Relativistic theory of particles in a scattering flow – I. Basic equations, diffusion, and drift

A. Achterberg^{1★} and C. A. Norman²

¹*Department of Astrophysics, IMAPP, Radboud University, PO Box 9010, 6500 GL Nijmegen, the Netherlands*

²*Department of Physics & Astronomy, Johns Hopkins University, Baltimore, MD 21218, USA*

Accepted 2018 May 30. Received 2018 May 24; in original form 2017 November 27

ABSTRACT

We reconsider the theory of particle transport in a scattering medium, allowing for relativistic flow velocities. The theory uses a mixed set of variables, with position and time measured in the laboratory frame, and particle energy and momentum measured in the fluid rest frame, the reference frame where scattering is assumed to be elastic. We give a new derivation for the fictitious force terms in the equation of motion that are present if the fluid rest frame is not an inertial frame. By using a $3 + 1$ notation we discuss the physical interpretation of the different terms in the fictitious force. It is shown that different approaches to the problem of particle propagation in a magnetized medium due to Skilling and Kulsrud are largely equivalent. We extend known results for non-relativistic flows to include the effects of cross-field diffusion for cosmic rays in a magnetized plasma. We also carefully consider the correct form of the diffusion approximation for scattering, and show that the resulting equations can be cast in conservation form.

Key words: diffusion – MHD – relativistic processes – scattering – cosmic rays.

1 INTRODUCTION

The transport theory of charged particles or photons in a scattering and moving medium has many applications in astrophysics. In the case of energetic charged particles (cosmic rays) the scattering is due to random magnetic fields (magnetohydrodynamic, MHD waves), as was realized a long time ago by Parker (1965), Dolginov & Toptygin (1967), and Gleeson & Axford (1967). A very systematic and general derivation of the cosmic ray transport equation was given by Skilling (1975), which included the effect of the wave motion with respect to a non-relativistic flow and *did* allow for relativistic particles with particle momentum $p \gg mc$.

More or less in parallel Kulsrud (1983) formulated a theory in the limit of the vanishing gyro-radius for the kinetic description of charged particles in a moving magnetized plasma, assuming non-relativistic particles and slow motion (bulk velocity much less than the speed of light).

The application of these two approaches (i.e. Skilling’s formulation versus Kulsrud’s approach) was mostly limited to the astrophysics and the plasma physics community, respectively, even though they are (in the appropriate limit) essentially equivalent in physical content, as we will show.

A generalization of Skilling’s approach to a relativistic flow was first formulated by Webb (1985). However, as was the case with the calculations of both Skilling (1975) and Kulsrud (1983), the final equations of Webb (1985) are obtained by taking an average over the gyro-phase, the phase of the rapid circular motion of a charged particle in the plane perpendicular to the magnetic field. This is allowed if the gyro-frequency $\Omega = qB/\gamma mc$ of a charge q of mass m and Lorentz factor γ in a magnetic field B is very large so that the gyration period is usually much smaller than any of the hydrodynamical time-scales. This averaging procedure is mathematically convenient and allows for a relatively simple final form of the transport equation. But as already noted by Skilling (1975), the gyro-average eliminates physical processes such as diffusion across the magnetic field and viscous interactions in a shear flow which both depend on anisotropies in the gyro-phase. Diffusion *along* the magnetic field is unaffected by the gyro-averaging and is obtained in a simple fashion. The importance of viscous effects of cosmic rays that are coupled to a flow by frequent scattering was already noted by Jokipii & Williams (1992) and Williams & Jokipii (1993).

* E-mail: a.achterberg@astro.ru.nl

For photons, radiation transport in a moving medium has been considered by many authors. The classical results and principles of radiation hydrodynamics can be found in the books of Pomraning (1973), Mihalas & Weibel Mihalas (1984), and Castor (2004). The viscous effects of a radiation fluid were considered by Weinberg (1971) and more recently by Coughlin & Begelman (2014).

In this paper we reconsider the general theory with an emphasis on charged particles in a magnetized medium. We extend existing calculations to include anisotropies in the gyro-phase, a necessary step for the calculation of the full diffusion tensor and of viscous effects. We will derive the fundamental equations and show the relation to known results for non-relativistic flows. We also cast the equations in a more intuitive form, employing a 3 + 1 split for space–time. This approach allows an easy comparison with known results for non-relativistic flows. We use the results to derive a transport equation for the particles, showing the effects of diffusion and drift with respect to the bulk fluid. We also consider (and prove) the general covariance of the final equations. The treatment of viscous effects and a detailed discussion of the form of the stress tensor for charged particles in a magnetized medium are deferred to the companion paper (Paper 2). Radiation transport is treated in Paper 3.

2 BASIC EQUATIONS: USE OF MIXED VARIABLES

In what follows we assume a flow with four-velocity $U^\mu = (\Gamma, \mathbf{U}) = (\sqrt{1+U^2}, \mathbf{U})$ in the laboratory frame (LF). This flow scatters particles (or photons). Scattering is assumed to be elastic in the fluid rest frame (FRF). We employ units with $c = 1$ throughout the paper, but reinstate c when it clarifies results. The assumption of elastic scattering in the FRF is a simplification in the case of scattering of cosmic rays by MHD waves: It neglects the phase velocity of the waves with respect to the fluid. As shown by Skilling (1975), this becomes important when the intensity of the waves responsible for the scattering (in particular Alfvén waves) is not equal in the two possible propagation directions: along and against the large-scale magnetic field.

Since scattering takes the simplest form in the fluid rest frame, it is mathematically convenient to use a set of mixed variables, where position and time are measured in the LF but particle momentum and energy are measured in the FRF. In what follows we will employ the following notation: The LF momentum four-vector is denoted using a Greek index notation and the four-momentum in the FRF is denoted using a hatted Roman index notation (e.g. $\hat{a}$, $\hat{b}$, ...):

$$p^\mu = (\mathcal{E}, \mathbf{p}), \quad p^{\hat{a}} = (\bar{\mathcal{E}}, \bar{\mathbf{p}}). \quad (1)$$

The particle energy $\bar{\mathcal{E}}$ and the momentum three-vector $\bar{\mathbf{p}}$ in the FRF are barred in order to distinguish them from the corresponding quantities in the LF. The same convention is used when dealing with other non-covariant quantities such as three-vectors. For most of the paper we employ a 3 + 1 split of the equations. This makes a direct comparison with previous results more easy. We will show however that the final results can always be expressed covariantly in terms of scalars, four-vectors, and four-tensors. That means that, even though the calculations are done neglecting gravity, they are straightforwardly generalized to the general-relativistic case.

As long as we neglect the effects of gravity (i.e. in special relativity), p^μ and $p^{\hat{a}}$ are connected by Lorentz transformations. The case with gravity is considered briefly in Appendix A. For technical reasons (again, see Appendix A) we employ the covariant momentum four-vectors $p_\mu = (\mathcal{E}, -\mathbf{p})$ and $p_{\hat{a}} = (\bar{\mathcal{E}}, -\bar{\mathbf{p}})$, defined formally through $p_\mu = \eta_{\mu\nu} p^\nu$, $p_{\hat{a}} = \eta_{\hat{a}\hat{b}} p^{\hat{b}}$ with $\eta_{\mu\nu} = \eta_{\hat{a}\hat{b}} = \text{diag}(1, -1, -1, -1)$ the Minkowski tensor, the metric tensor of flat space–time. Using the Einstein summation convention for double indices we can write:

$$p_\mu = E_{\mu}^{\hat{a}} p_{\hat{a}}, \quad p_{\hat{a}} = E_{\hat{a}}^{\mu} p_{\mu}. \quad (2)$$

Here $E_{\mu}^{\hat{a}}$ and $E_{\hat{a}}^{\mu}$ are Lorentz transformations (see for instance Landau & Lifshitz 1975, ch. 1) between the two frames. Represented as 4×4 matrices

$$E_{\mu}^{\hat{a}} = \begin{pmatrix} \Gamma & -U \\ -U & \mathbf{I} + \frac{\mathbf{U} \otimes \mathbf{U}}{\Gamma + 1} \end{pmatrix}, \quad E_{\hat{a}}^{\mu} = \begin{pmatrix} \Gamma & U \\ U & \mathbf{I} + \frac{\mathbf{U} \otimes \mathbf{U}}{\Gamma + 1} \end{pmatrix}. \quad (3)$$

They are each other's inverse, or alternatively $E_{\hat{a}}^{\mu}(U) = E_{\mu}^{\hat{a}}(-U)$. The notation $\mathbf{U} \otimes \mathbf{U}$ is the direct product, with components $(\mathbf{U} \otimes \mathbf{U})_{ij} = U^i U^j$. We use unhatted roman upper indices $i, j, \dots$, with $(i, j, \dots) = 1, 2, 3$, when considering the spatial components of four-vectors in the LF. The following relations hold:

$$E_{\mu}^{\hat{a}} E_{\hat{b}}^{\mu} = \eta_{\hat{b}}^{\hat{a}}, \quad E_{\mu}^{\hat{a}} E_{\hat{a}}^{\mu} = \eta_{\mu}^{\mu} = \delta_{\mu}^{\mu}, \quad \eta_{\mu\nu} E_{\hat{a}}^{\mu} E_{\hat{b}}^{\nu} = \eta_{\hat{a}\hat{b}}, \quad \eta_{\hat{a}\hat{b}} E_{\mu}^{\hat{a}} E_{\nu}^{\hat{b}} = \eta_{\mu\nu}. \quad (4)$$

The last two relations guarantee the invariance of the scalar product of two four-vectors, in particular $p_\mu p^\mu = \mathcal{E}^2 - p^2 = p_{\hat{a}} p^{\hat{a}} = \bar{\mathcal{E}}^2 - \bar{p}^2 = m^2$.

3 THE FICTITIOUS FORCE

Generally, the FRF is not an inertial frame. The LF on the other hand is assumed to be an inertial frame. The use of the FRF four-momentum $p_{\hat{a}}$ implies the appearance of, in classical terms, a *fictitious force term* in the equation of motion. In relativity the equivalence principle implies that this can also be thought of as space–time curvature in the accelerated frame. We will return to this point in Section 7.

3.1 Special-relativistic calculation

In special relativity the form of the fictitious force can be derived by using the generalization of Newton's law of inertia for a free particle. In the absence of gravity or other forces the LF four-momentum is conserved. This means we can write

$$\frac{dp_\mu}{d\lambda} = \frac{d}{d\lambda} (E_\mu^{\hat{a}} p_{\hat{a}}) = 0. \quad (5)$$

Here λ is the affine parameter along the orbit of a particle. This yields

$$E_\mu^{\hat{a}} \frac{dp_{\hat{a}}}{d\lambda} + \frac{dE_\mu^{\hat{a}}}{d\lambda} p_{\hat{a}} = 0. \quad (6)$$

Using the first relation in equation (4) one finds, after multiplication by E_c^μ and re-ordering:

$$E_c^\mu E_\mu^{\hat{a}} \frac{dp_{\hat{a}}}{d\lambda} = \frac{dp_c}{d\lambda} = -E_c^\mu \frac{dE_\mu^{\hat{a}}}{d\lambda} p_{\hat{a}}. \quad (7)$$

As $d/d\lambda$ is the proper time derivative along the particle's orbit it equals

$$\frac{d}{d\lambda} \equiv \frac{p^\nu}{m} \frac{\partial}{\partial x^\nu} = E_b^\nu \frac{p^{\hat{b}}}{m} \frac{\partial}{\partial x^\nu}. \quad (8)$$

One finds:

$$\frac{dp_c}{d\lambda} = -E_c^\mu E_b^\nu \frac{\partial E_\mu^{\hat{a}}}{\partial x^\nu} \frac{p_{\hat{a}} p^{\hat{b}}}{m} \equiv K_{\hat{c}}. \quad (9)$$

This gives the fictitious force on a free particle, now expressed entirely in terms of FRF momentum components. Additional force terms can now be simply added.

3.2 General-relativistic calculation

If one allows for space-time curvature with a metric $g_{\mu\nu}$, equation (5) is replaced by (see for instance Carroll 2004, ch.3 and defining $u_\alpha = p_\alpha/m$ and $d/d\lambda = u^\mu (\partial/\partial x^\mu)$)

$$\frac{dp_\mu}{d\lambda} = \Gamma_{\mu\beta}^\alpha u_\alpha p^\beta = -\frac{1}{2m} \left(\frac{\partial g^{\alpha\beta}}{\partial x^\mu} \right) p_\alpha p_\beta. \quad (10)$$

In this case the FRF coordinates can be put in quasi-Minkowskian form (so that *locally* $g_{\hat{a}\hat{b}} = \eta_{\hat{a}\hat{b}}$) by using the tetrad formalism (see Appendix A). The transformations $E_\mu^{\hat{a}}$ and the inverse $E_\mu^{\hat{a}}$ now satisfy relations analogous to those in equation (4), where one replaces $\eta^{\mu\nu}$ ($\eta_{\mu\nu}$) by $g^{\mu\nu}$ ($g_{\mu\nu}$). In particular one has $\eta_{\hat{a}\hat{b}} E_\mu^{\hat{a}} E_\nu^{\hat{b}} = g_{\mu\nu}$ and $g_{\mu\nu} E_a^\mu E_b^\nu = \eta_{\hat{a}\hat{b}}$. Then expressing all LF momentum variables in equation (10) in terms of FRF momentum variables one arrives at

$$\frac{dp_{\hat{c}}}{d\lambda} = E_c^\mu \left(\frac{\partial E_\mu^{\hat{a}}}{\partial x^\mu} - \frac{\partial E_\mu^{\hat{a}}}{\partial x^\nu} \right) \frac{p_{\hat{a}} p^{\hat{b}}}{m}. \quad (11)$$

A formal proof of this result is given in Appendix A, using a variational principle. The extra term in equation (11) with respect to special-relativistic result (9) (i.e. the first term inside the bracket) vanishes in flat space-time. In this paper we will use another route to obtain results that can be applied in general relativity, namely by proving the covariance of the final equations so that this transition is one from ordinary derivatives to covariant derivatives.

3.3 Fictitious force calculation for a non-inertial fluid rest frame

In flat space-time all variations of E_a^μ and its inverse stem from changes in the velocity V of the fluid in the laboratory frame. This means that we can write the derivatives that appear in equation (9) in terms of the derivatives with respect to the spatial components $U^i = \Gamma V^i$ (with $i = 1, 2, 3$) of the four-velocity of the FRF in the LF:

$$\frac{dE_\mu^{\hat{a}}}{d\lambda} = \frac{\partial E_\mu^{\hat{a}}}{\partial U^k} \frac{dU^k}{d\lambda}, \quad \frac{\partial E_\mu^{\hat{a}}}{\partial x^\nu} = \frac{\partial E_\mu^{\hat{a}}}{\partial U^k} \frac{\partial U^k}{\partial x^\nu} \quad (12)$$

The following relations hold:

$$\frac{\partial E_0^{\hat{0}}}{\partial U^k} = \frac{\partial \Gamma}{\partial U^k} = \frac{U^k}{\Gamma} = V^k, \quad \frac{\partial E_0^{\hat{i}}}{\partial U^k} = -\frac{\partial U^i}{\partial U^k} = -\delta_{ik}, \quad \frac{\partial E_i^{\hat{j}}}{\partial U^k} = \frac{1}{\Gamma + 1} \left(\delta_{ik} U^j + \delta_{jk} U^i - \frac{U^i U^j U^k}{\Gamma(\Gamma + 1)} \right). \quad (13)$$

We are somewhat sloppy with the index notation, replacing $(\hat{i}, \hat{j})$ by (i, j) when no confusion can arise. All unhatted roman indices $i, j, \dots$ run from 1 to 3.

3.3.1 Energy change

The change of FRF energy due to the fictitious force follows from $p_0 = \bar{\mathcal{E}}$ and equation (9) with $\hat{c} = \hat{0}$. Using equations (12) and (13), together with $\mathbf{U} = \Gamma \mathbf{V}$ and $\Gamma = \sqrt{1 + |\mathbf{U}|^2}$, one finds after a significant amount of algebra (see Appendix B)

$$\frac{d\bar{\mathcal{E}}}{d\lambda} = -(\Gamma + 1) \left[\bar{\mathbf{p}} \cdot \frac{d}{d\lambda} \left(\frac{\mathbf{U}}{\Gamma + 1} \right) \right]. \quad (14)$$

3.3.2 Momentum change

In a similar fashion, the momentum change due to the fictitious force is found to be (again see Appendix B for details)

$$\frac{d\bar{\mathbf{p}}}{d\lambda} = -(\Gamma + 1) \bar{\mathcal{E}} \frac{d}{d\lambda} \left(\frac{\mathbf{U}}{\Gamma + 1} \right) + \bar{\mathbf{p}} \times \boldsymbol{\omega}, \quad (15)$$

where

$$\boldsymbol{\omega} \equiv \frac{\mathbf{U}}{\Gamma + 1} \times \frac{d\mathbf{U}}{d\lambda}. \quad (16)$$

The second term on the right-hand side of equation (15) is reminiscent of the Coriolis term in a rotating reference frame in classical mechanics. Indeed it stems from the change in the direction of $\mathbf{U}$. In addition it is a purely relativistic effect. Writing $\mathbf{U} = U \hat{\ell}$, with $\hat{\ell}$ a unit vector along $\mathbf{U}$ and $U = |\mathbf{U}|$, one immediately finds using $U^2 = \Gamma^2 - 1$

$$\boldsymbol{\omega} = \frac{U^2}{\Gamma + 1} \left(\hat{\ell} \times \frac{d\hat{\ell}}{d\lambda} \right) = (\Gamma - 1) \left(\hat{\ell} \times \frac{d\hat{\ell}}{d\lambda} \right). \quad (17)$$

For non-relativistic flows with $U \ll 1$ we have $\Gamma - 1 \simeq U^2/2c^2 \ll 1$, where we reinstated c for clarity, and the $\bar{\mathbf{p}} \times \boldsymbol{\omega}$ term can be neglected.

It is easily checked that results of equations (14) and (15) maintain the Lorentz invariance of $\bar{\mathcal{E}}^2 - \bar{\mathbf{p}}^2$ since

$$\bar{\mathcal{E}} \frac{d\bar{\mathcal{E}}}{d\lambda} - \bar{\mathbf{p}} \cdot \frac{d\bar{\mathbf{p}}}{d\lambda} = 0. \quad (18)$$

3.3.3 The origin of the $\bar{\mathbf{p}} \times \boldsymbol{\omega}$ term

The appearance of this term in the fictitious force is related to the fact that two Lorentz boosts L_1 and L_2 , where the boost directions are not equal, do not commute, so $L_1 L_2 \neq L_2 L_1$. The difference $L_1 L_2 - L_2 L_1$ corresponds to a spatial rotation. The same effect is, for instance, responsible for the Thomas precession and the similar appearance of a cross-product in the theory of Fermi–Walker transport of a tetrad of quasi-Minkowskian unit vectors carried by an accelerated observer (e.g. Misner, Thorne, & Wheeler 1973, ch. 6) when the acceleration is not uniform.

A simple physical argument for the appearance and form of this term goes as follows: Consider a free particle with constant LF energy $\mathcal{E}$ and momentum $\mathbf{p}$. The FRF momentum can be expressed using the Lorentz transformation from the LF to the FRF:

$$\bar{\mathbf{p}} = \mathbf{p} - U \mathcal{E} \hat{\ell} + (\Gamma - 1) (\mathbf{p} \cdot \hat{\ell}) \hat{\ell}. \quad (19)$$

Now consider an infinitesimal change in the direction of $\mathbf{U} \equiv U \hat{\ell}$ with $U \equiv |\mathbf{U}|$:

$$\hat{\ell} \Rightarrow \hat{\ell} + \delta\hat{\ell}, \quad \hat{\ell} \cdot \delta\hat{\ell} = 0, \quad (20)$$

so that the vector $\hat{\ell} + \delta\hat{\ell}$ remains a unit vector to first order. The induced linear change in $\bar{\mathbf{p}}$ with U , $\mathcal{E}$, and $\mathbf{p}$ constant is

$$\delta\bar{\mathbf{p}} = -U \mathcal{E} \delta\hat{\ell} + (\Gamma - 1) \left[(\mathbf{p} \cdot \delta\hat{\ell}) \hat{\ell} + (\mathbf{p} \cdot \hat{\ell}) \delta\hat{\ell} \right]. \quad (21)$$

Next we write $\mathcal{E}$ and $\mathbf{p}$ in terms of the (unperturbed) $\bar{\mathcal{E}}$ and $\bar{\mathbf{p}}$ using the Lorentz transformation back to the LF:

$$\mathcal{E} = \Gamma \bar{\mathcal{E}} + U (\bar{\mathbf{p}} \cdot \hat{\ell}), \quad \mathbf{p} = \bar{\mathbf{p}} + U \bar{\mathcal{E}} \hat{\ell} + (\Gamma - 1) (\bar{\mathbf{p}} \cdot \hat{\ell}) \hat{\ell}. \quad (22)$$

Substituting this into relation (21) yields, again to first order in $\delta\hat{\ell}$:

$$\delta\bar{\mathbf{p}} = -U \bar{\mathcal{E}} \delta\hat{\ell} + (\Gamma - 1) \left[(\bar{\mathbf{p}} \cdot \delta\hat{\ell}) \hat{\ell} - (\bar{\mathbf{p}} \cdot \hat{\ell}) \delta\hat{\ell} \right] = -U \bar{\mathcal{E}} \delta\hat{\ell} + (\Gamma - 1) \left[\bar{\mathbf{p}} \times (\hat{\ell} \times \delta\hat{\ell}) \right]. \quad (23)$$

The $(\Gamma - 1) [\bar{\mathbf{p}} \times (\hat{\ell} \times \delta\hat{\ell})]$ term is the source of the cross-product term involving $\boldsymbol{\omega}$ in equation (15).

4 VLASOV EQUATION

Consider charged particles with rest mass m and charge q . The Vlasov equation for such particles in the laboratory frame reads (Klimontovich 1960; Hakim 2011; see also Debbasch & van Leeuwen 2009a, b, who discuss the general-relativistic case):

$$\frac{p^\mu}{m} \frac{\partial \mathcal{F}}{\partial x^\mu} + \frac{q}{m} F_\mu^\nu p_\nu \frac{\partial \mathcal{F}}{\partial p_\mu} = \text{St}(\mathcal{F}). \quad (24)$$

Here $\mathcal{F}(x, p)$ is the distribution function in terms of the eight phase-space variables (x^μ, p_μ) , defined in such a way that (see Klimontovich 1960) $dN^\mu = \mathcal{F}(x, p)(p^\mu/m)d^4x d^4p$ is the number flux four-vector of particles residing in the eight-dimensional phase-space element. It is related to the more familiar distribution $f(\mathbf{x}, t, \mathbf{p})$ in seven-dimensional phase space through

$$\mathcal{F}(x^\mu, p_\mu) = 2m\Theta(p_0)f(\mathbf{x}, t, \mathbf{p})\delta(p^2 - m^2), \quad (25)$$

with $\Theta(p_0) = \Theta(\mathcal{E})$ the Heaviside step function and $p^2 = p^\mu p_\mu = \mathcal{E}^2 - |\mathbf{p}|^2$. The delta function keeps particles on mass shell while the step function $\Theta(p_0)$ ensures positive energy. The tensor F_μ^ν is the electromagnetic Faraday tensor, see for instance Jackson (1998), and $(q/m)F_\mu^\nu p_\nu = \gamma(\mathbf{E} \cdot \mathbf{V}, -[\mathbf{E} + \mathbf{V} \times \mathbf{B}])$ is the covariant Lorentz four-force with $(\mathbf{E}, \mathbf{B})$ the electromagnetic field in the laboratory frame and $\mathbf{V} = \mathbf{p}/\gamma m$ the particle velocity. $\text{St}(\mathcal{F})$ is a scattering term that will be specified below.

