the one-fluid hydromagnetic equations in the absence of particle collisions. *Proc. R. Soc. Lond. Ser. A* **236**, 112–118.

- CHUST, T. & BELMONT, G. 2006 Closure of fluid equations in collisionless magnetoplasmas. *Phys. Plasmas* **13**, 012506.
- CONSTANTINESCU, O. D. 2002 Self-consistent model of mirror structures. *J. Atmos. Sol. Terr. Phys.* **64**, 645–649.
- DEL SARTO, D., PEGORARO, F. & CALIFANO, F. 2016 Pressure anisotropy and small spatial scales induced by velocity shear. *Phys. Rev. E* **93** (5), 053203.
- DEL SARTO, D., PEGORARO, F. & TENERANI, A. 2017 Magneto-elastic waves in an anisotropic magnetised plasma. *Plasma Phys. Control. Fusion* **59** (4), 045002.
- DRAKE, J. F., SWISDAK, M. & FERMO, R. 2013 The power-law spectra of energetic particles during multi-island magnetic reconnection. *Astrophys. J. Lett.* **763**, L5.
- DZHALILOV, N. S., KUZNETSOV, V. D. & STAUDE, J. 2011 Wave instabilities of a collisionless plasma in fluid approximation. *Contrib. Plasma Phys.* **51**, 621–638.
- FERRIÈRE, K. M. & ANDRÉ, N. 2002 A mixed magnetohydrodynamic-kinetic theory of low-frequency waves and instabilities in homogeneous, gyrotropic plasmas. *J. Geophys. Res.* **107**, 1349.
- GARY, S. P. 1993 *Theory of Space Plasma Microinstabilities*. Cambridge University Press.
- GHOSH, S. & PARASHAR, T. N. 2015 Linear versus nonlinear acceleration in plasma turbulence. II. Hall-finite-Larmor-radius magnetohydrodynamics. *Phys. Plasmas* **22** (4), 042303.
- GOLDSTEIN, M. L., ROBERTS, D. A. & MATTHAEUS, W. H. 1995 Magnetohydrodynamic turbulence in the solar wind. *Annu. Rev. Astron. Astrophys.* **33**, 283–325.
- GOSWAMI, P., PASSOT, T. & SULEM, P. L. 2005 A Landau fluid model for warm collisionless plasma. *Phys. Plasmas* **12**, 102109.
- GRAD, H. 1949 On the kinetic theory of rarified gases. *Commun. Pure Appl. Maths* **2**, 325.
- GURNETT, D. A. & BHATTACHARJEE, A. 2005 *Introduction to Plasma Physics: With Space and Laboratory Applications*. Cambridge University Press.
- HAMMETT, G. W. & PERKINS, W. F. 1990 Fluid moment models for Landau damping with application to the ion-temperature-gradient instability. *Phys. Rev. Lett.* **64**, 3019–3022.
- HAU, L.-N., PHAN, T.-D., SONNERUP, B. U. Ö. & PASCHMANN, G. 1993 Double-polytropic closure in the magnetosheath. *Geophys. Res. Lett.* **20**, 2255–2258.
- HAU, L. N. & SONNERUP, B. U. Ö. 1993 On slow-mode waves in an anisotropic plasma. *Geophys. Res. Lett.* **20**, 1763–1766.
- HELLINGER, P., TRÁVNÍČEK, P. M., KASPER, J. C. & LAZARUS, A. J. 2006 Solar wind proton temperature anisotropy: linear theory and WIND/SWE observations. *Geophys. Res. Lett.* **33**, L09101.
- HOWES, G. G. 2009 Limitations of Hall MHD as a model for turbulence in weakly collisional plasmas. *Nonlin. Processes Geophys.* **16**, 219–232.
- HSU, C. T., HAZELTINE, R. D. & MORRISON, P. J. 1986 A generalized reduced fluid model with finite ion-gyroradius effects. *Phys. Fluids* **29**, 1480–1487.
- HUNANA, P., GOLDSTEIN, M. L., PASSOT, T., SULEM, P. L., LAVEDER, D. & ZANK, G. P. 2013 Polarization and compressibility of oblique kinetic Alfvén waves. *Astrophys. J.* **766**, 93.
- HUNANA, P., LAVEDER, D., PASSOT, T., SULEM, P. L. & BORGOGNO, D. 2011 Reduction of compressibility and parallel transfer by Landau damping in turbulent magnetized plasmas. *Astrophys. J.* **743**, 128.
- HUNANA, P. & ZANK, G. P. 2017 On the parallel and oblique firehose instability in fluid models. *Astrophys. J.* **839**, 13.
