Notes on nonlinear gyrokinetic equation

BY YOUJUN HU

Institute of Plasma Physics, Chinese Academy of Sciences
Email: yjhu@ipp.cas.cn

Abstract

The nonlinear gyrokinetic equation given in Frieman-Chen's paper[3] is re-derived in this note, providing more details on the derivation. The aspects that are relevant to numerical implementation of the gyrokinetic model are also discussed, which are actually the emphasis of this note.

1 Introduction

Subtle things in a theory, which may be crucial for one to fully understand the theory, can be identified and understood only when one re-derives the theory by oneself.

Presently, I do not understand the modern phase-space-Lagrangian Lie perturbation method of deriving the gyrokinetic mode. So I am using the old style asymptotic expansion method in deriving the gyrokinetic model, as was done in the original Frieman-Chen's paper[3]. The motivation of deriving the gyrokinetic equation by myself is that I am developing a new kinetic model in the `GEM` code, which requires me to understand every details of the gyrokinetic theory.

2 Transform Vlasov equation from particle coordinates to guiding-center coordinates

To facilitate the derivation, we first need to choose some good variables to work with (why they are good can only be realized when we get the results, but some preliminary considerations can suggest that some variables be good).

The Vlasov equation in terms of particle coordinates $(\mathbf{x},\mathbf{v})$ is given by

$$\frac{\partial f_p}{\partial t}+\mathbf{v}\cdot\frac{\partial f_p}{\partial \mathbf{x}}+\frac{q}{m}(\mathbf{E}+\mathbf{v}\times\mathbf{B})\cdot\frac{\partial f_p}{\partial \mathbf{v}}=0, \tag{1}$$

where $f_p=f_p(\mathbf{x},\mathbf{v})$ is the particle distribution function, $\mathbf{x}$ and $\mathbf{v}$ are the location and velocity of particles. Next, we first define the guiding-center transformation. Then we transform the Vlasov equation to the guiding-center coordinates, i.e., express the gradient operators $\partial/\partial\mathbf{x}$ and $\partial/\partial\mathbf{v}$ in terms of the guiding-center variables.

2.1 Guiding-center transformation

The transformation from the particle variables $(\mathbf{x},\mathbf{v})$ to the guiding-center variables $\mathbf{X}$ is defined by[1]

$$\mathbf{X}(\mathbf{x},\mathbf{v})=\mathbf{x}+\mathbf{v}\times\frac{\boldsymbol{e}_{\parallel}(\mathbf{x})}{\Omega(\mathbf{x})}, \tag{2}$$

where $\mathbf{x}$ and $\mathbf{v}$ are particle position and velocity, $\boldsymbol{e}_{\parallel}=\mathbf{B}_0/B_0$, $\Omega=qB_0/m$, $\mathbf{B}_0=\mathbf{B}_0(\mathbf{x})$ is the equilibrium (macroscopic) magnetic field at the particle position. It is obvious that $\boldsymbol{\rho}\equiv-\mathbf{v}\times\mathbf{e}_{\parallel}/\Omega$ is the vector gyro-radius pointing from the the guiding-center to the particle position. This transformation involves both position and velocity of particles.

Given $(\mathbf{x}, \mathbf{v})$, it is straightforward to obtain $\mathbf{X}$ by using Eq. (2). On the other hand, the inverse transformation, i.e., given $(\mathbf{X}, \mathbf{v})$, to find $\mathbf{x}$. This is in principle not easy because Ω and $\mathbf{e}_\parallel$ depend on $\mathbf{x}$, which usually requires solving for the root of a nonlinear equation. Numerically, one can use

$$\mathbf{x}_{n+1} = \mathbf{X} - \mathbf{v} \times \frac{\boldsymbol{e}_\parallel(\mathbf{x}_n)}{\Omega(\mathbf{x}_n)}. \tag{3}$$

as an iteration scheme to compute $\mathbf{x}$. The initial guess can be chosen as $\mathbf{x}_0 = \mathbf{X}$. The equilibrium magnetic field we will consider has spatial scale length much larger than the thermal gyro radius $\boldsymbol{\rho}$. In this case the difference between the values of $\mathbf{e}_\parallel(\mathbf{x})/\Omega(\mathbf{x})$ and $\mathbf{e}_\parallel(\mathbf{X})/\Omega(\mathbf{X})$ is small and thus can be neglected. Then the inverse guiding-center transform can be written as

$$\mathbf{x} \approx \mathbf{X} - \mathbf{v} \times \frac{\boldsymbol{e}_\parallel(\mathbf{X})}{\Omega(\mathbf{X})}, \tag{4}$$

which can also be considered as we using the iterative scheme (3) to computer $\mathbf{x}$ and stopping at the first iteration. The difference between equilibrium field values evaluated at $\mathbf{X}$ and $\mathbf{x}$ is always neglected in gyrokinetic theory. Therefore it does not matter whether the above $\mathbf{e}_\parallel/\Omega$ is evaluated at $\mathbf{x}$ or $\mathbf{X}$. What matters is where the perturbed fields are evaluated: at $\mathbf{x}$ or at $\mathbf{X}$. The values of perturbed fields at $\mathbf{x}$ or at $\mathbf{X}$ are different and this is called the finite Larmor radius (FLR) effect, which is almost all that the gyrokinetic theory is about.

2.2 Gyro-angle

The forward and backward guiding-center transformations (2) and (4) involve the velocity vector $\mathbf{v}$. It is the cross product between $\mathbf{v}$ and $\boldsymbol{e}_\parallel(\mathbf{x})$ or $\boldsymbol{e}_\parallel(\mathbf{X})$ that is actually used. Therefore, only two of the three coordinates of $\mathbf{v}$ is actually needed in a local (at $\mathbf{x}$ or $\mathbf{X}$) cylindrical coordinate system, namely $v_\perp$ and α, where α is the azimuthal angle of the velocity in the local cylindrical system. Here α will be called the gyrophase or gyro-angle in the following.

2.3 Distribution functions in terms of guiding-center variables

Denote the distribution function in terms of the guiding-center variables $(\mathbf{X}, \mathbf{v})$ by $f_g(\mathbf{X}, \mathbf{v})$, then the relation between f_g and f_p is given by

$$f_g(\mathbf{X}, \mathbf{v}) = f_p(\mathbf{x}, \mathbf{v}), \tag{5}$$

where $\mathbf{X}$ and $\mathbf{x}$ are related to each other via Eq. (2). Equation (5) along with Eq. (2) can be considered as the definition of the guiding-center distribution function f_g. Using Eq. (2), the above relation can be equivalently written as

$$f_g(\mathbf{X}, \mathbf{v}) = f_p(\mathbf{X} + \boldsymbol{\rho}, \mathbf{v}), \tag{6}$$

or

$$f_p(\mathbf{x}, \mathbf{v}) = f_g(\mathbf{x} - \boldsymbol{\rho}, \mathbf{v}). \tag{7}$$

or $f_p(\mathbf{x}, \mathbf{v}) = f_p(\mathbf{X} + \boldsymbol{\rho}, \mathbf{v})$, or $f_g(\mathbf{X}, \mathbf{v}) = f_g(\mathbf{x} - \boldsymbol{\rho}, \mathbf{v})$.

As is conventionally adopted in multivariables calculus, both f_p and f_g are sometimes simply denoted by f. Which one is actually assumed depends on the context, i.e., depends on which independent variables are actually assumed: particle variables or guiding-center variables. This is one of the subtle (trivial?) things needed to be noted for gyrokinetic theory in particular and for multivariables calculus in general. This notation will cause confusions when, for example, $f_g(\mathbf{X}, \mathbf{v})$ is evaluated at $\mathbf{X} = \mathbf{x}$, i.e., $f_g(\mathbf{x}, \mathbf{v})$, which, if the subscript g is omitted and we rely on dependent variables to identify whether it is f_p or f_g, will be wrongly understood as f_p. Therefore it seems better to use the accurate notation. One example where this distinguishing is important is encountered when we try to express the diamagnetic flow in terms of f_g, which is discussed in Appendix A.

In practice, f_g is often called the guiding-center distribution function whereas f_p is called the particle distribution function. They are the same distribution function expressed in different variables.

In the above, we assume that $\mathbf{X}$ and $\mathbf{x}$ are always related to each other by the guiding-center transformation (2) or (4) , i.e., $\mathbf{x}$ and $\mathbf{X}$ are not independent. For some cases, it may be convienent to treat $\mathbf{x}$ and $\mathbf{X}$ as independent variables and express the guiding-center transformation via an integral of the Dirac delta function. For example, expression (7) can be written as

$$f_p(\mathbf{x},\mathbf{v}) = \int f_g(\mathbf{X},\mathbf{v})\delta^3(\mathbf{X}-\mathbf{x}+\boldsymbol{\rho})d\mathbf{X}, \tag{8}$$

where $\mathbf{x}$ and $\mathbf{X}$ are considered as independent variables, $\delta^3(\mathbf{x}-\mathbf{X}-\boldsymbol{\rho})$ is the three-dimensional Dirac delta function. [In terms of general coordinates (x_1,x_2,x_3), the three-dimensional Dirac delta function is defined via the 1D Dirac delta function as follows:

$$\delta^3(\mathbf{x}) = \frac{1}{|\mathcal{J}|}\delta(x_1)\delta(x_2)\delta(x_3), \tag{9}$$

where $\mathcal{J}$ is the the Jacobian of the general coordinate system. The Jacobian is included in order to make $\delta^3(\mathbf{x})$ satisfy the normalization condition $\int \delta^3(\mathbf{x})d\mathbf{x} = \int \delta^3(\mathbf{x})|\mathcal{J}|dx_1dx_2dx_3 = 1.$]

Expression (8) can be considered as a transformation that transforms an arbitrary function from the guiding-center coordinates to the particle coordinates. Similarly, equation (6) can be written as

$$f_g(\mathbf{X},\mathbf{v}) = \int f_p(\mathbf{x},\mathbf{v})\delta^3(\mathbf{x}-\mathbf{X}-\boldsymbol{\rho})d\mathbf{x}, \tag{10}$$

which can be considered as a transformation that transforms an arbitrary function from the the particle coordinates to the guiding-center coordinates.

2.4 Moments of distribution function expressed as integration over guiding-center variables

In terms of particle variables $(\mathbf{x},\mathbf{v})$, it is straightforward to calculate the moment of the distribution function. For example, the number density $n(\mathbf{x})$ is given by

$$n(\mathbf{x}) = \int f_p(\mathbf{x},\mathbf{v})d\mathbf{v}. \tag{11}$$

However, it is a little difficult to calculate $n(\mathbf{x})$ at real space location $\mathbf{x}$ by using the guiding-center variables $(\mathbf{X},\mathbf{v})$. This is because holding $\mathbf{x}$ constant and changing $\mathbf{v}$, as required by the integration in Eq. (11), means the guiding-center variable $\mathbf{X}$ is changing, according to Eq. (2). Using Eq. (5), expression (11) is written as

$$n(\mathbf{x}) = \int f_g(\mathbf{X}(\mathbf{x},\mathbf{v}),\mathbf{v})d\mathbf{v}, \tag{12}$$

As is mentioned above, the $d\mathbf{v}$ integration in Eq. (12) should be performed by holding $\mathbf{x}$ constant and changing $\mathbf{v}$, which means the guiding-center variable $\mathbf{X} = \mathbf{X}(\mathbf{x},\mathbf{v})$ is changing. This means that, in $(\mathbf{X},\mathbf{v})$ space, the above integration is a (generalized) line integral along the the line $\mathbf{X}(\mathbf{v}) = \mathbf{x} - \boldsymbol{\rho}(\mathbf{x},\mathbf{v})$ with $\mathbf{x}$ being constant. By using the Dirac delta function δ, this line integral can be written as the following two-dimensional integration over the independent variables $\mathbf{X}$ and $\mathbf{v}$:

$$n(\mathbf{x}) = \iint f_g(\mathbf{X},\mathbf{v})\delta^3(\mathbf{X}-\mathbf{x}+\boldsymbol{\rho})d\mathbf{v}d\mathbf{X}. \tag{13}$$

In gyrokinetic particle in cell (PIC) simulations, the integral (13) is evaluated by using Monte-Carlo markers. Actually evaluated in the simulation is the integral of a special f_g that is independent of the gyro-angle α. Using $v_\perp$ and $v_\parallel$ as the two remaining velocity coordinates, then f_g is written as $f_g = f_g(\mathbf{X},v_\perp,v_\parallel)$. In this case f_g is a function of five variables (rather than six variables).

In simulations, the position of a marker represents the guiding-center location (rather than the particle location), i.e., it is the guiding-center location that is directly sampled (the particle position is indirectly sampled, as is discussed below). For a marker with coordinators $(\mathbf{X}, v_\perp, v_\parallel)$, we can calculate particle positions by using the transformation (4). There are infinite number of particle positions associated with the marker since the direction of $v_\perp$ (i.e., gyro-agnle) is not specified. All these possible particle positions are on a circle around the guiding-center position $\mathbf{X}$. This circle is often called the gyro-ring. In particle simulations, the gyro-ring is descretized by several sampling points in calculating its integration contribution to grids. We choose N sampling points that are evenly distributed on the gyro-ring (N is usually 4 as a compromise between efficiency and accuracy). Denote the Mote-Carlo weight of the jth marker by w_j. Then the weight is evenly split by the N sub-markers on the gyro-ring since $f_g(\mathbf{X}, v_\perp, v_\parallel)$ is independent of the gyro-phase (put it another way, f_g is uniformed distributed in the gyro-phase) Therefore each sub-marker have a Monte-Carlo weight w_j/N. Then calculating the integration (13) at a grid corresponds to depositing all the N sub-markers associated with each guiding-center marker to the grid, as is illustrated in Fig. 1. A more accurate interpretation of why the charge deposition can be done this way is discussed in Sec. 2.5.

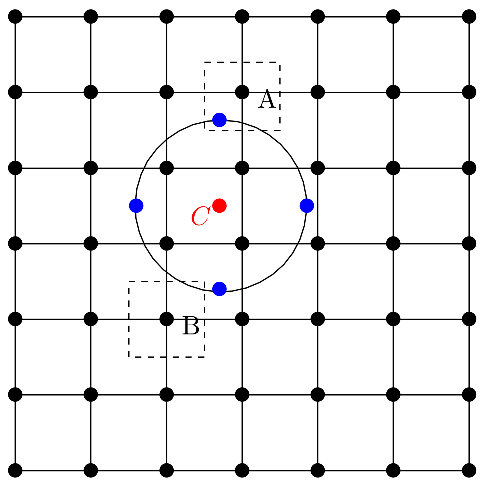

Figure 1. The spatial grids in the plane perpendicular to the equilibrium magnetic field and one gyro-ring with 4 sampling points (sub-markers) on it. For a guiding-center marker C with a Monte-Carlo weight of w_j, the 4 sub-markers are calculated by using the transformation (4) (rebuilding the gyro-ring). The Monte-Carlo weight of each sub-marker is $w_j/4$. The value of integration (13) at a grid point is approximated by $I/\Delta V$, where I is the Monte-Carlo integration of all sub-markers associated with all guiding-center markers in the cell, ΔV is the cell volume. The cell associated with a grid-point (e.g., A) is indicated by the dashed rectangle (this is for 2D case, for 3D, it is a cube). If the Dirac delta function is used as the shape function of the sub-markers, then calculating the contribution of a sub-marker to a grid corresponds to the nearest-point interpolation (for example, the 4 sub-markers will contribute noting to grid B since no-sub marker is located within the cell). In practice, the flat-top shape function with its support equal to the cell size is often used, then the depositing corresponds to linearly interpolating the weight of each sub-marker to the nearby grids.

The gyro-averaging of a perturbed field can be numerically calculated in a way similar to the above (discussed later).

2.5 Monte-Carlo sampling of 6D guiding-center phase-space

Suppose that the 6D guiding-center phase-space $(\mathbf{X}, \mathbf{v})$ is described by $(\psi, \theta, \phi, v_\parallel, v_\perp, \alpha)$ coordinates. The Jacobian of the coordinate system is given by $\mathcal{J} = \mathcal{J}_r v_\perp$, where $\mathcal{J}_r$ is the Jacobian of (ψ, θ, ϕ) coordinates. Suppose that we sample the 6D phase-space by using the following probability density function:

$$P(\psi, \theta, \phi, v_\parallel, v_\perp, \alpha) = \frac{1}{V_r}\left(\frac{m}{2\pi T}\right)^{3/2} \exp\left[-\frac{m\,(v_\parallel^2 + v_\perp^2)}{2T}\right], \tag{14}$$

where V_r is the volume of the spatial simulation box, T is a constant temperature. (P given above is actually independent of ψ, θ, ϕ and α.) It is ready to verify that P satisfies the following normalization condition:

$$\int_{V_r}\int P d\mathbf{v} d\mathbf{X} = \int_{V_r}\int_{-\infty}^{+\infty}\int_0^{\infty}\int_0^{2\pi} P v_\perp d\alpha dv_\perp dv_\parallel \mathcal{J}_r d\psi d\theta d\phi = 1. \tag{15}$$

We use the rejection method to numerically generate N_p markers that satisfy the above probability density function. [The effective probability density function actually used in the rejection method is P', which is related to P by

$$P' = |\mathcal{J}_r \mathcal{J}_v| P = |\mathcal{J}_r| v_\perp P. \tag{16}$$

Note that P' does not depend on the gyro-angle α.] Then the phase-space volume occupied by a marker is given by

$$V_{pj} = \frac{1}{g} = \frac{1}{N_p P} = \frac{V_r}{N_p \left(\frac{m}{2\pi T}\right)^{3/2} \exp\left[-\frac{m\,(v_{\parallel j}^2 + v_{\perp j}^2)}{2T}\right]}, \tag{17}$$

which is independent of ψ_j, θ_j, ϕ_j and α_j. The weight of a marker is then given by $w_j = V_{pj} \delta f_{gj}$, where δf_g is the perturbed distribution function evolved in simulations.

Although the distribution function δf_g to be sampled in gyrokinetic simulations is independent of the gyro-angle α, we still need to sample the gyro-angle because we need to use the inverse guiding-center transformation, which needs the gyro-angle. Each marker needs to have a specific gyro-angle value α_j so that we know how to transform its $\mathbf{X}_j$ to $\mathbf{x}_j$ and then do the charge deposition in $\mathbf{x}$ space. To increase the resolution over the gyro-angle, we need to increase the number of markers. However, thanks to the fact that both sampling probability density function P and δf_g are independent of α, the resolution over the gyro-angle can be increased in a simple way. Specifically, for a marker with $(\mathbf{X}_j, v_{\parallel j}, v_{\perp j})$, its gyro-angle value α_j can be adjusted arbitrary without changing the value of its weight w_j because both the phase-space volume V_{pj} and δf_{gj} are independent of α_j. In other words, we have the freedom of choosing the value of α_j and the independence of V_{pj} and δf_{gj} on α_j guarantees that the weight of the marker is still equal to the original value that this marker takes. Suppose that we choose 4 different values of α_j for the jth marker, denoted by α_{j1}, α_{j2}, α_{j3}, and α_{j4}. Then for each of the four gyro-angle values, we do the inverse guiding-center transformation and then do the charge deposition using their original weight w_j for each marker and then loop over all the markers. Denote a grid quantity (e.g. density) build by the deposition process by $n_1(\mathbf{x})$, $n_2(\mathbf{x})$, $n_3(\mathbf{x})$, and $n_4(\mathbf{x})$, corresponding to using the four gyro-angle values. Then the more accurate estimation of the grid quantity is given by

$$n(\mathbf{x}) = \frac{n_1(\mathbf{x}) + n_2(\mathbf{x}) + n_3(\mathbf{x}) + n_4(\mathbf{x})}{4}. \tag{18}$$

This corresponds to sampling the 6D phase-space 4 separate times (each time with identical sampling points in $(\mathbf{X}, v_\parallel, v_\perp)$ but different sampling points in α) and then using the averaging of the 4 Monte-Carlo integrals to estimate the exact value. This estimation can also be (roughly) considered as a Monte-Carlo estimation using 4 times larger number of markers as that is originally used (the Monte-Carlo estimation using truly 4 time larger number of markers is more accurate than the result we obtained above because the former also increase the resolution of $(\mathbf{X}, v_\parallel, v_\perp)$ while the latter keeps the resolution of $(\mathbf{X}, v_\parallel, v_\perp)$ unchanged.)