If we change momentum variables from LF four-momentum p_μ to FRF four-momentum $p_{\hat{a}}$, we have to use

$$\frac{\partial}{\partial x^\mu} \Rightarrow \frac{\partial}{\partial x^\mu} + \frac{\partial p_{\hat{a}}}{\partial x^\mu} \frac{\partial}{\partial p_{\hat{a}}} = \frac{\partial}{\partial x^\mu} + \frac{\partial E_{\hat{a}}^\nu}{\partial x^\mu} p_{\hat{a}}^\nu \frac{\partial}{\partial p_{\hat{a}}}, \quad (26)$$

where $\partial/\partial x^\mu$ is now with $(\bar{\mathcal{E}}, \bar{\mathbf{p}})$ fixed. Also

$$\frac{\partial}{\partial p_\mu} = \frac{\partial p_{\hat{a}}}{\partial p_\mu} \frac{\partial}{\partial p_{\hat{a}}} = E_{\hat{a}}^\mu \frac{\partial}{\partial p_{\hat{a}}}. \quad (27)$$

The Jacobian of the momentum transformation is $\det(E_{\hat{a}}^\mu) = 1$, so $\mathcal{F}$ is invariant under this particular change of momentum variables. Equation (24) then becomes

$$E_{\hat{a}}^\mu \left(\frac{p_{\hat{a}}}{m} \right) \frac{\partial \mathcal{F}}{\partial x^\mu} + \left(\frac{p^\mu}{m} \right) \frac{\partial E_{\hat{a}}^\nu}{\partial x^\mu} E_{\hat{a}}^\nu p_{\hat{b}} \frac{\partial \mathcal{F}}{\partial p_{\hat{a}}} + \frac{q}{m} E_{\hat{a}}^\nu E_{\hat{a}}^\mu F_\mu^\nu p_{\hat{b}} \frac{\partial \mathcal{F}}{\partial p_{\hat{a}}} = \text{St}(\mathcal{F}). \quad (28)$$

Here we have used transformation rules (equation 2) repeatedly. The Faraday tensor in the FRF is $F_{\hat{a}}^{\hat{b}} = E_{\hat{a}}^{\hat{b}} E_{\hat{a}}^\mu F_\mu^\nu$. The first relation in equation (4) allows us to write the factor in front of $\partial \mathcal{F}/\partial p_{\hat{a}}$ in the second term of equation (28) as:

$$\left(\frac{p^\mu}{m} \right) \frac{\partial E_{\hat{a}}^\nu}{\partial x^\mu} E_{\hat{a}}^\nu p_{\hat{b}} = - \left(\frac{p^\mu}{m} \right) E_{\hat{a}}^\nu \frac{\partial E_{\hat{a}}^\nu}{\partial x^\mu} p_{\hat{b}} = - E_{\hat{a}}^\mu E_{\hat{a}}^\nu \frac{\partial E_{\hat{a}}^\nu}{\partial x^\mu} \frac{p_{\hat{a}}^\mu p_{\hat{b}}}{m}. \quad (29)$$

A comparison with definition (9) shows immediately that this is the fictitious force $K_{\hat{a}}$ (apart from a trivial relabelling of indices) as is expected on physical grounds. This gives the final form of the Vlasov equation, now in mixed LF/FRF variables:

$$E_{\hat{a}}^\mu \left(\frac{p_{\hat{a}}}{m} \right) \frac{\partial \mathcal{F}}{\partial x^\mu} + \left(K_{\hat{a}} + \frac{q}{m} F_{\hat{a}}^{\hat{b}} p_{\hat{b}} \right) \frac{\partial \mathcal{F}}{\partial p_{\hat{a}}} = \text{St}(\mathcal{F}). \quad (30)$$

Here $\mathcal{F}$ should now be interpreted as $\mathcal{F} \equiv \mathcal{F}(x^\mu, p_{\hat{a}}) = 2m\Theta(\bar{\mathcal{E}})f(\mathbf{x}, t, \bar{\mathbf{p}})\delta(\bar{\mathcal{E}}^2 - |\bar{\mathbf{p}}|^2 - m^2)$. Using the properties of the δ -function and performing an integration over FRF energy $\bar{\mathcal{E}} = p_{\hat{0}}$ of the form $\int d\bar{\mathcal{E}} \bar{\mathcal{E}} \dots$ of the whole equation one can express equation (30) in terms of $f(\mathbf{x}, t, \bar{\mathbf{p}})$ as

$$E_{\hat{a}}^\mu \frac{p_{\hat{a}}}{m} \frac{\partial f}{\partial x^\mu} + \left\{ \bar{\gamma} q (\bar{\mathbf{E}} + \bar{\mathbf{V}} \times \bar{\mathbf{B}}) - (\Gamma + 1) \bar{\mathcal{E}} \frac{d}{d\lambda} \left(\frac{\mathbf{U}}{\Gamma + 1} \right) + \bar{\mathbf{p}} \times \boldsymbol{\omega} \right\} \cdot \frac{\partial f}{\partial \bar{\mathbf{p}}} = \text{St}(f). \quad (31)$$

Here $\text{St}(f)$ is the collision term acting on the ordinary Vlasov distribution $f(\mathbf{x}, t, \bar{\mathbf{p}})$. For pure scattering that only changes the direction of particle momenta we have $\text{St}(\mathcal{F}) = 2m\Theta(p_0)\delta(p^2 - m^2)\text{St}(f)$ (cf. equation 25). Also $\bar{\gamma} = \bar{\mathcal{E}}/m$ is the Lorentz factor of a particle in the FRF with velocity $\bar{\mathbf{V}} = \bar{\mathbf{p}}/\bar{\gamma}m$, and where $\bar{\mathcal{E}}$ must be interpreted as $\bar{\mathcal{E}} \equiv \sqrt{\bar{\mathbf{p}}^2 + m^2}$. The three-vectors $(\bar{\mathbf{E}}, \bar{\mathbf{B}})$ are the electromagnetic fields in the FRF. Definition (8) for $d/d\lambda$ together with the definition of $E_{\hat{a}}^\mu$ allows us to write

$$\frac{d}{d\lambda} = E_{\hat{a}}^\mu \frac{p_{\hat{a}}}{m} \frac{\partial}{\partial x^\mu} = \bar{\gamma} \left(\frac{d}{d\tau} + \bar{\mathbf{V}} \cdot \bar{\nabla} \right) = \bar{\gamma} \left\{ \Gamma [1 + (\bar{\mathbf{V}} \cdot \mathbf{V})] \frac{\partial}{\partial t} + (\bar{\mathbf{V}} + \mathbf{U}) \cdot \nabla + \frac{(\bar{\mathbf{V}} \cdot \mathbf{U})}{\Gamma + 1} (\mathbf{U} \cdot \nabla) \right\}. \quad (32)$$

Here we use $\mathbf{U} = \Gamma \mathbf{V}$ and $\bar{\mathbf{V}} \equiv \bar{\mathbf{p}}/\bar{\gamma}m$ and introduce the notation

$$\frac{d}{d\tau} = \frac{\partial}{\partial x^0} = \mathbf{U} \cdot \nabla = \Gamma \frac{\partial}{\partial t} + (\mathbf{U} \cdot \nabla), \quad (33)$$

for the comoving proper time derivative associated with the bulk plasma, and

$$\bar{\nabla} = \frac{\partial}{\partial \bar{\mathbf{x}}} = \nabla + \mathbf{U} \left(\frac{\partial}{\partial t} + \frac{(\mathbf{U} \cdot \nabla)}{\Gamma + 1} \right) \quad (34)$$

for the FRF gradient operator expressed in LF time and spatial derivatives.

4.1 Collision operator

To describe the effect of ‘collisions’ (for instance scattering by magnetic irregularities) we choose the Anderson–Witting form for the relaxation operator (Anderson & Witting 1974) with $p \cdot \mathbf{U} \equiv p_\mu U^\mu = \bar{\mathcal{E}} = \bar{\gamma}m$:

$$\text{St}(f) \equiv \bar{\gamma} \bar{\text{St}}(f) = - \frac{p \cdot \mathbf{U}}{m} \frac{f - f_0}{\bar{\tau}_r} = - \frac{\bar{\gamma} (f - f_0)}{\bar{\tau}_r}. \quad (35)$$

This operator is the proper relativistic version of the well-known Krook collision operator (see Bhatnagar, Gross, & Krook 1954). Here f_0 is an equilibrium distribution function (usually chosen to be isotropic in the FRF) and $\bar{\tau}_r$ the relaxation time in the FRF. For our present purposes this is the most convenient choice, but other choices (such as a Fokker–Planck type diffusion operator in momentum that describes scattering) are possible. In Section 11 below we briefly discuss the effect of using an alternative choice for the relaxation operator: diffusive (Fokker–Planck type) scattering.

If we use the Anderson–Witting relaxation operator the Vlasov equation becomes

$$\Gamma [1 + (\bar{\mathbf{V}} \cdot \mathbf{V})] \frac{\partial f}{\partial t} + (\bar{\mathbf{V}} + \mathbf{U}) \cdot \nabla f + \frac{(\bar{\mathbf{V}} \cdot \mathbf{U})}{\Gamma + 1} (\mathbf{U} \cdot \nabla) f + \{ \bar{\mathbf{K}}_f + \bar{\mathbf{p}} \times \bar{\boldsymbol{\omega}} + q (\bar{\mathbf{E}} + \bar{\mathbf{V}} \times \bar{\mathbf{B}}) \} \cdot \frac{\partial f}{\partial \bar{\mathbf{p}}} = - \frac{f - f_0}{\bar{\tau}_r}. \quad (36)$$

Here we have cancelled a common factor $\bar{\gamma}$ and defined the fictitious three-force as $(d\bar{\mathbf{p}}/dt)_{\text{fict}} \equiv \bar{\mathbf{K}}_f + \bar{\mathbf{p}} \times \bar{\boldsymbol{\omega}}$, where

$$\bar{\mathbf{K}}_f \equiv \frac{1}{\bar{\gamma}} \frac{d\bar{\mathbf{p}}}{d\lambda} = -(\Gamma + 1) \frac{d}{d\lambda} \left(\frac{m \mathbf{U}}{\Gamma + 1} \right) = -\bar{\gamma} m (\Gamma + 1) \left\{ \Gamma [1 + (\bar{\mathbf{V}} \cdot \mathbf{V})] \frac{\partial}{\partial t} + (\bar{\mathbf{V}} + \mathbf{U}) \cdot \nabla + \frac{(\bar{\mathbf{V}} \cdot \mathbf{U})}{\Gamma + 1} (\mathbf{U} \cdot \nabla) \right\} \left(\frac{\mathbf{U}}{\Gamma + 1} \right) \quad (37)$$

and

$$\bar{\boldsymbol{\omega}} \equiv \frac{\boldsymbol{\omega}}{\bar{\gamma}} = \frac{\Gamma - 1}{\bar{\gamma}} \left(\hat{\boldsymbol{\ell}} \times \frac{d\hat{\boldsymbol{\ell}}}{d\lambda} \right) = (\Gamma - 1) \hat{\boldsymbol{\ell}} \times \left\{ \Gamma [1 + (\bar{\mathbf{V}} \cdot \mathbf{V})] \frac{\partial \hat{\boldsymbol{\ell}}}{\partial t} + (\bar{\mathbf{V}} + \mathbf{U}) \cdot \nabla \hat{\boldsymbol{\ell}} + \frac{(\bar{\mathbf{V}} \cdot \mathbf{U})}{\Gamma + 1} (\mathbf{U} \cdot \nabla) \hat{\boldsymbol{\ell}} \right\}. \quad (38)$$

Note that our definition for $\bar{\mathbf{K}}_f$ agrees with the usual definition of the Minkowski force in special relativity (see for instance Goldstein, Poole, & Safko 2002). Defining $\bar{\mathbf{K}} = \bar{\mathbf{K}}_f + \bar{\mathbf{p}} \times \bar{\boldsymbol{\omega}}$, the change of the FRF four-momentum $p^{\hat{a}} = (\bar{\mathcal{E}}, \bar{\mathbf{p}})$ can be written in standard form:

$$K^{\hat{a}} \equiv \frac{dp^{\hat{a}}}{d\lambda} = \bar{\gamma} (\bar{\mathbf{V}} \cdot \bar{\mathbf{K}}, \bar{\mathbf{K}}) = \bar{\gamma} (\bar{\mathbf{V}} \cdot \bar{\mathbf{K}}_f, \bar{\mathbf{K}}_f + \bar{\mathbf{p}} \times \bar{\boldsymbol{\omega}}). \quad (39)$$

Here we used $\bar{\mathbf{V}} \cdot (\bar{\mathbf{p}} \times \bar{\boldsymbol{\omega}}) = 0$.

With definitions (33) and (34), Vlasov equation (36) and relations (37) and (38) can be written more compactly:

$$\frac{df}{d\tau} + (\bar{\mathbf{V}} \cdot \bar{\boldsymbol{\nabla}}) f + \{ \bar{\mathbf{K}}_f + \bar{\mathbf{p}} \times \bar{\boldsymbol{\omega}} + q (\bar{\mathbf{E}} + \bar{\mathbf{V}} \times \bar{\mathbf{B}}) \} \cdot \frac{\partial f}{\partial \bar{\mathbf{p}}} = - \frac{f - f_0}{\bar{\tau}_r}, \quad (40)$$

with

$$\bar{\mathbf{K}}_f = -\bar{\gamma} m (\Gamma + 1) \left(\frac{d}{d\tau} + \bar{\mathbf{V}} \cdot \bar{\boldsymbol{\nabla}} \right) \left(\frac{\mathbf{U}}{\Gamma + 1} \right), \quad \bar{\boldsymbol{\omega}} = (\Gamma - 1) \hat{\boldsymbol{\ell}} \times \left(\frac{d\hat{\boldsymbol{\ell}}}{d\tau} + (\bar{\mathbf{V}} \cdot \bar{\boldsymbol{\nabla}}) \hat{\boldsymbol{\ell}} \right). \quad (41)$$

Version (40) of the Vlasov equation is the equation that we will use in our further analysis, as it allows direct comparison with known (non-relativistic) results.

4.2 Liouville's theorem

By introducing the seven-dimensional state vector in phase-space $\tilde{\mathbf{Z}}$ as $\tilde{\mathbf{Z}} \equiv (x^\mu, -p_i) = (t, \mathbf{x}, \bar{\mathbf{p}})$, equation (40) can be written as

$$\dot{\tilde{\mathbf{Z}}} \cdot \frac{\partial f(\tilde{\mathbf{Z}})}{\partial \tilde{\mathbf{Z}}} = \bar{\text{St}}(f) = - \left(\frac{f - f_0}{\bar{\tau}_r} \right). \quad (42)$$

Here we define

$$\dot{\tilde{\mathbf{Z}}} = \frac{1}{\bar{\gamma}} \frac{d\tilde{\mathbf{Z}}}{d\lambda} = (i, \dot{\mathbf{x}}, \dot{\bar{\mathbf{p}}}) \quad (43)$$

with

$$\begin{aligned} i &= \frac{1}{\bar{\gamma}} \frac{dt}{d\lambda} = \Gamma [1 + (\bar{\mathbf{V}} \cdot \mathbf{V})] = \frac{\gamma}{\bar{\gamma}}, \\ \dot{\mathbf{x}} &= \frac{1}{\bar{\gamma}} \frac{d\mathbf{x}}{d\lambda} = \bar{\mathbf{V}} + \mathbf{U} + \frac{(\bar{\mathbf{V}} \cdot \mathbf{U})}{\Gamma + 1} \mathbf{U} = \frac{\mathbf{p}}{\bar{\gamma} m}, \\ \dot{\bar{\mathbf{p}}} &= \frac{1}{\bar{\gamma}} \frac{d\bar{\mathbf{p}}}{d\lambda} = \bar{\mathbf{K}}_f + \bar{\mathbf{p}} \times \bar{\boldsymbol{\omega}} + q (\bar{\mathbf{E}} + \bar{\mathbf{V}} \times \bar{\mathbf{B}}) \equiv \bar{\mathbf{K}} + q (\bar{\mathbf{E}} + \bar{\mathbf{V}} \times \bar{\mathbf{B}}). \end{aligned} \quad (44)$$

In order to ensure particle conservation after integration over the invariant phase-space element $d^3\bar{\mathbf{p}}/\bar{\mathcal{E}} = d^3\bar{\mathbf{p}}/\bar{\gamma}m$ one should be able to write equation (42) in conservation form:

$$\frac{\partial}{\partial \tilde{\mathbf{Z}}} \cdot \left(\dot{\tilde{\mathbf{Z}}} f(\tilde{\mathbf{Z}}) \right) = - \left(\frac{f - f_0}{\bar{\tau}_r} \right). \quad (45)$$

This implies that the flow in phase space is incompressible (Liouville's theorem):

$$\frac{\partial}{\partial \tilde{\mathbf{Z}}} \cdot \dot{\tilde{\mathbf{Z}}} \equiv \text{div}_{\tilde{\mathbf{Z}}} \dot{\tilde{\mathbf{Z}}} = 0. \quad (46)$$

For the original (covariant) equation (24) proving the equivalence of the two formulations is simple. The equations of motion in phase space can be derived from a Hamiltonian $\mathcal{H}(x, p)$ and the Vlasov equation has a symplectic structure, meaning that the transition from the non-conservative form to the conservation form generates a term that vanishes trivially:

$$-\left(\frac{\partial^2 \mathcal{H}}{\partial x^\mu \partial p_\mu} - \frac{\partial^2 \mathcal{H}}{\partial p_\mu \partial x^\mu}\right) \mathcal{F}(x, p) = 0 \quad (47)$$

When one uses mixed phase-space coordinates $\tilde{\mathbf{Z}} = (t, \mathbf{x}, \tilde{\mathbf{p}})$, the proof is more difficult. It is therefore shown in detail.

4.2.1 Liouville's theorem for mixed coordinates

The space-time part of $\text{div}_Z \dot{\tilde{\mathbf{Z}}}$ is

$$\left(\text{div}_Z \dot{\tilde{\mathbf{Z}}}\right)_{\text{st}} = \frac{\partial \dot{x}^\mu}{\partial x^\mu} = \frac{\partial \Gamma}{\partial t} + \nabla \cdot \mathbf{U} + \left(\frac{\partial}{\partial t} + \frac{\mathbf{U} \cdot \nabla}{\Gamma + 1}\right) (\tilde{\mathbf{V}} \cdot \mathbf{U}) + (\tilde{\mathbf{V}} \cdot \mathbf{U}) \left[\nabla \cdot \left(\frac{\mathbf{U}}{\Gamma + 1}\right)\right]. \quad (48)$$

Using the definition for $\tilde{\mathbf{V}}$, the relation

$$\frac{\partial \Gamma}{\partial t} + \nabla \cdot \mathbf{U} = (\Gamma + 1) \tilde{\mathbf{V}} \cdot \left(\frac{\mathbf{U}}{\Gamma + 1}\right) \quad (49)$$

and the relation valid for an infinitesimal change $\delta \mathbf{U}$

$$\mathbf{U} \cdot \delta \left(\frac{\mathbf{U}}{\Gamma + 1}\right) = \frac{\delta \Gamma}{\Gamma + 1}, \quad (50)$$

one finds that one can write:

$$\left(\text{div}_Z \dot{\tilde{\mathbf{Z}}}\right)_{\text{st}} = \left(1 + \frac{(\tilde{\mathbf{V}} \cdot \mathbf{U})}{\Gamma + 1}\right) \left\{(\Gamma + 1) \tilde{\mathbf{V}} \cdot \left(\frac{\mathbf{U}}{\Gamma + 1}\right)\right\} + (\Gamma + 1) \left(\frac{\partial}{\partial t} + \frac{\mathbf{U} \cdot \nabla}{\Gamma + 1}\right) \left(\frac{(\tilde{\mathbf{V}} \cdot \mathbf{U})}{\Gamma + 1}\right). \quad (51)$$

The momentum part of $\text{div}_Z \dot{\tilde{\mathbf{Z}}}$ is (in an obvious notation)

$$\left(\text{div}_Z \dot{\tilde{\mathbf{Z}}}\right)_{\text{mom}} = \frac{\partial \dot{p}_i}{\partial p_i} = \frac{\partial}{\partial \tilde{\mathbf{p}}} \cdot (\tilde{\mathbf{K}}_t + \tilde{\mathbf{p}} \times \tilde{\boldsymbol{\omega}}). \quad (52)$$

The term in $\tilde{\mathbf{p}}$ involving the Lorentz force does not contribute: Its momentum divergence vanishes trivially. We write $\tilde{\boldsymbol{\omega}}$ as

$$\tilde{\boldsymbol{\omega}} = \mathbf{U} \times \left(\frac{d}{d\tau} + \tilde{\mathbf{V}} \cdot \tilde{\nabla}\right) \left(\frac{\mathbf{U}}{\Gamma + 1}\right) = \left(\frac{\mathbf{U}}{\tilde{\gamma}}\right) \times \frac{d}{d\lambda} \left(\frac{\mathbf{U}}{\Gamma + 1}\right) \quad (53)$$

and expand the double cross-product:

$$\tilde{\mathbf{p}} \times \tilde{\boldsymbol{\omega}} = m \left[\tilde{\mathbf{V}} \cdot \left(\frac{d}{d\lambda} \left(\frac{\mathbf{U}}{\Gamma + 1}\right)\right)\right] \mathbf{U} - m (\tilde{\mathbf{V}} \cdot \mathbf{U}) \frac{d}{d\lambda} \left(\frac{\mathbf{U}}{\Gamma + 1}\right). \quad (54)$$

One then finds

$$\left(\text{div}_Z \dot{\tilde{\mathbf{Z}}}\right)_{\text{mom}} = - \left(1 + \frac{(\tilde{\mathbf{V}} \cdot \mathbf{U})}{\Gamma + 1}\right) \left\{(\Gamma + 1) \tilde{\mathbf{V}} \cdot \left(\frac{\mathbf{U}}{\Gamma + 1}\right)\right\} + \tilde{\mathbf{V}} \cdot \mathbf{U} : \tilde{\nabla} \left(\frac{\mathbf{U}}{\Gamma + 1}\right) - (\Gamma + 1) \tilde{\mathbf{V}} \cdot \frac{d}{d\tau} \left(\frac{\mathbf{U}}{\Gamma + 1}\right). \quad (55)$$

The first term in this expression cancels the corresponding term in the space-time part $\left(\text{div}_Z \dot{\tilde{\mathbf{Z}}}\right)_{\text{st}}$. For the two remaining terms we can proceed as follows. Use the definitions $d/d\tau = \Gamma \partial/\partial t + \mathbf{U} \cdot \nabla$, $d/d\lambda = \tilde{\gamma} d/d\tau + \tilde{\gamma} \tilde{\mathbf{V}} \cdot \tilde{\nabla}$ together with the fact that all derivatives are with $\tilde{\mathbf{p}}$ (and therefore $\tilde{\mathbf{V}}$) kept constant. This allows us to write

$$\tilde{\mathbf{V}} \cdot \mathbf{U} : \tilde{\nabla} \left(\frac{\mathbf{U}}{\Gamma + 1}\right) = \mathbf{U} \cdot \tilde{\nabla} (\tilde{\mathbf{V}} \cdot \mathbf{U}), \quad \tilde{\mathbf{V}} \cdot \frac{d}{d\tau} \left(\frac{\mathbf{U}}{\Gamma + 1}\right) = \frac{d}{d\tau} \left(\frac{(\tilde{\mathbf{V}} \cdot \mathbf{U})}{\Gamma + 1}\right). \quad (56)$$

The definition of $\tilde{\mathbf{V}}$ allows one to combine the second and third term in expression (55) and manipulate them into

$$\tilde{\mathbf{V}} \cdot \mathbf{U} : \tilde{\nabla} \left(\frac{\mathbf{U}}{\Gamma + 1}\right) - (\Gamma + 1) \tilde{\mathbf{V}} \cdot \frac{d}{d\tau} \left(\frac{\mathbf{U}}{\Gamma + 1}\right) = -(\Gamma + 1) \left(\frac{\partial}{\partial t} + \frac{\mathbf{U} \cdot \nabla}{\Gamma + 1}\right) \left(\frac{(\tilde{\mathbf{V}} \cdot \mathbf{U})}{\Gamma + 1}\right). \quad (57)$$

Substituting this into equation (55) quickly establishes that

$$\left(\text{div}_Z \dot{\tilde{\mathbf{Z}}}\right)_{\text{mom}} = - \left(\text{div}_Z \dot{\tilde{\mathbf{Z}}}\right)_{\text{st}} \iff \text{div}_Z \dot{\tilde{\mathbf{Z}}} = 0 \quad (58)$$

and Liouville's theorem holds for the Vlasov equation in mixed variables. The fact that the Vlasov equation in mixed variables can be written in conservation form will prove useful later, in particular in Paper 2.