- HUNANA, P., ZANK, G. P., LAURENZA, M., TENERANI, A., WEBB, G. M., GOLDSTEIN, M. L., VELLI, M. & ADHIKARI, L. 2018 New closures for more precise modeling of Landau damping in the fluid framework. *Phys. Rev. Lett.* **121**, 135101.
- KHOMENKO, E., COLLADOS, M., DÍAZ, A. & VITAS, N. 2014 Fluid description of multi-component solar partially ionized plasma. *Phys. Plasmas* **21** (9), 092901.
- KHOMENKO, E., VITAS, N., COLLADOS, M. & DE VICENTE, A. 2017 Numerical simulations of quiet Sun magnetic fields seeded by the Biermann battery. *Astron. Astrophys.* **604**, A66.
- KIVELSON, M. G. & SOUTHWOOD, D. J. 1996 Mirror instability II: the mechanism of nonlinear saturation. *J. Geophys. Res.* **101**, 17365–17372.

- KOBAYASHI, S., SAHRAOUI, F., PASSOT, T., LAVEDER, D., SULEM, P. L., HUANG, S. Y., HENRI, P. & SMETS, R. 2017 Three-dimensional simulations and spacecraft observations of sub-ion scale turbulence in the solar wind: influence of Landau damping. *Astrophys. J.* **839**, 122.
- KULSRUD, R. M. 1983 MHD description of plasma. In *Handbook of Plasma Physics*, Eds. M. N. Rosenbluth and R. Z. Sagdeev. Volume 1: *Basic Plasma Physics 1* (ed. A. A. Galeev & R. N. Sudan), pp. 115–145. North-Holland.
- KULSRUD, R. M., CEN, R., OSTRICKER, J. P. & RYU, D. 1997 The protogalactic origin for cosmic magnetic fields. *Astrophys. J.* **480**, 481–491.
- LAVEDER, D., PASSOT, T. & SULEM, P. L. 2013 Fluid simulations of non-resonant anisotropic ion heating. *Ann. Geophys.* **31**, 1195–1204.
- LIVADIOTIS, G. & MCCOMAS, D. J. 2013 Understanding kappa distributions: a toolbox for space science and astrophysics. *Space Sci. Rev.* **175**, 183–214.
- MACMAHON, A. 1965 Finite gyro-radius corrections to the hydromagnetic equations for a Vlasov plasma. *Phys. Fluids* **8**, 1840–1845.
- MATTEINI, L., HELLINGER, P., GOLDSTEIN, B. E., LANDI, S., VELLI, M. & NEUGEBAUER, M. 2013 Signatures of kinetic instabilities in the solar wind. *J. Geophys. Res.* **118**, 2771–2782.
- MATTEINI, L., LANDI, S., HELLINGER, P., PANTELLINI, F., MAKSIMOVIC, M., VELLI, M., GOLDSTEIN, B. E. & MARSCH, E. 2007 Evolution of the solar wind temperature anisotropy from 0.3 to 2.5 AU. *Geophys. Res. Lett.* **34**, L20105.
- MEYRAND, R., KANEKAR, A., DORLAND, W. & SCHEKOCIHIN, A. A. 2019 Fluidization of collisionless plasma turbulence. *Proc. Natl Acad. Sci. USA* **116**, 1185–1194.
- MIKHAILOVSKII, A. B. & SMOLYAKOV, A. I. 1985 Theory of low-frequency magnetosonic solitons. *Sov. Phys. JETP* **61**, 109.
- MJØLHUS, E. 2009 Finite Larmor radius influence on MHD solitary waves. *Nonlinear Process. Geophys.* **16**, 251–264.
- NG, J., HAKIM, A., BHATTACHARJEE, A., STANIER, A. & DAUGHTON, W. 2017 Simulations of anti-parallel reconnection using a nonlocal heat flux closure. *Phys. Plasmas* **24** (8), 082112.
- ORAEVSKII, V., CHODURA, R. & FENEBERG, W. 1968 Hydrodynamic equations for plasmas in strong magnetic fields. I. Collisionless approximation. *Plasma Phys.* **10**, 819–828.
- PARKER, E. N. 1958a Dynamical instability in an anisotropic ionized gas of low density. *Phys. Rev.* **109**, 1874–1876.
- PARKER, E. N. 1958b Dynamics of the interplanetary gas and magnetic fields. *Astrophys. J.* **128**, 664–676.