In numerical code, one may split the weight w_j as $w_j / 4$ in doing the deposition and then add the contributions from all the 4 gyro-angles when doing the deposition and then one does not need to take the averaging at the end. However, interpreting in this way is confusing to me because, with a single sampling of the phase-space, the phase-space volume or weight can not be easily split. I prefer the above interpretation that the 6D phase space is sampled 4 separate times and thus we get 4 estimations and finally we take the averaging of these 4 estimations. It took me several days to finally find this way of understanding.

In summary, the phase-space to be sampled in gyrokinetic simulations are still 6D rather than 5D. In this sense, the statement that gyrokinetic simulation works in a 5D phase space is misleading. We are still working in the 6D phase-space. The only subtle thing is that the sixth dimension, i.e., gyro-angle, can be sampled in a easy way that is independent of the other 5 variables.

In numerical implementation, the gyro-angle may not be explicitly used. We just try to calculate 4 arbitrary points on the gyro-ring that are easy to calculate. Some codes (e.g. ORB5) may introduce a random variable to rotate these 4 points for different markers so that the gyro-angle is sampled more unbiased.

2.6 Choosing velocity coordinates

Coordinates of the velocity space can be chosen as $(v, v_\parallel, \alpha)$, or $(v, v_\perp, \alpha)$, or $(v_\parallel, v_\perp, \alpha)$, where $v_\parallel$ and $v_\perp$ are the parallel and perpendicular (with respect to $\mathbf{B}_0(\mathbf{x})$) velocity, respectively, α is the azimuthal angle in a cylindrical/spherical velocity coordinate system. Here α is usually called the "gyro-angle", which is the most important variable we need because we need to directly perform averaging over this variable in deriving the gyrokinetic equation.

(**incorrect**The parallel direction is fully determined by $\mathbf{B}_0(\mathbf{x})$, but there are degrees of freedom in choosing one of the two perpendicular basis vectors. It seems that, in order to make the azimuthal angle α fully defined, we need to choose a way to define one of the two perpendicular directions. It turns out that this is not necessary because the gyro-angle is defined only locally in space. At each spatial point, the perpendicular directions can be arbitrarily chosen. In other words, for a given $\mathbf{v}$, the corresponding gyro-angles α at different spatial locations are unrelated, i.e., $\partial\alpha/\partial\mathbf{x}|_\mathbf{v}$ is not defined.** Update: we DO need to define the perpendicular direction at each spatial location to make $\partial\alpha/\partial\mathbf{x}|_\mathbf{v}$ defined, which is needed in the Vlasov differential operators. However, it seems that terms related to $\partial\alpha/\partial\mathbf{x}|_\mathbf{v}$ are finally dropped due to that they are of higher order, check***)

The gyro-angle is a velocity variable that we will stick to. The other velocity variables, $v_\parallel$, $v_\perp$, and v, can be replaced by other choices. In Frieman-Chen's paper, these variables are replaced by

$$\varepsilon = \varepsilon(v, \mathbf{x}) \equiv \frac{v^2}{2} + \frac{q\Phi_0(\mathbf{x})}{m}, \tag{19}$$

and

$$\mu = \mu(v_\perp, \mathbf{x}) \equiv \frac{v_\perp^2}{2B_0(\mathbf{x})}, \tag{20}$$

where $\Phi_0(\mathbf{x})$ is the equilibrium (macroscopic) electrical potential, $v_\perp = |\mathbf{v}_\perp|$, and $\mathbf{v}_\perp = \mathbf{v} - \mathbf{v}\cdot\mathbf{e}_\parallel$ is the perpendicular velocity. Note that $(\varepsilon, \mu, \alpha)$ is not sufficient in uniquely determining a velocity vector. An additional parameter $\sigma = \mathrm{sign}(v_\parallel)$ is needed to determine the sign of $v_\parallel = \mathbf{v}\cdot\mathbf{e}_\parallel$. In the following, the dependence of the distribution function on σ is often not explicitly shown in the variable list (i.e., σ is hidden/suppressed), which, however, does not mean that the distribution function is independent of σ. Another frequently used velocity coordinates are $(\mu, v_\parallel, \alpha)$. In the following, I will derive the gyrokinetic equation in $(\varepsilon, \mu, \alpha)$ coordinates. After that, I transform it to $(\mu, v_\parallel, \alpha)$ coordinates.

2.7 Spatial gradient operator in terms of guiding-center variables

Using the chain-rule, the spatial gradient $\partial f/\partial\mathbf{x}$ is written

$$\frac{\partial f_p}{\partial \mathbf{x}}\Big|_\mathbf{v} = \frac{\partial \mathbf{X}}{\partial \mathbf{x}}\cdot\frac{\partial f_g}{\partial \mathbf{X}} + \frac{\partial \varepsilon}{\partial \mathbf{x}}\frac{\partial f_g}{\partial \varepsilon} + \frac{\partial \mu}{\partial \mathbf{x}}\frac{\partial f_g}{\partial \mu} + \frac{\partial \alpha}{\partial \mathbf{x}}\frac{\partial f_g}{\partial \alpha}. \tag{21}$$

From the definition of $\mathbf{X}$, Eq. (2), we obtain

$$\frac{\partial \mathbf{X}}{\partial \mathbf{x}} = \mathbf{I} + \mathbf{v}\times\frac{\partial}{\partial \mathbf{x}}\left(\frac{\mathbf{e}_\parallel}{\Omega}\right), \tag{22}$$

where $\mathbf{I}$ is the unit dyad. From the definition of ε, we obtain

$$\frac{\partial \varepsilon}{\partial \mathbf{x}} = -\frac{q}{m}\mathbf{E}_0, \tag{23}$$

where $\mathbf{E}_0 = -\partial\Phi_0/\partial\mathbf{x}$. [From the definition of μ, we obtain

$$\frac{\partial \mu}{\partial \mathbf{x}} = -\frac{v_\perp^2}{2B_0^2}\frac{\partial B_0}{\partial \mathbf{x}} + \frac{1}{2B_0}\frac{\partial v_\perp^2}{\partial \mathbf{x}} = -\frac{\mu}{B_0}\frac{\partial B_0}{\partial \mathbf{x}} + \frac{\partial \mathbf{v}_\perp}{\partial \mathbf{x}}\cdot\frac{\mathbf{v}_\perp}{B_0} \tag{24}$$

Using

$$\frac{\partial \mathbf{v}_\perp}{\partial \mathbf{x}} = \frac{\partial[\mathbf{v} - v_\parallel \mathbf{e}_\parallel]}{\partial \mathbf{x}} = -v_\parallel \frac{\partial \mathbf{e}_\parallel}{\partial \mathbf{x}} - \frac{\partial v_\parallel}{\partial \mathbf{x}} \mathbf{e}_\parallel, \tag{25}$$

expression (24) is written as

$$\frac{\partial \mu}{\partial \mathbf{x}} = -\frac{\mu}{B_0} \frac{\partial B_0}{\partial \mathbf{x}} - v_\parallel \frac{\partial \mathbf{e}_\parallel}{\partial \mathbf{x}} \cdot \frac{\mathbf{v}_\perp}{B_0}, \tag{26}$$

which agrees with Eq. (10) in Frieman-Chen's paper[3]. (Note that the partial derivative $\partial/\partial \mathbf{x}$ is taken by holding $\mathbf{v}$ constant. Since the direction of $\mathbf{B}_0$ is spatially dependent and thus $\mathbf{v}_\perp$ and $\mathbf{v}_\parallel$ are also spatially dependent when holding $\mathbf{v}$ constant.) To obtain an explicit formula for $\partial\alpha/\partial \mathbf{x}$, we need to choose a perpendicular (to $\mathbf{B}_0$) direction. Then an explicit formula can be derived (not given here), which involves the gradient of $\mathbf{B}_0$.]

Using the above results, equation (21) is written as

$$\frac{\partial f_p}{\partial \mathbf{x}}\Big|_{\mathbf{v}} = \frac{\partial f_g}{\partial \mathbf{X}} + \mathbf{v} \times \frac{\partial}{\partial \mathbf{x}}\left(\frac{\boldsymbol{e}_\parallel}{\Omega}\right) \cdot \frac{\partial f_g}{\partial \mathbf{X}} - \frac{q}{m} \mathbf{E}_0 \frac{\partial f_g}{\partial \varepsilon} + \frac{\partial \mu}{\partial \mathbf{x}} \frac{\partial f_g}{\partial \mu} + \frac{\partial \alpha}{\partial \mathbf{x}} \frac{\partial f_g}{\partial \alpha}. \tag{27}$$

For notation ease, define

$$\boldsymbol{\lambda}_{B1} = \mathbf{v} \times \frac{\partial}{\partial \mathbf{x}}\left(\frac{\boldsymbol{e}_\parallel}{\Omega}\right) \cdot \frac{\partial}{\partial \mathbf{X}}, \tag{28}$$

and

$$\boldsymbol{\lambda}_{B2} = \frac{\partial \mu}{\partial \mathbf{x}} \frac{\partial}{\partial \mu} + \frac{\partial \alpha}{\partial \mathbf{x}} \frac{\partial}{\partial \alpha}, \tag{29}$$

then the above expression is written as

$$\frac{\partial f_p}{\partial \mathbf{x}}\Big|_{\mathbf{v}} = \frac{\partial f_g}{\partial \mathbf{X}} + [\boldsymbol{\lambda}_{B1} + \boldsymbol{\lambda}_{B2}] f_g - \frac{q}{m} \mathbf{E}_0 \frac{\partial f_g}{\partial \varepsilon}. \tag{30}$$

2.8 Velocity gradient operator in terms of guiding-center variables

Next, consider the form of the velocity gradient $\partial f/\partial \mathbf{v}$ in terms of the guiding-center variables. Using the chain rule, $\partial f/\partial \mathbf{v}$ is written

$$\frac{\partial f_p}{\partial \mathbf{v}}\Big|_{\mathbf{x}} = \frac{\partial \mathbf{X}}{\partial \mathbf{v}} \cdot \frac{\partial f_g}{\partial \mathbf{X}} + \frac{\partial \varepsilon}{\partial \mathbf{v}} \frac{\partial f_g}{\partial \varepsilon} + \frac{\partial \mu}{\partial \mathbf{v}} \frac{\partial f_g}{\partial \mu} + \frac{\partial \alpha}{\partial \mathbf{v}} \frac{\partial f_g}{\partial \alpha}. \tag{31}$$

From the definition of $\mathbf{X}$, we obtain

$$\begin{aligned} \frac{\partial \mathbf{X}}{\partial \mathbf{v}} &= \frac{\partial}{\partial \mathbf{v}}\left(\frac{\boldsymbol{v} \times \boldsymbol{e}_\parallel}{\Omega}\right) \\ &= \frac{\partial \mathbf{v}}{\partial \mathbf{v}} \times \frac{\boldsymbol{e}_\parallel}{\Omega} \\ &= \mathbf{I} \times \frac{\boldsymbol{e}_\parallel}{\Omega}. \end{aligned} \tag{32}$$

From the definition of ε, we obtain

$$\frac{\partial \varepsilon}{\partial \mathbf{v}} = \mathbf{v}, \tag{33}$$

From the definition of μ, we obtain

$$\frac{\partial \mu}{\partial \mathbf{v}} = \frac{\mathbf{v}_\perp}{B_0}, \tag{34}$$

From the definition of α, we obtain

$$\frac{\partial \alpha}{\partial \mathbf{v}} = \frac{1}{v_\perp}\left(\boldsymbol{e}_\parallel \times \frac{\mathbf{v}_\perp}{v_\perp}\right) = \frac{\boldsymbol{e}_\alpha}{v_\perp}, \tag{35}$$

where $\mathbf{e}_\alpha$ is defined by

$$\boldsymbol{e}_\alpha = \boldsymbol{e}_\parallel \times \frac{\mathbf{v}_\perp}{v_\perp}. \tag{36}$$

Using the above results, expression (31) is written

$$\frac{\partial f_p}{\partial \mathbf{v}}\Big|_{\mathbf{x}} = \frac{\mathbf{I} \times \boldsymbol{e}_\parallel}{\Omega} \cdot \frac{\partial f_g}{\partial \mathbf{X}} + \mathbf{v} \frac{\partial f_g}{\partial \varepsilon} + \frac{\mathbf{v}_\perp}{B_0} \frac{\partial f_g}{\partial \mu} + \frac{\boldsymbol{e}_\alpha}{v_\perp} \frac{\partial f_g}{\partial \alpha}. \tag{37}$$

2.9 Time derivatives in terms of guiding-center variables

In the guiding-center variables, the time partial derivative $\partial f / \partial t$ appearing in Vlasov equation is written as

$$\frac{\partial f_p}{\partial t}|_{\mathbf{x},\mathbf{v}} = \frac{\partial f_g}{\partial t}|_{\mathbf{X},\mathbf{V}} + \frac{\partial \mathbf{X}}{\partial t} \cdot \frac{\partial f_g}{\partial \mathbf{X}} + \frac{\partial \mathbf{V}}{\partial t} \cdot \frac{\partial f_g}{\partial \mathbf{V}}, \tag{38}$$

where $\mathbf{V} = (\varepsilon, \mu, \alpha)$. Here $\partial \mathbf{X} / \partial t$ and $\partial \mathbf{V} / \partial t$ are not necessarily zero because the equilibrium quantities involved in the definition of the guiding-center transformation are in general time dependent. This time dependence is assumed to be evolving slow in the gyrokinetic ordering discussed later.

2.10 Final form of Vlasov equation in guiding-center coordinates

Using the above results, the Vlasov equation in guiding-center variables is written

$$\begin{aligned} & \frac{\partial f_g}{\partial t} + \frac{\partial \mathbf{X}}{\partial t} \cdot \frac{\partial f_g}{\partial \mathbf{X}} + \frac{\partial \mathbf{V}}{\partial t} \cdot \frac{\partial f_g}{\partial \mathbf{V}} \\ + & \mathbf{v} \cdot \left[\frac{\partial f_g}{\partial \mathbf{X}} + [\boldsymbol{\lambda}_{B1} + \boldsymbol{\lambda}_{B2}] f_g - \frac{q}{m} \mathbf{E}_0 \frac{\partial f_g}{\partial \varepsilon} \right] \\ + & \frac{q}{m} (\mathbf{E} + \mathbf{v} \times \mathbf{B}) \cdot \left(\frac{\mathbf{I} \times \boldsymbol{e}_\parallel}{\Omega} \cdot \frac{\partial f_g}{\partial \mathbf{X}} + \boldsymbol{v} \frac{\partial f_g}{\partial \varepsilon} + \frac{\boldsymbol{v}_\perp}{B_0} \frac{\partial f_g}{\partial \mu} + \frac{\boldsymbol{e}_\alpha}{v_\perp} \frac{\partial f_g}{\partial \alpha} \right) \\ = & 0, \end{aligned} \tag{39}$$

Using tensor identity $\mathbf{a} \cdot \boldsymbol{I} \times \mathbf{b} = \mathbf{a} \times \mathbf{b}$, equation (39) is written as

$$\begin{aligned} & \frac{\partial f_g}{\partial t} + \frac{\partial \mathbf{X}}{\partial t} \cdot \frac{\partial f_g}{\partial \mathbf{X}} + \frac{\partial \mathbf{V}}{\partial t} \cdot \frac{\partial f_g}{\partial \mathbf{V}} \\ + & \mathbf{v} \cdot \left[\frac{\partial f_g}{\partial \mathbf{X}} + [\boldsymbol{\lambda}_{B1} + \boldsymbol{\lambda}_{B2}] f_g - \frac{q}{m} \mathbf{E}_0 \frac{\partial f_g}{\partial \varepsilon} \right] \\ + & \frac{q}{m} (\mathbf{E} + \mathbf{v} \times \mathbf{B}) \times \left(\frac{\boldsymbol{e}_\parallel}{\Omega} \right) \cdot \frac{\partial f_g}{\partial \mathbf{X}} + \frac{q}{m} (\mathbf{v} \times \mathbf{B}) \cdot \left(\frac{\boldsymbol{e}_\alpha}{v_\perp} \frac{\partial f_g}{\partial \alpha} \right) \\ + & \frac{q}{m} \mathbf{E} \cdot \left(\mathbf{v} \frac{\partial f_g}{\partial \varepsilon} + \frac{\mathbf{v}_\perp}{B_0} \frac{\partial f_g}{\partial \mu} + \frac{\boldsymbol{e}_\alpha}{v_\perp} \frac{\partial f_g}{\partial \alpha} \right) \\ = & 0, \end{aligned} \tag{40}$$

This is the Vlasov equation in terms of guiding-center variables.