5 GYRO-AVERAGED EQUATION

In order to compare our results with the results obtained for a non-relativistic bulk flow by Skilling (1975; see also Jones 1990) and Kulsrud (1983) we first consider the gyro-averaged Vlasov equation. The basic idea is that the magnetic term in the Lorentz force (due to a large-scale regular magnetic field) is dominant. If the field equals $\bar{\mathbf{B}}_0$ in the FRF, the corresponding term in the Lorentz force is, defining the gyro-frequency $\bar{\Omega} = q\bar{B}_0/\gamma m$, including sign of charge, and the gyro-angle $\bar{\phi}$,

$$(\bar{\mathbf{V}} \times \bar{\mathbf{B}}_0) \cdot \frac{\partial f}{\partial \bar{\mathbf{p}}} = -\bar{\Omega} \frac{\partial f}{\partial \bar{\phi}}. \quad (59)$$

If the gyro-period $P_g = 2\pi/|\bar{\Omega}|$ is much shorter than all other relevant time-scales and the gyro-radius $r_g \sim \bar{v}/|\bar{\Omega}|$ is much smaller than all relevant length-scales, one may assume that the distribution is *gyrotropic* with $\partial\langle f \rangle/\partial\bar{\phi}$ vanishingly small to lowest order in $1/\bar{\Omega}$. One can then average the Vlasov equation over gyro-angle. The averaging is most easily performed by erecting a Cartesian coordinate system (x, y, z) with unit base vectors $\hat{\mathbf{x}}, \hat{\mathbf{y}}$, and $\hat{\mathbf{z}}$ in the FRF, choosing $\bar{\mathbf{B}}_0$ locally along that z -axis and writing the FRF three-momentum as

$$\bar{\mathbf{p}} = \bar{p}_\perp (\cos \bar{\phi} \hat{\mathbf{x}} + \sin \bar{\phi} \hat{\mathbf{y}}) + \bar{p}_\parallel \hat{\mathbf{z}}. \quad (60)$$

The momentum pitch angle $\bar{\theta}$ is defined in the usual manner by $\bar{\mathbf{p}} \cdot \hat{\mathbf{b}} = \bar{p}_\parallel = \bar{p} \bar{\mu}$, with $\bar{\mu} = \cos \bar{\theta}$, so that $\bar{p}_\perp = \bar{p} \sin \bar{\theta} = \bar{p} \sqrt{1 - \bar{\mu}^2}$. The gyro-averaged distribution is

$$\bar{f}(\mathbf{x}, t, \bar{p}, \bar{\mu}) \equiv \frac{1}{2\pi} \int_0^{2\pi} d\bar{\phi} f(\mathbf{x}, t, \bar{\mathbf{p}}). \quad (61)$$

The variable change, from $\bar{\mathbf{p}}$ to $\bar{p}, \bar{\mu}$, and $\bar{\phi}$, implies that one has to take account of the variation of $\hat{\mathbf{b}} = \bar{\mathbf{B}}_0/\bar{B}_0$, the unit vector along the FRF magnetic field. A change in $\hat{\mathbf{b}}$ affects the definition of the pitch angle cosine $\bar{\mu}$. In particular:

$$\left(\frac{\partial}{\partial t} \right)_{\bar{\mathbf{p}}} \implies \left(\frac{\partial}{\partial t} \right)_{\bar{p}, \bar{\mu}} + \left(\frac{\bar{\mathbf{p}} \cdot \partial \hat{\mathbf{b}}}{\bar{p}} \right) \frac{\partial}{\partial \bar{\mu}}, \quad \nabla_{\bar{\mathbf{p}}} \implies \nabla_{\bar{p}, \bar{\mu}} + \frac{(\bar{\mathbf{p}} \cdot \nabla) \hat{\mathbf{b}}}{\bar{p}} \frac{\partial}{\partial \bar{\mu}}. \quad (62)$$

The gyro-averaged Vlasov equation for $\bar{f}(\mathbf{x}, t, \bar{p}, \bar{\mu})$ reads:

$$\frac{d\bar{f}}{d\tau} + \bar{v} \bar{\mu} (\hat{\mathbf{b}} \cdot \tilde{\nabla}) \bar{f} + \bar{\mu} \frac{\partial \bar{f}}{\partial \bar{\mu}} + \dot{\bar{p}} \frac{\partial \bar{f}}{\partial \bar{p}} = -\frac{\bar{f} - f_0}{\bar{\tau}_r}. \quad (63)$$

Here

$$\begin{aligned} \dot{\bar{\mu}} \equiv & \frac{1 - \bar{\mu}^2}{2} \left\{ \bar{v} (\tilde{\nabla} \cdot \hat{\mathbf{b}}) - \frac{2(\Gamma + 1)}{\bar{v}} \left[\hat{\mathbf{b}} \cdot \frac{d}{d\tau} \left(\frac{\mathbf{U}}{\Gamma + 1} \right) \right] - (\Gamma + 1) \bar{\mu} \left[3 \hat{\mathbf{b}} \hat{\mathbf{b}} : \tilde{\nabla} \left(\frac{\mathbf{U}}{\Gamma + 1} \right) - \tilde{\nabla} \cdot \left(\frac{\mathbf{U}}{\Gamma + 1} \right) \right] \right. \\ & \left. + (\Gamma - 1) \bar{v} \left[(\hat{\mathbf{b}} \cdot \hat{\mathbf{e}}) (\tilde{\nabla} \cdot \hat{\mathbf{e}}) - \hat{\mathbf{b}} \hat{\mathbf{e}} : \tilde{\nabla} \hat{\mathbf{e}} \right] \right\} \end{aligned} \quad (64)$$

and

$$\dot{\bar{p}} \equiv (\Gamma + 1) \left\{ -\frac{\bar{\mu}}{\bar{v}} \left[\hat{\mathbf{b}} \cdot \frac{d}{d\tau} \left(\frac{\mathbf{U}}{\Gamma + 1} \right) \right] + \frac{1 - 3\bar{\mu}^2}{2} \left[\hat{\mathbf{b}} \hat{\mathbf{b}} : \tilde{\nabla} \left(\frac{\mathbf{U}}{\Gamma + 1} \right) - \frac{1}{3} \tilde{\nabla} \cdot \left(\frac{\mathbf{U}}{\Gamma + 1} \right) \right] - \frac{1}{3} \left[\tilde{\nabla} \cdot \left(\frac{\mathbf{U}}{\Gamma + 1} \right) \right] \right\} \bar{p}. \quad (65)$$

We employ the notation $\hat{\mathbf{b}} \hat{\mathbf{b}} : \tilde{\nabla} \mathbf{A} = \hat{b}^i \hat{b}^j \tilde{\nabla}_j A^i$ for a double contraction between three-vectors and the gradient operator. We assumed for simplicity that there are no electromagnetic fields save for the large-scale magnetic field $\bar{\mathbf{B}}_0$ in the FRF. This corresponds to the limit of magnetohydrodynamics (see below).

This result generalizes the result of Skilling (1975) to a relativistic flow. His result can be obtained by taking the limit $|\mathbf{U}| \simeq |\mathbf{V}| \ll 1$, $\Gamma \simeq 1$ and neglecting all terms of order $|\mathbf{V}|^2$ and higher. This implies

$$\tilde{\nabla} \simeq \nabla + \mathbf{V} \frac{\partial}{\partial t}, \quad \frac{d}{d\tau} \simeq \frac{\partial}{\partial t} + \mathbf{V} \cdot \nabla \equiv \frac{d}{dt}. \quad (66)$$

Using these approximations in the expressions (64) and (65) and substituting the results into (63) leads to

$$\begin{aligned} & \left\{ 1 + \bar{v} \bar{\mu} (\hat{\mathbf{b}} \cdot \mathbf{V}) \right\} \frac{\partial \bar{f}}{\partial t} + (\mathbf{V} + \bar{v} \bar{\mu} \hat{\mathbf{b}}) \cdot \nabla \bar{f} \\ & + \left\{ \bar{v} \left(\nabla \cdot \hat{\mathbf{b}} + \mathbf{V} \cdot \frac{\partial \hat{\mathbf{b}}}{\partial t} \right) + \bar{\mu} (\nabla \cdot \mathbf{V}) + \frac{1 - 3\bar{\mu}^2}{2} (\hat{\mathbf{b}} \hat{\mathbf{b}} : \nabla \mathbf{V}) - \frac{2}{\bar{v}} \left(\hat{\mathbf{b}} \cdot \frac{d\mathbf{V}}{dt} \right) \right\} \frac{1 - \bar{\mu}^2}{2} \frac{\partial \bar{f}}{\partial \bar{\mu}} \\ & + \left\{ \frac{1 - 3\bar{\mu}^2}{2} (\hat{\mathbf{b}} \hat{\mathbf{b}} : \nabla \mathbf{V}) - \frac{1 - \bar{\mu}^2}{2} (\nabla \cdot \mathbf{V}) - \frac{\bar{\mu}}{\bar{v}} \left(\hat{\mathbf{b}} \cdot \frac{d\mathbf{V}}{dt} \right) \right\} \bar{p} \frac{\partial \bar{f}}{\partial \bar{p}} = -\frac{\bar{f} - f_0}{\bar{\tau}_r}. \end{aligned} \quad (67)$$

This agrees exactly with Skilling's result (equation 4 in Skilling 1975), except for a different choice for the scattering operator on the right-hand side of this equation. In the term in front of $\bar{p} (\partial \bar{f} / \partial \bar{p})$ Skilling makes the additional approximation $d\mathbf{V}/dt \simeq \partial \mathbf{V} / \partial t$. While this is formally correct since the convective $(\mathbf{V} \cdot \nabla) \mathbf{V}$ term in $d\mathbf{V}/dt \equiv \partial \mathbf{V} / \partial t + (\mathbf{V} \cdot \nabla) \mathbf{V}$ is of order $|\mathbf{V}|^2$, both terms in $d\mathbf{V}/dt$ are usually of similar magnitude in realistic (magneto)hydrodynamic flows. Therefore, in practical applications both should be retained (or discarded).

6 RELATIVISTIC DRIFT KINETIC EQUATION

Kulsrud (1983) employs $\bar{p}_\perp$ and $\bar{p}_\parallel$ as fundamental momentum variables. In addition he allows for a *weak* quasi-static electric field $\bar{E}_\parallel$ along the magnetic field in the FRF, corresponding to a *small* deviation from the MHD approximation for the bulk flow. The gyro-averaged equation for $\bar{f}(\bar{p}_\perp, \bar{p}_\parallel, \mathbf{x}, t)$ in these variables formally reads (with $\bar{v}_\parallel = \bar{v}\bar{\mu}$)

$$\frac{d\bar{f}}{d\tau} + \bar{v}_\parallel (\hat{\mathbf{b}} \cdot \bar{\nabla}) \bar{f} + \dot{\bar{p}}_\perp \frac{\partial \bar{f}}{\partial \bar{p}_\perp} + (\dot{\bar{p}}_\parallel + q \bar{E}_\parallel) \frac{\partial \bar{f}}{\partial \bar{p}_\parallel} = -\frac{\bar{f} - f_0}{\bar{\tau}_r}. \quad (68)$$

We have separated out the Coulomb force $q \bar{E}_\parallel$ in the parallel force term for clarity.

The quantities $\dot{\bar{p}}_\perp$ and $\dot{\bar{p}}_\parallel$ follow from equations (64) and (65) as

$$\begin{aligned} \dot{\bar{p}}_\perp &= \sqrt{1 - \bar{\mu}^2} \dot{\bar{p}} - \frac{\bar{p}_\perp \dot{\bar{\mu}}}{\sqrt{1 - \bar{\mu}^2}} = (\Gamma + 1) \frac{\bar{p}_\perp}{2} \left\{ \hat{\mathbf{b}} \hat{\mathbf{b}} : \bar{\nabla} \left(\frac{\mathbf{U}}{\Gamma + 1} \right) - \bar{\nabla} \cdot \left(\frac{\mathbf{U}}{\Gamma + 1} \right) \right\} \\ &\quad - \frac{\bar{p}_\perp \bar{v}_\parallel}{2} \left\{ (\bar{\nabla} \cdot \hat{\mathbf{b}}) + (\Gamma - 1) [(\hat{\mathbf{b}} \cdot \hat{\mathbf{l}})(\bar{\nabla} \cdot \hat{\mathbf{l}}) - \hat{\mathbf{b}} \hat{\mathbf{l}} : \bar{\nabla} \hat{\mathbf{l}}] \right\}. \end{aligned} \quad (69)$$

and

$$\begin{aligned} \dot{\bar{p}}_\parallel &\equiv \bar{\mu} \dot{\bar{p}} + \bar{p} \dot{\bar{\mu}} = -(\Gamma + 1) \bar{\gamma} m \left\{ \hat{\mathbf{b}} \cdot \left[\frac{d}{d\tau} \left(\frac{\mathbf{U}}{\Gamma + 1} \right) + \bar{v}_\parallel (\hat{\mathbf{b}} \cdot \bar{\nabla}) \left(\frac{\mathbf{U}}{\Gamma + 1} \right) \right] \right\} \\ &\quad + \frac{\bar{p}_\perp \bar{v}_\perp}{2} \left\{ (\bar{\nabla} \cdot \hat{\mathbf{b}}) + (\Gamma - 1) [(\hat{\mathbf{b}} \cdot \hat{\mathbf{l}})(\bar{\nabla} \cdot \hat{\mathbf{l}}) - \hat{\mathbf{b}} \hat{\mathbf{l}} : \bar{\nabla} \hat{\mathbf{l}}] \right\}. \end{aligned} \quad (70)$$

These results can be simplified if we assume that – to lowest order – relativistic MHD applies when describing the large-scale magnetic field. In that case one can derive an equation for the evolution of $B = |\bar{\mathbf{B}}_0|$, the strength of the magnetic field in the FRF. This is in fact a Lorentz-invariant quantity that follows from $F^{\mu\nu} F_{\mu\nu} = 2(|\mathbf{E}|^2 - |\mathbf{B}|^2) \simeq -2B^2$ as long as $|\bar{E}_\parallel| \ll B$ as is assumed here. As shown in Appendix C we have

$$\frac{1}{B} \frac{dB}{d\tau} = (\Gamma + 1) \left\{ \hat{\mathbf{b}} \hat{\mathbf{b}} : \bar{\nabla} \left(\frac{\mathbf{U}}{\Gamma + 1} \right) - \bar{\nabla} \cdot \left(\frac{\mathbf{U}}{\Gamma + 1} \right) \right\}, \quad (71)$$

and

$$\bar{\nabla} \cdot \hat{\mathbf{b}} = -\frac{(\hat{\mathbf{b}} \cdot \bar{\nabla})B}{B} + (\Gamma + 1) \left\{ \hat{\mathbf{b}} \cdot \frac{d}{d\tau} \left(\frac{\mathbf{U}}{\Gamma + 1} \right) \right\}. \quad (72)$$

The second term in this last expression is a relativistic correction that occurs when the FRF is not an inertial frame. This allows us to write

$$\dot{\bar{p}}_\perp = \frac{\bar{p}_\perp}{2B} \left(\frac{dB}{d\tau} + \bar{v}_\parallel (\hat{\mathbf{b}} \cdot \bar{\nabla})B \right) - \nu_p \frac{\bar{p}_\perp \bar{v}_\perp}{2}, \quad (73)$$

and

$$\dot{\bar{p}}_\parallel = -\frac{\bar{p}_\perp \bar{v}_\perp}{2B} (\hat{\mathbf{b}} \cdot \bar{\nabla})B - (\Gamma + 1) \bar{\gamma} m \hat{\mathbf{b}} \cdot \left\{ \frac{d}{d\tau} \left(\frac{\mathbf{U}}{\Gamma + 1} \right) + \bar{v}_\parallel (\hat{\mathbf{b}} \cdot \bar{\nabla}) \left(\frac{\mathbf{U}}{\Gamma + 1} \right) \right\} + \nu_p \frac{\bar{p}_\perp \bar{v}_\perp}{2}. \quad (74)$$

Here ν_p is a frequency that is essentially a relativistic correction, given by

$$\nu_p \equiv (\Gamma + 1) \left[\hat{\mathbf{b}} \cdot \frac{d}{d\tau} \left(\frac{\mathbf{U}}{\Gamma + 1} \right) \right] + (\Gamma - 1) [(\hat{\mathbf{b}} \cdot \hat{\mathbf{l}})(\bar{\nabla} \cdot \hat{\mathbf{l}}) - \hat{\mathbf{b}} \hat{\mathbf{l}} : \bar{\nabla} \hat{\mathbf{l}}]. \quad (75)$$

6.1 Non-relativistic limit and conservation of magnetic moment

If we now take the limit of a non-relativistic flow and assume non-relativistic particles so that $\bar{v} \ll 1$, we can make the replacements (equation 66) and in addition neglect the terms involving ν_p . In that case equation (68) takes the form derived by Kulsrud (1983), in his equation (51), after substituting the non-relativistic equivalent of (72),

$$\nabla \cdot \hat{\mathbf{b}} = -\frac{(\hat{\mathbf{b}} \cdot \nabla)B}{B}, \quad (76)$$

and adding the scattering term. One finds

$$\frac{D\bar{f}}{Dt} + \left(\frac{\bar{p}_\perp}{2B} \frac{DB}{Dt} \right) \frac{\partial \bar{f}}{\partial \bar{p}_\perp} + \left(q \bar{E}_\parallel - \frac{\bar{p}_\perp \bar{v}_\perp}{2B} (\hat{\mathbf{b}} \cdot \nabla)B - m \hat{\mathbf{b}} \cdot \frac{D\mathbf{V}}{Dt} \right) \frac{\partial \bar{f}}{\partial \bar{p}_\parallel} = -\frac{\bar{f} - f_0}{\bar{\tau}_r}. \quad (77)$$

Here the operator

$$\frac{D}{Dt} \equiv \frac{\partial}{\partial t} + (\mathbf{V} + \bar{v}_\parallel \hat{\mathbf{b}}) \cdot \nabla \quad (78)$$

is the comoving time derivative along the orbit of the guiding centre of a particle, which includes the motion along a field line.

The fully relativistic version can be cast in an analogous form by defining the relativistic generalization of (78):

$$\frac{D}{D\tau} \equiv \frac{d}{d\tau} + \bar{v}_{\parallel} (\hat{\mathbf{b}} \cdot \tilde{\nabla}). \quad (79)$$

Equation (68) then becomes

$$\frac{D\tilde{f}}{D\tau} + \left\{ \frac{\bar{p}_{\perp}}{2B} \frac{DB}{D\tau} - \nu_p \frac{\bar{p}_{\perp} \bar{v}_{\parallel}}{2} \right\} \frac{\partial \tilde{f}}{\partial \bar{p}_{\perp}} + \left\{ -\frac{\bar{p}_{\perp} \bar{v}_{\perp}}{2B} (\hat{\mathbf{b}} \cdot \tilde{\nabla})B - (\Gamma + 1) \tilde{\gamma} m \hat{\mathbf{b}} \cdot \frac{D}{D\tau} \left(\frac{\mathbf{U}}{\Gamma + 1} \right) + \nu_p \frac{\bar{p}_{\perp} \bar{v}_{\perp}}{2} + q \bar{E}_{\parallel} \right\} \frac{\partial \tilde{f}}{\partial \bar{p}_{\parallel}} = -\frac{\tilde{f} - f_0}{\bar{\tau}_r}. \quad (80)$$

In the special case that the FRF is an inertial frame so that $\nu_p = 0$ and $D(\mathbf{U}/\Gamma + 1)/D\tau = 0$ and in the absence of scattering equation (80) simplifies to

$$\frac{D\tilde{f}}{D\tau} + \left(\frac{\bar{p}_{\perp}}{2B} \frac{DB}{D\tau} \right) \frac{\partial \tilde{f}}{\partial \bar{p}_{\perp}} + \left(q \bar{E}_{\parallel} - \frac{\bar{p}_{\perp} \bar{v}_{\perp}}{2B} (\hat{\mathbf{b}} \cdot \tilde{\nabla})B \right) \frac{\partial \tilde{f}}{\partial \bar{p}_{\parallel}} = 0. \quad (81)$$

This is equivalent with

$$\frac{D\bar{p}_{\perp}}{D\tau} = \frac{\bar{p}_{\perp}}{2B} \frac{DB}{D\tau} \iff \frac{\bar{p}_{\perp}^2}{2B} = \text{constant}, \quad \frac{D\bar{p}_{\parallel}}{D\tau} = q \bar{E}_{\parallel} - \frac{\bar{p}_{\perp} \bar{v}_{\perp}}{2B} (\hat{\mathbf{b}} \cdot \tilde{\nabla})B. \quad (82)$$

The first relation respectively corresponds with the adiabatic invariance of $\bar{p}_{\perp}^2/2B$ if gyro-motion is sufficiently rapid as assumed here. The second term on the right-hand side of the second relation is the magnetic mirror force that occurs under the same circumstances if the strength of the large-scale field varies (see for instance Melrose 1989, ch. 12).