- PASSOT, T., HENRI, P., LAVEDER, D. & SULEM, P.-L. 2014 Fluid simulations of ion scale plasmas with weakly distorted magnetic fields. FLR-Landau fluid simulations. *Eur. Phys. J. D* **68**, 207.
- PASSOT, T., RUBAN, V. & SULEM, P. L. 2006 Fluid description of trains of stationary mirror structures in a magnetized plasma. *Phys. Plasmas* **13**, 102310.
- PASSOT, T. & SULEM, P. L. 2004 A Landau fluid model for dispersive magnetohydrodynamics. *Phys. Plasmas* **11**, 5173–5189.
- PASSOT, T. & SULEM, P. L. 2007 Collisionless magnetohydrodynamics with gyrokinetic effects. *Phys. Plasmas* **14**, 082502.
- PASSOT, T., SULEM, P. L. & HUNANA, P. 2012 Extending magnetohydrodynamics to the slow dynamics of collisionless plasmas. *Phys. Plasmas* **19**, 082113.
- PERRONE, D., PASSOT, T., LAVEDER, D., VALENTINI, F., SULEM, P. L., ZOUGANELIS, I., VELTRI, P. & SERVIDIO, S. 2018 Fluid simulations of plasma turbulence at ion scales: comparison with Vlasov–Maxwell simulations. *Phys. Plasmas* **25**, 052302.
- PIERRARD, V. & LAZAR, M. 2010 Kappa distributions: theory and applications in space plasmas. *Solar Phys.* **267**, 153–174.
- PIERRARD, V., LAZAR, M. & SCHLICKEISER, R. 2011 Evolution of the electron distribution function in the whistler wave turbulence of the solar wind. *Solar Phys.* **269**, 421–438.
- RAMOS, J. J. 2005 Fluid formalism for collisionless magnetized plasmas. *Phys. Plasmas* **12**, 052102.
- ROSENBLUTH, M. N. 1956 Stability of the pinch. *LANL Report No. LA-2030*.



- ROSIN, M. S., SCHEKOCHIHIN, A. A., RINCON, F. & COWLEY, S. C. 2011 A non-linear theory of the parallel firehose and gyrothermal instabilities in a weakly collisional plasma. *Mon. Not. R. Astron. Soc.* **413**, 7–38.
- SAHRAOUI, F., BELMONT, G. & GOLDSTEIN, M. L. 2012 New insight into short-wavelength solar wind fluctuations from Vlasov theory. *Astrophys. J.* **748**, 100.
- DOS SANTOS, M. S., ZIEBELL, L. F. & GAELZER, R. 2015 Ion-cyclotron instability in plasmas described by product-bi-kappa distributions. *Phys. Plasmas* **22** (12), 122107.
- SCHEKOCHIHIN, A. A., COWLEY, S. C., DORLAND, W., HAMMETT, G. W., HOWES, G. G., QUATAERT, E. & TATSUNG, T. 2009 Astrophysical Gyrokinetics: kinetic and fluid turbulent cascades in magnetized weakly collisional plasmas. *Astrophys. J. Suppl. Ser.* **182**, 310–377.
- SCHEKOCHIHIN, A. A., COWLEY, S. C., RINCON, F. & ROSIN, M. S. 2010 Magnetofluid dynamics of magnetized cosmic plasma: firehose and gyrothermal instabilities. *Mon. Not. R. Astron. Soc.* **405**, 291–300.
- SNYDER, P. B., HAMMETT, G. W. & DORLAND, W. 1997 Landau fluid models of collisionless magnetohydrodynamics. *Phys. Plasmas* **4**, 3974.
- SOUTHWOOD, D. J. & KIVELSON, M. G. 1993 Mirror instability. I. Physical mechanism of linear instability. *J. Geophys. Res.* **98**, 9181–9187.
- SULEM, P. L. & PASSOT, T. 2012 Fluid modeling of magnetized plasmas with anisotropic temperatures. *AIP Conf. Proc.* **1439**, 94.
- SULEM, P. L. & PASSOT, T. 2015 Landau fluid closures with nonlinear large-scale finite Larmor radius corrections for collisionless plasmas. *J. Plasma Phys.* **81**, 325810103.
- SUMMERS, D. & THORNE, R. M. 1992 A new tool for analyzing microinstabilities in space plasmas modeled by a generalized lorentzian (kappa) distribution. *J. Geophys. Res.* **97**, 16827–16832.
- TAJIRI, M. 1967 Propagation of hydromagnetic waves in collisionless plasma. II. Kinetic approach. *J. Phys. Soc. Japan* **22**, 1482–1494.