3 Perturbed Vlasov equation in guiding-center variables

Since the definition of the guiding-center variables $(\mathbf{X}, \varepsilon, \mu, \alpha)$ involves the macroscopic (equilibrium) fields $\mathbf{B}_0$ and $\mathbf{E}_0$, to further simplify Eq. (40), we need to separate electromagnetic field into equilibrium and perturbation parts. Writing the electromagnetic field as

$$\mathbf{E} = \mathbf{E}_0 + \delta \mathbf{E} \tag{41}$$

and

$$\mathbf{B} = \mathbf{B}_0 + \delta \mathbf{B}, \tag{42}$$

then substituting these expressions into equation (40) and moving all terms involving the perturbed fields to the right-hand side, we obtain

$$\begin{aligned} & \frac{\partial f_g}{\partial t} + \frac{\partial \mathbf{X}}{\partial t} \cdot \frac{\partial f_g}{\partial \mathbf{X}} + \frac{\partial \mathbf{V}}{\partial t} \cdot \frac{\partial f_g}{\partial \mathbf{V}} \\ + & \mathbf{v} \cdot \frac{\partial f_g}{\partial \mathbf{X}} + \mathbf{v} \cdot \left[[\boldsymbol{\lambda}_{B1} + \boldsymbol{\lambda}_{B2}] f_g - \frac{q}{m} \mathbf{E}_0 \frac{\partial f_g}{\partial \varepsilon} \right] \\ + & \frac{q}{m} (\mathbf{E}_0 + \mathbf{v} \times \mathbf{B}_0) \times \left(\frac{\boldsymbol{e}_\parallel}{\Omega} \right) \cdot \frac{\partial f_g}{\partial \mathbf{X}} + \frac{q}{m} (\mathbf{v} \times \mathbf{B}_0) \cdot \left(\frac{\boldsymbol{e}_\alpha}{v_\perp} \frac{\partial f_g}{\partial \alpha} \right) \\ + & \frac{q}{m} \mathbf{E}_0 \cdot \left(\mathbf{v} \frac{\partial f_g}{\partial \varepsilon} + \frac{\mathbf{v}_\perp}{B_0} \frac{\partial f_g}{\partial \mu} + \frac{\boldsymbol{e}_\alpha}{v_\perp} \frac{\partial f_g}{\partial \alpha} \right) \\ = & \delta R f_g, \end{aligned} \tag{43}$$

where δR is defined by

$$\begin{aligned} \delta R &= -\frac{q}{m}(\delta\mathbf{E}+\mathbf{v}\times\delta\mathbf{B})\times\left(\frac{\boldsymbol{e}_\parallel}{\Omega}\right)\cdot\frac{\partial}{\partial\mathbf{X}}-\frac{q}{m}(\mathbf{v}\times\delta\mathbf{B})\cdot\left(\frac{\boldsymbol{e}_\alpha}{v_\perp}\frac{\partial}{\partial\alpha}\right) \\ &\quad -\frac{q}{m}\delta\mathbf{E}\cdot\left(\mathbf{v}\frac{\partial}{\partial\varepsilon}+\frac{\mathbf{v}_\perp}{B_0}\frac{\partial}{\partial\mu}+\frac{\boldsymbol{e}_\alpha}{v_\perp}\frac{\partial}{\partial\alpha}\right). \end{aligned} \tag{44}$$

Next, let us simplify the left-hand side of Eq. (43). Note that

$$\frac{q}{m}\mathbf{E}_0\times\left(\frac{\boldsymbol{e}_\parallel}{\Omega}\right)\cdot\frac{\partial f_g}{\partial\mathbf{X}}=c\left(\frac{\mathbf{E}_0\times\mathbf{e}_\parallel}{B_0}\right)\cdot\frac{\partial f_g}{\partial\mathbf{X}}=\mathbf{v}_{\mathbf{E}0}\cdot\frac{\partial f_g}{\partial\mathbf{X}}, \tag{45}$$

where $\mathbf{v}_{\mathbf{E}0}$ is defined by $\mathbf{v}_{\mathbf{E}0}=c\mathbf{E}_0\times\mathbf{e}_\parallel/B_0$, which is the macroscopic (equilibrium) flow due to $\mathbf{E}_0\times\mathbf{B}_0$ drift. Further note that

$$\begin{aligned} \frac{q}{m}\frac{\mathbf{v}\times\mathbf{B}_0}{c}\times\left(\frac{\boldsymbol{e}_\parallel}{\Omega}\right)\cdot\frac{\partial f_g}{\partial\mathbf{X}} &= [(\mathbf{v}\times\mathbf{e}_\parallel)\times\mathbf{e}_\parallel]\cdot\frac{\partial f_g}{\partial\mathbf{X}} \\ &= [v_\parallel\mathbf{e}_\parallel-\mathbf{v}]\cdot\frac{\partial f_g}{\partial\mathbf{X}}, \end{aligned} \tag{46}$$

which can be combined with $\mathbf{v}\cdot\partial f_g/\partial\mathbf{X}$ term, yielding $v_\parallel\mathbf{e}_\parallel\cdot\partial f_g/\partial\mathbf{X}$. Finally note that

$$\begin{aligned} &\frac{q}{m}(\mathbf{v}\times\mathbf{B}_0)\cdot\left(\frac{\mathbf{e}_\alpha}{v_\perp}\frac{\partial f_g}{\partial\alpha}\right) \\ &=\frac{q}{m}(\mathbf{v}\times\mathbf{B}_0)\cdot\frac{\boldsymbol{e}_\parallel\times\mathbf{v}_\perp}{v_\perp^2}\frac{\partial f_g}{\partial\alpha} \\ &=-\Omega\frac{\partial f_g}{\partial\alpha}. \end{aligned} \tag{47}$$

Using Eqs. (45), (46), and (47), the left-hand side of equation (43) is written as

$$\begin{aligned} &\frac{\partial f_g}{\partial t}+\frac{\partial\mathbf{X}}{\partial t}\cdot\frac{\partial f_g}{\partial\mathbf{X}}+\frac{\partial\mathbf{V}}{\partial t}\cdot\frac{\partial f_g}{\partial\mathbf{V}} \\ &+\ (v_\parallel\mathbf{e}_\parallel+\mathbf{V}_{\mathbf{E}0})\cdot\frac{\partial f_g}{\partial\mathbf{X}}+\mathbf{v}\cdot[\boldsymbol{\lambda}_{B1}+\boldsymbol{\lambda}_{B2}]f_g-\Omega\frac{\partial f_g}{\partial\alpha} \\ &+\ \frac{q}{m}\mathbf{E}_0\cdot\left(\frac{\mathbf{v}_\perp}{B_0}\frac{\partial f_g}{\partial\mu}+\frac{\boldsymbol{e}_\alpha}{v_\perp}\frac{\partial f_g}{\partial\alpha}\right)\equiv L_g f_g \end{aligned} \tag{48}$$

which corresponds to Eq. (7) in Frieman-Chen's paper[3]. (In Frieman-Chen's equation (7), there is a term

$$\frac{q}{m}(\mathbf{E}-\mathbf{E}_0)\cdot\mathbf{v}\frac{\partial}{\partial\varepsilon}$$

where $\mathbf{E}$ is the macroscopic electric field and is in general different from the $\mathbf{E}_0$ introduced when defining the guiding-center transformation. In my derivation $\mathbf{E}_0$ is chosen to be equal to the macroscopic electric field thus the above term does not appear.) In expression (48), L_g is often called the unperturbed Vlasov propagator in guiding-center coordinates $(\mathbf{X},\varepsilon,\mu,\alpha)$.

Using the above results, Eq. (43) is written as

$$L_g f_g=\delta R f_g, \tag{49}$$

i.e.

$$\begin{aligned} &\frac{\partial f_g}{\partial t}+\frac{\partial\mathbf{X}}{\partial t}\cdot\frac{\partial f_g}{\partial\mathbf{X}}+\frac{\partial\mathbf{V}}{\partial t}\cdot\frac{\partial f_g}{\partial\mathbf{V}} \\ &+(v_\parallel\mathbf{e}_\parallel+\mathbf{V}_{\mathbf{E}0})\cdot\frac{\partial f_g}{\partial\mathbf{X}}+\mathbf{v}\cdot\left[\mathbf{v}\times\frac{\partial}{\partial\mathbf{x}}\left(\frac{\boldsymbol{e}_\parallel}{\Omega}\right)\cdot\frac{\partial f_g}{\partial\mathbf{X}}+\frac{\partial\mu}{\partial\mathbf{x}}\frac{\partial f_g}{\partial\mu}+\frac{\partial\alpha}{\partial\mathbf{x}}\frac{\partial f_g}{\partial\alpha}\right]-\Omega\frac{\partial f_g}{\partial\alpha} \\ &+\frac{q}{m}\mathbf{E}_0\cdot\left(\frac{\mathbf{v}_\perp}{B_0}\frac{\partial f_g}{\partial\mu}+\frac{\boldsymbol{e}_\alpha}{v_\perp}\frac{\partial f_g}{\partial\alpha}\right) \\ &=-\frac{q}{m}(\delta\mathbf{E}+\mathbf{v}\times\delta\mathbf{B})\times\left(\frac{\boldsymbol{e}_\parallel}{\Omega}\right)\cdot\frac{\partial f_g}{\partial\mathbf{X}}-\frac{q}{m}(\mathbf{v}\times\delta\mathbf{B})\cdot\left(\frac{\boldsymbol{e}_\alpha}{v_\perp}\frac{\partial f_g}{\partial\alpha}\right) \\ &-\frac{q}{m}\delta\mathbf{E}\cdot\left(\mathbf{v}\frac{\partial f_g}{\partial\varepsilon}+\frac{\mathbf{v}_\perp}{B_0}\frac{\partial f_g}{\partial\mu}+\frac{\boldsymbol{e}_\alpha}{v_\perp}\frac{\partial f_g}{\partial\alpha}\right). \end{aligned} \tag{50}$$

It is instructive to consider some special cases of the above complicated equation. Consider the case that the equilibrium magnetic field $\mathbf{B}_0$ is uniform and time-independent, $\mathbf{E}_0 = 0$, and the electrostatic limit $\delta\mathbf{B} = 0$, then equation (50) is simplified as

$$\frac{\partial f_g}{\partial t} + v_\parallel \mathbf{e}_\parallel \cdot \frac{\partial f_g}{\partial \mathbf{X}} - \Omega \frac{\partial f_g}{\partial \alpha}$$

$$= -\frac{q}{m}(\delta\mathbf{E}) \times \left(\frac{\boldsymbol{e}_\parallel}{\Omega}\right) \cdot \frac{\partial f_g}{\partial \mathbf{X}} \longrightarrow \text{spatial gradient drive} \tag{51}$$

$$-\frac{q}{m}\delta\mathbf{E} \cdot \left(\mathbf{v}\frac{\partial f_g}{\partial \varepsilon} + \frac{\mathbf{v}_\perp}{B_0}\frac{\partial f_g}{\partial \mu} + \frac{\boldsymbol{e}_\alpha}{v_\perp}\frac{\partial f_g}{\partial \alpha}\right) \longrightarrow \text{velocity space damping.} \tag{52}$$

3.1 Scale separation and two-scales expansion

Turbulence in tokamaks are usually of short perpendicular (to $\mathbf{B}_0$) spatial scale lengths of order ρ_i, which is much smaller than the macroscopic scale length L_0 (i.e., $\lambda = \rho_i / L_0$ is a small parameter). Therefor physical quantities f_g can be separated into macroscopic and microscopic parts as

$$f_g = F_g + \delta F_g, \tag{53}$$

where F is defined by

$$F_g \equiv \langle f_g \rangle_{\mathbf{X}_\perp} \equiv \frac{\int f_g d^2\mathbf{X}_\perp}{\int d^2\mathbf{X}_\perp}, \tag{54}$$

which is the averaging of f_g over (several times of) the short-scale perpendicular spatial scale. This is to say, F_g is constant over the short scale length ρ_i in the perpendicular direction, i.e., the perpendicular spatial scale length of F_g is much larger than ρ_i. This long perpendicular scale length of F_g is denoted by L_0. Equations (53) and (54) imply that

$$\langle \delta F_g \rangle_{\mathbf{X}_\perp} = 0. \tag{55}$$

An example of this two-scales expansion in one-dimension case is given in Fig. 2.

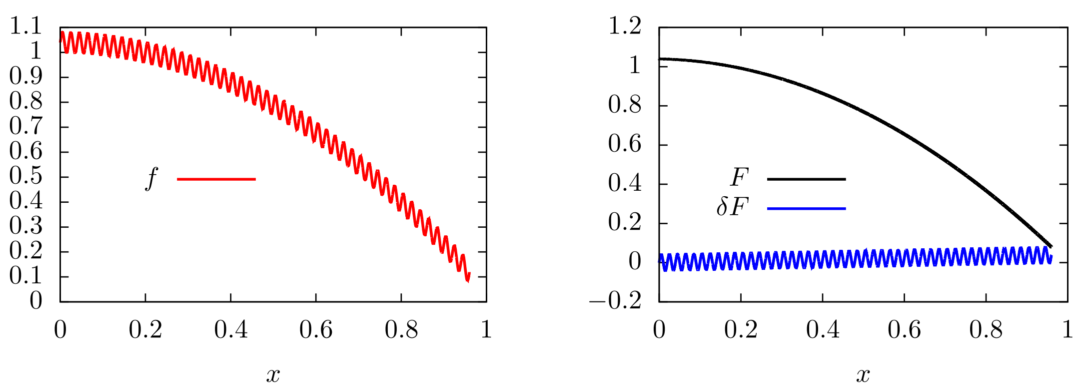

Figure 2. Given a f, then F is obtained from $F = \langle f \rangle_x = \int_x^{x+l} f(x)dx / \int_x^{x+l} dx$ and δF is obtained from $\delta F = f - F$. Here l is a length comparable to the small scale-length of f.

The expansion for the electromagnetic field given by Eqs. (41) and (42) should also be considered to be in this two-spatial-scale expansion. Using this expression in Eq. (49), we obtain

$$L_g F_g + L_g \delta F_g = \delta R F_g + \delta R \delta F_g. \tag{56}$$

Performing the perpendicular short scale length averaging on both sides of the above equation, we obtain

$$L_g F_g = \langle \delta R \delta F_g \rangle_{\mathbf{X}_\perp}, \tag{57}$$

where use has been made of $\langle L_g \delta F_g \rangle_{\mathbf{X}_\perp} = 0$ and $\langle \delta R F_g \rangle_{\mathbf{X}_\perp} = 0$. Subtracting the above equation from Eq. (56), we obtain

$$L_g \delta F_g = (\delta R F_g + \delta R \delta F_g - \langle \delta R \delta F_g \rangle_{\mathbf{X}_\perp}). \tag{58}$$

3.2 Gyrokinetic orderings

The following orderings assumed is often called the standard gyrokinetic orderings.

3.2.1 Ordering of macroscopic quantities

The spatial scale length L_0 of macroscopic quantities is assumed to be much larger than the thermal ion gyro-radius ρ_i, i.e., $\lambda \equiv \rho_i / L$ is a small parameter. Define the spatial scale length L_0 as $L_0 \approx F_g / |\nabla_X F_g|$, then we have

$$\rho_i |\nabla_X F_g| \sim O(\lambda^1) F_g. \tag{59}$$

This ordering is the result of the above two-scales expansion.

The macroscopic quantities evolve on the long transport time scale, which is assumed to be

$$\frac{1}{\Omega} \frac{\partial F_g}{\partial t} \sim O(\lambda^3) F_g, \tag{60}$$

i.e., transport time scale is $O(\lambda^{-3})$ longer than gyro-period.

The equilibrium (macroscopic) $\mathbf{E}_0 \times \mathbf{B}_0$ flow, $\mathbf{v}_{\mathbf{E}0} = \mathbf{E}_0 \times \mathbf{e}_\parallel / B_0 = -\nabla \Phi_0 \times \mathbf{e}_\parallel / B_0$, is assumed to be weak with

$$\frac{|\mathbf{v}_{E0}|}{v_t} \sim O(\lambda^1), \tag{61}$$

where v_t is the thermal velocity.

3.2.2 Orderings of microscopic quantities

We consider low frequency perturbations with $\omega / \Omega \sim O(\lambda^1)$, then

$$\frac{1}{\Omega} \frac{\partial \delta F_g}{\partial t} \sim O(\lambda^1) \delta F. \tag{62}$$

We consider small amplitude perturbations and adopt the following ordering for the microscopic fluctuations:

$$\frac{\delta F_g}{F_g} \sim \frac{q \delta \Phi}{T} \sim \frac{|\delta \mathbf{B}|}{B} \sim O(\lambda^1), \tag{63}$$

where $\delta \Phi$ is the perturbed scalar potential defined later in Eq. (68).

The perturbation is a assumed to have a long wavelength (compared with ρ_i) in the parallel direction

$$|\rho_i \mathbf{e}_\parallel \cdot \nabla_X \delta F_g| \sim O(\lambda^1) \delta F_g, \tag{64}$$

[and have a short wavelength comparable to the ion gyro-radius in the perpendicular direction ** seems to be not used in the derivation***

$$|\rho_i \nabla_{X\perp} \delta F_g| \sim O(\lambda^0) \delta F_g. \tag{65}$$

Combining Eq. (64) and (65), we obtain

$$\frac{k_\parallel}{k_\perp} \approx \frac{\mathbf{e}_\parallel \cdot \nabla_X}{\nabla_{X\perp}} \sim O(\lambda), \tag{66}$$

i.e., the parallel wave number is one order smaller than the perpendicular wave-number.]

In terms of the scalar and vector potentials $\delta \Phi$ and $\delta \mathbf{A}$, the perturbed electromagnetic field is written as

$$\delta \mathbf{B} = \nabla_x \times \delta \mathbf{A}, \tag{67}$$

and

$$\delta \mathbf{E} = -\nabla_x \delta \Phi - \frac{\partial \delta \mathbf{A}}{\partial t}. \tag{68}$$

Using the above orderings, it is ready see that $\delta E_{\|}$ is one order smaller than $\delta E_{\perp}$, i.e.,

$$\frac{\delta E_{\|}}{\delta E_{\perp}} = O(\lambda^1). \tag{69}$$

Further note that

$$\delta E_{\perp} = -\nabla_{\perp}\delta\Phi - \left(\frac{\partial \delta \mathbf{A}}{\partial t}\right)_{\perp}, \tag{70}$$

where the second term on the right-hand side is one order smaller than the first one.

3.3 Equation for macroscopic distribution function F_g

The evolution of the macroscopic quantity F_g is governed by Eq. (57), i.e.,

$$L_g F_g = \langle \delta R \delta F_g \rangle_{\mathbf{X}_{\perp}}, \tag{71}$$

where the left-hand side is written as

$$\begin{aligned} L_g F_g &= \frac{\partial F_g}{\partial t} + \frac{\partial \mathbf{X}}{\partial t} \cdot \frac{\partial F_g}{\partial \mathbf{X}} + \frac{\partial \mathbf{V}}{\partial t} \cdot \frac{\partial F_g}{\partial \mathbf{V}} \\ &+ (v_{\|}\mathbf{e}_{\|} + \mathbf{V}_{\mathbf{E}0}) \cdot \frac{\partial F_g}{\partial \mathbf{X}} + \mathbf{v} \cdot [(\boldsymbol{\lambda}_{B1} + \boldsymbol{\lambda}_{B2}) F_g] - \Omega \frac{\partial F_g}{\partial \alpha} \\ &+ \frac{q}{m} \mathbf{E}_0 \cdot \left(\frac{\mathbf{v}_{\perp}}{B_0} \frac{\partial F_g}{\partial \mu} + \frac{\boldsymbol{e}_{\alpha}}{v_{\perp}} \frac{\partial F_g}{\partial \alpha}\right) \end{aligned}$$

and the right-hand side is written as

$$\begin{aligned} \langle \delta R \delta F_g \rangle_{\mathbf{X}_{\perp}} &= -\frac{q}{m} \left\langle (\delta \mathbf{E} + \mathbf{v} \times \delta \mathbf{B}) \times \left(\frac{\boldsymbol{e}_{\|}}{\Omega}\right) \cdot \frac{\partial \delta F_g}{\partial \mathbf{X}} \right\rangle_{\mathbf{X}_{\perp}} - \frac{q}{m} \left\langle (\mathbf{v} \times \delta \mathbf{B}) \cdot \left(\frac{\boldsymbol{e}_{\alpha}}{v_{\perp}} \frac{\partial \delta F_g}{\partial \alpha}\right) \right\rangle_{\mathbf{X}_{\perp}} \\ & -\frac{q}{m} \left\langle \delta \mathbf{E} \cdot \left(\mathbf{v} \frac{\partial \delta F_g}{\partial \varepsilon} + \frac{\mathbf{v}_{\perp}}{B_0} \frac{\partial \delta F_g}{\partial \mu} + \frac{\boldsymbol{e}_{\alpha}}{v_{\perp}} \frac{\partial \delta F_g}{\partial \alpha}\right) \right\rangle_{\mathbf{X}_{\perp}}, \end{aligned} \tag{72}$$

which is of $O(\lambda^2)$ (or $O(\lambda^3)$?), which indicates that the turbulence effects on the macroscopic quantities enters through terms of $O(\lambda^2)$.