7 LINK WITH WEBB'S APPROACH

The above derivation of the transport equation (40) is more intuitive than that of Webb (1985). In addition there are a number of notational differences with Webb's paper that are designed to make a direct comparison with the previous results of Skilling (1975) and Kulsrud (1983) more simple. Other than that our results are equivalent with Webb's, as we show briefly here.

In both cases the calculation is based on the evaluation of the fictitious force (9). That expression (for flat space-time) can be written as

$$\frac{dp_{\hat{c}}}{d\lambda} = \omega_{\hat{b}\hat{c}}^{\hat{a}} \frac{p_{\hat{a}} p^{\hat{b}}}{m}, \quad \omega_{\hat{b}\hat{c}}^{\hat{a}} \equiv -E_{\hat{c}}^{\mu} E_{\hat{b}}^{\nu} \frac{\partial E_{\mu}^{\hat{a}}}{\partial x^{\nu}}, \quad (83)$$

which is equivalent to equation (6.4) in Webb (1985), who uses the notation $\Gamma_{\hat{c}\hat{b}}^{\hat{a}}$ for $\omega_{\hat{b}\hat{c}}^{\hat{a}}$, essentially taking the approach mentioned in Section 3 that treats the effects of acceleration as space-time curvature in the accelerated frame. The $\omega_{\hat{b}\hat{c}}^{\hat{a}}$ indeed play the role of the connection coefficients, as follows rigorously if one derives these results using the so-called tetrad formalism (see Appendix A). Using $p_{\hat{i}} = -p^{\hat{i}}$ for $\hat{i} = 1, 2, 3$ with the $p^{\hat{i}}$ the components of $\bar{\mathbf{p}}$, a straightforward calculation using equation (32) or equations (14) and (15) yields

$$\begin{aligned} \omega_{00}^{\hat{0}} &= 0, \\ \omega_{00}^{\hat{i}} &= \omega_{0i}^{\hat{0}} = (\Gamma + 1) \frac{d}{d\tau} \left(\frac{U^i}{\Gamma + 1} \right), \\ \omega_{j0}^{\hat{i}} &= \omega_{ji}^{\hat{0}} = (\Gamma + 1) \tilde{\nabla}_j \left(\frac{U^i}{\Gamma + 1} \right), \\ \omega_{0i}^{\hat{j}} &= \frac{1}{\Gamma + 1} \left(U^i \frac{dU^j}{d\tau} - U^j \frac{dU^i}{d\tau} \right) = (\Gamma - 1) \left(\hat{\ell}^i \frac{d\hat{\ell}^j}{d\tau} - \hat{\ell}^j \frac{d\hat{\ell}^i}{d\tau} \right) \\ \omega_{ki}^{\hat{j}} &= \frac{1}{\Gamma + 1} (U^i \tilde{\nabla}_k U^j - U^j \tilde{\nabla}_k U^i) = (\Gamma - 1) (\hat{\ell}^i \tilde{\nabla}_k \hat{\ell}^j - \hat{\ell}^j \tilde{\nabla}_k \hat{\ell}^i). \end{aligned} \quad (84)$$

Here we write the components of the unit three-vector $\hat{\ell} = \mathbf{U}/U$ as $\hat{\ell}^i$ with indices (i, j) and $(\hat{i}, \hat{j})$ run from 1 to 3. Using the definitions of $d/d\tau = \Gamma(d/dt)$, $\tilde{\nabla}$ (see equations 33 and 34) and $U^i = \Gamma V^i$ one can show that these expressions are equivalent to Webb's results, equation (6.9) in Webb (1985) which incidentally does not list all coefficients. We provide two examples: Webb lists $\omega_{0i}^{\hat{j}}$ (our notation) by the equivalent expression that follows from $U^i = \Gamma V^i$ and $d/d\tau = \Gamma(d/dt)$:

$$\omega_{0i}^{\hat{j}} = \frac{\Gamma^3}{(\Gamma + 1)c^3} \left(V^i \frac{dV^j}{dt} - V^j \frac{dV^i}{dt} \right), \quad (85)$$

where we reinstated c for the comparison. Webb also uses the relation (with the Einstein summation convention for double indices)

$$\omega_{i0}^{\hat{j}} = (\Gamma + 1) \left[\tilde{\nabla} \cdot \left(\frac{\mathbf{U}}{\Gamma + 1} \right) \right] = \frac{1}{c} \left(\frac{\partial \Gamma}{\partial t} + \nabla \cdot (\Gamma \mathbf{V}) \right). \quad (86)$$

8 THE GENERAL CASE: INCLUDING THE GYRO-PHASE DEPENDENCE

In order to fully describe such effects as cosmic ray (or photon) diffusion and cosmic ray viscosity we must calculate the *full* deviations of $f(\tilde{\mathbf{Z}})$ from isotropy in $\tilde{\mathbf{p}}$ -space, now including the dependence on gyro-phase that was neglected in earlier calculations and in Sections 5 and 6. For photons the concept of gyro-averaging makes no sense to begin with. To that end one must solve the Vlasov equation

$$\dot{\tilde{\mathbf{Z}}} \cdot \frac{\partial f(\tilde{\mathbf{Z}})}{\partial \tilde{\mathbf{Z}}} = i \frac{\partial f}{\partial t} + \dot{\mathbf{x}} \cdot \nabla f + \tilde{\mathbf{K}} \cdot \frac{\partial f}{\partial \tilde{\mathbf{p}}} - \bar{\Omega} \frac{\partial f}{\partial \tilde{\phi}} = - \frac{f - f_0}{\bar{\tau}_r}, \quad (87)$$

making an expansion for small anisotropies in the fluid rest frame. Here $\tilde{\mathbf{K}} = \tilde{\mathbf{K}}_f + \tilde{\mathbf{p}} \times \tilde{\boldsymbol{\omega}}$ (see equation 44). This is a natural assumption in a relativistic flow of a scattering medium. Generally speaking, this assumption does *not* mean that the anisotropy in the distribution is small in any other frame. If the fluid moves at relativistic speeds, this is true in particular in the laboratory frame, due to the effect of relativistic beaming.

8.1 The diffusion approximation and the choice for $f_0(\mathbf{x}, t, \tilde{\mathbf{p}})$

We assume that the gradient scale L and the time-scale T of the variations in $f(\tilde{\mathbf{Z}})$ and $\mathbf{U}(\mathbf{x}, t)$ satisfy

$$\frac{V \bar{\tau}_r}{L}, \quad \frac{\bar{v} \bar{\tau}_r}{L}, \quad \frac{V}{\bar{\Omega} L}, \quad \frac{\bar{v}}{\bar{\Omega} L}, \quad \frac{\bar{\tau}_r}{T}, \quad \frac{1}{\bar{\Omega} T} = \mathcal{O}(\epsilon) \ll 1. \quad (88)$$

This can be formalized by introducing ϵ as an ordering parameter (to be put equal to unity at the end of the calculation) and by making the replacements

$$\bar{\tau}_r \Rightarrow \epsilon \bar{\tau}_r, \quad \bar{\Omega} \Rightarrow \bar{\Omega}/\epsilon. \quad (89)$$

Note that this leaves the dimensionless parameter $\bar{\Omega} \bar{\tau}_r$ unchanged so that this parameter can take any value. Then equation (87) takes the form

$$\epsilon \left(i \frac{\partial f}{\partial t} + \dot{\mathbf{x}} \cdot \nabla f + \tilde{\mathbf{K}} \cdot \frac{\partial f}{\partial \tilde{\mathbf{p}}} \right) = \bar{\Omega} \frac{\partial f}{\partial \tilde{\phi}} + \bar{\text{St}}(f) = \bar{\Omega} \frac{\partial f}{\partial \tilde{\phi}} - \frac{f - f_0}{\bar{\tau}_r}. \quad (90)$$

One expands the Vlasov distribution $f(\tilde{\mathbf{Z}})$ as

$$f(\tilde{\mathbf{Z}}) = f_0(t, \mathbf{x}, |\tilde{\mathbf{p}}|) + \epsilon f_1(\tilde{\mathbf{Z}}) + \epsilon^2 f_2(\tilde{\mathbf{Z}}) + \dots \quad (91)$$

The leading term f_0 is isotropic in $\tilde{\mathbf{p}}$ -space while f_1, f_2 , etc. contain the effect of anisotropies. In the diffusion approximation (as is employed here) one breaks off expansion (equation 91) after the first-order term and never calculates $f_2(\tilde{\mathbf{Z}})$.

Equation (90) implies that $f_0(t, \mathbf{x}, |\tilde{\mathbf{p}}|)$ should be chosen as the average of the Vlasov distribution over solid angle in $\tilde{\mathbf{p}}$ -space to all orders in ϵ . It is the most general ‘equilibrium distribution’ that (i) annihilates the combined gyration/scattering operator on the right-hand side equation (90) and (ii) is consistent with the exact Vlasov equation (90). This implies

$$f_0(\mathbf{x}, t, \tilde{\mathbf{p}}) = \frac{1}{4\pi} \int_{-1}^{+1} d\bar{\mu} \int_0^{2\pi} d\bar{\phi} f(\mathbf{x}, t, \tilde{\mathbf{p}}) \equiv \langle f(\tilde{\mathbf{Z}}) \rangle, \quad (92)$$

where we introduce the notation $\langle \dots \rangle$ for the average over momentum solid angle. In the case of pure scattering, which is the case we consider here, relation (92) also follows from the fact that scattering merely redistributes particles over the solid angle $\bar{\Omega}_p$ in momentum space, so that the scattering operator $\bar{\text{St}}(f)$ (or $\text{St}(f) = \bar{\gamma} \bar{\text{St}}(f)$) must satisfy

$$\int_{4\pi} d\bar{\Omega}_p \bar{\text{St}}(f) = 0, \quad (93)$$

regardless of the precise form of the relaxation term $\bar{\text{St}}(f)$. We note in passing that, because of condition (93), the use of the algebraic Anderson–Witting term for the scattering operator does *not* suffer from the problem that it violates particle conservation. This problem does occur in relativistic theories such as this whenever one wants to use such a term to force the distribution into a certain distribution in $\tilde{\mathbf{E}} = \sqrt{\tilde{\mathbf{p}}^2 + m^2}$ or $\tilde{\mathbf{p}}$, such as a relativistic Maxwellian.

In the terminology of perturbation theory the evolution equation for f_0 (derived below) should contain all ‘secular’ effects. The higher-order terms f_1, f_2 , etc. are ‘oscillating terms’ that contain all anisotropies in momentum space. Because of equation (92) these terms satisfy $\langle f_n(\tilde{\mathbf{Z}}) \rangle = 0$ for $n \geq 1$: Their average over solid angle in $\tilde{\mathbf{p}}$ space vanishes. In particular one has $\langle f_1(\tilde{\mathbf{Z}}) \rangle = 0$, which will prove important in our discussion of particle diffusion (this paper) and viscosity (Paper 2) and even more so in the discussion of viscous stresses due to radiation in Paper 3.

As a result $f_0 = \langle f(\tilde{\mathbf{Z}}) \rangle$ will in principle include $\mathcal{O}(\epsilon)$ and higher-order corrections that can only be calculated by solving the self-consistent transport equation for f_0 . The subscript ‘0’ should therefore *not* be interpreted to mean that f_0 is of order $\mathcal{O}(\epsilon^0)$: It can be expanded as $f_0 = f_{00} + \epsilon f_{01} + \dots$. In practice this expansion is never needed.

One version of this transport equation for the average distribution (to first order in ϵ) is obtained by averaging equation (90) over solid angle:

$$\Gamma \left[\frac{\partial f_0}{\partial t} + (\mathbf{V} \cdot \nabla) f_0 \right] + \langle \tilde{\mathbf{K}}_f \cdot \hat{\mathbf{n}} \rangle \frac{\partial f_0}{\partial \tilde{\mathbf{p}}} = -\epsilon \left\{ \langle \tilde{\mathbf{V}} \cdot \tilde{\nabla} f_1 \rangle + \left\langle \tilde{\mathbf{K}} \cdot \frac{\partial f_1}{\partial \tilde{\mathbf{p}}} \right\rangle \right\} + \mathcal{O}(\epsilon^2). \quad (94)$$

Here we have transported all *explicit* order ϵ terms to the right-hand side, and defined $\tilde{\mathbf{K}} \equiv \bar{\mathbf{K}} - \langle \bar{\mathbf{K}} \rangle$, the oscillating part of the fictitious force. Also $\hat{\mathbf{n}} = \bar{\mathbf{p}}/\bar{p}$ is the unit vector along the FRF momentum and we have used $\partial f_0/\partial \bar{\mathbf{p}} = \hat{\mathbf{n}} (\partial f_0/\partial \bar{p})$. The second term on the left-hand side of equation (94) is the average change of $\bar{p}$ due to scattering by the flow. This term describes adiabatic losses/gains and was also found by Webb (1985). Using equation (49) it can be written as

$$\langle \bar{\mathbf{K}}_{\text{f}} \cdot \hat{\mathbf{n}} \rangle = -\frac{(\Gamma+1)\bar{p}}{3} \left[\tilde{\mathbf{v}} \cdot \left(\frac{\mathbf{U}}{\Gamma+1} \right) \right] = -\frac{\bar{p}}{3} \left(\frac{\partial \Gamma}{\partial t} + \nabla \cdot (\Gamma \mathbf{V}) \right) = -\frac{\bar{p}}{3} (\nabla \cdot \mathbf{U}). \quad (95)$$

The two terms $\mathcal{O}(\epsilon)$ terms on the right-hand side of equation (94) respectively lead to spatial diffusion plus a drift velocity (see Section 9) and to irreversible (viscous) heating of the scattered particles, as shown in Paper 2.

8.2 Solution for $f_1(\tilde{\mathbf{Z}})$

To first order in ϵ Vlasov equation (90) yields the following equation for f_1 :

$$\bar{\Omega} \frac{\partial f_1}{\partial \bar{\phi}} - \frac{f_1}{\bar{\tau}_r} = \frac{df_0}{d\tau} + \tilde{\mathbf{V}} \cdot \tilde{\mathbf{v}} f_0 + (\bar{\mathbf{K}}_{\text{f}} \cdot \hat{\mathbf{n}}) \frac{\partial f_0}{\partial \bar{p}}. \quad (96)$$

Using relation (94) to eliminate $df_0/d\tau$, neglecting the $\mathcal{O}(\epsilon)$ term so that

$$\frac{df_0}{d\tau} = \frac{(\Gamma+1)\bar{p}}{3} \left[\tilde{\mathbf{v}} \cdot \left(\frac{\mathbf{U}}{\Gamma+1} \right) \right] \bar{p} \frac{\partial f_0}{\partial \bar{p}} + \mathcal{O}(\epsilon), \quad (97)$$

one finds

$$\bar{\Omega} \frac{\partial f_1}{\partial \bar{\phi}} - \frac{f_1}{\bar{\tau}_r} = \tilde{\mathbf{V}} \cdot \tilde{\mathbf{v}} f_0 + \tilde{K}(\tilde{\mathbf{Z}}) \frac{\partial f_0}{\partial \bar{p}} + \mathcal{O}(\epsilon). \quad (98)$$

Here $\tilde{K}(\tilde{\mathbf{Z}})$ is the ‘oscillating’ part of the rate of change of $\bar{p}$, formally defined as $\tilde{K}(\tilde{\mathbf{Z}}) = \dot{\bar{p}} - \langle \dot{\bar{p}} \rangle$:

$$\tilde{K}(\tilde{\mathbf{Z}}) \equiv \bar{\mathbf{K}}_{\text{f}} \cdot \hat{\mathbf{n}} - \langle \bar{\mathbf{K}}_{\text{f}} \cdot \hat{\mathbf{n}} \rangle = \frac{\bar{p}}{3} \left[(\Gamma+1) \tilde{\mathbf{v}} \cdot \left(\frac{\mathbf{U}}{\Gamma+1} \right) \right] - (\Gamma+1)m \left(\hat{\mathbf{n}} \cdot \frac{d}{d\lambda} \left(\frac{\mathbf{U}}{\Gamma+1} \right) \right). \quad (99)$$

When written out in terms of the local Cartesian xyz coordinates introduced in Section 5, equation (98) reads

$$\begin{aligned} \bar{\Omega} \frac{\partial f_1}{\partial \bar{\phi}} - \frac{f_1}{\bar{\tau}_r} = & \bar{v} \left\{ \sqrt{1-\bar{\mu}^2} [\cos \bar{\phi} \hat{\mathbf{x}} + \sin \bar{\phi} \hat{\mathbf{y}}] + \bar{\mu} \hat{\mathbf{b}} \right\} \cdot \tilde{\mathbf{v}} f_0 \\ & - \bar{\gamma} m (\Gamma+1) \left\{ \left(\sqrt{1-\bar{\mu}^2} [\cos \bar{\phi} \hat{\mathbf{x}} + \sin \bar{\phi} \hat{\mathbf{y}}] + \bar{\mu} \hat{\mathbf{b}} \right) \cdot \frac{d}{d\tau} \left(\frac{\mathbf{U}}{\Gamma+1} \right) \right\} \frac{\partial f_0}{\partial \bar{p}} \\ & - (\Gamma+1) \left\{ \left[(1-\bar{\mu}^2) \mathbf{M}_1 + \bar{\mu} \sqrt{1-\bar{\mu}^2} \mathbf{M}_2 \right] : \tilde{\mathbf{v}} \left(\frac{\mathbf{U}}{\Gamma+1} \right) \right\} \bar{p} \frac{\partial f_0}{\partial \bar{p}} \\ & - (\Gamma+1) \left\{ \frac{3\bar{\mu}^2-1}{2} \left[\hat{\mathbf{b}} \hat{\mathbf{b}} : \tilde{\mathbf{v}} \left(\frac{\mathbf{U}}{\Gamma+1} \right) - \frac{1}{3} \tilde{\mathbf{v}} \cdot \left(\frac{\mathbf{U}}{\Gamma+1} \right) \right] \right\} \bar{p} \frac{\partial f_0}{\partial \bar{p}}. \end{aligned} \quad (100)$$

Here we have put $\bar{\mathbf{E}}_{\parallel} = 0$ for simplicity, used $\tilde{\mathbf{V}} = \bar{v} \left(\sqrt{1-\bar{\mu}^2} [\cos \bar{\phi} \hat{\mathbf{x}} + \sin \bar{\phi} \hat{\mathbf{y}}] + \bar{\mu} \hat{\mathbf{b}} \right)$ with $\hat{\mathbf{b}} = \hat{\mathbf{z}}$, together with the fact that f_0 is isotropic. The Coriolis term in the fictitious force does not contribute at this order as $|\bar{\omega}|/\bar{\Omega}$ and $|\bar{\omega}|/\bar{\tau}_r$ are small (first-order in ϵ) quantities under the assumptions of this derivation, and since f_0 is isotropic. The two 3×3 matrices $\mathbf{M}_1$ and $\mathbf{M}_2$ are formally defined as

$$\mathbf{M}_1 = \frac{1}{2} \begin{pmatrix} 2 \cos^2 \bar{\phi} - 1 & 2 \cos \bar{\phi} \sin \bar{\phi} & 0 \\ 2 \cos \bar{\phi} \sin \bar{\phi} & 2 \sin^2 \bar{\phi} - 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} \cos 2\bar{\phi} & \sin 2\bar{\phi} & 0 \\ \sin 2\bar{\phi} & -\cos 2\bar{\phi} & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \mathbf{M}_2 = \begin{pmatrix} 0 & 0 & \cos \bar{\phi} \\ 0 & 0 & \sin \bar{\phi} \\ \cos \bar{\phi} & \sin \bar{\phi} & 0 \end{pmatrix}. \quad (101)$$

The double contractions $\mathbf{M}_{1,2} : \tilde{\mathbf{v}}(\mathbf{U}/\Gamma+1)$ stand for $M_{1,2}^{ij} \tilde{v}_j(U^i/\Gamma+1)$. The gyro-average of $\mathbf{M}_{1,2}$ vanishes. This means that all terms on the right-hand side of equation (100) vanish when one averages over momentum solid angle in the FRF, which implies $\langle f_1 \rangle = 0$, as required.