- TENERANI, A. & VELLI, M. 2018 Nonlinear firehose relaxation and constant-B field fluctuations. *Astrophys. J. Lett.* **867**, L26.
- TENERANI, A., VELLI, M. & HELLINGER, P. 2017 The parametric instability of Alfvén waves: effects of temperature anisotropy. *Astrophys. J.* **851**, 99.
- TU, C. Y. & MARSCH, E. 1995 MHD structures, waves and turbulence in the solar wind: observations and theories. *Space Sci. Rev.* **73**, 1–210.
- VEDENOV, A. A. & SAGDEEV, R. Z. 1958 On some properties of a plasma with an anisotropic ion-velocity distribution in a magnetic field. *Sov. Phys. Dokl.* **3**, 278.
- VOCKS, C., MANN, G. & RAUSCHE, G. 2008 Formation of suprathermal electron distributions in the quiet solar corona. *Astron. Astrophys.* **480**, 527–536.
- VOCKS, C., SALEM, C., LIN, R. P. & MANN, G. 2005 Electron Halo and Strahl formation in the solar wind by resonant interaction with whistler waves. *Astrophys. J.* **627**, 540–549.
- WAELEBROECK, F. L. 2010 Models for sub-Alfvénic magnetodynamics of fusion plasmas. *Fusion Sci. Technol.* **59**, 499.
- WANG, B.-J. & HAU, L.-N. 2003 MHD aspects of fire-hose type instabilities. *J. Geophys. Res.* **108**, 1463.
- WANG, B.-J. & HAU, L.-N. 2010 Parallel proton fire hose instability in gyrotropic Hall MHD model. *J. Geophys. Res.* **115**, A04105.
- WANG, L., HAKIM, A. H., BHATTACHARJEE, A. & GERMASCHESKI, K. 2015 Comparison of multi-fluid moment models with particle-in-cell simulations of collisionless magnetic reconnection. *Phys. Plasmas* **22** (1), 012108.
- WEBB, G. M., MCKENZIE, J. F., MACE, R. L., KO, C. M. & ZANK, G. P. 2007 Dual variational principles for nonlinear traveling waves in multifluid plasmas. *Phys. Plasmas* **14** (8), 082318.
- YAJIMA, N. 1966 The effect of finite ion Larmor radius on the propagation of magnetoacoustic waves. *Prog. Theoret. Phys.* **36**, 1.
- YANG, Y., MATTHAEUS, W. H., PARASHAR, T. N., HAGGERTY, C. C., ROYTERTSHTYEN, V., DAUGHTON, W., WAN, M., SHI, Y. & CHEN, S. 2017a Energy transfer, pressure tensor, and heating of kinetic plasma. *Phys. Plasmas* **24** (7), 072306.

- YANG, Y., MATTHAEUS, W. H., PARASHAR, T. N., WU, P., WAN, M., SHI, Y., CHEN, S., ROYTERTSHTEYN, V. & DAUGHTON, W. 2017*b* Energy transfer channels and turbulence cascade in Vlasov–Maxwell turbulence. *Phys. Rev. E* **95** (6), 061201.
- ZANK, G. P. 1999 Interaction of the solar wind with the local interstellar medium: a theoretical perspective. *Space Sci. Rev.* **89**, 413–688.
- ZANK, G. P. 2014 *Transport Processes in Space Physics and Astrophysics*, Lecture Notes in Physics, vol. 877. Springer.
- ZANK, G. P., ADHIKARI, L., HUNANA, P., SHIOTA, D., BRUNO, R. & TELLONI, D. 2017 Theory and transport of nearly incompressible magnetohydrodynamic turbulence. *Astrophys. J.* **835**, 147.
- ZANK, G. P., HUNANA, P., MOSTAFAVI, P. & GOLDSTEIN, M. L. 2014 Pickup ion mediated plasmas. I. Basic model and linear waves in the solar wind and local interstellar medium. *Astrophys. J.* **797**, 87.
- ZANK, G. P., LE ROUX, J. A., WEBB, G. M., DOSCH, A. & KHABAROVA, O. 2014 Particle acceleration via reconnection processes in the supersonic solar wind. *Astrophys. J.* **797**, 28.
- ZANK, G. P. & MATTHAEUS, W. H. 1993 Nearly incompressible fluids. II. Magnetohydrodynamics, turbulence, and waves. *Phys. Fluids A* **5**, 257–273.
- ZHOU, Y., MATTHAEUS, W. H. & DMITRUK, P. 2004 Colloquium: magnetohydrodynamic turbulence and time scales in astrophysical and space plasmas. *Rev. Mod. Phys.* **76**, 1015–1035.