Expand F_g as $F_g = F_{g0} + F_{g1} + \ldots.$, where $F_{gi} \sim F_{g0} O(\lambda^i)$. Then, the balance on order $O(\lambda^0)$ gives

$$\frac{\partial F_{g0}}{\partial \alpha} = 0 \tag{73}$$

i.e., F_{g0} is independent of the gyro-angle α. The balance on $O(\lambda^1)$ gives

$$v_{\|}\mathbf{e}_{\|} \cdot \frac{\partial F_{g0}}{\partial \mathbf{X}} + \frac{q}{m} \mathbf{E}_0 \cdot \left(\frac{\mathbf{v}_{\perp}}{B_0} \frac{\partial F_{g0}}{\partial \mu}\right) = \Omega \frac{\partial F_{g1}}{\partial \alpha}. \tag{74}$$

Performing averaging over α, $\int_0^{2\pi} (\ldots) d\alpha$, on the above equation and noting that F_{g0} is independent of α, we obtain

$$\left(\int_0^{2\pi} d\alpha v_{\|}\mathbf{e}_{\|}\right) \cdot \frac{\partial F_{g0}}{\partial \mathbf{X}} + \frac{q}{m} \frac{\partial F_{g0}}{\partial \mu} \int_0^{2\pi} d\alpha \mathbf{E}_0 \cdot \left(\frac{\mathbf{v}_{\perp}}{B_0}\right) = \int_0^{2\pi} d\alpha \Omega \frac{\partial F_{g1}}{\partial \alpha} \tag{75}$$

Note that a quantity $A = A(\mathbf{x})$ that is independent of $\mathbf{v}$ will depend on $\mathbf{v}$ when transformed to guiding-center coordinates, i.e., $A(\mathbf{x}) = A_g(\mathbf{X}, \mathbf{v})$. Therefore A_g depends on gyro-angle α. However, since $\rho_i / L \ll 1$ for equilibrium quantities, the gyro-angle dependence of the equilibrium quantities can be neglected. Specifically, $\mathbf{e}_{\|}$, B_0 and Ω can be considered to be independent of α. As to $v_{\|}$, we have $v_{\|} = \pm\sqrt{2(\varepsilon - B_0\mu)}$. Since B_0 is considered independent of α, so does $v_{\|}$. Using these results, equation (75) is written

$$v_{\|}\mathbf{e}_{\|} \cdot \frac{\partial F_{g0}}{\partial \mathbf{X}} + \frac{q}{m} \frac{\partial F_{g0}}{\partial \mu} \int_0^{2\pi} d\alpha \mathbf{E}_0 \cdot \left(\frac{\mathbf{v}_{\perp}}{B_0}\right) = 0. \tag{76}$$

Using $\mathbf{E}_0 = -\nabla\Phi_0$, the above equation is written as

$$v_{\|}\mathbf{e}_{\|} \cdot \frac{\partial F_{g0}}{\partial \mathbf{X}} + \frac{q}{m} \frac{\partial F_{g0}}{\partial \mu} \int_0^{2\pi} d\alpha \left(\frac{-\mathbf{v}_{\perp} \cdot \nabla\Phi_0}{B_0}\right) = 0, \tag{77}$$

Note that

$$\int_0^{2\pi} d\alpha \frac{1}{B_0} \mathbf{v}_\perp \cdot \nabla_X \Phi_0 \approx 0, \tag{78}$$

where the error is of $O(\lambda^2)\Phi_0$, and thus, accurate to $O(\lambda)$, the last term of equation (77) is zero. Then equation (77) is written as

$$v_\parallel \mathbf{e}_\parallel \cdot \frac{\partial F_{g0}}{\partial \mathbf{X}} = 0, \tag{79}$$

which implies that F_{g0} is constant along a magnetic field line.

3.4 Equation for perturbed distribution function δF_g

Using $F_g \approx F_{g0}$, equation (58) is written as

$$L_g \delta F_g = \delta R F_{g0} + \underbrace{\delta R \delta F_g - \langle \delta R \delta F_g \rangle_{\mathbf{X}_\perp}}_{\text{nonlinear terms}}, \tag{80}$$

where $\delta R \delta F_g - \langle \delta R \delta F_g \rangle_{\mathbf{X}_\perp}$are nonlinear terms which are of order $O(\lambda^2)$ or higher, $L_g \delta F_g$ and $\delta R F_{g0}$ are linear terms which are of order $O(\lambda^1)$ or higher. The linear term $\delta R F_{g0}$ is given by

$$\delta R F_{g0} = \underbrace{-\frac{q}{m}\left(\delta \mathbf{E} + \frac{\mathbf{v} \times \delta \mathbf{B}}{c}\right) \times \left(\frac{\mathbf{e}_\parallel}{\Omega}\right) \cdot \frac{\partial F_{g0}}{\partial \mathbf{X}}}_{O(\lambda^2)} \underbrace{- \frac{q}{m} \delta \mathbf{E} \cdot \left(\mathbf{v} \frac{\partial F_{g0}}{\partial \varepsilon} + \frac{\mathbf{v}_\perp}{B_0} \frac{\partial F_{g0}}{\partial \mu}\right)}_{O(\lambda^1)}, \tag{81}$$

where use has been made of $\partial F_{g0} / \partial \alpha = 0$. Note that the first two terms are of order $O(\lambda^2)$ while the second two term are of order $O(\lambda^1)$. The linear term $L_g \delta F_g$ is written as

$$\begin{aligned} L_g \delta F_g \approx L_{gf} \delta F_g &= \frac{\partial \delta F_g}{\partial t} + (v_\parallel \mathbf{e}_\parallel + \mathbf{V}_{\mathbf{E}0}) \cdot \frac{\partial \delta F_g}{\partial \mathbf{X}} + \mathbf{v} \cdot [\boldsymbol{\lambda}_{B1} + \boldsymbol{\lambda}_{B2}] \delta F_g - \underbrace{\Omega \frac{\partial \delta F_g}{\partial \alpha}}_{O(\lambda^1)} \\ &+ \frac{q}{m} \mathbf{E}_0 \cdot \left(\frac{\mathbf{v}_\perp}{B_0} \frac{\partial \delta F_g}{\partial \mu} + \frac{\boldsymbol{e}_\alpha}{v_\perp} \frac{\partial \delta F_g}{\partial \alpha}\right), \end{aligned} \tag{82}$$

where use has been made of the assumption that the time partial derivative of the macroscopic quantities are of order $O(\lambda^3)$ and thus can be dropped since we want to derive an equation for δF_g that is accurate to $O(\lambda^2)$. The sub-index f in the notation L_{gf} means "frozen", indicating the effect of equilibrium evolution has be neglected. Specifically, L_{gf} operator is given by

$$L_{gf} = \frac{\partial}{\partial t} + (v_\parallel \mathbf{e}_\parallel + \mathbf{V}_{\mathbf{E}0}) \cdot \frac{\partial}{\partial \mathbf{X}} + \mathbf{v} \cdot [\boldsymbol{\lambda}_{B1} + \boldsymbol{\lambda}_{B2}] - \Omega \frac{\partial}{\partial \alpha} + \frac{q}{m} \mathbf{E}_0 \cdot \left(\frac{\mathbf{v}_\perp}{B_0} \frac{\partial}{\partial \mu} + \frac{\boldsymbol{e}_\alpha}{v_\perp} \frac{\partial}{\partial \alpha}\right). \tag{83}$$

where $\boldsymbol{\lambda}_{B1}$ and $\boldsymbol{\lambda}_{B2}$ operators are given by expression (28) and (29). In Eq. (82), $\Omega \partial \delta F_g / \partial \alpha$ is of order $O(\lambda^1)$ and all the other terms are of order $O(\lambda^2)$.

Next, to reduce the complexity of algebra, we consider the easier case in which $\partial F_{g0} / \partial \mu = 0$.

3.4.1 Separate δF_g into adiabatic and non-adiabatic parts

Note that the term $\Omega \partial \delta F_g / \partial \alpha$ in Eq. (82) and the last two terms in Eq. (81) are of $O(\lambda^1)$. Write δF_g as

$$\delta F_g = \delta F_a + \delta G, \tag{84}$$

where

$$\delta F_a = \frac{q}{m} \delta \Phi \frac{\partial F_{g0}}{\partial \varepsilon}, \tag{85}$$

which depends on the gyro-angle via $\delta \Phi$ and this term is often called adiabatic term. Plugging expression (84) into equation (80), we obtain

$$L_{gf} \delta G = \underbrace{\delta R F_{g0} - L_{gf} \delta F_a}_{\text{linear terms}} + \underbrace{\delta R \delta F_g - \langle \delta R \delta F_g \rangle_{\mathbf{X}_\perp}}_{\text{nonlinear terms}}. \tag{86}$$

Next, let us focus on the linear term on the right-hand side, i.e, $\delta R F_{g0} - L_{gf}\delta F_a$, and prove this term is of $O(\lambda^2)$ or higher (this means that $\Omega\partial\delta F_a/\partial\alpha$ cancels all the $O(\lambda^1)$ terms in $\delta R F_{g0}$). $L_{gf}\delta F_a$ is written

$$\begin{aligned} L_{gf}\delta F_a &= \frac{q}{m}\frac{\partial F_{g0}}{\partial\varepsilon}L_{gf}\delta\Phi + \frac{q}{m}\delta\Phi L_{gf}\frac{\partial F_{g0}}{\partial\varepsilon} \\ &\approx \frac{q}{m}\frac{\partial F_{g0}}{\partial\varepsilon}L_{gf}\delta\Phi, \end{aligned} \tag{87}$$

where the error is of order $O(\lambda^3)$. In obtaining the above expression, use has been made of $\mathbf{e}_\parallel \cdot \partial F_{g0}/\partial\mathbf{X} = 0$, $\partial F_{g0}/\partial\mathbf{X} = O(\lambda^1)F_{g0}$, $\partial F_{g0}/\partial\alpha = 0$, $\partial F_{g0}/\partial\mu = 0$, and the definition of $\boldsymbol{\lambda}_{B1}$ and $\boldsymbol{\lambda}_{B2}$ given in expressions (28) and (29). The expression (87) involves $\delta\Phi$ imposed by the Vlasov propagator L_{gf}. Since $\delta\Phi$ takes the most simple form when expressed in particle coordinates (if in guiding-center coordinates, $\delta\Phi(\mathbf{x}) = \delta\Phi(\mathbf{X} - \mathbf{v}\times\mathbf{e}_\parallel/\Omega)$, which depends on velocity coordinates and thus more complicated), it is convenient to use the Vlasov propagator L_{gf} expressed in particle coordinates. Transforming L_{gf} back to the particle coordinates, expression (87) is written

$$\begin{aligned} L_{gf}\delta F_a &= \frac{q}{m}\frac{\partial F_{g0}}{\partial\varepsilon}\left[\frac{\partial\delta\Phi}{\partial t}|_{\mathbf{x},\mathbf{v}} + \mathbf{v}\cdot\nabla_{\mathbf{x}}\delta\Phi + \frac{q}{m}(\mathbf{E}_0 + \mathbf{v}\times\mathbf{B}_0)\cdot\frac{\partial\Phi}{\partial\mathbf{v}}|_{\mathbf{x}}\right] \\ &= \frac{q}{m}\frac{\partial F_{g0}}{\partial\varepsilon}\left[\frac{\partial\delta\Phi}{\partial t}|_{\mathbf{x},\mathbf{v}} + \mathbf{v}\cdot\nabla_{\mathbf{x}}\delta\Phi\right] \qquad (88) \\ &= \frac{q}{m}\frac{\partial F_{g0}}{\partial\varepsilon}\left[\frac{\partial\delta\Phi}{\partial t}|_{\mathbf{x},\mathbf{v}} + \mathbf{v}\cdot\left(-\delta\mathbf{E} - \frac{\partial\delta\mathbf{A}}{\partial t}|_{\mathbf{x},\mathbf{v}}\right)\right] \\ &= \frac{q}{m}\frac{\partial F_{g0}}{\partial\varepsilon}\left[\frac{\partial\delta\Phi}{\partial t}|_{\mathbf{x},\mathbf{v}} - \mathbf{v}\cdot\delta\mathbf{E} - \frac{\partial\mathbf{v}\cdot\delta\mathbf{A}}{\partial t}|_{\mathbf{x},\mathbf{v}}\right]. \qquad (89) \\ &= \frac{q}{m}\frac{\partial F_{g0}}{\partial\varepsilon}\left[\frac{\partial\delta\Phi}{\partial t} - \mathbf{v}\cdot\delta\mathbf{E} - \frac{\partial\mathbf{v}\cdot\delta\mathbf{A}}{\partial t}\right]. \qquad (90) \end{aligned}$$

Using this and expression (81), $\delta R F_{g0} - L_g\delta F_a$ is written as

$$\begin{aligned} \delta R F_{g0} - L_{gf}\delta F_a &= -\frac{q}{m}(\delta\mathbf{E} + \mathbf{v}\times\delta\mathbf{B})\times\left(\frac{\mathbf{e}_\parallel}{\Omega}\right)\cdot\frac{\partial F_{g0}}{\partial\mathbf{X}} - \frac{q}{m}\delta\mathbf{E}\cdot\left(\mathbf{v}\frac{\partial F_{g0}}{\partial\varepsilon}\right) \\ &- \frac{q}{m}\frac{\partial F_{g0}}{\partial\varepsilon}\left[\frac{\partial\delta\Phi}{\partial t} - \mathbf{v}\cdot\delta\mathbf{E} - \frac{\partial\mathbf{v}\cdot\delta\mathbf{A}}{\partial t}\right] \\ &= -\frac{q}{m}(\delta\mathbf{E} + \mathbf{v}\times\delta\mathbf{B})\times\left(\frac{\mathbf{e}_\parallel}{\Omega}\right)\cdot\frac{\partial F_{g0}}{\partial\mathbf{X}} - \frac{q}{m}\frac{\partial F_{g0}}{\partial\varepsilon}\left[\frac{\partial\Phi}{\partial t} - \frac{\partial\mathbf{v}\cdot\delta\mathbf{A}}{\partial t}\right], \end{aligned} \tag{91}$$

where the two terms of $O(\lambda^1)$ (the terms in blue and red) cancel each other, with the remain terms being all of $O(\lambda^2)$, i.e, the contribution of the adiabatic term cancels the leading order terms of $O(\lambda^1)$ on the RHS of Eq. (86). The consequence of this is that, as will see in Sec. 3.5.1, on order $O(\lambda^1)$, δG is independent of the gyro-angle. Therefore separating δF into adiabatic and non-adiabatic parts also means separating δF into gyro-angle dependent and gyro-angle independent parts.

3.4.2 To Frieman-Chen form

Let us write the linear terms in expression (91) into the Frieman-Chen form. The $\delta\mathbf{E} + \mathbf{v}\times\delta\mathbf{B}$ term in Eq. (91) can be further written as

$$\begin{aligned} \delta\mathbf{E} + \mathbf{v}\times\delta\mathbf{B} &= -\nabla_x\delta\Phi - \frac{\partial\delta\mathbf{A}}{\partial t} + \mathbf{v}\times\nabla_x\times\delta\mathbf{A} \\ &\approx -\nabla_x\delta\Phi + \mathbf{v}\times\nabla_x\times\delta\mathbf{A}, \end{aligned}$$

where the error is of $O(\lambda^2)$. Using the vector identity $\mathbf{v}\times\nabla_x\times\delta\mathbf{A} = (\nabla\delta\mathbf{A})\cdot\mathbf{v} - (\mathbf{v}\cdot\nabla)\delta\mathbf{A}$ and noting $\mathbf{v}$ is constant for ∇_x operator, the above equation is written

$$\delta\mathbf{E} + \mathbf{v}\times\delta\mathbf{B} = -\nabla_x\delta\Phi + \nabla_x(\delta\mathbf{A}\cdot\mathbf{v}) - (\mathbf{v}\cdot\nabla_x)\delta\mathbf{A} \tag{92}$$

Note that Eq. (30) indicates that $\nabla_x\delta\phi \approx \nabla_X\delta\phi$, where the error is of $O(\lambda^2)$, then the above equation is written

$$\delta\mathbf{E} + \mathbf{v}\times\delta\mathbf{B} = -\nabla_X\delta\Phi + \nabla_X(\delta\mathbf{A}\cdot\mathbf{v}) - (\mathbf{v}\cdot\nabla_X)\delta\mathbf{A} \tag{93}$$

Further note that the parallel gradients in the above equation are of $O(\lambda^2)$ and thus can be dropped. Then Eq. (93) is written

$$\begin{aligned}&\delta\mathbf{E}+\mathbf{v}\times\delta\mathbf{B}\\&=-\nabla_{X\perp}\delta\Phi+\nabla_{X\perp}(\delta\mathbf{A}\cdot\mathbf{v})-(\mathbf{v}_\perp\cdot\nabla_{X\perp})\delta\mathbf{A}.\\&=-\nabla_{X\perp}\delta L-\mathbf{v}_\perp\cdot\nabla_{X\perp}\delta\mathbf{A},\end{aligned}\tag{94}$$

where δL is defined by

$$\delta L=\delta\Phi-\mathbf{v}\cdot\delta\mathbf{A}.\tag{95}$$

Using expression (94), equation (91) is written

$$\delta RF_0-L_g\delta F_a=-\frac{q}{m}\Big[(-\nabla_{X\perp}\delta L-\mathbf{v}_\perp\cdot\nabla_{X\perp}\delta\mathbf{A})\times\frac{\boldsymbol{e}_\parallel}{\Omega}\Big]\cdot\frac{\partial F_{g0}}{\partial\mathbf{X}}-\frac{q}{m}\frac{\partial\delta L}{\partial t}\frac{\partial F_{g0}}{\partial\varepsilon},\tag{96}$$

where all terms are of $O(\lambda^2)$.