Integration of equation (100) for $f_1(\mathbf{x}, t, \bar{\mathbf{p}})$ effectively corresponds to the replacement

$$\cos(m\bar{\phi}) \implies -\frac{\bar{\tau}_r}{1+m^2\bar{\Omega}^2\bar{\tau}_r^2} [\cos(m\bar{\phi}) - m\bar{\Omega}\bar{\tau}_r \sin(m\bar{\phi})], \quad \sin(m\bar{\phi}) \implies -\frac{\bar{\tau}_r}{1+m^2\bar{\Omega}^2\bar{\tau}_r^2} [\sin(m\bar{\phi}) + m\bar{\Omega}\bar{\tau}_r \cos(m\bar{\phi})], \quad (102)$$

for $m = 0, 1, 2$. Furthermore, the matrices $\mathbf{M}_1$ and $\mathbf{M}_2$ are both traceless and symmetric. This implies that in the double contraction with $\tilde{\mathbf{v}}[\mathbf{U}/(\Gamma+1)]$ we can make the replacement

$$\tilde{\mathbf{v}} \left(\frac{\mathbf{U}}{\Gamma+1} \right) \implies \frac{1}{2} \left\{ \tilde{\mathbf{v}} \left(\frac{\mathbf{U}}{\Gamma+1} \right) + \left[\tilde{\mathbf{v}} \left(\frac{\mathbf{U}}{\Gamma+1} \right) \right]^{\text{tr}} - \frac{2}{3} \left[\tilde{\mathbf{v}} \cdot \left(\frac{\mathbf{U}}{\Gamma+1} \right) \right] \mathbf{I} \right\}. \quad (103)$$

Here the superscript ‘tr’ stands for ‘transpose’ and $\mathbf{I} \equiv \text{diag}(1, 1, 1)$ is the unit 3×3 tensor in three-space. To facilitate notation we will define the symmetrized and traceless rate-of-strain tensor $\mathbf{W}$:

$$\mathbf{W} \equiv (\Gamma + 1) \left\{ \tilde{\nabla} \left(\frac{\mathbf{U}}{\Gamma + 1} \right) + \left[\tilde{\nabla} \left(\frac{\mathbf{U}}{\Gamma + 1} \right) \right]^{\text{tr}} - \frac{2}{3} \left[\tilde{\nabla} \cdot \left(\frac{\mathbf{U}}{\Gamma + 1} \right) \right] \mathbf{I} \right\}. \quad (104)$$

For a non-relativistic flow ($\Gamma \simeq 1$, $\mathbf{U} \simeq \mathbf{V}$ with $|\mathbf{V}| \ll 1$, $\tilde{\nabla} \simeq \nabla$) this reduces to the rate of strain tensor as defined by Landau & Lifshitz (1959) and Braginskii (1965):

$$\mathbf{W}_{\text{nr}} = \nabla \mathbf{V} + (\nabla \mathbf{V})^{\text{tr}} - \frac{2}{3} (\nabla \cdot \mathbf{V}) \mathbf{I}. \quad (105)$$

With definition (104) one has:

$$\mathbf{M}_{1,2} : \tilde{\nabla} \left(\frac{\mathbf{U}}{\Gamma + 1} \right) = \frac{1}{2} \mathbf{M}_{1,2} : \mathbf{W}, \quad \hat{\mathbf{b}} \hat{\mathbf{b}} : \tilde{\nabla} \left(\frac{\mathbf{U}}{\Gamma + 1} \right) - \frac{1}{3} \tilde{\nabla} \cdot \left(\frac{\mathbf{U}}{\Gamma + 1} \right) = \frac{1}{2} \hat{\mathbf{b}} \hat{\mathbf{b}} : \mathbf{W}. \quad (106)$$

It is also useful to define

$$\bar{\mathbf{a}} \equiv (\Gamma + 1) \frac{d}{d\tau} \left(\frac{\mathbf{U}}{\Gamma + 1} \right), \quad (107)$$

which corresponds to the non-vanishing spatial part of the four-acceleration of the fluid in the FRF: $a^{\hat{a}} = (0, , \bar{\mathbf{a}})$ (see Appendix C). One finds that

$$f_1(\mathbf{x}, t, \bar{\mathbf{p}}) = -\bar{\tau}_r (\bar{\mathbf{v}} \cdot \tilde{\nabla}) f_0 + \bar{\tau}_r (\tilde{\nabla} \cdot \bar{\mathbf{a}}) \bar{\gamma} m \frac{\partial f_0}{\partial \bar{p}} + \bar{\tau}_r (\hat{\mathbf{M}} : \mathbf{W}) \bar{p} \frac{\partial f_0}{\partial \bar{p}}. \quad (108)$$

The FRF vector $\tilde{\nabla}$ and FRF three-tensor $\hat{\mathbf{M}}$ contain all the magnetic effects. When expressed in terms of components in the local Cartesian xyz frame in the FRF the three-vector $\tilde{\nabla}$ equals

$$\tilde{\nabla} = \sqrt{1 - \bar{\mu}^2} \left\{ \frac{\cos \bar{\phi} \hat{\mathbf{x}} + \sin \bar{\phi} \hat{\mathbf{y}}}{1 + \bar{\Omega}^2 \bar{\tau}_r^2} + \bar{\Omega} \bar{\tau}_r \left(\frac{\cos \bar{\phi} \hat{\mathbf{y}} - \sin \bar{\phi} \hat{\mathbf{x}}}{1 + \bar{\Omega}^2 \bar{\tau}_r^2} \right) \right\} + \bar{\mu} \hat{\mathbf{b}}. \quad (109)$$

If we write $\bar{\mathbf{p}} = \bar{\mathbf{p}}_{\perp} + \bar{p}_{\parallel} \hat{\mathbf{b}}$, it is easily seen that equation (109) is equivalent with

$$\tilde{\nabla} = \frac{1}{\bar{p}} \left[\frac{\bar{\mathbf{p}}_{\perp}}{1 + \bar{\Omega}^2 \bar{\tau}_r^2} + \frac{\bar{\Omega} \bar{\tau}_r}{1 + \bar{\Omega}^2 \bar{\tau}_r^2} (\hat{\mathbf{b}} \times \bar{\mathbf{p}}_{\perp}) + \bar{p}_{\parallel} \hat{\mathbf{b}} \right]. \quad (110)$$

The symmetric and traceless 3×3 matrix $\hat{\mathbf{M}}$ equals

$$\hat{\mathbf{M}} \equiv \frac{1 - \bar{\mu}^2}{2} \frac{\mathbf{M}_1 - 2\bar{\Omega} \bar{\tau}_r \bar{\mathbf{M}}_1}{1 + 4\bar{\Omega}^2 \bar{\tau}_r^2} + \frac{\bar{\mu} \sqrt{1 - \bar{\mu}^2}}{2} \frac{\mathbf{M}_2 - \bar{\Omega} \bar{\tau}_r \bar{\mathbf{M}}_2}{1 + \bar{\Omega}^2 \bar{\tau}_r^2} + \frac{3\bar{\mu}^2 - 1}{4} \hat{\mathbf{b}} \hat{\mathbf{b}}. \quad (111)$$

The matrices $\bar{\mathbf{M}}_1$ and $\bar{\mathbf{M}}_2$ are defined as

$$\bar{\mathbf{M}}_1 = -\frac{d\mathbf{M}_1}{d\bar{\phi}} = \frac{1}{2} \begin{pmatrix} +\sin 2\bar{\phi} & -\cos 2\bar{\phi} & 0 \\ -\cos 2\bar{\phi} & -\sin 2\bar{\phi} & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \bar{\mathbf{M}}_2 = -\frac{d\mathbf{M}_2}{d\bar{\phi}} = \begin{pmatrix} 0 & 0 & \sin \bar{\phi} \\ 0 & 0 & -\cos \bar{\phi} \\ \sin \bar{\phi} & -\cos \bar{\phi} & 0 \end{pmatrix}. \quad (112)$$

8.3 The unmagnetized case

In the absence of a magnetic field (or for uncharged particles or photons scattered by the flow) the expression for $f_1(\mathbf{x}, t, \bar{\mathbf{p}})$ simplifies considerably. Equation (94) for the isotropic part of the distribution remains valid in this case. If we write the particle three-momentum in the FRF as $\bar{\mathbf{p}} = \bar{p} \hat{\mathbf{n}}$, $f_1(\mathbf{x}, t, \bar{\mathbf{p}})$ is given by equation (108) with

$$\tilde{\nabla} = \frac{\bar{\mathbf{p}}}{\bar{p}} = \hat{\mathbf{n}}, \quad \hat{\mathbf{M}} = \frac{1}{2} (\hat{\mathbf{n}} \hat{\mathbf{n}} - \langle \hat{\mathbf{n}} \hat{\mathbf{n}} \rangle) = \frac{1}{2} \left(\hat{\mathbf{n}} \hat{\mathbf{n}} - \frac{1}{3} \mathbf{I} \right). \quad (113)$$

This yields, using $\bar{\mathbf{V}} = \bar{v} \hat{\mathbf{n}}$,

$$f_1(\mathbf{x}, t, \bar{\mathbf{p}}) = -\bar{\tau}_r \bar{v} (\hat{\mathbf{n}} \cdot \tilde{\nabla}) f_0 + \bar{\tau}_r (\hat{\mathbf{n}} \cdot \bar{\mathbf{a}}) \bar{\gamma} m \frac{\partial f_0}{\partial \bar{p}} + \frac{\bar{\tau}_r}{2} (\hat{\mathbf{n}} \hat{\mathbf{n}} : \mathbf{W}) \bar{p} \frac{\partial f_0}{\partial \bar{p}}. \quad (114)$$

In the last term we have used that $\mathbf{W}$ is traceless, which implies $\mathbf{I} : \mathbf{W} = \text{trace}(\mathbf{W}) = 0$.

9 APPLICATION: DIFFUSION AND DRIFT

As a first application of the theory presented here we consider spatial diffusion and particle drift in a relativistic flow. The calculation of the viscosity and its effects is rather more involved and needs a more careful discussion of the assumptions when we compare with the existing results from radiation hydrodynamics for a non-relativistic flow. For this reason the calculation of the stress tensor is relegated to Paper 2.

Since we are using FRF momentum variables, it is easiest to first calculate the number flux four-vector in the FRF. With $c = 1$ and $\bar{\mathbf{v}} \equiv \bar{\mathbf{p}}/\bar{\gamma}m$ the FRF particle velocity the number flux four-vector $N^{\hat{a}}$ is given by:

$$N^{\hat{a}}(x^{\mu}) = \int \frac{d^3\bar{\mathbf{p}}}{\bar{\mathcal{E}}} f(\bar{\mathbf{Z}}) p^{\hat{a}} = \int d^3\bar{\mathbf{p}} f(\bar{\mathbf{Z}}) (1, -\bar{\mathbf{v}}). \quad (115)$$

The corresponding four-vector in the LF follows from the Lorentz transformation: $N^{\mu} = E_a^{\mu} N^{\hat{a}}$. The quantity $d^3\bar{\mathbf{p}}/\bar{\mathcal{E}}$ is a Lorentz invariant (see for instance Landau & Lifshitz 1975, ch. 2). Making the diffusion approximation, $f \simeq f_0 + \varepsilon f_1$, one finds (in obvious notation) $N^{\hat{a}} = N_0^{\hat{a}} + \varepsilon N_1^{\hat{a}}$, with

$$N_0^{\hat{a}} = (\bar{n}, 0, 0, 0), \bar{n}(\mathbf{x}, t) \equiv \int d^3\bar{\mathbf{p}} f_0 = \int_0^{\infty} d\bar{p} 4\pi\bar{p}^2 f_0(\mathbf{x}, t, \bar{p}). \quad (116)$$

The first-order particle flux four-vector $N_1^{\hat{a}}$ contains the diffusive flux as well as a drift term. Since the angular average of $f_1(\bar{\mathbf{Z}})$ satisfies $\langle f_1 \rangle = 0$ we have $N_1^{\hat{0}} = 0$, $N_1^{\mu} U_{\mu} = 0$ in a general Lorentz frame. So $N_1^{\hat{a}}$ can be written as

$$N_1^{\hat{a}}(\mathbf{x}, t) = \int_0^{\infty} d\bar{p} (0, \bar{\mathbf{S}}(\mathbf{x}, t, \bar{p})), \quad (117)$$

with

$$\bar{\mathbf{S}}(\mathbf{x}, t, \bar{p}) \equiv \bar{p}^2 \int_0^{2\pi} d\bar{\phi} \int_{-1}^1 d\bar{\mu} \bar{\mathbf{v}} f_1(\mathbf{x}, t, \bar{p}) = 4\pi\bar{p}^2 \langle \bar{\mathbf{v}} f_1(\mathbf{x}, t, \bar{p}) \rangle. \quad (118)$$

Using $\bar{\mathbf{v}} = \bar{v}\sqrt{1-\bar{\mu}^2}(\cos\bar{\phi}\hat{\mathbf{x}} + \sin\bar{\phi}\hat{\mathbf{y}}) + \bar{v}\bar{\mu}\hat{\mathbf{b}}$ in the local Cartesian coordinate system, substituting expression (108) for $f_1(\mathbf{x}, t, \bar{p})$ and performing the integration over momentum solid angle one finds that only the first two terms in the expression for $f_1(\mathbf{x}, t, \bar{p})$ contribute to the flux: The term involving $\hat{\mathbf{M}} : \mathbf{W}$ has an angular dependence such that it integrates to zero when performing the integration over $\bar{\mu}$ and $\bar{\phi}$. That term will become important when we discuss viscosity in Paper 2.

9.1 Diffusion

The first term ($\propto \tilde{\nabla} f_0$) in expression (108) gives the diffusive contribution to the flux. It takes a form equivalent to Fick's law:

$$\bar{\mathbf{S}}_{\text{diff}}(\mathbf{x}, t, \bar{p}) = -4\pi\bar{p}^2 [\bar{\mathbf{D}} \cdot \tilde{\nabla} f_0(\bar{\mathbf{Z}})], \quad (119)$$

with the three-tensor $\bar{\mathbf{D}}$ (diffusion tensor) defined in the FRF as:

$$\bar{\mathbf{D}} = \frac{\bar{v}^2 \bar{\tau}_r}{3} \begin{pmatrix} \frac{1}{1 + \bar{\Omega}^2 \bar{\tau}_r^2} & -\frac{\bar{\Omega} \bar{\tau}_r}{1 + \bar{\Omega}^2 \bar{\tau}_r^2} & 0 \\ \frac{\bar{\Omega} \bar{\tau}_r}{1 + \bar{\Omega}^2 \bar{\tau}_r^2} & \frac{1}{1 + \bar{\Omega}^2 \bar{\tau}_r^2} & 0 \\ 0 & 0 & 1 \end{pmatrix}. \quad (120)$$

As in the non-relativistic case (e.g. Miyamoto 1980, ch. 7.3), diffusion in a magnetic field is characterized entirely by three coefficients for parallel, perpendicular, and cross diffusion (Hall diffusion) with respect to the plasma rest frame magnetic field. They are respectively given by

$$\bar{D}_{\parallel} = \frac{\bar{v}^2 \bar{\tau}_r}{3}, \bar{D}_{\perp} = \frac{\bar{v}^2}{3} \frac{\bar{\tau}_r}{1 + \bar{\Omega}^2 \bar{\tau}_r^2}, \bar{D}_{\times} = \frac{\bar{v}^2}{3} \frac{\bar{\Omega} \bar{\tau}_r^2}{1 + \bar{\Omega}^2 \bar{\tau}_r^2}. \quad (121)$$

Note that these three diffusion coefficients are all defined entirely in terms of quantities in the FRF: $\bar{v}$, $\bar{\tau}_r$, and $\bar{\Omega}$. The relaxation time $\bar{\tau}_r$ appearing here was introduced in the Anderson–Witting collision operator (equation 35). While algebraically convenient that collision operator is not what one expects when charged particles are scattered by MHD turbulence in the FRF. In Section 11 below we briefly consider the case of a Fokker–Planck type relaxation operator describing diffusive elastic scattering in the FRF that resembles what one finds in the case of MHD turbulence. There we show that the results for diffusion and drift are formally the same as obtained here. In the viscosity calculation the results differ slightly, as explained in Paper 2.

9.2 Drift velocity

The term $\propto \tilde{\mathbf{S}} \cdot \bar{\mathbf{a}}$ in the expression for $f_1(\mathbf{x}, t, \bar{p})$, due to the acceleration felt in the non-inertial FRF, leads to a drift term of the form

$$\bar{\mathbf{S}}_{\text{drift}}(\mathbf{x}, t, \bar{p}) = 4\pi\bar{p}^2 \bar{\mathbf{V}}_{\text{D}}(\mathbf{x}, t, \bar{p}) f_0(\mathbf{x}, t, \bar{p}), \quad (122)$$

with

$$\bar{\mathbf{V}}_{\text{D}}(\mathbf{x}, t, \bar{p}) = -\frac{1}{3\bar{p}^2} \frac{\partial}{\partial \bar{p}} \left\{ \bar{p}^3 \bar{\tau}_r \left[\frac{\bar{\mathbf{a}}_{\perp}}{1 + \bar{\Omega}^2 \bar{\tau}_r^2} + \frac{\bar{\Omega} \bar{\tau}_r}{1 + \bar{\Omega}^2 \bar{\tau}_r^2} (\bar{\mathbf{a}} \times \hat{\mathbf{b}}) + (\hat{\mathbf{b}} \cdot \bar{\mathbf{a}}) \hat{\mathbf{b}} \right] \right\}. \quad (123)$$

Here $\bar{\mathbf{a}}_{\perp} \equiv \bar{\mathbf{a}} - (\hat{\mathbf{b}} \cdot \bar{\mathbf{a}}) \hat{\mathbf{b}}$ is the component of the FRF three-vector $\bar{\mathbf{a}}$ perpendicular to the FRF magnetic field. This result is most easily obtained by (i) using expression (110) for $\bar{\Sigma}$, (ii) employing that the integration over angles gives 4π times the average over solid angle, with $\langle \bar{p}_x^2 \rangle = \langle \bar{p}_y^2 \rangle = \langle \bar{p}_z^2 \rangle = \bar{p}^2/3$, and (iii) followed by a partial integration over $\bar{p}$. The drift term is small compared to the diffusion term in the non-relativistic limit, and is therefore absent in the results of Kulsrud (1983) and Skilling (1975): Formally it is of the order $V^2 \bar{\tau}_r/L$ with L the gradient scale of the flow. The diffusion term is typically of the order $\bar{v}^2 \bar{\tau}_r/L$, that is a factor $(\bar{v}/V)^2 \gg 1$ larger. In the limit $\bar{\Omega} \bar{\tau}_r \gg 1$ the term involving $\bar{\mathbf{a}} \times \hat{\mathbf{b}}$ has a simple physical interpretation: It is independent of the collisional relaxation time $\bar{\tau}_r$ and equals (reinstating c for clarity)

$$\bar{\mathbf{V}}_{\times} = -\frac{1}{3\bar{p}^2} \frac{\partial}{\partial \bar{p}} \left[\bar{p}^3 \left(\frac{c}{qB} (\bar{\gamma} m \bar{\mathbf{a}} \times \hat{\mathbf{b}}) \right) \right]. \quad (124)$$

This is the effect of the well-known $\mathbf{F} \times \mathbf{B}$ drift that particles experience in a strong magnetic field, where the force here is the average fictitious force $-\bar{\gamma} m \bar{\mathbf{a}}$ due to the three-acceleration $\bar{\mathbf{a}}$ of the FRF.

If one defines the average drift velocity as

$$\bar{\mathbf{V}}_{\text{drift}}(\mathbf{x}, t) \equiv \frac{1}{\bar{n}} \int_0^{\infty} d\bar{p} 4\pi \bar{p}^2 f_0(\bar{\mathbf{Z}}) \bar{\mathbf{V}}_{\text{D}}(\mathbf{x}, t, \bar{p}) \quad (125)$$

the FRF number flux four-vector is (putting $\epsilon = 1$)

$$N^{\hat{\mathbf{a}}}(\mathbf{x}, t) = \bar{n} (1, \bar{\mathbf{V}}_{\text{drift}}) - \int_0^{\infty} d\bar{p} 4\pi \bar{p}^2 [\bar{\mathbf{D}} \cdot \bar{\mathbf{V}} f_0(\bar{\mathbf{Z}})]. \quad (126)$$

10 COVARIANCE OF THE RESULTS

One can always find the flux $N^{\mu}(\mathbf{x}, t, \bar{p})$ in the LF by using the Lorentz transformation from the FRF: $N^{\mu} = E_{\hat{\mathbf{a}}}^{\mu} N^{\hat{\mathbf{a}}}$. There is however a quicker and more general way that uses the fact that physical relations can always be expressed covariantly. The covariance requirement also serves as a check on the algebra of the previous sections that was done in a 3 + 1 split. In what follows we put the ordering parameter ϵ equal to unity.

It is useful to introduce the two symmetric projection four-tensors $h^{\mu\nu}$ and $H^{\mu\nu}$, as well as the anti-symmetric tensor $\hat{f}^{\mu\nu}$:

$$h^{\mu\nu} = \eta^{\mu\nu} - U^{\mu} U^{\nu}, \quad H^{\mu\nu} = h^{\mu\nu} + b^{\mu} b^{\nu}, \quad \hat{f}^{\mu\nu} = e^{\mu\nu\sigma\tau} U_{\sigma} b_{\tau} = F^{\mu\nu}/B. \quad (127)$$

Here $e^{\mu\nu\sigma\tau}$ is the totally antisymmetric Levi (pseudo-)tensor. In the FRF, using the local xyz spatial coordinates introduced in Section 5, the first two tensors take the form $h^{\hat{\mathbf{a}}\hat{\mathbf{b}}} = \text{diag}(0, -1, -1, -1)$, $H^{\hat{\mathbf{a}}\hat{\mathbf{b}}} = \text{diag}(0, -1, -1, 0)$. The only non-vanishing components of the anti-symmetric tensor $\hat{f}^{\hat{\mathbf{a}}\hat{\mathbf{b}}}$ in the FRF are $\hat{f}^{12} = -\hat{f}^{21} = -1$, which implies that in the FRF

$$\hat{f}^{\hat{\mathbf{a}}\hat{\mathbf{b}}} A_{\hat{\mathbf{b}}} = (0, \bar{\mathbf{A}} \times \hat{\mathbf{b}}). \quad (128)$$

These tensors satisfy $h^{\mu\nu} U_{\nu} = 0$, $H^{\mu\nu} U_{\nu} = H^{\mu\nu} b_{\nu} = 0$ and $\hat{f}^{\mu\nu} U_{\nu} = \hat{f}^{\mu\nu} b_{\nu} = 0$.

It is possible to write the terms involving the vector $\bar{\Sigma}$ in expression (108) covariantly. To that end we define the two projection tensors

$$G_{\pm}^{\mu\nu} \equiv \frac{H^{\mu\nu}}{1 + \bar{\Omega}^2 \bar{\tau}_r^2} \pm \frac{\bar{\Omega} \bar{\tau}_r}{1 + \bar{\Omega}^2 \bar{\tau}_r^2} \hat{f}^{\mu\nu} - b^{\mu} b^{\nu}. \quad (129)$$

The antisymmetry of $\hat{f}^{\mu\nu}$ and the symmetry of the two other tensors in this definition imply that $G_{-}^{\mu\nu} = G_{+}^{\nu\mu}$. The vector $\bar{\Sigma}$ was defined in (110) using FRF momentum variables. We introduce the FRF four-vector $\Sigma^{\hat{\mathbf{a}}} \equiv (0, \bar{\Sigma})$. The same vector is given as

$$\Sigma^{\mu} \equiv \frac{1}{\bar{p}} G_{-}^{\mu\nu} p_{\nu} \quad (130)$$

in an arbitrary Lorentz frame. One can express $\bar{p} = |\bar{\mathbf{p}}|$ as the Lorentz-invariant quantity

$$\bar{p} = \sqrt{(p \cdot U)^2 - (p \cdot p)}. \quad (131)$$

Here we use the notation $A \cdot B \equiv A_{\mu} B^{\mu} = A^0 B^0 - \mathbf{A} \cdot \mathbf{B}$ for the scalar product of two four-vectors. One has $G_{\pm}^{\mu\nu} U_{\nu} = U_{\nu} G_{\pm}^{\mu\nu} = 0$ and $G_{\pm}^{\mu\nu} b_{\nu} = b^{\mu}$, $b_{\mu} G_{\pm}^{\mu\nu} = b^{\nu}$. When there is no magnetic field (or for uncharged particles), the relation $G_{\pm}^{\mu\nu} = h^{\mu\nu}$ applies.