3.5 Equation for the non-adiabatic part δG

Plugging expression (96) into Eq. (86), we obtain

$$L_{gf}\delta G=-\frac{q}{m}\Big[(-\nabla_{X\perp}\delta L-\mathbf{v}_\perp\cdot\nabla_{X\perp}\delta\mathbf{A})\times\frac{\mathbf{e}_\parallel}{\Omega}\Big]\cdot\frac{\partial F_{g0}}{\partial\mathbf{X}}-\frac{q}{m}\frac{\partial\delta L}{\partial t}\frac{\partial F_{g0}}{\partial\varepsilon}+\delta R\delta F_g-\langle\delta R\delta F_g\rangle_{\mathbf{X}_\perp}\tag{97}$$

3.5.1 Expansion of δG

Expand δG as

$$\delta G=\delta G_0+\delta G_1+...,$$

where $\delta G_i\sim O(\lambda^{i+1})F_{g0}$, and note that the right-hand side of Eq. (97) is of $O(\lambda^2)$, then, the balance on order $O(\lambda^1)$ requires

$$\frac{\partial\delta G_0}{\partial\alpha}=0,\tag{98}$$

i.e., δG_0 is gyro-phase independent. The balance on order $O(\lambda^2)$ requires

$$\begin{aligned}&\frac{\partial\delta G_0}{\partial t}+v_\parallel\mathbf{e}_\parallel\cdot\frac{\partial\delta G_0}{\partial\mathbf{X}}+\mathbf{v}\cdot[\boldsymbol{\lambda}_{B1}+\boldsymbol{\lambda}_{B2}]\delta G_0\\=\;&-\frac{q}{m}\Big[(-\nabla_{X\perp}\delta L-\mathbf{v}_\perp\cdot\nabla_X\delta\mathbf{A})\times\frac{\mathbf{e}_\parallel}{\Omega}\Big]\cdot\frac{\partial F_{g0}}{\partial\mathbf{X}}-\frac{q}{m}\frac{\partial\delta L}{\partial t}\frac{\partial F_{g0}}{\partial\varepsilon}+\delta R\delta F_g-\langle\delta R\delta F_g\rangle_{\mathbf{X}_\perp}.\end{aligned}\tag{99}$$

3.5.2 Gyro-averaging

Define the gyro-average operator $\langle...\rangle_\alpha$ by

$$\langle h\rangle_\alpha=(2\pi)^{-1}\int_0^{2\pi}h\,d\alpha,\tag{100}$$

where $h=h(\mathbf{X},\,\alpha,\,\varepsilon,\,\mu)$ is an arbitrary function of guiding-center variables. [For a field quantity that is independent of the velocity in particle coordinates, i.e., $h=h(\mathbf{x})$, it is ready to see that the above averaging is a spatial averaging over a gyro-ring]

Gyro-averaging Eq. (99), we obtain

$$\begin{aligned}&\frac{\partial\delta G_0}{\partial t}+\left\langle v_\parallel\mathbf{e}_\parallel\cdot\frac{\partial\delta G_0}{\partial\mathbf{X}}\right\rangle+\langle\mathbf{v}\cdot[\boldsymbol{\lambda}_{B1}+\boldsymbol{\lambda}_{B2}]\delta G_0\rangle_\alpha\\&=-\frac{q}{m}\Big[-\nabla_{X\perp}\langle\delta L\rangle_\alpha\times\frac{\mathbf{e}_\parallel}{\Omega}\Big]\cdot\frac{\partial F_{g0}}{\partial\mathbf{X}}-\frac{q}{m}\frac{\partial\langle\delta L\rangle_\alpha}{\partial t}\frac{\partial F_{g0}}{\partial\varepsilon}+\langle\delta R\delta F_g\rangle_\alpha-\langle\langle\delta R\delta F_g\rangle_{\mathbf{X}_\perp}\rangle_\alpha,\end{aligned}\tag{101}$$

where use has been made of $\langle(\mathbf{v}_\perp\cdot\nabla_X)\delta\mathbf{A}\rangle_\alpha\approx0$, where the error is of order higher than $O(\lambda^2)$. Note that $v_\parallel=\pm\sqrt{2(\varepsilon-B_0\mu)}$. Since B_0 is approximately independent of α, so does $v_\parallel$. Using this, the first gyro-averaging on the left-hand side of the above equation is written

$$\left\langle v_\parallel\mathbf{e}_\parallel\cdot\frac{\partial\delta G_0}{\partial\mathbf{X}}\right\rangle_\alpha=\langle v_\parallel\mathbf{e}_\parallel\rangle\cdot\frac{\partial\delta G_0}{\partial\mathbf{X}}=v_\parallel\mathbf{e}_\parallel\cdot\frac{\partial\delta G_0}{\partial\mathbf{X}}\tag{102}$$

The second gyro-averaging appearing on the left-hand side can be written

$$\langle \mathbf{v}\cdot[\boldsymbol{\lambda}_{B1}+\boldsymbol{\lambda}_{B2}]\delta G_0\rangle_\alpha = \mathbf{V}_D\cdot\nabla_X\delta G_0, \tag{103}$$

where $\mathbf{V}_D$ is the magnetic curvature and gradient drift (I do not derive Eq. (103), to be done). Then Eq. (101) is written

$$\begin{aligned}&\left[\frac{\partial}{\partial t}+v_\parallel\mathbf{e}_\parallel\cdot\nabla_X+\mathbf{V}_D\cdot\nabla_X\right]\delta G_0\\ &= -\frac{q}{m}\left[-\nabla_{X\perp}\langle\delta L\rangle_\alpha\times\frac{\boldsymbol{e}_\parallel}{\Omega}\right]\cdot\frac{\partial F_{g0}}{\partial\mathbf{X}}-\frac{q}{m}\frac{\partial\langle\delta L\rangle_\alpha}{\partial t}\frac{\partial F_{g0}}{\partial\varepsilon}+\langle\delta R\delta F_g\rangle_\alpha-\langle\langle\delta R\delta F_g\rangle_{\mathbf{x}_\perp}\rangle_\alpha.\end{aligned} \tag{104}$$

3.5.3 Simplification of the nonlinear term

Next, we try to simplify the nonlinear term $\langle\delta R\delta F_g\rangle_\alpha$ appearing in Eq. (104), which is written as

$$\begin{aligned}\langle\delta R\delta F_g\rangle_\alpha &= \left\langle\delta R\left(\frac{q}{m}\delta\Phi\frac{\partial F_{g0}}{\partial\varepsilon}+\delta G_0\right)\right\rangle_\alpha\\ &= \left\langle\frac{q}{m}\delta R\left(\delta\Phi\frac{\partial F_{g0}}{\partial\varepsilon}\right)\right\rangle_\alpha+\langle\delta R\delta G_0\rangle_\alpha\end{aligned} \tag{105}$$

First, let us focus on the first term, which can be written as

$$\begin{aligned}\delta R\left(\delta\Phi\frac{\partial F_{g0}}{\partial\varepsilon}\right) &\approx -\frac{q}{m}\frac{\partial F_{g0}}{\partial\varepsilon}\left(\delta\mathbf{E}+\frac{\mathbf{v}\times\delta\mathbf{B}}{c}\right)\times\left(\frac{\boldsymbol{e}_\parallel}{\Omega}\right)\cdot\frac{\partial\delta\Phi}{\partial\mathbf{X}}-\frac{q}{m}\frac{\partial F_{g0}}{\partial\varepsilon}\left(\frac{\mathbf{v}\times\delta\mathbf{B}}{c}\right)\cdot\left(\frac{\boldsymbol{e}_\alpha}{v_\perp}\frac{\partial\delta\Phi}{\partial\alpha}\right)\\ &\quad-\frac{q}{m}\frac{\partial F_{g0}}{\partial\varepsilon}\delta\mathbf{E}\cdot\left(\mathbf{v}\frac{\partial\delta\Phi}{\partial\varepsilon}+\frac{\mathbf{v}_\perp}{B_0}\frac{\partial\delta\Phi}{\partial\mu}+\frac{\boldsymbol{e}_\alpha}{v_\perp}\frac{\partial\delta\Phi}{\partial\alpha}\right)+\frac{q}{m}\delta\Phi\delta\mathbf{E}\cdot\mathbf{v}\frac{\partial^2F_{g0}}{\partial\varepsilon^2}\\ &= -\frac{q}{m}\left(\delta\mathbf{E}+\frac{\mathbf{v}\times\delta\mathbf{B}}{c}\right)\cdot\nabla_v\delta\Phi+\frac{q}{m}\delta\Phi\delta\mathbf{E}\cdot\mathbf{v}\frac{\partial^2F_{g0}}{\partial\varepsilon^2}\\ &= \frac{q}{m}\delta\Phi\delta\mathbf{E}\cdot\mathbf{v}\frac{\partial^2F_{g0}}{\partial\varepsilon^2}\end{aligned} \tag{106}$$

Using the above results, the nonlinear term $\langle\delta R\delta F\rangle_\alpha$ is written as

$$\langle\delta R\delta F\rangle_\alpha = \frac{q}{m}\left\langle\delta\Phi\delta\mathbf{E}\cdot\mathbf{v}\frac{\partial^2F_{g0}}{\partial\varepsilon^2}\right\rangle_\alpha+\langle\delta R\delta G_0\rangle_\alpha \tag{107}$$

Accurate to $O(\lambda^2)$, the first term on the right-hand side of the above is zero. [Proof:

$$\begin{aligned}\left\langle\delta\Phi\delta\mathbf{E}\cdot\mathbf{v}\frac{\partial^2F_{g0}}{\partial\varepsilon^2}\right\rangle_\alpha &= \left\langle\frac{\partial^2F_{g0}}{\partial\varepsilon^2}\delta\Phi\nabla\delta\Phi\cdot\mathbf{v}\right\rangle_\alpha\\ &= \frac{\partial^2F_{g0}}{\partial\varepsilon^2}\langle\mathbf{v}\cdot\nabla(\delta\Phi)^2\rangle_\alpha\\ &\approx \frac{\partial^2F_{g0}}{\partial\varepsilon^2}\langle\mathbf{v}_\perp\cdot\nabla(\delta\Phi)^2\rangle_\alpha\\ &\approx 0,\end{aligned} \tag{108}$$

where use has been made of $\langle\mathbf{v}_\perp\cdot\nabla_X\delta\Phi\rangle_\alpha\approx 0$, where the error is of $O(\lambda^2)$. Using the above results, expression (107) is written as

$$\langle\delta R\delta F_g\rangle_\alpha = \langle\delta R\delta G_0\rangle_\alpha. \tag{109}$$

Using the expression of δR given by Eq. (44), the above expression is written as

$$\begin{aligned}\langle\delta R\delta G_0\rangle_\alpha &= -\frac{q}{m}\left\langle\left(\delta\mathbf{E}+\frac{\mathbf{v}\times\delta\mathbf{B}}{c}\right)\times\left(\frac{\mathbf{e}_\parallel}{\Omega}\right)\right\rangle_\alpha\cdot\frac{\partial\delta G_0}{\partial\mathbf{X}}\\ &\quad-\frac{q}{m}\frac{\partial\delta G_0}{\partial\varepsilon}\langle\delta\mathbf{E}\cdot\mathbf{v}\rangle_\alpha-\frac{q}{m}\frac{\partial\delta G_0}{\partial\mu}\left\langle\delta\mathbf{E}\cdot\frac{\mathbf{v}_\perp}{B_0}\right\rangle_\alpha\end{aligned} \tag{110}$$

where use has been made of $\partial\delta G_0/\partial\alpha=0$. Using Eq. (94), we obtain

$$-\frac{q}{m}\left\langle(\delta\mathbf{E}+\mathbf{v}\times\delta\mathbf{B})\times\left(\frac{\boldsymbol{e}_\parallel}{\Omega}\right)\right\rangle_\alpha=\frac{q}{m}\nabla_{X\perp}\langle\delta L\rangle_\alpha\times\frac{\boldsymbol{e}_\parallel}{\Omega}. \tag{111}$$

The other two terms in Eq. (110) can be proved to be zero. [Proof:

$$\begin{aligned}-\frac{q}{m}\frac{\partial\delta G_0}{\partial\varepsilon}\langle\delta\mathbf{E}\cdot\mathbf{v}\rangle_\alpha &= \frac{q}{m}\frac{\partial\delta G_0}{\partial\varepsilon}\langle\mathbf{v}\cdot\nabla_x\Phi\rangle_\alpha\\ &\approx \frac{q}{m}\frac{\partial\delta G_0}{\partial\varepsilon}\langle\mathbf{v}_\perp\cdot\nabla_x\Phi\rangle_\alpha\\ &\approx \frac{q}{m}\frac{\partial\delta G_0}{\partial\varepsilon}\langle\mathbf{v}_\perp\cdot\nabla_X\Phi\rangle_\alpha\\ &\approx 0\end{aligned} \tag{112}$$

$$\begin{aligned}-\frac{q}{m}\frac{\partial\delta G_0}{\partial\mu}\left\langle\delta\mathbf{E}\cdot\frac{\mathbf{v}_\perp}{B_0}\right\rangle_\alpha &= \frac{q}{m}\frac{\partial\delta G_0}{\partial\mu}\left\langle\frac{1}{B_0}\mathbf{v}_\perp\cdot\nabla_x\Phi\right\rangle_\alpha\\ &\approx \frac{q}{m}\frac{\partial\delta G_0}{\partial\mu}\left\langle\frac{1}{B_0}\mathbf{v}_\perp\cdot\nabla_X\Phi\right\rangle_\alpha\\ &\approx 0\end{aligned} \tag{113}$$

] Using the above results, the nonlinear term is finally written as

$$\langle\delta R\delta G_0\rangle_\alpha=\frac{q}{m}\nabla_{X\perp}\langle\delta L\rangle_\alpha\times\frac{\boldsymbol{e}_\parallel}{\Omega}\cdot\nabla_X\delta G_0. \tag{114}$$

Using this in Eq. (109), we obtain

$$\langle\delta R\delta F_g\rangle_\alpha=\frac{q}{m}\nabla_{X\perp}\langle\delta L\rangle_\alpha\times\frac{\boldsymbol{e}_\parallel}{\Omega}\cdot\nabla_X\delta G_0, \tag{115}$$

which is of $O(\lambda^2)$. Using this form, it can be proved that $\langle\langle\delta R\delta F_g\rangle_\alpha\rangle_{X\perp}$ is of $O(\lambda^3)$ (to be proved later).

3.5.4 Final equation for the non-adiabatic part of the perturbed distribution function

Using the above results, the gyro-averaged kinetic equation for δG_0 is finally written as

$$\begin{aligned}&\left[\frac{\partial}{\partial t}+\left(v_\parallel\mathbf{e}_\parallel+\mathbf{V}_D-\frac{q}{m}\nabla_{X\perp}\langle\delta L\rangle_\alpha\times\frac{\boldsymbol{e}_\parallel}{\Omega}\right)\cdot\nabla_X\right]\delta G_0\\ &=\frac{q}{m}\left[\nabla_{X\perp}\langle\delta L\rangle_\alpha\times\frac{\boldsymbol{e}_\parallel}{\Omega}\right]\cdot\nabla_X F_{g0}-\frac{q}{m}\frac{\partial\langle\delta L\rangle_\alpha}{\partial t}\frac{\partial F_{g0}}{\partial\varepsilon}.\end{aligned} \tag{116}$$

where $\delta L=\delta\Phi-\mathbf{v}\cdot\delta\mathbf{A}$, $\langle\ldots\rangle_\alpha$ is the gyro-phase averaging operator, $\mathbf{V}_D$ is the equilibrium guiding-center drift velocity, and $\delta G_0=\delta G_0(\mathbf{X},\varepsilon,\mu)$ is gyro-angle independent and is related to the perturbed distribution function δF_g by

$$\delta F_g=\frac{q}{m}\delta\Phi\frac{\partial F_{g0}}{\partial\varepsilon}+\delta G_0, \tag{117}$$

where the first term is called the adiabatic term and this term depends on the gyro-phase α via $\delta\Phi$. Equation (116) is the special case ($\partial F_{g0}/\partial\mu|_\varepsilon=0$) of the Frieman-Chen nonlinear gyrokinetic equation given in Ref. [3]. Note that the nonlinear terms only appear on the left-hand side of this equation and all the terms on the right-hand side are linear.

4 Characteristic curves of Frieman-Chen nonlinear gyrokinetic equation

The Frieman-Chen nonlinear gyrokinetic equation takes the following form:

$$\frac{\partial\delta G_0}{\partial t}+\left(v_\parallel\mathbf{e}_\parallel+\mathbf{V}_D-\frac{q}{m}\nabla_X\langle\delta L\rangle_\alpha\times\frac{\mathbf{e}_\parallel}{\Omega}\right)\cdot\nabla_X\delta G_0=\frac{q}{m}\left[\nabla_X\langle\delta L\rangle_\alpha\times\frac{\mathbf{e}_\parallel}{\Omega}\right]\cdot\nabla_X F_0-\frac{q}{m}\frac{\partial\langle\delta L\rangle_\alpha}{\partial t}\frac{\partial F_0}{\partial\varepsilon}-\frac{q}{m}S_{l2}, \tag{118}$$

where $\delta G_0 = \delta G_0(\mathbf{X}, \varepsilon, \mu)$, which is gyro-phase independent, $\delta L = \delta\Phi - \mathbf{v}\cdot\delta\mathbf{A}$, and the term S_{l2} is due to the μ dependence of F_0, which is not considered in this note. Examining the left-hand side of Eq. (118), it is ready to find that the characteristic curves of this equation are given by the following equations:

$$\frac{d\mathbf{X}}{dt} = v_\parallel \mathbf{e}_\parallel + \mathbf{V}_D - \frac{q}{m}\nabla_X \langle \delta L\rangle_\alpha \times \frac{\mathbf{e}_\parallel}{\Omega}, \tag{119}$$

$$\frac{d\varepsilon}{dt} = 0, \tag{120}$$

$$\frac{d\mu}{dt} = 0. \tag{121}$$

(It is instructive to notice that the kinetic energy ε is conserved along the characteristic curves while the real kinetic energy of a particle is usually not conserved in a perturbed electromagnetic field. This may be an indication that Frieman-Chen equation neglects the velocity space nonlinearity.) For notation ease, we define the perturbed drift $\delta\mathbf{V}_D$ by

$$\delta\mathbf{V}_D = -\frac{q}{m}\nabla_X \langle \delta L\rangle_\alpha \times \frac{\mathbf{e}_\parallel}{\Omega}. \tag{122}$$

and the total guiding-center velocity $\mathbf{V}_G$ by

$$\mathbf{V}_G = \frac{d\mathbf{X}}{dt} = v_\parallel \mathbf{e}_\parallel + \mathbf{V}_D + \delta\mathbf{V}_D. \tag{123}$$

Equations (119)-(121) for the characteristics are, however, not in the form that can be readily evolved numerically because there is no time evolution equation for $v_\parallel$, which appears explicitly in Eq. (119). It is ready to realize that the equation for $v_\parallel$ is implicitly contained in the combination of the equations for ε and μ. Next, we derive the equation for $v_\parallel$.

4.1 Time evolution equation for $\mathbf{v}_\parallel$

Using the definition $\mu = v_\perp^2/(2B_0)$, equation (121), i.e., $d\mu/dt = 0$, is written as

$$\frac{d}{dt}\left(\frac{v_\perp^2}{2B_0}\right) = 0, \tag{124}$$

which is written as

$$\frac{1}{B}\frac{d}{dt}(v_\perp^2) + v_\perp^2 \frac{d}{dt}(\frac{1}{B_0}) = 0, \tag{125}$$

which can be further written as

$$\frac{d}{dt}(v_\perp^2) = 2\mu \frac{d}{dt}(B_0). \tag{126}$$

Using the definition of the characteristics, the right-hand side of the above equation can be expanded, giving

$$\frac{d}{dt}(v_\perp^2) = 2\mu\left(\frac{\partial B_0}{\partial t} + \frac{d\mathbf{X}}{dt}\cdot\nabla_X B_0 + \frac{d\varepsilon}{dt}\frac{\partial B_0}{\partial\varepsilon} + \frac{d\mu}{dt}\frac{\partial B_0}{\partial\mu}\right), \tag{127}$$

where $d\mathbf{X}/dt$, $d\varepsilon/dt$, and $d\mu/dt$ are given by Eq. (119), (120), and (121), respectively. Using Eqs. (119)-(121) and $\partial B_0/\partial t = 0$, equation (127) is reduced to

$$\frac{d}{dt}(v_\perp^2) = 2\mu\left(\frac{d\mathbf{X}}{dt}\cdot\nabla_X B_0\right). \tag{128}$$

On the other hand, equation (120), i.e., $d\varepsilon/dt = 0$, is written as

$$\frac{d}{dt}(v^2) = 0, \tag{129}$$

which can be further written as

$$\frac{d}{dt}(v_\parallel^2) = -\frac{d}{dt}(v_\perp^2). \tag{130}$$

Using Eq. (128), the above equation is written as

$$\frac{d}{dt}(v_\parallel) = -\frac{\mu}{v_\parallel}\left(\frac{d\mathbf{X}}{dt}\cdot\nabla_X B_0\right), \tag{131}$$

which is the equation for the time evolution of $v_\parallel$. This equation involves $d\mathbf{X}/dt$, i,e., the guiding-center drift, which is given by Eq. (119). Equation (131) for $v_\parallel$ can be simplified by noting that the Frieman-Chen equation is correct only to the second order, $O(\lambda^2)$, and thus the characteristics need to be correct only to the first order $O(\lambda)$ and higher order terms can be dropped. Note that, in the guiding-center drift $d\mathbf{X}/dt$ given by Eq. (119), only the $v_\parallel \mathbf{e}_\parallel$ term in is of order $O(\lambda^0)$, all the other terms are of $O(\lambda^1)$. Using this, accurate to order $O(\lambda^1)$, equation (131) is written as

$$\frac{d}{dt}(v_\parallel) = -\mu \mathbf{e}_\parallel \cdot \nabla_X B_0, \tag{132}$$

which is the time evolution equation ready to be used for numerically advancing $v_\parallel$. Note that only the mirror force $-\mu \mathbf{e}_\parallel \cdot \nabla B$ appears in Eq. (132) and there is no parallel acceleration term $q v_\parallel \delta E_\parallel / m$ in Eq. (132). This is because $\delta E_\parallel = -\mathbf{b} \cdot \nabla \delta\Phi - \partial \delta A_\parallel / \partial t$ is of order $O(\lambda^2)$ and (**check**the terms involving $E_\parallel$ are of $O(\lambda^3)$ or higher and thus have been dropped in the process of deriving Frieman-Chen equation.)

5 Gyrokinetic equation in forms amenable to numerical simulation

In the case of isotropic F_0, Frieman-Chen's gyrokinetic equation for δG_0 is given by

$$\begin{aligned} &\left[\frac{\partial}{\partial t} + (v_\parallel \mathbf{e}_\parallel + \mathbf{V}_D + \delta \mathbf{V}_D) \cdot \nabla_X\right] \delta G_0 \\ &= -\delta \mathbf{V}_D \cdot \nabla_X F_0 - \frac{q}{m} \frac{\partial \langle \delta L \rangle_\alpha}{\partial t} \frac{\partial F_0}{\partial \varepsilon}, \end{aligned} \tag{133}$$

where δG_0 is the gyro-phase independent part of the perturbed distribution δF, and is related to δF by

$$\delta F = \frac{q}{m} \delta\Phi \frac{\partial F_0}{\partial \varepsilon} + \delta G_0, \tag{134}$$

where the first term is called "the adiabatic term", which depends on gyro-phase α via $\delta\phi$. In Eq. (133), $\delta L = \delta\Phi - \mathbf{v} \cdot \delta \mathbf{A}$.

5.1 Eliminate $\partial \langle \delta\phi \rangle_\alpha / \partial t$ term on the right-hand side of GK equation

Note that the coefficient before $\partial F_0 / \partial \varepsilon$ in Eq. (133) involves the time derivative of $\langle \delta\phi \rangle_\alpha$, which is problematic if treated by using explicit finite difference in particle simulations (I test the algorithm that treats this term by implicit scheme, the result roughly agrees with the standard method discussed in Sec. 5.6). It turns out that $\partial \langle \delta\phi \rangle_\alpha / \partial t$ can be readily eliminated by defining another gyro-phase independent function δf by

$$\delta f = \frac{q}{m} \langle \delta\Phi \rangle_\alpha \frac{\partial F_0}{\partial \varepsilon} + \delta G_0. \tag{135}$$

Then, in terms of δf, the perturbed distribution function δF is written as

$$\delta F = \frac{q}{m} (\delta\Phi - \langle \delta\Phi \rangle_\alpha) \frac{\partial F_0}{\partial \varepsilon} + \delta f. \tag{136}$$

Using Eq. (135) and Eq. (133), the equation for δf is written as

$$\begin{aligned} &\left[\frac{\partial}{\partial t} + (v_\parallel \mathbf{e}_\parallel + \mathbf{V}_D + \delta \mathbf{V}_D) \cdot \nabla_X\right] \delta f \\ &- \frac{q}{m} \frac{\partial F_0}{\partial \varepsilon} \left[\frac{\partial}{\partial t} + (v_\parallel \mathbf{e}_\parallel + \mathbf{V}_D + \delta \mathbf{V}_D) \cdot \nabla_X\right] \langle \delta\phi \rangle_\alpha \\ &- \frac{q}{m} \langle \delta\phi \rangle_\alpha \left[\frac{\partial}{\partial t} + (v_\parallel \mathbf{e}_\parallel + \mathbf{V}_D + \delta \mathbf{V}_D) \cdot \nabla_X\right] \frac{\partial F_0}{\partial \varepsilon} \\ &= -\delta \mathbf{V}_D \cdot \nabla_X F_0 - \frac{q}{m} \frac{\partial \langle \delta L \rangle_\alpha}{\partial t} \frac{\partial F_0}{\partial \varepsilon} \end{aligned} \tag{137}$$

Noting that $\partial F_0 / \partial t = 0$, $\mathbf{e}_\parallel \cdot \nabla F_0 = 0$, $\nabla F_0 \sim O(\lambda^1) F_0$, we find that the third line of the above equation is of order $O(\lambda^3)$ and thus can be dropped. Moving the second line to the right-hand side and noting that $\langle \delta L \rangle_\alpha = \langle \delta\phi - \mathbf{v} \cdot \delta\mathbf{A} \rangle_\alpha$, the above equation is written as

$$\begin{aligned}&\left[\frac{\partial}{\partial t} + (v_\parallel \mathbf{e}_\parallel + \mathbf{V}_D + \delta\mathbf{V}_D) \cdot \nabla_X\right] \delta f \\ =&-\delta\mathbf{V}_D \cdot \nabla_X F_0 \\ &-\frac{q}{m}\left[-\frac{\partial \langle \mathbf{v} \cdot \delta\mathbf{A} \rangle_\alpha}{\partial t} - (v_\parallel \mathbf{e}_\parallel + \mathbf{V}_D + \delta\mathbf{V}_D) \cdot \nabla_X \langle \delta\Phi \rangle_\alpha\right] \frac{\partial F_0}{\partial \varepsilon},\end{aligned} \tag{138}$$

where two $\partial \langle \phi \rangle_\alpha / \partial t$ terms cancel each other. Equation (138) corresponds to Eq. (A8) in Yang Chen's paper[2] (where the first minus on the right-hand side is wrong and should be replaced with q/m; one q is missing before $\partial(\mathbf{v} \cdot \delta\mathbf{A})/\partial t$ in A9).

The blue term in expression (136) will give rise to the so-called "polarization density" when integrated in the velocity space (discussed in Sec. 5.6). The reason for the name "polarization" is that $(\delta\phi - \langle \delta\phi \rangle_\alpha)$ is the difference between the local value and the averaged value on a gyro-ring, expressing a kind of "separation". This "separation" is obviously proportional to the gyro-radius and hence larger for massive particles with large velocity. In tokamak plasma, thermal ion gyro-radius is $\sqrt{m_i / m_e} \approx 60$ times larger than thermal electron gyro-radius. Therefore the ion polarization density is much larger than the electron polarization density and the latter is often dropped in simulations except for studying electron temperature gradient (ETG) turbulence.

5.2 Eliminate $\partial \langle \delta \mathbf{v} \cdot \delta \mathbf{A} \rangle_\alpha / \partial t$ term on the right-hand side of GK equation

Similar to the method of eliminating $\partial \langle \delta\phi \rangle_\alpha / \partial t$, we define another gyro-phase independent function δh by

$$\delta h = \delta f - \frac{q}{m} \langle \mathbf{v} \cdot \delta\mathbf{A} \rangle_\alpha \frac{\partial F_0}{\partial \varepsilon}. \tag{139}$$

then Eq. (138) is written in terms of δh as

$$\begin{aligned}&\left[\frac{\partial}{\partial t} + (v_\parallel \mathbf{e}_\parallel + \mathbf{V}_D + \delta\mathbf{V}_D) \cdot \nabla_X\right] \delta h \\ &+\frac{q}{m} \frac{\partial F_0}{\partial \varepsilon}\left[\left(\frac{\partial}{\partial t} + v_\parallel \mathbf{e}_\parallel + \mathbf{V}_D + \delta\mathbf{V}_D\right) \cdot \nabla_X\right] \langle \mathbf{v} \cdot \delta\mathbf{A} \rangle_\alpha \\ &+\frac{q}{m} \langle \mathbf{v} \cdot \delta\mathbf{A} \rangle_\alpha \left[\left(\frac{\partial}{\partial t} + v_\parallel \mathbf{e}_\parallel + \mathbf{V}_D + \delta\mathbf{V}_D\right) \cdot \nabla_X\right] \left(\frac{\partial F_0}{\partial \varepsilon}\right) \\ =&-\delta\mathbf{V}_D \cdot \nabla_X F_0 \\ &-\frac{q}{m}\left[-\frac{\partial \langle \mathbf{v} \cdot \delta\mathbf{A} \rangle_\alpha}{\partial t} - (v_\parallel \mathbf{e}_\parallel + \mathbf{V}_D + \delta\mathbf{V}_D) \cdot \nabla_X \langle \delta\Phi \rangle_\alpha\right] \frac{\partial F_0}{\partial \varepsilon},\end{aligned} \tag{140}$$

Noting that $\partial F_0 / \partial t = 0$, $\mathbf{e}_\parallel \cdot \nabla F_0 = 0$, $\nabla F_0 \sim O(\lambda^1) F_0$, we find that the third line of the above equation is of order $O(\lambda^3)$ and thus can be dropped. Moving the second line to the right-hand side and noting that $\langle \delta L \rangle_\alpha = \langle \delta\phi - \mathbf{v} \cdot \delta\mathbf{A} \rangle_\alpha$, the above equation is written as

$$\begin{aligned}&\left[\frac{\partial}{\partial t} + (v_\parallel \mathbf{e}_\parallel + \mathbf{V}_D + \delta\mathbf{V}_D) \cdot \nabla_X\right] \delta h \\ =&-\delta\mathbf{V}_D \cdot \nabla_X F_0 \\ &-\frac{q}{m}[(v_\parallel \mathbf{e}_\parallel + \mathbf{V}_D + \delta\mathbf{V}_D) \cdot \nabla_X (\langle \mathbf{v} \cdot \delta\mathbf{A} - \delta\Phi \rangle_\alpha)] \frac{\partial F_0}{\partial \varepsilon},\end{aligned} \tag{141}$$

where two $\partial \langle \mathbf{v} \cdot \delta\mathbf{A} \rangle_\alpha / \partial t$ terms cancel each other and no time derivatives of the perturbed fields appear on the right-hand side. Noting that $\delta\mathbf{V}_D$ given by Eq. (122) is perpendicular to $\nabla_X \langle \mathbf{v} \cdot \delta\mathbf{A} - \delta\Phi \rangle_\alpha$ and thus the blue term in Eq. (141) is zero, then Eq. (141) simplifies to

$$\begin{aligned}&\left[\frac{\partial}{\partial t} + (v_\parallel \mathbf{e}_\parallel + \mathbf{V}_D + \delta\mathbf{V}_D) \cdot \nabla_X\right] \delta h \\ =&-\delta\mathbf{V}_D \cdot \nabla_X F_0 \\ &-\frac{q}{m}[(v_\parallel \mathbf{e}_\parallel + \mathbf{V}_D) \cdot \nabla_X (\langle \mathbf{v} \cdot \delta\mathbf{A} - \delta\Phi \rangle_\alpha)] \frac{\partial F_0}{\partial \varepsilon}.\end{aligned} \tag{142}$$

5.2.1 For special case $\delta\mathbf{A}\approx\delta A_{\parallel}\mathbf{e}_{\parallel}$

Most gyrokinetic simulations approximate the vector potential as $\delta\mathbf{A}\approx\delta A_{\parallel}\mathbf{e}_{\parallel}$. In this case, accurate to $O(\lambda^2)$, equation (142) can be further written as

$$\begin{aligned}&\left[\frac{\partial}{\partial t}+(v_{\parallel}\mathbf{e}_{\parallel}+\mathbf{V}_D+\delta\mathbf{V}_D)\cdot\nabla_X\right]\delta h\\ &=-\delta\mathbf{V}_D\cdot\nabla_X F_0\\ &-\frac{q}{m}[-\mathbf{V}_G\cdot\nabla_X\langle\delta\Phi\rangle_\alpha+\langle\delta A_{\parallel}\rangle_\alpha v_{\parallel}\mathbf{e}_{\parallel}\cdot\nabla_X(v_{\parallel})+v_{\parallel}\mathbf{V}_G\cdot\nabla_X(\langle\delta A_{\parallel}\rangle_\alpha)]\frac{\partial F_0}{\partial\varepsilon},\end{aligned}\tag{143}$$

Using Eq. (222), we obtain $\nabla_X(v_{\parallel})=-\mu(\nabla B_0)/v_{\parallel}$. Using this, equation (143) is further written as

$$\begin{aligned}&\left[\frac{\partial}{\partial t}+(v_{\parallel}\mathbf{e}_{\parallel}+\mathbf{V}_D+\delta\mathbf{V}_D)\cdot\nabla_X\right]\delta h\\ &=-\delta\mathbf{V}_D\cdot\nabla_X F_0\\ &-\frac{q}{m}[-\mathbf{V}_G\cdot\nabla_X\langle\delta\Phi\rangle_\alpha-\langle\delta A_{\parallel}\rangle_\alpha\mu\mathbf{e}_{\parallel}\cdot\nabla B_0+v_{\parallel}\mathbf{V}_G\cdot\nabla_X(\langle\delta A_{\parallel}\rangle_\alpha)]\frac{\partial F_0}{\partial\varepsilon},\end{aligned}\tag{144}$$

which agrees with the so-called $p_{\parallel}$ formulation given in GEM code manual (the first line of Eq. 28), which uses $p_{\parallel}=v_{\parallel}+q\langle A_{\parallel}\rangle_\alpha/m$ as an independent variable.

5.3 Summary of split of the distribution function

In the above, the perturbed part of the distribution function, δF, is split at least three times in order to (1) simplify the gyrokinetic equation by splitting out the adiabatic response and (2) eliminate the time derivatives, $\partial\delta\phi/\partial t$ and $\partial\delta\mathbf{A}/\partial t$, on the right-hand. To avoid confusion, I summarize the split of the distribution function here. The total distribution function F is split as

$$F=F_0+\delta F,\tag{145}$$

where F_0 is the equilibrium distribution function and δF is the perturbed part of the total distribution function. δF is further split as

$$\delta F=\delta h+\frac{q}{m}(\delta\Phi-\langle\delta\Phi\rangle_\alpha)\frac{\partial F_0}{\partial\varepsilon}+\frac{q}{m}\langle\mathbf{v}\cdot\delta\mathbf{A}\rangle_\alpha\frac{\partial F_0}{\partial\varepsilon},\tag{146}$$

where δh satisfies the gyrokinetic equation (142) or (144). In Eq. (146), the red term gives rise to the so-called polarization density (discussed in Sec. 5.6). The analytic dependence of this term on $\delta\Phi$ is utilized in solving the Poisson equation. The blue term also has an analytic dependence on $\delta\mathbf{A}$, which, however, will cause numerical problems in particle simulations (so-called "cancellation problem" in gyrokinetic simulations) if it is utilized in solving the Ampere equation.

5.3.1 Velocity space moment of $\frac{q}{m}\langle\mathbf{v}\cdot\delta\mathbf{A}\rangle_\alpha\frac{\partial F_0}{\partial\varepsilon}$

Consider the approximation $\delta\mathbf{A}\approx\delta A_{\parallel}\mathbf{e}_{\parallel}$, then the blue term in Eq. (146) is written as

$$\frac{q}{m}\langle v_{\parallel}\delta A_{\parallel}\rangle_\alpha\frac{\partial F_0}{\partial\varepsilon}.\tag{147}$$

For electrons, the FLR effect can be neglected and then the above expression is written

$$\frac{q}{m}v_{\parallel}\delta A_{\parallel}\frac{\partial F_0}{\partial\varepsilon}.\tag{148}$$

The zeroth order moment (number density) is then written as

$$\frac{q}{m}\delta A_{\parallel}\int v_{\parallel}\frac{\partial F_0}{\partial\varepsilon}d\mathbf{v},\tag{149}$$

which is zero if F_0 is Maxwellian. The parallel current is given by

$$\delta j_{\parallel}=\frac{q^2}{m}\delta A_{\parallel}\int v_{\parallel}^2\frac{\partial F_0}{\partial\varepsilon}d\mathbf{v}.\tag{150}$$

If F_0 is a Maxwellian distribution, then

5.4 Generalized split-weight scheme for electrons

For electrons, due to their small Larmor radius, the difference $\delta\Phi - \langle\delta\Phi\rangle_\alpha$ can be neglected. Then Eq. (146) is reduced to

$$\delta F = \delta h + \frac{q}{m}\langle \mathbf{v}\cdot\delta\mathbf{A}\rangle_\alpha \frac{\partial F_0}{\partial\varepsilon}, \tag{151}$$

where the adiabatic response is cancelled out. The generalized split-weight scheme introduces again the adiabatic response but with a free small parameter ϵ_g. Specifically, δh is split as

$$\delta h = \delta h_s + \epsilon_g \frac{q}{m}\delta\Phi \frac{\partial F_0}{\partial\varepsilon}. \tag{152}$$

Substituting this expression into Eq. (142), we obtain the following equation for δh_s:

$$\begin{aligned}
&\left[\frac{\partial}{\partial t} + (v_\parallel \mathbf{e}_\parallel + \mathbf{V}_D + \delta\mathbf{V}_D)\cdot\nabla_X\right]\delta h_s \\
&+\epsilon_g \frac{q}{m}\frac{\partial F_0}{\partial\varepsilon}\left[\frac{\partial}{\partial t} + (v_\parallel \mathbf{e}_\parallel + \mathbf{V}_D + \delta\mathbf{V}_D)\cdot\nabla_X\right]\delta\Phi \\
&+\epsilon_g \frac{q}{m}\delta\Phi\left[\frac{\partial}{\partial t} + (v_\parallel \mathbf{e}_\parallel + \mathbf{V}_D + \delta\mathbf{V}_D)\cdot\nabla_X\right]\frac{\partial F_0}{\partial\varepsilon} \\
&= -\delta\mathbf{V}_D\cdot\nabla_X F_0 \\
&-\frac{q}{m}[(v_\parallel \mathbf{e}_\parallel + \mathbf{V}_D)\cdot\nabla_X(\langle \mathbf{v}\cdot\delta\mathbf{A} - \delta\Phi\rangle_\alpha)]\frac{\partial F_0}{\partial\varepsilon}.
\end{aligned} \tag{153}$$