The first two terms in expression (108) for $f_1(\mathbf{x}, t, \bar{\mathbf{p}})$, which respectively lead to the diffusive flux and the drift, can be written quasi-covariantly in terms of Σ^{μ} and the four-acceleration $a^{\mu} = dU^{\mu}/d\tau$:

$$f_1^{\text{diff}}(\mathbf{x}, t, \bar{\mathbf{p}}) = -\bar{v} \bar{\tau}_r \Sigma^{\mu} \nabla_{\mu} f_0, \quad f_1^{\text{drift}}(\mathbf{x}, t, \bar{\mathbf{p}}) = -\bar{\tau}_r (\Sigma^{\mu} a_{\mu}) \bar{\gamma} m \frac{\partial f_0}{\partial \bar{p}}. \quad (132)$$

Using equation (34) it is easily seen that one can write the FRF diffusive flux as

$$N_{\text{diff}}^{\hat{\mathbf{a}}} = -4\pi \bar{p}^2 \left[-\bar{D}_{\perp} H^{\hat{\mathbf{a}}\hat{\mathbf{b}}} \frac{\partial f_0}{\partial x^{\hat{\mathbf{b}}}} + \bar{D}_{\parallel} b^{\hat{\mathbf{a}}} b^{\hat{\mathbf{b}}} \frac{\partial f_0}{\partial x^{\hat{\mathbf{b}}}} + \bar{D}_{\times} \hat{f}^{\hat{\mathbf{a}}\hat{\mathbf{b}}} \frac{\partial f_0}{\partial x^{\hat{\mathbf{b}}}} \right] \equiv -4\pi \bar{p}^2 D^{\hat{\mathbf{a}}\hat{\mathbf{b}}} \frac{\partial f_0}{\partial x^{\hat{\mathbf{b}}}}. \quad (133)$$

This defines a diffusion tensor $D^{\hat{a}\hat{b}} \equiv -\bar{D}_\perp H^{\hat{a}\hat{b}} + \bar{D}_\parallel b^{\hat{a}} b^{\hat{b}} + \bar{D}_\times \hat{f}^{\hat{a}\hat{b}} = -\bar{D} G_-^{\hat{a}\hat{b}}$, using definition (129) for $G_-^{\mu\nu}$. The scalar diffusion coefficient $\bar{D}$ that appears in this definition equals $\bar{D} \equiv \bar{v}^2 \bar{\tau}_r/3$. This tensor has the FRF components $D^{\hat{0}\hat{0}} = D^{\hat{0}\hat{i}} = D^{\hat{i}\hat{0}} = 0$, $D^{\hat{i}\hat{j}} = (\bar{\mathbf{D}})_{ij}$. Since relation (133) is formally covariant, the corresponding relation

$$N_{\text{diff}}^\mu = -4\pi\bar{p}^2 D^{\mu\nu} \frac{\partial f_0(\mathbf{x}, t, \bar{p})}{\partial x^\nu} \quad (134)$$

holds in *any* frame, with $D^{\mu\nu}$ defined as

$$D^{\mu\nu} \equiv -\bar{D}_\perp H^{\mu\nu} + \bar{D}_\parallel b^\mu b^\nu + \bar{D}_\times \hat{f}^{\mu\nu} = -\bar{D} G_-^{\mu\nu}. \quad (135)$$

It should be noted that we continue to use FRF momentum variables.

The drift term can similarly be cast into a covariant form. Relation (123) in the FRF corresponds to

$$V_{\text{drift}}^{\hat{a}} = -\frac{1}{3\bar{p}^2} \frac{\partial}{\partial \bar{p}} \left[\bar{p}^3 \bar{\tau}_r \left\{ \frac{H^{\hat{a}\hat{b}} a_{\hat{b}}}{1 + \bar{\Omega}^2 \bar{\tau}_r^2} + \frac{\bar{\Omega} \bar{\tau}_r}{1 + \bar{\Omega}^2 \bar{\tau}_r^2} \left(\hat{f}^{\hat{a}\hat{b}} a_{\hat{b}} \right) - (b \cdot a) b^{\hat{a}} \right\} \right]. \quad (136)$$

The FRF fluid four-acceleration is $a_{\hat{a}} = (0, -, -\bar{\mathbf{a}})$ (see Appendix C, equation 107). We also use $b \cdot a \equiv b^{\hat{a}} a_{\hat{a}} = -\hat{\mathbf{b}} \cdot \bar{\mathbf{a}}$. This implies that one can write the drift velocity in any frame covariantly in terms of the four-acceleration $a_\mu = dU_\mu/d\tau$ and the projection tensor $G_+^{\mu\nu}$ as

$$V_{\text{drift}}^\mu = -\frac{1}{3\bar{p}^2} \frac{\partial}{\partial \bar{p}} \left(\bar{p}^3 \bar{\tau}_r G_+^{\mu\nu} a_\nu \right). \quad (137)$$

Note that $V^\mu U_\mu = 0$. The fact that $G_+^{\mu\nu}$ appears here while $G_-^{\mu\nu}$ appears in the definition of Σ^μ and $D^{\mu\nu}$ is the result of isotropy of f_0 so that $\langle p^{\hat{i}} p^{\hat{j}} \rangle = \bar{p}^2 \delta_{\hat{i}\hat{j}}/3$ and the properties of the cross-product in $f^{\hat{i}\hat{j}} p_{\hat{j}} = \bar{\mathbf{p}} \times \hat{\mathbf{b}}$. This leads to the following FRF relation, needed for the solid-angle average of the Hall term $\propto \bar{\Omega} \bar{\tau}_r / (1 + \bar{\Omega}^2 \bar{\tau}_r^2)$ in V_{drift}^μ :

$$\left\langle \left[(\hat{\mathbf{b}} \times \bar{\mathbf{p}}) \cdot \bar{\mathbf{a}} \right] \bar{\mathbf{p}} \right\rangle^{\hat{i}} = \epsilon_{\hat{k}\hat{l}\hat{m}} b^{\hat{l}} \bar{a}^{\hat{k}} \langle p^{\hat{m}} p^{\hat{i}} \rangle = \frac{\bar{p}^2}{3} \epsilon_{\hat{k}\hat{l}\hat{i}} b^{\hat{l}} \bar{a}^{\hat{k}} = \frac{\bar{p}^2}{3} (\bar{\mathbf{a}} \times \hat{\mathbf{b}})^{\hat{i}}. \quad (138)$$

10.1 The unmagnetized case

If there is no magnetic field or for uncharged particles (in particular photons) scattered by the medium one has $\bar{D}_\perp = \bar{D}_\parallel = \bar{v}^2 \bar{\tau}_r/3 = \bar{D}$, $\bar{D}_\times = 0$, and $G_\pm^{\mu\nu} = h^{\mu\nu}$. Then $D^{\mu\nu} = -\bar{D} h^{\mu\nu}$ and relation (134) for the diffusive part of the particle flux then simplifies to the covariant version of Fick's law:

$$N_{\text{diff}}^\mu = 4\pi\bar{p}^2 \bar{D} h^{\mu\nu} \frac{\partial f_0(\mathbf{x}, t, \bar{p})}{\partial x^\nu} \quad (\text{for } B = 0 \text{ or } q = 0). \quad (139)$$

The drift term also simplifies as the drift velocity now equals, with $U^\mu a_\mu = 0$, so that $h^{\mu\nu} a_\mu = a^\nu$:

$$V_{\text{drift}}^\mu = -\frac{1}{3\bar{p}^2} \frac{\partial}{\partial \bar{p}} \left(\bar{p}^3 \bar{\tau}_r a^\mu \right) \quad (\text{for } B = 0 \text{ or } q = 0). \quad (140)$$

10.2 Conservation form of the transport equation

The number flux four-vector per unit momentum (including the contribution from the zero-order term $N_0^{\hat{a}}$, equation 116) can be written in terms of $\mathcal{N}(\mathbf{x}, t, \bar{p}) \equiv 4\pi\bar{p}^2 f_0(\mathbf{x}, t, \bar{p}) = d\bar{n}/d\bar{p}$ as

$$N^\mu(\mathbf{x}, t, \bar{p}) = (U^\mu + V_{\text{drift}}^\mu) \mathcal{N} - D^{\mu\nu} \frac{\partial \mathcal{N}}{\partial x^\nu}. \quad (141)$$

Because $U_\mu V_{\text{drift}}^\mu = 0$ and $U_\mu D^{\mu\nu} = 0$ one has

$$N^\mu U_\mu = \mathcal{N}(\mathbf{x}, t, \bar{p}) = \frac{d\bar{n}}{d\bar{p}}, \quad (142)$$

so the drift and diffusion terms in $N^\mu(\mathbf{x}, t, \bar{p})$ do not change the ‘proper density per unit momentum’ of the particles.

We have shown in Section 4 that Liouville's theorem holds for mixed variables. One can average the Vlasov equation in conservation form over solid angle. With $f_0 \equiv \langle f(\tilde{\mathbf{Z}}) \rangle$ this leads to

$$\left\langle (\partial/\partial \tilde{\mathbf{Z}}) \cdot [\dot{\tilde{\mathbf{Z}}} f(\tilde{\mathbf{Z}})] \right\rangle = \langle \text{St}(f) \rangle \equiv -\frac{\langle f \rangle - \langle f \rangle}{\bar{\tau}_r} = 0. \quad (143)$$

After multiplication of this relation by $4\pi\bar{p}^2$ one finds an equation for $\mathcal{N}(\mathbf{x}, t, \bar{p})$:

$$\nabla_\mu \left\{ (U^\mu + V_{\text{drift}}^\mu) \mathcal{N} - D^{\mu\nu} \nabla_\nu \mathcal{N} \right\} - \frac{\partial}{\partial \bar{p}} \left(\frac{\bar{p}}{3} (\nabla \cdot U) \mathcal{N} \right) + \frac{\partial}{\partial \bar{p}} (4\pi\bar{p}^2 \langle \tilde{K}(\tilde{\mathbf{Z}}) f_1(\tilde{\mathbf{Z}}) \rangle) = 0. \quad (144)$$

This is the transport equation in conservation form in the diffusion approximation. Here $\nabla_\mu = \partial/\partial x^\mu$ with FRF momentum fixed, and we have used $d/d\tau = U \cdot \nabla$ as well as relations (94) and (95). We have put $\epsilon = 1$ in the last term on the left-hand side. The quantity $\tilde{K}(\tilde{\mathbf{Z}})$ was defined in equation (99).

The last term on the left-hand side of equation (144) represents viscous heating. This term is usually absent in calculations for non-relativistic flows since it is (by assumption) small in these calculations. The works of Jokipii & Williams (1992) and Williams & Jokipii (1993) are the exception: They calculate this term for a non-relativistic flow. In Paper 2 we evaluate this term explicitly. There we show that it takes the form of momentum diffusion in $\bar{p}$ and that it indeed represents the *irreversible heating* of the scattered particles in a thermodynamic sense. As already mentioned the second term ($\propto \nabla \cdot \mathbf{U}$) in equation (144) corresponds to adiabatic momentum gains/losses (see equation 95).

Equation (144) contains *all* terms formally of order ϵ : spatial drift and spatial diffusion as well as the momentum diffusion (heating) term $\propto \langle \tilde{K}(\tilde{\mathbf{Z}}) f_1(\tilde{\mathbf{Z}}) \rangle$, consistent with the diffusion approximation. It shows for the first time that the transport equation for particles scattered in a relativistic flow is an almost straightforward generalization of the non-relativistic result in the case of simple *elastic* scattering in the FRF. The main differences are [1] the appearance of the drift term $\propto V_{\text{drift}}^\mu$, and [2] the consistent use of the FRF momentum rather than the LF momentum so that the spatial derivatives are taken with the FRF four-momentum fixed.

In a non-relativistic flow, with $|\mathbf{U}| \ll 1$ and $V/\bar{v} \ll 1$, one can neglect the drift term $\propto V_{\text{drift}}^\mu$ and (as long as one works to first order in $V/\bar{v}$) one can replace the FRF momentum $\bar{p}$ by the LF momentum p . Then equation (144) is equivalent to the results of Skilling (1975), *without* the momentum diffusion terms due to acceleration by Alfvén waves that are treated in that paper but that do not appear here since we assumed elastic scattering in the FRF. As discussed we do calculate the additional effects of the diffusion across the large-scale magnetic field and (in Paper 2) of irreversible heating.

The simple form (equation 144) of the transport equation is implicit in the results of Webb (1985), but there it is disguised by his $3 + 1$ split in the notation in terms of $\mathbf{x}$, t (and the corresponding derivatives), Γ , and $\mathbf{V}$, and the different (more complicated) choice for the scattering operator. Webb also uses gyro-averaging in most of his results, which eliminates the terms associated with cross-field diffusion and possible viscous effects from consideration, as argued above.

11 THE EFFECT OF AN ALTERNATIVE CHOICE FOR THE SCATTERING OPERATOR

In the calculation of $f_1(\mathbf{x}, t, \bar{\mathbf{p}})$, which involves solving equation (100), the precise form of the relaxation operator $\text{St}(f)$ does influence results. So far we have chosen to employ the algebraic Anderson–Witting form, the relativistic generalization of the Krook collision term. This operator damps all anisotropies (dipole, quadrupole, etc.) at the same rate determined by the relaxation time $\bar{\tau}_r$.

In this section we briefly discuss the effect of employing an alternative choice: a Fokker–Plank type collision operator. In terms of the pitch angle $\bar{\theta} = \cos^{-1} \bar{\mu}$ and the gyro-phase angle $\bar{\phi}$ one has in terms of $\bar{\text{St}}(f) = \bar{\gamma}^{-1} \text{St}(f)$

$$\bar{\text{St}}_{\text{FP}}(f) \equiv \frac{1}{2\bar{\tau}_r} \left[\frac{1}{\sin \bar{\theta}} \frac{\partial}{\partial \bar{\theta}} \left(\sin \bar{\theta} \frac{\partial f}{\partial \bar{\theta}} \right) + \frac{1}{\sin^2 \bar{\theta}} \frac{\partial^2 f}{\partial \bar{\phi}^2} \right], \quad (145)$$

which corresponds to isotropic diffusion of the momentum direction in the fluid frame. The factor $\bar{\gamma} = (p \cdot \mathbf{U})/m$ takes care of time dilation and the factor $1/2$ is a convention so that the relaxation time for dipole anisotropies equals $\bar{\tau}_r$.

The equivalent of equation (100) becomes

$$\bar{\Omega} \frac{\partial f_1}{\partial \bar{\phi}} + \frac{1}{2\bar{\tau}_r} \left[\frac{1}{\sin \bar{\theta}} \frac{\partial}{\partial \bar{\theta}} \left(\sin \bar{\theta} \frac{\partial f}{\partial \bar{\theta}} \right) + \frac{1}{\sin^2 \bar{\theta}} \frac{\partial^2 f}{\partial \bar{\phi}^2} \right] = \text{source term}, \quad (146)$$

where the source term is the entire term on the right-hand side of equation (100). The obvious approach in this case is to use spherical harmonics $Y_\ell^m(\bar{\theta}, \bar{\phi})$, the eigenfunctions of the diffusive relaxation operator $\bar{\text{St}}_{\text{FP}}(f)$ that satisfy

$$\bar{\text{St}}_{\text{FP}}(Y_\ell^m(\bar{\theta}, \bar{\phi})) = -\frac{\ell(\ell+1)}{2\bar{\tau}_r} Y_\ell^m(\bar{\theta}, \bar{\phi}). \quad (147)$$

Here $\ell = 0, 1, 2, \dots$ and $-\ell \leq m \leq +\ell$. In what follows we employ the form of $Y_\ell^m(\bar{\theta}, \bar{\phi})$ that includes the Condon–Shortley phase factor (c.f. Arfken & Weber 2005, ch.12.6):

$$Y_\ell^m(\bar{\theta}, \bar{\phi}) = (-1)^m \sqrt{\frac{2\ell+1}{4\pi} \frac{(\ell-m)!}{(\ell+m)!}} P_\ell^m(\cos \bar{\theta}) e^{im\bar{\phi}}. \quad (148)$$

The $P_\ell^m(\cos \bar{\theta})$ are the associated Legendre polynomials. We expand $f_1(\mathbf{x}, t, \bar{\mathbf{p}})$ in spherical harmonics,

$$f_1(\mathbf{x}, t, \bar{\mathbf{p}}) = \sum_{\ell=1}^{\infty} \sum_{m=-\ell}^{m=+\ell} a_\ell^m(\mathbf{x}, t, \bar{\mathbf{p}}) Y_\ell^m(\bar{\theta}, \bar{\phi}), \quad (149)$$

which implies that (as before) $\langle f_1(\mathbf{x}, t, \bar{\mathbf{p}}) \rangle = 0$. In a similar fashion we expand the (known) source term on the right-hand side of equation (146), which by definition has a vanishing momentum space solid-angle average:

$$\text{source term} = \sum_{\ell=1}^{\infty} \sum_{m=-\ell}^{m=+\ell} s_\ell^m(\mathbf{x}, t, \bar{\mathbf{p}}) Y_\ell^m(\bar{\theta}, \bar{\phi}). \quad (150)$$

After substitution of these two expansions into equation (146) one finds the following relation that determines the coefficients a_ℓ^m in terms of the known source-term coefficients s_ℓ^m :

$$\left(im\bar{\Omega} - \frac{\ell(\ell+1)}{2\bar{\tau}_r} \right) a_\ell^m = s_\ell^m. \quad (151)$$

This follows from the fact that the spherical harmonics are orthogonal eigenfunctions of the Fokker–Planck relaxation operator. Using the identities

$$\begin{pmatrix} \bar{\mu} \\ \sqrt{1-\bar{\mu}^2} \cos \bar{\phi} \\ \sqrt{1-\bar{\mu}^2} \sin \bar{\phi} \end{pmatrix} = \sqrt{\frac{4\pi}{3}} \begin{pmatrix} Y_1^0(\bar{\theta}, \bar{\phi}) \\ \frac{1}{\sqrt{2}} (Y_1^{-1}(\bar{\theta}, \bar{\phi}) - Y_1^1(\bar{\theta}, \bar{\phi})) \\ \frac{i}{\sqrt{2}} (Y_1^{-1}(\bar{\theta}, \bar{\phi}) + Y_1^1(\bar{\theta}, \bar{\phi})) \end{pmatrix} \quad (152)$$

it is easily established that the terms in equation (100) that lead to the first two terms in expression (108) for $f_1(\mathbf{x}, t, \bar{\mathbf{p}})$, the terms that involve the three-vector $\bar{\mathbf{\Sigma}}$ defined in (109), *all* come exclusively from the $\ell = 1$ terms (dipole terms) in expansion (150). Solving (151) for a_1^0 and $a_1^{\pm 1}$ one finds that the resulting contributions to $f_1(\mathbf{x}, t, \bar{\mathbf{p}})$ in the case of Fokker–Planck type scattering are *unchanged* with respect to the result found using the simpler Anderson–Witting relaxation term, i.e. $\bar{\mathbf{\Sigma}}_{\text{FP}} = \bar{\mathbf{\Sigma}}$. This result is essentially due to the fact that the Fokker–Planck relaxation operator has the eigenvalue $-1/\bar{\tau}_r$ for the spherical harmonics with $\ell = 1$, just as in the Anderson–Witting case.

The remaining terms in the source term in equation (100) are the terms that involve the strain tensor $\mathbf{W}$ via the replacement rule (equation 103). These terms consist exclusively of terms in the expansion from spherical harmonics with $\ell = 2$: They are quadrupole terms. This follows from the identities

$$\begin{pmatrix} \frac{3\bar{\mu}^2 - 1}{2} \\ \bar{\mu} \sqrt{1-\bar{\mu}^2} \cos \bar{\phi} \\ \bar{\mu} \sqrt{1-\bar{\mu}^2} \sin \bar{\phi} \end{pmatrix} = \sqrt{\frac{4\pi}{5}} \begin{pmatrix} Y_2^0(\bar{\theta}, \bar{\phi}) \\ \frac{1}{\sqrt{6}} (Y_2^{-1}(\bar{\theta}, \bar{\phi}) - Y_2^1(\bar{\theta}, \bar{\phi})) \\ \frac{i}{\sqrt{6}} (Y_2^{-1}(\bar{\theta}, \bar{\phi}) + Y_2^1(\bar{\theta}, \bar{\phi})) \end{pmatrix} \quad (153)$$

and

$$(1 - \bar{\mu}^2) \begin{pmatrix} \cos 2\bar{\phi} \\ \sin 2\bar{\phi} \end{pmatrix} = \sqrt{\frac{8\pi}{3}} \begin{pmatrix} Y_2^2(\bar{\theta}, \bar{\phi}) + Y_2^{-2}(\bar{\theta}, \bar{\phi}) \\ i (Y_2^2(\bar{\theta}, \bar{\phi}) - Y_2^{-2}(\bar{\theta}, \bar{\phi})) \end{pmatrix}. \quad (154)$$

The eigenvalue for the $\ell = 2$ under diffusive scattering is $-3/\bar{\tau}_r$, corresponding to a relaxation time $\bar{\tau}_r/3$: The Fokker–Planck relaxation term damps the $\ell = 2$ spherical harmonics (quadrupole terms) three times faster than the Anderson–Witting relaxation term, where all terms with $\ell \neq 0$ have the same relaxation time, equal to $\bar{\tau}_r$. As a result, when dealing with Fokker–Planck (diffusive) scattering one must replace $\bar{\tau}_r$ by $\bar{\tau}_r/3$ in the terms in quadrupole terms in $f_1(\mathbf{x}, t, \bar{\mathbf{p}})$. The last term in expression (108), the term that governs viscous effects as we will show in Paper 2, becomes

$$\frac{\bar{\tau}_r}{3} (\hat{\mathbf{M}}_{\text{FP}} : \mathbf{W}) \bar{p} \frac{\partial f_0}{\partial \bar{p}}. \quad (155)$$

Here the 3×3 matrix $\hat{\mathbf{M}}_{\text{FP}}$ equals (analogous to definition 111 for $\hat{\mathbf{M}}$ in the Anderson–Witting relaxation)

$$\hat{\mathbf{M}}_{\text{FP}} \equiv \frac{1 - \bar{\mu}^2}{2} \frac{\mathbf{M}_1 - \frac{2}{3} \bar{\Omega} \bar{\tau}_r \bar{\mathbf{M}}_1}{1 + \frac{4}{9} \bar{\Omega}^2 \bar{\tau}_r^2} + \frac{\bar{\mu} \sqrt{1 - \bar{\mu}^2}}{2} \frac{\mathbf{M}_2 - \frac{1}{3} \bar{\Omega} \bar{\tau}_r \bar{\mathbf{M}}_2}{1 + \frac{1}{9} \bar{\Omega}^2 \bar{\tau}_r^2} + \frac{3\bar{\mu}^2 - 1}{4} \hat{\mathbf{b}} \hat{\mathbf{b}}. \quad (156)$$

Otherwise, the functional form of $f_1(\mathbf{x}, t, \bar{\mathbf{p}})$ is unchanged. We therefore find in the Fokker–Planck case

$$f_{\text{IFP}}(\mathbf{x}, t, \bar{\mathbf{p}}) = -\bar{\tau}_r (\bar{\mathbf{v}} \cdot \bar{\mathbf{\Sigma}} \cdot \bar{\mathbf{\nabla}}) f_0 + \bar{\tau}_r (\bar{\mathbf{\Sigma}} \cdot \bar{\mathbf{a}}) \bar{\gamma} m \frac{\partial f_0}{\partial \bar{p}} + \frac{\bar{\tau}_r}{3} (\hat{\mathbf{M}}_{\text{FP}} : \mathbf{W}) \bar{p} \frac{\partial f_0}{\partial \bar{p}}, \quad (157)$$

with the vector $\bar{\mathbf{\Sigma}}$ still given by equation (109). All results of the previous sections are now easily transcribed to this case. In particular, the results obtained for diffusion and drift in Section 9 are unchanged.