Noting that $\partial F_0/\partial t = 0$, $\mathbf{e}_\parallel\cdot\nabla F_0 = 0$, $\nabla F_0 \sim O(\lambda^1)F_0$, we find that the third line of the above equation is of order $O(\lambda^3)$ and thus can be dropped. Moving the second line to the right-hand side, the above equation is written as

$$\begin{aligned}
&\left[\frac{\partial}{\partial t} + (v_\parallel \mathbf{e}_\parallel + \mathbf{V}_D + \delta\mathbf{V}_D)\cdot\nabla_X\right]\delta h_s \\
&= -\delta\mathbf{V}_D\cdot\nabla_X F_0 \\
&-\frac{q}{m}\left\{(v_\parallel \mathbf{e}_\parallel + \mathbf{V}_D)\cdot\nabla_X\langle \mathbf{v}\cdot\delta\mathbf{A} - \delta\Phi\rangle_\alpha + \epsilon_g\left[\frac{\partial\delta\Phi}{\partial t} + \mathbf{V}_G\cdot\nabla_X\delta\Phi\right]\right\}\frac{\partial F_0}{\partial\varepsilon}.
\end{aligned} \tag{154}$$

Equation (154) agrees with Eq. (39) in the GEM manual. Note that the right-hand side of Eq. (154) contains a nonlinear term $\mathbf{V}_G\cdot\nabla_X\delta\Phi$, which is different from the original Frieman-Chen equation, where all nonlinear terms appear on the left-hand side. For the special case of $\epsilon_g = 1$ (the default and most used case in GEM code, Yang Chen said $\epsilon_g \neq 1$ cases are sometimes not accurate, so he gave up using it since 2009), the above equation can be simplified as (again nelecting the difference between $\delta\Phi$ and $\langle\delta\Phi\rangle_\alpha$):

$$\begin{aligned}
&\left[\frac{\partial}{\partial t} + (v_\parallel \mathbf{e}_\parallel + \mathbf{V}_D + \delta\mathbf{V}_D)\cdot\nabla_X\right]\delta h_s \\
&= -\delta\mathbf{V}_D\cdot\nabla_X F_0 \\
&-\frac{q}{m}\left\{\mathbf{V}_G\cdot\nabla_X\langle \mathbf{v}\cdot\delta\mathbf{A}\rangle_\alpha + \frac{\partial\delta\Phi}{\partial t}\right\}\frac{\partial F_0}{\partial\varepsilon}.
\end{aligned} \tag{155}$$

Since the original Frieman-Chen equation already splits the perturbed distribution into adiabatic part and nonadiabatic part, equation (155) actually goes back to the original Frieman-Chen equation. The only difference is that $\frac{q}{m}\langle \mathbf{v}\cdot\delta\mathbf{A}\rangle_\alpha \frac{\partial F_0}{\partial\varepsilon}$ is further split from the perturbed distribution function. Considering this, equation (155) can also be obtained from the original Frieman-Chen equation (133) by write δG_0 as

$$\delta G_0 = \delta h_s + \frac{q}{m}\langle \mathbf{v}\cdot\delta\mathbf{A}\rangle_\alpha \frac{\partial F_0}{\partial\varepsilon}, \tag{156}$$

[In this case δF is written as

$$\delta F = \delta h_s + \frac{q}{m}\delta\Phi\frac{\partial F_{g0}}{\partial\varepsilon} + \frac{q}{m}\langle \mathbf{v}\cdot\delta\mathbf{A}\rangle_\alpha \frac{\partial F_0}{\partial\varepsilon}, \tag{157}$$

] Substituting expression (156) into equation (133), we obtain the following equation for δh_s:

$$\begin{aligned}
&\left[\frac{\partial}{\partial t}+(v_{\parallel}\mathbf{e}_{\parallel}+\mathbf{V}_D+\delta\mathbf{V}_D)\cdot\nabla_X\right]\delta h_s\\
&+\frac{q}{m}\frac{\partial F_0}{\partial\varepsilon}\left[\frac{\partial}{\partial t}+(v_{\parallel}\mathbf{e}_{\parallel}+\mathbf{V}_D+\delta\mathbf{V}_D)\cdot\nabla_X\right]\langle\mathbf{v}\cdot\delta\mathbf{A}\rangle_{\alpha}\\
&+\frac{q}{m}\langle\mathbf{v}\cdot\delta\mathbf{A}\rangle_{\alpha}\left[\frac{\partial}{\partial t}+(v_{\parallel}\mathbf{e}_{\parallel}+\mathbf{V}_D+\delta\mathbf{V}_D)\cdot\nabla_X\right]\frac{\partial F_0}{\partial\varepsilon}\\
&=-\delta\mathbf{V}_D\cdot\nabla_X F_0-\frac{q}{m}\frac{\partial\langle\delta\Phi-\mathbf{v}\cdot\delta\mathbf{A}\rangle_{\alpha}}{\partial t}\frac{\partial F_0}{\partial\varepsilon},
\end{aligned}\tag{158}$$

Noting that $\partial F_0/\partial t=0$, $\mathbf{e}_{\parallel}\cdot\nabla F_0=0$, $\nabla F_0\sim O(\lambda^1)F_0$, we find that the third line of the above equation is of order $O(\lambda^3)$ and thus can be dropped. Moving the second line to the right-hand side, the above equation is written as

$$\begin{aligned}
&\left[\frac{\partial}{\partial t}+(v_{\parallel}\mathbf{e}_{\parallel}+\mathbf{V}_D+\delta\mathbf{V}_D)\cdot\nabla_X\right]\delta h_s\\
&=-\delta\mathbf{V}_D\cdot\nabla_X F_0\\
&-\frac{q}{m}\left[\frac{\partial\langle\delta\Phi\rangle_{\alpha}}{\partial t}+[(v_{\parallel}\mathbf{e}_{\parallel}+\mathbf{V}_D+\delta\mathbf{V}_D)\cdot\nabla_X]\langle\mathbf{v}\cdot\delta\mathbf{A}\rangle_{\alpha}\right]\frac{\partial F_0}{\partial\varepsilon},
\end{aligned}\tag{159}$$

which agrees with Eq. (155).

5.5 Comments on how to split the distribution function

In particle simulations, the seemingly trivial thing on how to split the distribution function is always considered to be a big deal. Separating the perturbed part from the equilibrium part is considered to be a big deal and gives rise to the famous name "δf particle method", in contrast to the conventional particle method which is now called full-f particle method. Summarizing the above result, we know that the total distribution function F is split in the following form:

$$\begin{aligned}
F &= F_0+\delta F\\
&= F_0+\delta h+\frac{q}{m}(\delta\Phi-\langle\delta\Phi\rangle_{\alpha})\frac{\partial F_0}{\partial\varepsilon}+\frac{q}{m}\langle\mathbf{v}\cdot\delta\mathbf{A}\rangle_{\alpha}\frac{\partial F_0}{\partial\varepsilon},
\end{aligned}\tag{160}$$

and only δh is actually evolved by using markers and its moment in the phase-space is evaluated via Monte-Carlo integration. The blue and red terms in the above expression depends on the perturbed field and their value can be obtained from the values on the grid and the velocity integration can be performed analytically. However, in some cases, the phase space integration of the blue terms must be evaluated using markers, i.e., using Monte-Carlo method, to avoid the inaccurate cancellation between the integration of these parts and the integration of δh (which are computed using Monte-Carlo method). When will the inaccurate cancellation is significant depends on the problem being investigated and thus can only be determined by actual numerical experiments. Many electromagnetic particle simulation experiments indicate that the parallel current carried by the blue term must be evaluated via Monte-Carlo method, otherwise inaccurate cancellation between this term and δh will give rise to numerical instabilities.

5.6 Poisson's equation and polarization density

Poisson's equation is written as

$$-\varepsilon_0\nabla^2\delta\Phi=q_i\delta n_i+q_e\delta n_e,\tag{161}$$

where $-\varepsilon_0\nabla^2\delta\Phi$ is called the space-charge term. Since we consider modes with $k_{\parallel}\ll k_{\perp}$, the space-charge term is approximated as $\nabla^2\delta\Phi\equiv\nabla_{\perp}^2\delta\Phi+\nabla_{\parallel}^2\delta\Phi\approx\nabla_{\perp}^2\delta\Phi$. Then Eq. (161) is written as

$$-\varepsilon_0\nabla_{\perp}^2\delta\Phi=q_i\delta n_i+q_e\delta n_e.\tag{162}$$

This approximation eliminates the parallel plasma oscillation from the system. The perpendicular plasma oscillations seem to be partially eliminated in system with gyrokinetic ions and drift-kinetic electrons. There are the so-called Ω_H modes that appear in the gyrokinetic system which have some similarity with the plasma oscillations but with a much smaller frequency $\Omega_H\sim(k_{\parallel}/k_{\perp})(\lambda_D/\rho_s)\omega_{pe}$. (electrostatic shear Alfven wave)

Using expression (146), the perturbed ion density δn_i is written as

$$\begin{aligned}\delta n_i &= \int \delta F d\mathbf{v} \\ &= \int \delta h\, d\mathbf{v} + \int \left[\frac{q}{m}(\delta\Phi - \langle\delta\Phi\rangle_\alpha)\frac{\partial F_0}{\partial\varepsilon}\right] d\mathbf{v} + \int \left[\frac{q}{m}\langle\mathbf{v}\cdot\delta\mathbf{A}\rangle_\alpha \frac{\partial F_0}{\partial\varepsilon}\right] d\mathbf{v},\end{aligned} \tag{163}$$

where the last term (in blue) is approximately zero for isotropic F_0 and this term is usually dropped in simulations that assume isotropic F_0 and approximate $\delta\mathbf{A}$ as $\delta A_\parallel \mathbf{e}_\parallel$. The second term (in red) in expression (163) is the so-called the polarization density n_p, i.e.,

$$\delta n_{pi}(\mathbf{x}) = \int \frac{q_i}{m_i}(\delta\Phi - \langle\delta\Phi\rangle_\alpha)\frac{\partial F_{i0}}{\partial\varepsilon} d\mathbf{v}, \tag{164}$$

which has an explicit dependence on $\delta\Phi$ and is usually moved to the left hand of Poisson's equation when constructing the numerical solver of the Poisson equation, i.e., equation (162) is written as

$$-\varepsilon_0 \nabla_\perp^2 \delta\Phi - q_i \int \frac{q_i}{m_i}(\delta\Phi - \langle\delta\Phi\rangle_\alpha)\frac{\partial F_{i0}}{\partial\varepsilon} d\mathbf{v} = q_i \delta n_i' + q_e \delta n_e, \tag{165}$$

where $\delta n_i' = \delta n_i - \delta n_{pi} = \int \delta h_i\, d\mathbf{v}$, which is evaluated by using Monte-Carlo markers. Since some parts depending on $\delta\Phi$ are moved from the right-hand side to the left-hand side of the field equation, numerical solvers (for $\delta\Phi$) based on the left-hand side of Eq. (165) probably behaves better than the one that is based on the left-hand side of Eq. (162), i.e., $-\varepsilon_0 \nabla_\perp^2 \delta\Phi$.

The polarization density is part of the perturbed density that is extracted from the right-hand side and moved to the left-hand side. The polarization density will be evaluated analytically, which is independent of Monte-Carlo markers, whereas the remained density on the right-hand side will be evaluated using Mote-Carlo markers. The two different methods of evaluating the two different parts of the total perturbed density can possibly introduce significant systematic errors especially if the two terms are expected to cancel each other and give a small quantity that is much smaller than either of the two terms. This is one pitfall for PIC simulations that extract some parts from the source term and move them to the left-hand side. To remedy this, we can introduce an additional term on the right-hand that gives the difference between the polarization density evaluated by the two methods. This is often called the cancellation scheme. It turns out the cancellation scheme is not necessary for Eq. (165), but for the field solver for Ampere's equation (discussed later), this cancellation scheme is necessary in order to obtain stable results.

Next, let us perform the gyro-averaging and velocity integration in expression(164). Since $\delta\Phi$ is independent of velocity coordinates, the first term (adiabatic term) in expression (164) is trivial and the velocity integration can be readily performed (assume F_0 is Maxwellian), giving

$$\begin{aligned}&\int \frac{q}{m}(\delta\Phi)\frac{\partial F_0}{\partial\varepsilon} d\mathbf{v}. \\ =&\frac{q}{m}(\delta\Phi)\int \left(-\frac{m}{T} f_M\right) d\mathbf{v}. \\ =&-\frac{q n_0}{T}\delta\Phi.\end{aligned} \tag{166}$$

Next, we try to perform the gyro-averaging and the velocity integration of the second term in expression (164). In order to perform the gyro-averaging, we expand $\delta\Phi$ in wave-number space as

$$\delta\Phi(\mathbf{x}) = \int \delta\Phi_k \exp(i\mathbf{k}\cdot\mathbf{x}) \frac{d\mathbf{k}}{(2\pi)^3}, \tag{167}$$

and we need to express $\mathbf{x}$ in terms of the guiding center variables $(\mathbf{X}, \mathbf{v})$ since the gyro-averaging is taken by holding $\mathbf{X}$ rather than $\mathbf{x}$ constant. The guiding-center transformation gives

$$\mathbf{x} = \mathbf{X} + \boldsymbol{\rho}(\mathbf{x}, \mathbf{v}) \approx \mathbf{X} - \mathbf{v} \times \frac{\mathbf{e}_\parallel(\mathbf{X})}{\Omega(\mathbf{X})} \tag{168}$$

Using this, the gyro-average of $\delta\Phi$ is written as

$$
\begin{aligned}
\langle\delta\Phi\rangle_\alpha &= \left\langle\int \delta\Phi_k \exp(i\mathbf{k}\cdot\mathbf{x})\frac{d\mathbf{k}}{(2\pi)^3}\right\rangle_\alpha \\
&= \left\langle\int \delta\Phi_k \exp(i\mathbf{k}\cdot(\mathbf{X}+\boldsymbol{\rho}))\frac{d\mathbf{k}}{(2\pi)^3}\right\rangle_\alpha \\
&= \int \delta\Phi_k \exp(i\mathbf{k}\cdot\mathbf{X})\langle\exp(i\mathbf{k}\cdot\boldsymbol{\rho})\rangle_\alpha\frac{d\mathbf{k}}{(2\pi)^3} \\
&= \int \delta\Phi_k \exp(i\mathbf{k}\cdot\mathbf{X})\left\langle\exp\left(-i\mathbf{k}\cdot\mathbf{v}\times\frac{\mathbf{e}_\parallel(\mathbf{X})}{\Omega(\mathbf{X})}\right)\right\rangle_\alpha\frac{d\mathbf{k}}{(2\pi)^3}. \qquad (169)
\end{aligned}
$$

When doing the gyro-averaging, $\mathbf{X}$ is hold constant and thus $\mathbf{e}_\parallel(\mathbf{X})$ is also constant. Then it is straightforward to define the gyro-angle α. Let $\mathbf{k}_\perp$ define one of the perpendicular direction $\hat{\mathbf{e}}_1$, i.e., $\mathbf{k}_\perp = k_\perp\hat{\mathbf{e}}_1$. Then another perpendicular basis vector is defined by $\hat{\mathbf{e}}_2 = \mathbf{e}_\parallel\times\hat{\mathbf{e}}_1$. Then $\mathbf{v}_\perp$ is written as $\mathbf{v}_\perp = v_\perp(\hat{\mathbf{e}}_1\cos\alpha + \hat{\mathbf{e}}_2\sin\alpha)$, which defines the gyro-angle α. Then the expression in Eq. (169) is written as

$$
\begin{aligned}
-i\mathbf{k}\cdot\mathbf{v}\times\frac{\mathbf{e}_\parallel(\mathbf{X})}{\Omega(\mathbf{X})} &= -i\mathbf{k}\cdot v_\perp(\hat{\mathbf{e}}_1\cos\alpha+\hat{\mathbf{e}}_2\sin\alpha)\times\frac{\mathbf{e}_\parallel(\mathbf{X})}{\Omega(\mathbf{X})} \\
&= -i\mathbf{k}\cdot\frac{v_\perp}{\Omega(\mathbf{X})}(-\hat{\mathbf{e}}_2\cos\alpha+\hat{\mathbf{e}}_1\sin\alpha) \\
&= -i\frac{k_\perp v_\perp}{\Omega}\sin\alpha. \qquad (170)
\end{aligned}
$$

Then the gyro-averaging in expression (169) is written as

$$
\begin{aligned}
\left\langle\exp\left(-i\mathbf{k}\cdot\mathbf{v}\times\frac{\mathbf{e}_\parallel(\mathbf{X})}{\Omega(\mathbf{X})}\right)\right\rangle_\alpha &= \left\langle\exp\left(-i\frac{k_\perp v_\perp}{\Omega}\sin\alpha\right)\right\rangle_\alpha \\
&= \frac{1}{2\pi}\int_0^{2\pi}\exp\left(-i\frac{k_\perp v_\perp}{\Omega}\sin\alpha\right)d\alpha \\
&= J_0\left(\frac{k_\perp v_\perp}{\Omega}\right). \qquad (171)
\end{aligned}
$$

where use has been made of the definition of the zeroth Bessel function of the first kind. Then $\langle\delta\Phi\rangle_\alpha$ in expression (169) is written as

$$
\langle\delta\Phi\rangle_\alpha = \int \delta\Phi_k\exp(i\mathbf{k}\cdot\mathbf{X})J_0\left(\frac{k_\perp v_\perp}{\Omega}\right)\frac{d\mathbf{k}}{(2\pi)^3}. \qquad (172)
$$

Next, we need to perform the integration in velocity space, which is done by holding $\mathbf{x}$ rather than $\mathbf{X}$ constant. Therefore, it is convenient to transform back to particle coordinates. Using $\mathbf{X} = \mathbf{x} + \mathbf{v}\times\frac{\mathbf{e}_\parallel(\mathbf{x})}{\Omega(\mathbf{x})}$, expression (172) is written as

$$
\langle\delta\Phi\rangle_\alpha = \int \delta\Phi_k\exp(i\mathbf{k}\cdot\mathbf{x})J_0\left(\frac{k_\perp v_\perp}{\Omega}\right)\exp\left(i\mathbf{k}\cdot\mathbf{v}\times\frac{\mathbf{e}_\parallel}{\Omega}\right)\frac{d\mathbf{k}}{(2\pi)^3}. \qquad (173)
$$

Then the velocity integration is written as

$$
\int\langle\delta\Phi\rangle_\alpha\frac{\partial F_0}{\partial\varepsilon}d\mathbf{v} = \int \delta\Phi_k\exp(i\mathbf{k}\cdot\mathbf{x})\int J_0\left(\frac{k_\perp v_\perp}{\Omega}\right)\exp\left(i\mathbf{k}\cdot\mathbf{v}\times\frac{\mathbf{e}_\parallel}{\Omega}\right)\frac{\partial F_0}{\partial\varepsilon}d\mathbf{v}\frac{d\mathbf{k}}{(2\pi)^3}. \qquad (174)
$$

Similar to Eq. (170), except for now at $\mathbf{x}$ rather than $\mathbf{X}$, $i\mathbf{k}\cdot\mathbf{v}\times\frac{\mathbf{e}_\parallel}{\Omega}$ is written as

$$
i\mathbf{k}\cdot\mathbf{v}\times\frac{\mathbf{e}_\parallel}{\Omega} = i\frac{k_\perp v_\perp}{\Omega}\sin\alpha. \qquad (175)
$$

Since this is at $\mathbf{x}$ rather than $\mathbf{X}$, $k_\perp$, $v_\perp$, and Ω are different from those appearing in expression (170). However, since this difference is due to the variation of the equilibrium quantity $\mathbf{e}_\parallel/\Omega$ in a Larmor radius, and thus is small and is ignored in the following.