12 DISCUSSION

In this paper we have considered the transport of particles that are being scattered by a relativistic flow. We rederive and extend the results of Webb (1985). In Section 3 we gave a more intuitive derivation of the fictitious four-force that occurs if the fluid rest frame is not an inertial (Lorentz) frame. This fictitious force is the source of all terms associated with collisional transport: diffusion, drift, and viscosity. In Section 7 we explicitly showed the correspondence with Webb’s results.

In Section 4 we demonstrated that it is possible to write all final equations in a quasi-covariant form and proved that Liouville’s theorem, i.e. incompressible particle flow in seven-dimensional phase space $(x^\mu, \bar{\mathbf{p}})$, still holds if one uses mixed laboratory frame/fluid rest frame variables.

We generalized the approaches of Skilling (1975) and Kulsrud (1983) to a relativistic flow, showing in the process (Sections 5 and 6) that these two approaches are largely equivalent.

In Section 7 we extended existing results by allowing a dependence on gyro-phase, which opens the possibility of discussing the effect of the large-scale magnetic field on particle transport properties.

We derived in Section 8 that the perturbation expansion for $f(\mathbf{x}, t, \bar{\mathbf{p}}) = f(\tilde{\mathbf{Z}})$ that one employs in the diffusion approximation with the Anderson–Witting relaxation term $\text{St}(f) = -\bar{\gamma}(f - f_0)/\bar{\tau}_r$ should have $f_0 = \langle f \rangle$ with $\langle f \rangle(\mathbf{x}, t, \bar{\mathbf{p}})$ the angular average of the *exact* Vlasov distribution. The perturbation expansion therefore takes the form

$$f(\tilde{\mathbf{Z}}) = \langle f(\tilde{\mathbf{Z}}) \rangle + \varepsilon f_1(\mathbf{x}, t, \bar{\mathbf{p}}) + \dots \quad (158)$$

In particular, this is true for the case of pure scattering. This implies $\langle f_n(\tilde{\mathbf{Z}}) \rangle = 0$ for $n \geq 1$.

It is perhaps good to point out once again that this particular conclusion is *independent* of the form of the relaxation term $\text{St}(f)$, as long as this term corresponds to simple elastic scattering in the FRF. Our use of the algebraic Anderson–Witting scattering term in Sections 4–10 yields the conditions $f_0 = \langle f \rangle$ and $\langle f_1(\tilde{\mathbf{Z}}) \rangle = 0$ in an almost trivial manner. In Section 11 we discussed an alternative choice for the relaxation term: diffusive scattering in the form of the Fokker–Planck scattering operator. Then the same conclusion holds because of the properties (in particular $\langle Y_\ell^m \rangle = 0$ for $\ell \neq 0$) and the orthogonality of the spherical harmonics used in the perturbation expansion that leads to the final expression (equation 157) for $f_1(\mathbf{x}, t, \bar{\mathbf{p}})$.

In Section 9 we derived – as a first application – the cross-field diffusion of the scattered particles by explicitly calculating the gyro-phase-dependent terms in the distribution function $f(\mathbf{x}, t, \bar{\mathbf{p}})$ in a perturbation scheme valid for small anisotropies in the fluid rest frame (i.e. for frequent scattering). For relativistic flow speeds there is an extra drift term that does not appear in the non-relativistic treatment.

In the follow-up paper (Paper 2) we show that this argument has consequences for the calculation of the energy-momentum tensor in the diffusion approximation. In particular it implies that the viscous terms in the energy-momentum tensor correspond to the effect of shear viscosity, and that there is no contribution from a bulk viscosity. We will show in Paper 2, by explicitly calculating the term on the right-hand side of equation (144), that the back-reaction of the flow on the scattered particles (or photons) leads to an irreversible heating term in the transport equation.

In a second follow-up paper (Paper 3) we explore radiation transport in the diffusion approximation and the radiation stress tensor in a relativistic flow. Published results (e.g. Weinberg 1971; Mihalas & Weibel Mihalas 1984; and Castor 2004), derived for the case $|U| \ll 1$ ($|V| \ll c$) give diffusive corrections to the comoving number – and energy density of the photons. In Paper 3 we argue that the condition $\langle f_n(\tilde{\mathbf{Z}}) \rangle = 0$ for $n \geq 1$ implies that these corrections should *not* be included: The leading-order solution of the self-consistent radiation transfer equation in the diffusion approximation shows that these extra terms will vanish to (at least) second order in $\varepsilon \sim \bar{\lambda}/L \ll 1$ with $\bar{\lambda} = \bar{v}\bar{\tau}_r$ the FRF scattering mean free path and L the gradient scale in the plasma. Our approach, which identifies $f_0(\mathbf{x}, t, \bar{\mathbf{p}})$ with $\langle f(\tilde{\mathbf{Z}}) \rangle$ by including the leading-order transport equation in the expansion for $f(\tilde{\mathbf{Z}})$, avoids this issue entirely, and is the mathematically correct procedure.

These considerations also mean that our result for the viscous stresses from the scattered particles, when applied to photons as discussed in Paper 3, are different from the recent results of Coughlin & Begelman (2014). We demonstrate that for a gas of photons only shear viscosity (and therefore no bulk viscosity) is present, in agreement with the results of Weinberg (1971).

ACKNOWLEDGEMENTS

12 ACKNOWLEDGEMENTS

We thank: M.Begelman, A. Bell, R. Blandford, J. Kirk, the late D. Lynden-Bell, M. Rees and G. Peletier for stimulating discussions. This work was done in Nijmegen, Oxford, and Cambridge. We thank P. Groot, G. Nelemans, S. Balbus, and A. Fabian for the hospitality and support.

REFERENCES

- Achterberg A., 1983, *Phys. Rev. A*, 28, 2449
- Anderson J. L., Witting H. R., 1974, *Physica*, 74, 466
- Arfken G. B., Weber H. J., 2005, *Mathematical Methods for Physicists*, 6th ed. Elsevier, Amsterdam, Boston
- Bekenstein J. D., Oron E., 1978, *Phys. Rev. D*, 18, 1809
- Bhatnagar P. L., Gross E. P., Krook M., 1954, *Phys. Rev.*, 94, 511
- Braginskii S. I., 1965, *Reviews of Plasma Physics*, 1, 205
- Carroll S. M., 2004, *Spacetime and Geometry. An Introduction to General Relativity*. Addison-Wesley, San Francisco
- Cary J. R., Littlejohn R. G., 1983, *Ann. Phys. NY*, 151, 1
- Castor J. I., 2014, *Radiation Hydrodynamics*. Cambridge Univ. Press, Cambridge
- Coughlin E. R., Begelman M. C., 2014, *ApJ*, 797, 103
- Debbasch F., van Leeuwen W. A., 2009a, *Physica A*, 388, 1079
- Debbasch F., van Leeuwen W. A., 2009b, *Physica A*, 388, 1818

- Dolginov A. Z., Toptygin I. M., 1967, *Geomagn. Aeronomy*, 7, 785
- Gleeson L. J., Axford W. I., 1967, *ApJ*, 149, L115
- Goldstein H., Poole C., Saffko J., 2002, *Classical Mechanics*. Addison-Wesley, San Francisco
- Hakim R. J., 2011, *Introduction to Relativistic Statistical Mechanics: Classical and Quantum*. World Scientific, Singapore
- Jackson J. D., 1998, *Classical Electrodynamics*, 3rd edn. J. Wiley & Sons, New York
- Jokipii J. R., Williams L. L., 1992, *ApJ*, 394, 184
- Jones F. C., 1990, *ApJ*, 361, 162
- Klimontovich Y. L., 1960, *Sov. Phys. JETP*, 37, 524
- Kulsrud R. M., 1983, in Galeev A. A., Sudan R. N., eds, *Basic Plasma Physics: Selected Chapters, Handbook of Plasma Physics, Volume 1*. North Holland Publ Co., Amsterdam, p. 1
- Landau L. D., Lifshitz E. M., 1959, *Fluid Mechanics, Course of Theoretical Physics, Vol. 6*. Pergamon Press, Oxford
- Landau L. D., Lifshitz E. M., 1975, *The Classical Theory of Fields, Course of Theoretical Physics, Vol. 2*. Pergamon Press, Oxford
- Lichnerowicz A., 1967, *Relativistic Hydrodynamics and Magnetohydrodynamics, The Mathematical Physics Monograph Series*. Benjamin, New York
- Melrose D. B., 1989, *Instabilities in Space and Laboratory Plasmas*. Cambridge Univ. Press, Cambridge,
- Mihalas D., Weibel Mihalas B., 1984, *Foundations of Radiation Hydrodynamics*. Oxford University Press, New York
- Misner C. W., Thorne K. S., Wheeler J. A., 1973, *Gravitation*. W.H. Freeman and Co., San Francisco
- Miyamoto K., 1980, *Plasma Physics for Nuclear Fusion*. MIT Press, Cambridge, Mass.
- Parker E. N., 1965, *Planet. Space Sci.*, 13, 9
- Pomraning G. C., 1973, *The Equations of Radiation Hydrodynamics, International Series of Monographs in Natural Philosophy* Pergamon Press, Oxford
- Skilling J., 1975, *MNRAS*, 172, 557
- Webb G. M., 1985, *ApJ*, 296, 319
- Weinberg S., 1971, *ApJ*, 168, 175
- Williams L. L., Jokipii J. R., 1993, *ApJ*, 417, 725
- Zee A., 2013, *Einstein Gravity in a Nutshell*. Princeton Univ. Press, Princeton, NJ

APPENDIX A: EQUATION OF MOTION IN MIXED VARIABLES

The equation of motion using mixed variables (x^μ , p_a) can be derived using a variational principle. We use lab frame time and position and fluid rest frame momentum and energy, which are not canonically conjugate variables. In that situation one can employ the approach of Cary & Littlejohn (1983) for non-canonical Hamiltonian mechanics. To that end we quite generally define an action $\mathcal{A}$ for a freely falling particle in curved space-time with the metric $g_{\mu\nu}$:

$$\mathcal{A} = \int_{\lambda_1}^{\lambda_2} [p_\mu dx^\mu - \mathcal{H}(p, x) d\lambda] = \int_{\lambda_1}^{\lambda_2} \left[p_\mu \frac{dx^\mu}{d\lambda} - \mathcal{H}(p, x) \right] d\lambda. \quad (\text{A1})$$

Here λ is an affine parameter that parametrizes the four-distance along the particle world line, and $p_\mu \equiv g_{\mu\sigma}(x) p^\sigma$. For a free-falling particle the Hamiltonian $\mathcal{H}(p, x)$ is

$$\mathcal{H}(p, x) = \frac{1}{2m} p_\mu p^\mu = \frac{1}{2m} g^{\mu\nu}(x) p_\mu p_\nu. \quad (\text{A2})$$

The quantity

$$\mathcal{L}(x, p) = p_\mu \frac{dx^\mu}{d\lambda} - \mathcal{H}(p, x) \quad (\text{A3})$$

is known as the phase-space Lagrangian. The principle of stationary action is $\delta\mathcal{A} = 0$, where the variations in p_μ and x^μ vanish at the end points of the integration over λ : $\delta x^\mu(\lambda_{1,2}) = \delta p_\mu(\lambda_{1,2}) = 0$. We have

$$\delta\mathcal{A} = \int_{\lambda_1}^{\lambda_2} \delta\mathcal{L}(x, p) d\lambda = \int_{\lambda_1}^{\lambda_2} \left\{ \delta p_\mu \frac{dx^\mu}{d\lambda} + p_\mu \delta \left(\frac{dx^\mu}{d\lambda} \right) - \left(\frac{\partial\mathcal{H}}{\partial p_\mu} \delta p_\mu + \frac{\partial\mathcal{H}}{\partial x^\mu} \delta x^\mu \right) \right\} d\lambda. \quad (\text{A4})$$

Since $\delta(dx^\mu/d\lambda) = d\delta x^\mu/d\lambda$ we can partially integrate the second term:

$$\delta\mathcal{A} = \int_{\lambda_1}^{\lambda_2} \left\{ \delta p_\mu \left(\frac{dx^\mu}{d\lambda} - \frac{\partial\mathcal{H}}{\partial p_\mu} \right) - \delta x^\mu \left(\frac{dp_\mu}{d\lambda} + \frac{\partial\mathcal{H}}{\partial x^\mu} \right) \right\} d\lambda = 0. \quad (\text{A5})$$

With δp_μ and δx^μ arbitrary (except at the end points) we find the equations of motion (Euler–Lagrange equations):

$$\frac{dx^\mu}{d\lambda} = \frac{\partial\mathcal{H}}{\partial p_\mu} = \frac{p^\mu}{m}, \quad \frac{dp_\mu}{d\lambda} = -\frac{\partial\mathcal{H}}{\partial x^\mu} = -\frac{1}{2m} \left(\frac{\partial g^{\alpha\beta}}{\partial x^\mu} \right) p_\alpha p_\beta. \quad (\text{A6})$$

These are essentially the Hamiltonian equations and p_μ and x^μ are canonically conjugate variables. The first equation is simply $dx^\mu/d\lambda = u^\mu = p^\mu/m$, the correct definition for the four-velocity. The second equation can be cast in a somewhat more familiar form by employing the relation

$$\frac{\partial g^{\sigma\tau}}{\partial x^\mu} = -g^{\sigma\alpha} g^{\tau\beta} \frac{\partial g_{\alpha\beta}}{\partial x^\mu}. \quad (\text{A7})$$

This follows from differentiation of the relation $g^{\sigma\tau} = g^{\sigma\alpha} g^{\tau\beta} g_{\alpha\beta}$, the symmetry of the metric tensor, and relations like $g_{\alpha\beta} g^{\tau\beta} = \delta_\alpha^\tau$. This allows us to write the second equation in equation (A6) in terms of $u_\mu = p_\mu/m$ and the well-known connection coefficients $\Gamma_{\beta\gamma}^\alpha$ (see for

instance Carroll 2004, chapter 3):

$$\frac{du_\mu}{d\lambda} = \frac{1}{2} \left(\frac{\partial g_{\alpha\beta}}{\partial x^\mu} \right) u^\alpha u^\beta = \Gamma_{\mu\beta}^\tau u^\beta u_\tau. \quad (\text{A8})$$

This is the correct form for the parallel transport of the covariant momentum vector of a freely falling particle in curved space–time.

We now change momentum variables from the LF four-momentum p_μ to the FRF four-momentum $p_{\hat{a}} = E_{\hat{a}}^\mu(x) p_\mu$. We require that the new momentum variables are locally Lorentzian, i.e. the base vectors $\hat{e}_i$ employed in the FRF satisfy $\hat{e}_{\hat{a}} \cdot \hat{e}_{\hat{b}} = \eta_{\hat{a}\hat{b}}$, where $\hat{e}_{\hat{a}} = E_{\hat{a}}^\mu(x) e_\mu$ with the e_μ the base vectors of a coordinated basis that satisfy $e_\mu \cdot e_\nu = g_{\mu\nu}(x)$. Then $p_\mu p^\mu = g_{\mu\nu} p^\mu p^\nu = p_{\hat{a}} p^{\hat{a}} = \eta_{\hat{a}\hat{b}} p^{\hat{a}} p^{\hat{b}}$. In addition, suppressing the space–time dependence in $E_{\hat{a}}^\mu(x)$ and $g_{\mu\nu}$,

$$g^{\mu\nu} E_{\hat{a}}^\mu E_{\hat{b}}^\nu = \eta^{\hat{a}\hat{b}}, \eta_{\hat{a}\hat{b}} E_{\hat{a}}^\mu E_{\hat{b}}^\nu = g_{\mu\nu}, E_{\hat{a}}^\mu E_{\hat{b}}^\nu = g_{\mu\nu}^\nu = \delta_\mu^\nu, E_{\hat{a}}^\mu E_{\hat{b}}^\nu = \eta_{\hat{b}}^{\hat{a}} = \delta_{\hat{b}}^{\hat{a}}. \quad (\text{A9})$$

These equations are the generalization to curved space–time of the relations in equation (4). Together they form the basis of the *tetrad formalism* employed by Webb (1985) in his derivation of the transport equation. A discussion of this formalism can be found in Carroll (2004, appendix J), whose notation we largely follow, and in Zee (2013, ch. IX.7), who employs the term *Vielbein* rather than tetrad. For our derivation of the equation of motion in mixed variables we will only need the relations given above. At the end of this section we show that the results are equivalent to those obtained using the full tetrad formalism. Using the FRF momentum variables the phase-space Lagrangian is

$$\mathcal{L}(x, p) = p_{\hat{a}} E_{\hat{a}}^\mu(x) \frac{dx^\mu}{d\lambda} - \frac{p_{\hat{a}} p^{\hat{a}}}{2m} \equiv p_{\hat{a}} E_{\hat{a}}^\mu(x) \frac{dx^\mu}{d\lambda} - \hat{\mathcal{H}}(p). \quad (\text{A10})$$

The ‘new’ Hamiltonian $\hat{\mathcal{H}}(p) = p^{\hat{a}} p_{\hat{a}}/2m = \eta^{\hat{a}\hat{b}} p_{\hat{a}} p_{\hat{b}}/2m$ is now formally independent of space–time position. The principle of stationary action under the variation $(\delta x^\mu, \delta p_{\hat{a}})$ yields the following set of Euler–Lagrange equations:

$$E_{\hat{a}}^\mu \frac{dx^\mu}{d\lambda} = \frac{\partial \hat{\mathcal{H}}}{\partial p_{\hat{a}}} = \frac{p^{\hat{a}}}{m}, \quad \frac{d}{d\lambda} (E_{\hat{a}}^\mu p_{\hat{a}}) - p_{\hat{a}} \left(\frac{\partial E_{\hat{a}}^\mu}{\partial x^\mu} \right) \frac{dx^\mu}{d\lambda} = -\frac{\partial \hat{\mathcal{H}}}{\partial x^\mu} = 0. \quad (\text{A11})$$

Here we used $p^{\hat{a}} = \eta^{\hat{a}\hat{b}} p_{\hat{b}}$ and $\delta E_{\hat{a}}^\mu = (\partial E_{\hat{a}}^\mu / \partial x^\mu) \delta x^\mu$. With equation (A9) the Euler–Lagrange equations can be written as

$$\frac{dx^\nu}{d\lambda} = E_{\hat{a}}^\nu \frac{p^{\hat{a}}}{m} = u^\nu, \quad \frac{dp_{\hat{a}}}{d\lambda} = \omega_{\hat{a}\hat{b}}^{\hat{c}} \frac{p_{\hat{a}} p^{\hat{b}}}{m}. \quad (\text{A12})$$

Defining $\partial_{\hat{a}} = E_{\hat{a}}^\mu (\partial / \partial x^\mu)$ the $\omega_{\hat{a}\hat{b}}^{\hat{c}}$ are given by

$$\omega_{\hat{a}\hat{b}}^{\hat{c}} \equiv E_{\hat{b}}^\nu E_{\hat{c}}^\mu \left(\frac{\partial E_{\hat{a}}^\nu}{\partial x^\mu} - \frac{\partial E_{\hat{a}}^\mu}{\partial x^\nu} \right) = E_{\hat{b}}^\nu \partial_{\hat{c}} E_{\hat{a}}^\nu - E_{\hat{c}}^\nu \partial_{\hat{b}} E_{\hat{a}}^\nu. \quad (\text{A13})$$

These equations *do not* take the canonical form: x^μ and $p_{\hat{a}}$ are not canonically conjugate variables.

Result (A12) is in agreement with the results of the full tetrad formalism, as we now show briefly. The $\omega_{\hat{a}\hat{b}}^{\hat{c}}$ are the commutation coefficients of the orthonormal tetrad base vectors: $\hat{e}_i: [\hat{e}_{\hat{a}}, \hat{e}_{\hat{b}}] = \omega_{\hat{a}\hat{b}}^{\hat{c}} \hat{e}_{\hat{c}}$. The second equation in equation (A12) is equivalent with the standard form in terms of $u_{\hat{a}} = p_{\hat{a}}/m$:

$$\frac{du_{\hat{a}}}{d\lambda} + \Gamma_{\hat{a}\hat{b}\hat{c}} u^{\hat{b}} u^{\hat{c}} = 0. \quad (\text{A14})$$

Here $\Gamma_{\hat{a}\hat{b}\hat{c}} \equiv \eta_{\hat{a}\hat{d}} \Gamma_{\hat{c}\hat{b}}^{\hat{d}}$ and the tetrad connection coefficients $\Gamma_{\hat{c}\hat{b}}^{\hat{d}}$ are defined in terms of the covariant derivative of the $E_{\hat{a}}^\mu$ as

$$\Gamma_{\hat{c}\hat{b}}^{\hat{d}} \equiv E_{\hat{b}}^\nu E_{\hat{c}}^\mu \left(\frac{\partial E_{\hat{d}}^\nu}{\partial x^\mu} + \Gamma_{\mu\lambda}^\nu E_{\hat{d}}^\lambda \right) = -E_{\hat{b}}^\mu E_{\hat{c}}^\nu \left(\frac{\partial E_{\hat{d}}^\mu}{\partial x^\nu} - \Gamma_{\mu\nu}^\lambda E_{\hat{d}}^\lambda \right). \quad (\text{A15})$$

Defining $\omega_{\hat{a}\hat{b}\hat{c}} = \eta_{\hat{a}\hat{d}} \omega_{\hat{b}\hat{c}}^{\hat{d}}$ and using that in general relativity space is torsion free with $\Gamma_{\mu\nu}^\lambda = \Gamma_{\nu\mu}^\lambda$, so that in the definition (A13) for the $\omega_{\hat{a}\hat{b}}^{\hat{c}}$ ordinary derivatives can be replaced by covariant derivatives, and one can show (e.g. Misner et al. 1973, p. 358)

$$\Gamma_{\hat{a}\hat{b}\hat{c}} = -\Gamma_{\hat{b}\hat{a}\hat{c}} = \frac{1}{2} (\omega_{\hat{b}\hat{a}\hat{c}} + \omega_{\hat{c}\hat{a}\hat{b}} - \omega_{\hat{a}\hat{b}\hat{c}}), \quad \omega_{\hat{a}\hat{b}\hat{c}} = -\omega_{\hat{a}\hat{c}\hat{b}}. \quad (\text{A16})$$

With these relations one can prove $\Gamma_{\hat{a}\hat{b}\hat{c}} u^{\hat{b}} u^{\hat{c}} = \frac{1}{2} (\Gamma_{\hat{a}\hat{b}\hat{c}} + \Gamma_{\hat{a}\hat{c}\hat{b}}) u^{\hat{b}} u^{\hat{c}} = \omega_{\hat{b}\hat{a}\hat{c}} u^{\hat{b}} u^{\hat{c}} = -\omega_{\hat{c}\hat{a}\hat{b}} u^{\hat{b}} u^{\hat{c}} = -\omega_{\hat{c}\hat{a}}^{\hat{b}} u_{\hat{b}} u^{\hat{c}}$, and the second relation in equation (A12) follows after a relabelling of indices and writing everything in terms of momenta rather than four-velocities.

Finally, in flat space–time the term involving $E_{\hat{b}}^\nu (\partial E_{\hat{a}}^\nu / \partial x^\mu)$ in the equation for $p_{\hat{c}}$ does not contribute. From the above definitions it follows generally that

$$E_{\hat{b}}^\nu \left(\frac{\partial E_{\hat{a}}^\nu}{\partial x^\mu} \right) \frac{p_{\hat{a}} p^{\hat{b}}}{m} = \frac{1}{2m} \left(\frac{\partial g_{\alpha\beta}}{\partial x^\mu} \right) p^\alpha p^\beta, \quad (\text{A17})$$

which vanishes in flat space–time where $g_{\alpha\beta} = \eta_{\alpha\beta}$. In that case the second equation in equation (A12) reduces to equation (9) of the main paper.

APPENDIX B: CALCULATION OF THE FICTITIOUS FORCE TERMS**B1 Energy change**

The change in the FRF energy $\bar{\mathcal{E}}$ follows from using $p_0 = \bar{\mathcal{E}}$ and $E_0^\mu = U^\mu$ in equation (9) with $\hat{c} = 0$:

$$\frac{dp_0}{d\lambda} = \frac{d\bar{\mathcal{E}}}{d\lambda} = -E_0^\mu \frac{dE_\mu^{\hat{a}}}{d\lambda} p_{\hat{a}} = -U^\mu \left(\frac{d\mathbf{U}}{d\lambda} \cdot \frac{\partial E_\mu^{\hat{a}}}{\partial \mathbf{U}} \right) p_{\hat{a}}. \quad (\text{B1})$$

Here we employ the three-vector notation (with implied summation over double roman indices)

$$\frac{d\mathbf{U}}{d\lambda} \cdot \frac{\partial E_\mu^{\hat{a}}}{\partial \mathbf{U}} \equiv \frac{dU^k}{d\lambda} \frac{\partial E_\mu^{\hat{a}}}{\partial U^k}. \quad (\text{B2})$$

Since $p_{\hat{a}} = (\bar{\mathcal{E}}, -\bar{\mathbf{p}})$ one has

$$p_{\hat{a}} \frac{\partial E_\mu^{\hat{a}}}{\partial \mathbf{U}} = \bar{\mathcal{E}} \frac{\partial E_0^{\hat{0}}}{\partial \mathbf{U}} - \bar{p}^i \frac{\partial E_\mu^{\hat{i}}}{\partial \mathbf{U}} \quad (\text{B3})$$

and

$$\frac{d\bar{\mathcal{E}}}{d\lambda} = -U^\mu \left[\bar{\mathcal{E}} \left(\frac{d\mathbf{U}}{d\lambda} \cdot \frac{\partial E_0^{\hat{0}}}{\partial \mathbf{U}} \right) - \bar{p}^i \left(\frac{d\mathbf{U}}{d\lambda} \cdot \frac{\partial E_\mu^{\hat{i}}}{\partial \mathbf{U}} \right) \right]. \quad (\text{B4})$$

It is convenient to look at the contribution from $\mu = 0$ and from $\mu = 1, 2, 3$ separately. Using the relations (13) from the main paper in the form

$$\frac{\partial E_0^{\hat{0}}}{\partial \mathbf{U}} = \mathbf{U} / \Gamma, \quad \bar{p}^i \frac{\partial E_0^{\hat{i}}}{\partial \mathbf{U}} = -\bar{\mathbf{p}}, \quad (\text{B5})$$

one sees that the $\mu = 0$ term gives a contribution

$$\left(\frac{d\bar{\mathcal{E}}}{d\lambda} \right)_{\mu=0} = -(\bar{\mathcal{E}} \mathbf{U} + \Gamma \bar{\mathbf{p}}) \cdot \frac{d\mathbf{U}}{d\lambda}. \quad (\text{B6})$$

The terms with $\mu = 1, 2, 3$ follow from the second and third relation in equation (13) and add up to

$$\left(\frac{d\bar{\mathcal{E}}}{d\lambda} \right)_{\mu=1,2,3} = -U^j \left[\bar{\mathcal{E}} (-\delta_{jk}) - \frac{\bar{p}^i}{\Gamma + 1} \left(\delta_{jk} U^i + \delta_{ik} U^j - \frac{U^i U^j U^k}{\Gamma(\Gamma + 1)} \right) \right] \frac{dU^k}{d\lambda}. \quad (\text{B7})$$

In three-vector notation

$$\left(\frac{d\bar{\mathcal{E}}}{d\lambda} \right)_{\mu=1,2,3} = \left[\left(\bar{\mathcal{E}} + \frac{\bar{\mathbf{p}} \cdot \mathbf{U}}{\Gamma + 1} - \frac{U^2}{\Gamma(\Gamma + 1)^2} (\bar{\mathbf{p}} \cdot \mathbf{U}) \right) \mathbf{U} + \left(\frac{U^2}{\Gamma + 1} \right) \bar{\mathbf{p}} \right] \cdot \frac{d\mathbf{U}}{d\lambda}. \quad (\text{B8})$$

Adding contributions (B6) and (B8) one finds that the terms $\propto \bar{\mathcal{E}}$ cancel. Then by using $U^2 = \Gamma^2 - 1$ the remaining terms can be written as

$$\frac{d\bar{\mathcal{E}}}{d\lambda} = - \left(\bar{\mathbf{p}} - \frac{(\bar{\mathbf{p}} \cdot \mathbf{U})}{\Gamma(\Gamma + 1)} \mathbf{U} \right) \cdot \frac{d\mathbf{U}}{d\lambda} = -(\Gamma + 1) \left[\bar{\mathbf{p}} \cdot \frac{d}{d\lambda} \left(\frac{\mathbf{U}}{\Gamma + 1} \right) \right]. \quad (\text{B9})$$

The second equality follows from $\mathbf{U} \cdot (d\mathbf{U}/d\lambda) = \Gamma(d\Gamma/d\lambda)$.

B2 Momentum change

The change in FRF three-momentum $\bar{\mathbf{p}}$ is calculated from equation (9) using $p_i = -\hat{p}^i$ so that

$$\frac{d\hat{p}^i}{d\lambda} = -\frac{d\bar{p}_i}{d\lambda} = E_i^\mu \frac{dE_\mu^{\hat{a}}}{d\lambda} p_{\hat{a}} = E_i^\mu \left(\bar{\mathcal{E}} \frac{dE_\mu^{\hat{0}}}{d\lambda} - \bar{p}^j \frac{dE_\mu^{\hat{j}}}{d\lambda} \right) = E_i^\mu \left(\bar{\mathcal{E}} \frac{\partial E_\mu^{\hat{0}}}{\partial \mathbf{U}} - \bar{p}^j \frac{\partial E_\mu^{\hat{j}}}{\partial \mathbf{U}} \right) \cdot \frac{d\mathbf{U}}{d\lambda}. \quad (\text{B10})$$

Again using the first two relations in equation (13) (or equivalently equation B2) it follows that the contribution of the $\mu = 0$ term is (in three-vector notation)

$$\left(\frac{d\bar{\mathbf{p}}}{d\lambda} \right)_{\mu=0} = \left\{ (\bar{\mathcal{E}} \mathbf{V} + \bar{\mathbf{p}}) \cdot \frac{d\mathbf{U}}{d\lambda} \right\} \mathbf{U}. \quad (\text{B11})$$

The second and third relation in equation (13) yields for the terms from $\mu = 1, 2, 3$

$$\left(\frac{d\bar{p}^i}{d\lambda} \right)_{\mu=1,2,3} = \left(\delta_{si} + \frac{U^s U^i}{\Gamma + 1} \right) \left\{ \bar{\mathcal{E}} (-\delta_{sk}) - \frac{\bar{p}^j}{\Gamma + 1} \left(\delta_{sk} U^j + \delta_{jk} U^s - \frac{U^s U^j U^k}{\Gamma(\Gamma + 1)} \right) \right\} \frac{dU^k}{d\lambda}. \quad (\text{B12})$$

The terms involving the Kronecker delta function δ_{si} can be written in three-vector notation as

$$\left(\frac{d\bar{\mathbf{p}}}{d\lambda} \right)_{\mu=1,2,3}^{(1)} = - \left(\bar{\mathcal{E}} + \frac{\bar{\mathbf{p}} \cdot \mathbf{U}}{\Gamma + 1} \right) \frac{d\mathbf{U}}{d\lambda} - \frac{\mathbf{U}}{\Gamma + 1} \left\{ \left(\bar{\mathbf{p}} \cdot \frac{d\mathbf{U}}{d\lambda} \right) - \frac{\bar{\mathbf{p}} \cdot \mathbf{U}}{\Gamma(\Gamma + 1)} \left(\mathbf{U} \cdot \frac{d\mathbf{U}}{d\lambda} \right) \right\}. \quad (\text{B13})$$

The terms involving the dyadic product $U^s U^i$ yield:

$$\left(\frac{d\bar{p}^i}{d\lambda} \right)_{\mu=1,2,3}^{(2)} = -\frac{U^i}{\Gamma+1} \left(\bar{\mathcal{E}} + \frac{\bar{\mathbf{p}} \cdot \mathbf{U}}{\Gamma+1} \right) \left(\mathbf{U} \cdot \frac{d\mathbf{U}}{d\lambda} \right) - \frac{U^2 U^i}{(\Gamma+1)^2} \left\{ \left(\bar{\mathbf{p}} \cdot \frac{d\mathbf{U}}{d\lambda} \right) - \frac{\bar{\mathbf{p}} \cdot \mathbf{U}}{\Gamma(\Gamma+1)} \left(\mathbf{U} \cdot \frac{d\mathbf{U}}{d\lambda} \right) \right\}. \quad (\text{B14})$$

Using $U^2 = \Gamma^2 - 1 = (\Gamma - 1)(\Gamma + 1)$ and changing to three-vector notation this becomes:

$$\left(\frac{d\bar{\mathbf{p}}}{d\lambda} \right)_{\mu=1,2,3}^{(2)} = -\frac{\mathbf{U}}{\Gamma+1} \left\{ \left(\bar{\mathcal{E}} + \frac{\bar{\mathbf{p}} \cdot \mathbf{U}}{\Gamma(\Gamma+1)} \right) \left(\mathbf{U} \cdot \frac{d\mathbf{U}}{d\lambda} \right) + (\Gamma - 1) \left(\bar{\mathbf{p}} \cdot \frac{d\mathbf{U}}{d\lambda} \right) \right\}. \quad (\text{B15})$$

Adding the two terms (B13) and (B15) leads to a number of cancellations and one finds for the term from $\mu = 1, 2, 3$:

$$\left(\frac{d\bar{\mathbf{p}}}{d\lambda} \right)_{\mu=1,2,3} = -\left(\bar{\mathcal{E}} + \frac{\bar{\mathbf{p}} \cdot \mathbf{U}}{\Gamma+1} \right) \frac{d\mathbf{U}}{d\lambda} - \frac{\mathbf{U}}{\Gamma+1} \left\{ \Gamma \left(\bar{\mathbf{p}} \cdot \frac{d\mathbf{U}}{d\lambda} \right) + \bar{\mathcal{E}} \left(\mathbf{U} \cdot \frac{d\mathbf{U}}{d\lambda} \right) \right\}. \quad (\text{B16})$$

Finally, adding equation (B11) and (B16) and using $\mathbf{U} = \Gamma \mathbf{V}$ one finds the total FRF three-momentum change due to the fictitious force:

$$\frac{d\bar{\mathbf{p}}}{d\lambda} = -\left(\bar{\mathcal{E}} + \frac{\bar{\mathbf{p}} \cdot \mathbf{U}}{\Gamma+1} \right) \frac{d\mathbf{U}}{d\lambda} + \left((\bar{\mathbf{p}} + \bar{\mathcal{E}} \mathbf{V}) \cdot \frac{d\mathbf{U}}{d\lambda} \right) \frac{\mathbf{U}}{\Gamma+1}. \quad (\text{B17})$$

The relation $\mathbf{U} \cdot (d\mathbf{U}/d\lambda) = \Gamma(d\Gamma/d\lambda)$ allows one to cast this into the form

$$\frac{d\bar{\mathbf{p}}}{d\lambda} = -(\Gamma+1) \bar{\mathcal{E}} \frac{d}{d\lambda} \left(\frac{\mathbf{U}}{\Gamma+1} \right) + \frac{1}{\Gamma+1} \left[\left(\bar{\mathbf{p}} \cdot \frac{d\mathbf{U}}{d\lambda} \right) \mathbf{U} - (\bar{\mathbf{p}} \cdot \mathbf{U}) \frac{d\mathbf{U}}{d\lambda} \right]. \quad (\text{B18})$$

For our purposes this is most conveniently written as

$$\frac{d\bar{\mathbf{p}}}{d\lambda} = -(\Gamma+1) \bar{\mathcal{E}} \frac{d}{d\lambda} \left(\frac{\mathbf{U}}{\Gamma+1} \right) + \bar{\mathbf{p}} \times \left(\frac{\mathbf{U} \times \frac{d\mathbf{U}}{d\lambda}}{\Gamma+1} \right) \equiv \bar{\gamma} (\mathbf{K}_f + \bar{\mathbf{p}} \times \bar{\boldsymbol{\omega}}). \quad (\text{B19})$$

In this expression we introduce $\mathbf{K}_f$ and $\bar{\boldsymbol{\omega}}$ that are defined as

$$\mathbf{K}_f \equiv -(\Gamma+1) \frac{d}{d\lambda} \left(\frac{m\mathbf{U}}{\Gamma+1} \right), \quad \bar{\boldsymbol{\omega}} \equiv \frac{\mathbf{U} \times \frac{d\mathbf{U}}{d\lambda}}{\bar{\gamma}(\Gamma+1)}. \quad (\text{B20})$$

Relations (B9) and (B19/B20) are used in the main paper. It is easily checked using (B9) and (B19) that the fictitious four-force respects the Lorentz-invariant $p^\mu p_\mu = p^{\hat{a}} p_{\hat{a}} = \bar{\mathcal{E}}^2 - |\bar{\mathbf{p}}|^2 = m^2$, as $\bar{\mathcal{E}}(d\bar{\mathcal{E}}/d\lambda) - \bar{\mathbf{p}} \cdot (d\bar{\mathbf{p}}/d\lambda) = 0$. This serves as an independent check on the calculation presented here.

APPENDIX C: RELATIVISTIC MHD

We briefly derive a number of relations from relativistic MHD that are needed in the main paper. Additional details can be found in the book by Lichnerowicz (1967), the seminal paper of Bekenstein & Oron (1978), and in Achterberg (1983). Quite generally, one can define electric and magnetic four-vectors using the four-velocity U^μ of the fluid, Faraday tensor $F^{\mu\nu}$ and its dual $*F^{\mu\nu}$:

$$\mathcal{E}^\mu \equiv F^{\mu\nu} U_\nu, \quad \mathcal{B}^\mu \equiv *F^{\mu\nu} U_\nu. \quad (\text{C1})$$

Here $F^{\mu\nu}$ is the Faraday tensor and $*F^{\mu\nu} = \frac{1}{2} \epsilon^{\mu\nu\sigma\tau} F_{\sigma\tau}$ is its Hodge dual. In the FRF one has $\mathcal{E}^\mu = (0, \bar{\mathbf{E}})$ and $\mathcal{B}^\mu = (0, \bar{\mathbf{B}})$ with $(\bar{\mathbf{E}}, \bar{\mathbf{B}})$ the electromagnetic fields in that frame. The antisymmetry of $F^{\mu\nu}$ and its Hodge dual $*F^{\mu\nu}$ implies that $\mathcal{E}^\mu U_\mu = \mathcal{B}^\mu U_\mu = 0$. Ideal MHD holds for the plasma at large when

$$\mathcal{E}^\mu = 0, \quad (\text{C2})$$

which is equivalent with the well-known condition $\mathbf{E} + \mathbf{V} \times \mathbf{B} = 0$, valid in any frame. In that case the Faraday tensor and its dual are given by

$$F^{\mu\nu} = \epsilon^{\mu\nu\sigma\tau} U_\sigma B_\tau, \quad *F^{\mu\nu} = U^\nu \mathcal{B}^\mu - U^\mu \mathcal{B}^\nu. \quad (\text{C3})$$

The induction equation is $\nabla_\mu (*F^{\mu\nu}) = \nabla_\mu (U^\nu \mathcal{B}^\mu - U^\mu \mathcal{B}^\nu) = 0$. The resulting set of equations for the magnetic field can be written as

$$(U \cdot \nabla) \left(\frac{\mathcal{B}^\mu}{n} \right) \equiv \frac{d}{d\tau} \left(\frac{\mathcal{B}^\mu}{n} \right) = \left(\frac{\mathcal{B}}{n} \cdot \nabla \right) U^\mu - \left(a \cdot \frac{\mathcal{B}}{n} \right) U^\mu, \quad \nabla \cdot \mathcal{B} = -a \cdot \mathcal{B}. \quad (\text{C4})$$

Here n is the proper density of the fluid that satisfies the particle conservation law

$$dn/d\tau = -n(\nabla \cdot \mathbf{U}), \quad (\text{C5})$$

and $a^\mu \equiv dU^\mu/d\tau$ is the fluid four-acceleration that satisfies $U \cdot a = 0$. We employ the notation $A \cdot B \equiv A_\mu B^\mu$ throughout. The perhaps unfamiliar terms in these two equations (in a direct comparison with the corresponding equations in non-relativistic MHD) involve the

four-acceleration a^μ . They are needed to maintain the orthogonality between $\mathcal{B}^\mu$ and U^μ , (i.e. $\mathcal{B} \cdot U = 0$). These terms can be neglected in a non-relativistic flow.

Contracting the first equation in equation (C4) with $\mathcal{B}_\mu/n$ it follows with the help of equation (C5) that

$$\frac{1}{B} \frac{dB}{d\tau} = -(b \cdot b : \nabla U + \nabla \cdot U). \quad (C6)$$

Here $B \equiv \sqrt{-\mathcal{B}^\mu \mathcal{B}_\mu}$ is the magnetic field strength in the FRF, where $\mathcal{B}^\mu = (0, \cdot, \bar{\mathcal{B}})$ so that $\mathcal{B}^\mu \mathcal{B}_\mu = -|\bar{\mathcal{B}}|^2$. The vector b^μ is defined as $b^\mu \equiv \mathcal{B}^\mu/B$ so that $b^\mu b_\mu = -1$. It takes the form $b^\mu = (0, \hat{\mathbf{b}})$ in the FRF. We employ the notation $b \cdot b : \nabla U \equiv b_\mu b^\nu \nabla_\nu U^\mu$. The second relation in equation (C4) leads, after substituting $\mathcal{B}^\mu = B b^\mu$ to

$$\nabla \cdot b = -\frac{(b \cdot \nabla)B}{B} - a \cdot b = -\frac{(b \cdot \nabla)B}{B} - b \cdot U : \nabla U. \quad (C7)$$

We will convert the above results to a mixed form employing FRF magnetic field variables B and $\hat{\mathbf{b}}$ and LF time t and position $\mathbf{x}$. In the FRF one has *locally* $U^{\hat{a}} = (1, 0, 0, 0)$ and the covariant gradient operator is (in column notation)

$$\nabla_{\hat{a}} = E_{\hat{a}}^\mu \nabla_\mu = \begin{pmatrix} \Gamma \frac{\partial}{\partial t} + (\mathbf{U} \cdot \nabla) \\ \nabla + \mathbf{U} \left(\frac{\partial}{\partial t} + \frac{\mathbf{U} \cdot \nabla}{\Gamma + 1} \right) \end{pmatrix} = \begin{pmatrix} \frac{d}{d\tau} \\ \tilde{\nabla} \end{pmatrix}. \quad (C8)$$

The velocity gradient tensor evaluated in the FRF is

$$\nabla_{\hat{a}} U^{\hat{b}} = E_{\hat{a}}^\mu E_{\hat{b}}^\nu (\nabla_\mu U^\nu) = E_{\hat{b}}^\nu \nabla_{\hat{a}} U^\nu = \begin{pmatrix} 0 & (\Gamma + 1) \frac{d}{d\tau} \left(\frac{\mathbf{U}}{\Gamma + 1} \right) \\ \mathbf{0} & (\Gamma + 1) \tilde{\nabla} \left(\frac{\mathbf{U}}{\Gamma + 1} \right) \end{pmatrix}, \quad (C9)$$

using a 4×4 matrix notation in the last equality. This implies for the four-divergence of U^μ :

$$\nabla \cdot U = \frac{\partial \Gamma}{\partial t} + \nabla \cdot \mathbf{U} = \nabla_{\hat{a}} U^{\hat{a}} = (\Gamma + 1) \tilde{\nabla} \cdot \left(\frac{\mathbf{U}}{\Gamma + 1} \right), \quad (C10)$$

an identity that is easily checked using the definition of $\tilde{\nabla}$ (equation 34). The FRF four-acceleration is $a^{\hat{a}} = E_{\hat{a}}^\mu (dU^\mu/d\tau) = (0, \bar{\mathbf{a}})$, with

$$\bar{\mathbf{a}} = (\Gamma + 1) \frac{d}{d\tau} \left(\frac{\mathbf{U}}{\Gamma + 1} \right). \quad (C11)$$

Using these relations equation (C6) becomes

$$\frac{1}{B} \frac{dB}{d\tau} = (\Gamma + 1) \left\{ \hat{\mathbf{b}} \hat{\mathbf{b}} : \tilde{\nabla} \left(\frac{\mathbf{U}}{\Gamma + 1} \right) - \tilde{\nabla} \cdot \left(\frac{\mathbf{U}}{\Gamma + 1} \right) \right\}, \quad (C12)$$

and relation (C7) can be written as

$$\tilde{\nabla} \cdot \hat{\mathbf{b}} = -\frac{(\hat{\mathbf{b}} \cdot \tilde{\nabla})B}{B} + (\Gamma + 1) \left\{ \hat{\mathbf{b}} \cdot \frac{d}{d\tau} \left(\frac{\mathbf{U}}{\Gamma + 1} \right) \right\}. \quad (C13)$$

The second term on the right-hand side is a relativistic correction that occurs when the comoving frame is not an inertial frame. It can be neglected in a non-relativistic flow with $|\mathbf{U}| \ll 1$. In that same limit equations (107) through (72) reduce to known results. Defining $\mathbf{B} = B \hat{\mathbf{b}}$ and $d/dt = \partial/\partial t + (\mathbf{V} \cdot \nabla)$, the comoving derivative

$$\bar{\mathbf{a}} = \frac{d\mathbf{V}}{dt}, \quad \frac{dB}{dt} = B \left(\hat{\mathbf{b}} \hat{\mathbf{b}} : \nabla \mathbf{V} - \nabla \cdot \mathbf{V} \right), \quad \nabla \cdot \hat{\mathbf{b}} = -\frac{(\hat{\mathbf{b}} \cdot \nabla)B}{B} \iff \nabla \cdot \mathbf{B} = 0. \quad (C14)$$

Relations (C10), (C12), and (C13) are used repeatedly in the main paper.

This paper has been typeset from a $\text{\LaTeX}$ file prepared by the author.