Plugging expression (175) into expression (174) and using $d\mathbf{v} = v_\perp dv_\perp dv_\parallel d\alpha$, we obtain

$$
\int\langle\delta\Phi\rangle_\alpha\frac{\partial F_0}{\partial\varepsilon}d\mathbf{v} = \int \delta\Phi_k\exp(i\mathbf{k}\cdot\mathbf{x})\int J_0\left(\frac{k_\perp v_\perp}{\Omega}\right)\exp\left(i\frac{k_\perp v_\perp}{\Omega}\sin\alpha\right)\frac{\partial F_0}{\partial\varepsilon}v_\perp dv_\perp dv_\parallel d\alpha\frac{d\mathbf{k}}{(2\pi)^3}. \qquad (176)
$$

Note that $\partial F_0/\partial\varepsilon$ is independent of the gyro-angle α in terms of guiding-center variables. When transformed back to particle coordinates, $\mathbf{X}$ contained in $\partial F_0/\partial\varepsilon$ will introduce α dependence via $\mathbf{X}=\mathbf{x}+\mathbf{v}\times\frac{\mathbf{e}_\parallel}{\Omega}$. This dependence on α is weak since the equilibrium quantities can be considered constant over a Larmor radius distance evaluated at the thermal velocity. Therefore this dependence can be ignored when performing the integration over α, i.e., in terms of particle coordinates, $\partial F_0/\partial\varepsilon$ is approximately independent of the gyro-angle α. Then the integration over α in Eq. (176) can be performed, yielding

$$\begin{aligned}\int\langle\delta\Phi\rangle_\alpha\frac{\partial F_0}{\partial\varepsilon}d\mathbf{v} &= \int\delta\Phi_k\exp(i\mathbf{k}\cdot\mathbf{x})\iint J_0\left(\frac{k_\perp v_\perp}{\Omega}\right)\left[\int_0^{2\pi}\exp\left(i\frac{k_\perp v_\perp}{\Omega}\sin\alpha\right)d\alpha\right]\frac{\partial F_0}{\partial\varepsilon}v_\perp dv_\perp dv_\parallel\frac{d\mathbf{k}}{(2\pi)^3}\\ &= \int\delta\Phi_k\exp(i\mathbf{k}\cdot\mathbf{x})\left[\iint J_0\left(\frac{k_\perp v_\perp}{\Omega}\right)2\pi J_0\left(\frac{k_\perp v_\perp}{\Omega}\right)\frac{\partial F_0}{\partial\varepsilon}v_\perp dv_\perp dv_\parallel\right]\frac{d\mathbf{k}}{(2\pi)^3},\end{aligned}\tag{177}$$

where again use has been made of the definition of the Bessel function.

5.6.1 Special case: F_0 is Maxwellian

In order to perform the remaining velocity integration in expression (177), we assume that F_0 is a Maxwellian distribution given by

$$F_0=f_M=\frac{n_0(\mathbf{X})}{(2\pi T(\mathbf{X})/m)^{3/2}}\exp\left(\frac{-mv^2}{2T(\mathbf{X})}\right)=\frac{n_0}{(2\pi)^{3/2}v_t^3}\exp\left(\frac{-v^2}{2v_t^2}\right),\tag{178}$$

where $v_t=\sqrt{T/m}$, then

$$\frac{\partial F_0}{\partial\varepsilon}=-\frac{m}{T}f_M.\tag{179}$$

Again we will ignore the weak dependence of $n_0(\mathbf{X})$ and $\mathbf{T}_0(\mathbf{X})$ on v introduced by $\mathbf{X}=\mathbf{x}+\mathbf{v}\times\mathbf{e}_\parallel/\Omega$ when transformed back to particle coordinates (for sufficiently large velocity, the corresponding Larmor radius will be large enough to make the equilibrium undergo substantial variation. Since the velocity integration limit is to infinite, this will definitely occur. However, F_0 is exponentially decreasing with velocity, making those particles with velocity much larger than the thermal velocity negligibly few and thus can be neglected).

Plugging (179) into expression (177), we obtain

$$\begin{aligned}\int\langle\delta\Phi\rangle_\alpha\frac{\partial F_0}{\partial\varepsilon}d\mathbf{v} &= -\frac{m}{T}\frac{n_0}{(2\pi)^{3/2}}\int\delta\Phi_k\exp(i\mathbf{k}\cdot\mathbf{x})\iint 2\pi J_0^2\left(\frac{k_\perp v_\perp}{\Omega}\right)\exp\left(-\frac{v_\parallel^2+v_\perp^2}{2v_t^2}\right)\frac{1}{v_t^3}v_\perp dv_\perp dv_\parallel\frac{d\mathbf{k}}{(2\pi)^3},\\ &= -\frac{m}{T}\frac{n_0}{(2\pi)^{1/2}}\int\delta\Phi_k\exp(i\mathbf{k}\cdot\mathbf{x})\int_0^\infty\int_{-\infty}^{-\infty}J_0^2\left(\frac{k_\perp v_t}{\Omega}\,\overline{v}_\perp\right)\exp\left(-\frac{\overline{v}_\parallel^2+\overline{v}_\perp^2}{2}\right)\overline{v}_\perp\,d\,\overline{v}_\perp\,d\,\overline{v}_\parallel\frac{d\mathbf{k}}{(2\pi)^3},\end{aligned}\tag{180}$$

where $\overline{v}_\parallel=v_\parallel/v_t$, $\overline{v}_\perp=v_\perp/v_t$. Using

$$\int_{-\infty}^\infty\exp\left(-\frac{x^2}{2}\right)dx=\sqrt{2\pi}\tag{181}$$

the integration over $\overline{v}_\parallel$ can be performed, yielding

$$\int\langle\delta\Phi\rangle_\alpha\frac{\partial F_0}{\partial\varepsilon}d\mathbf{v} = -\frac{m}{T}n_0\int\delta\Phi_k\exp(i\mathbf{k}\cdot\mathbf{x})\int_0^\infty J_0^2\left(\frac{k_\perp v_t}{\Omega}\overline{v}_\perp\right)\exp\left(-\frac{\overline{v}_\perp^2}{2}\right)\overline{v}_\perp\,d\,\overline{v}_\perp\frac{d\mathbf{k}}{(2\pi)^3}\tag{182}$$

Using (I verified this by using Sympy)

$$\int_0^\infty J_0^2(ax)\exp\left(-\frac{x^2}{2}\right)x\,dx=\exp(-a^2)I_0(a^2),\tag{183}$$

where $I_0(a)$ is the zeroth modified Bessel function of the first kind, expression (182) is written

$$\int\langle\delta\Phi\rangle_\alpha\frac{\partial F_0}{\partial\varepsilon}d\mathbf{v}=-\frac{m}{T}n_0\int\delta\Phi_k\exp(i\mathbf{k}\cdot\mathbf{x})\exp(-b)I_0(b)\frac{d\mathbf{k}}{(2\pi)^3},\tag{184}$$

where $b=k_\perp^2v_t^2/\Omega^2$. Then the corresponding polarization density is written as

$$\int\frac{q}{m}\langle\delta\Phi\rangle_\alpha\frac{\partial F_0}{\partial\varepsilon}d\mathbf{v}=-\frac{qn_0}{T}\int\delta\Phi_k\exp(i\mathbf{k}\cdot\mathbf{x})\exp(-b)I_0(b)\frac{d\mathbf{k}}{(2\pi)^3}.\tag{185}$$

In Fourier space, the adiabatic term in expression (166) is written as

$$-\frac{qn_0}{T}\delta\Phi = -\frac{qn_0}{T}\int \delta\Phi_k \exp(i\mathbf{k}\cdot\mathbf{x})\frac{d\mathbf{k}}{(2\pi)^3}. \tag{186}$$

Plugging expression (184) and (186) into expression (164), the polarization density n_p is finally written as

$$\begin{aligned} n_p &= -\frac{qn_0}{T}\int \delta\Phi_k \exp(i\mathbf{k}\cdot\mathbf{x})[1-\exp(-b)I_0(b)]\frac{d\mathbf{k}}{(2\pi)^3}. & (187)\\ &= -\frac{qn_0}{T}\int \delta\Phi_k \exp(i\mathbf{k}\cdot\mathbf{x})[1-\Gamma_0]\frac{d\mathbf{k}}{(2\pi)^3}, & (188)\end{aligned}$$

where

$$\Gamma_0 = \exp(-b)I_0(b). \tag{189}$$

Expression (188) agrees with the result given in Yang Chen's notes.

5.6.2 Pade approximation

Γ_0 defined in Eq. (189) can be approximated by the Pade approximation as

$$\Gamma_0 \approx \frac{1}{1+b}. \tag{190}$$

The comparison between the exact value of Γ_0 and the above Pade approximation is shown in Fig. 3.

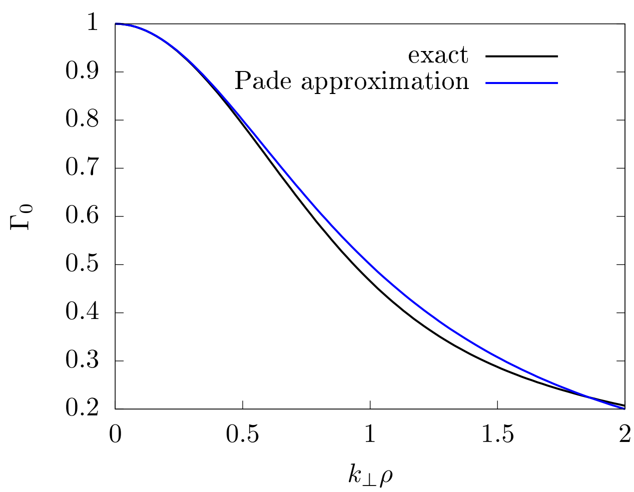

Figure 3. Comparison between the exact value of $\Gamma_0 = \exp(-(k_\perp\rho)^2)I_0((k_\perp\rho)^2)$ and the corresponding Pade approximation $1/(1+(k_\perp\rho)^2)$.

Using the Pade approximation (190), the polarization density n_p in expression (188) can be written as

$$n_p \approx -\frac{qn_0}{T}\int \delta\Phi_k \exp(i\mathbf{k}\cdot\mathbf{x})\frac{k_\perp^2\rho^2}{1+k_\perp^2\rho^2}\frac{d\mathbf{k}}{(2\pi)^3}. \tag{191}$$

5.6.3 Long wavelength approximation of the polarization density

In the long wavelength limit, $k_\perp\rho \ll 1$, expression (191) can be further approximated as

$$\begin{aligned} n_p &\approx -\frac{qn_0}{T}\int \delta\Phi_k \exp(i\mathbf{k}\cdot\mathbf{x})k_\perp^2\rho^2\frac{d\mathbf{k}}{(2\pi)^3}, \\ &= \frac{qn_0}{T}\rho^2\nabla_\perp^2\delta\Phi. & (192)\end{aligned}$$

Then the corresponding term in the Poisson equation is written as

$$\begin{aligned} \frac{q}{\varepsilon_0}n_p &= \frac{q^2 n_0}{\varepsilon_0 T}\rho^2\nabla_\perp^2\delta\Phi \\ &= \frac{\rho^2}{\lambda_D^2}\nabla_\perp^2\delta\Phi, & (193)\end{aligned}$$

where λ_D is the Debye length defined by $\lambda_D^2 = T\varepsilon_0/(n_0 q^2)$. For typical tokamak plasmas, the thermal ion gyroradius ρ_i is much larger than λ_D. This indicates that the term in expression (193) for ions is much larger than the space charge term $\nabla^2\delta\Phi \equiv \nabla_\perp^2\delta\Phi + \nabla_\parallel^2\delta\Phi \approx \nabla_\perp^2\delta\Phi$ in the Poisson equation. Therefore the space charge term can be neglected for ion-scale modes, indicating that gyrokinetic plasmas are intrinsically charge-neutral on the ion-scale. In other words, the ion polarization shielding is dominant compared with the Debye shielding for ion-scale low-frequency micro-instabilities. For electron-scale modes, since $\rho_e \approx \lambda_D$, the space charge term is comparable to the polarization density and should be included.

5.6.4 Polarization density expressed in terms of Laplacian operator

The polarization density expression (192) is for the long wavelength limit, which partially neglects FLR effect. Let us go back to the more general expression (191). In the wave-number space, the Poisson equation is written

$$-\varepsilon_0\nabla_\perp^2\delta\Phi = q_i\delta n_i + q_e\delta n_e. \tag{194}$$

Write $\delta n_i = n_{pi} + \delta n_i'$, where δn_{pi} is the ion polarization density, then the above expression is written

$$-\varepsilon_0\nabla_\perp^2\delta\Phi - q_i n_{pi} = q_i\delta n_i' + q_e\delta n_e. \tag{195}$$

The Fourier transforming in space, the above equation is written

$$-\varepsilon_0 k_\perp^2\delta\hat{\Phi} - q_i\hat{n}_{pi} = q_i\delta\hat{n}_i' + q_e\delta\hat{n}_e, \tag{196}$$

where $\hat{n}_{pi}$ is the Fourier transformation (in space) of the polarization density n_{pi} and similar meanings for $\delta\hat{\Phi}$, $\delta\hat{n}_i'$, and $\delta\hat{n}_e$. Expression (191) implies that $\hat{n}_{pi}$ is given by

$$\hat{n}_{pi} = -\frac{q_i n_{i0}}{T_i}\delta\hat{\Phi}\frac{k_\perp^2\rho_i^2}{1+k_\perp^2\rho_i^2}. \tag{197}$$

Using this, equation (196) is written

$$-\varepsilon_0 k_\perp^2\delta\hat{\Phi} - q_i\left(-\frac{q_i n_{i0}}{T_i}\frac{k_\perp^2\rho_i^2}{1+k_\perp^2\rho_i^2}\delta\hat{\Phi}\right) = q_i\delta\hat{n}_i' + q_e\delta\hat{n}_e, \tag{198}$$

Multiplying both sides by $(1+k_\perp^2\rho_i^2)/\varepsilon_0$, the above equation is written

$$-(1+k_\perp^2\rho_i^2)k_\perp^2\delta\hat{\Phi} - \frac{q_i}{\varepsilon_0}\left(-\frac{q_i n_{i0}}{T_i}(k_\perp^2\rho_i^2)\delta\hat{\Phi}\right) = \frac{1}{\varepsilon_0}(1+k_\perp^2\rho_i^2)(q_i\delta\hat{n}_i' + q_e\delta\hat{n}_e). \tag{199}$$

Next, transforming the above equation back to the real space, we obtain

$$-(1-\rho_i^2\nabla_\perp^2)\nabla_\perp^2\delta\Phi - \frac{q_i}{\varepsilon_0}\left(\frac{q_i n_{i0}}{T_i}\rho_i^2\nabla_\perp^2\delta\Phi\right) = \frac{1}{\varepsilon_0}(1-\rho_i^2\nabla_\perp^2)(q_i\delta n_i' + q_e\delta n_e). \tag{200}$$

Neglecting the Debye shielding term, the above equation is written

$$-\left(\frac{\rho_i^2}{\lambda_{Di}^2}\nabla_\perp^2\delta\Phi\right) = \frac{1}{\varepsilon_0}(1-\rho_i^2\nabla_\perp^2)(q_i\delta n_i' + q_e\delta n_e), \tag{201}$$

which is the equation actually solved in many gyrokinetic codes, where $\lambda_{Di}^2 = \varepsilon_0 T_i/(q_i^2 n_{i0})$.

6 Summary

Equations (142) and (188) are the primary results derived in this note. For reference ease, let us summarize the results. The total distribution function F is split as

$$F = F_0 + \delta F, \tag{202}$$

where F_0 and δF are the equilibrium part and perturbed part of the distribution function, respectively. δF is further split as

$$\delta F = \delta h + \frac{q}{m}(\delta\Phi - \langle\delta\Phi\rangle_\alpha)\frac{\partial F_0}{\partial\varepsilon} + \frac{q}{m}\langle\mathbf{v}\cdot\delta\mathbf{A}\rangle_\alpha\frac{\partial F_0}{\partial\varepsilon}, \tag{203}$$

where $\delta h = \delta h(\mathbf{X}, \mu, \varepsilon)$ satisfies the following nonlinear gyrokinetic equation (142), i.e.,

$$\begin{aligned} & \left[\frac{\partial}{\partial t} + (v_\parallel \mathbf{e}_\parallel + \mathbf{V}_D + \delta \mathbf{V}_D) \cdot \nabla_X\right] \delta h \\ = & -\delta \mathbf{V}_D \cdot \nabla_X F_0 \\ - & \frac{q}{m}[(v_\parallel \mathbf{e}_\parallel + \mathbf{V}_D) \cdot \nabla_X (\langle \mathbf{v} \cdot \delta \mathbf{A} - \delta \Phi \rangle_\alpha)] \frac{\partial F_0}{\partial \varepsilon}. \end{aligned} \tag{204}$$

where $\mathbf{V}_D$ is the equilibrium guiding-center drift velocity and $\delta \mathbf{V}_D$ is the perturbed drift velocity given by (122), i.e.,

$$\delta \mathbf{V}_D = \frac{q}{m} \nabla_X \langle \mathbf{v} \cdot \delta \mathbf{A} - \delta \Phi \rangle_\alpha \times \frac{\mathbf{e}_\parallel}{\Omega}. \tag{205}$$

Here the independent variables for the distribution functions F_0 and δh are $(\mathbf{X}, \mu, \varepsilon)$ with $\mathbf{X}$ being the guiding-center position, $\mu = v_\perp^2 / (2B_0)$, and $\varepsilon = v^2/2$. In the above, $\delta\Phi$ and $\delta\mathbf{A}$ is the perturbed electric potential and vector potential, $\langle ... \rangle_\alpha$ is the gyro-averaging operator over the gyro-phase α, $\mathbf{e}_\parallel = \mathbf{B}_0 / B_0$, $\Omega = B_0 q / m$ with q and m being the particle charge and mass, respectively.

When integrated in velocity space, the red term in expression (203) gives rise to the polarization density n_p, i.e.,

$$n_p = \int \frac{q}{m} (\delta\Phi - \langle \delta\Phi \rangle_\alpha) \frac{\partial F_0}{\partial \varepsilon} d\mathbf{v}, \tag{206}$$

which, for Maxwellian equilibrium distribution, can be written in wave-number space as

$$n_p = -\frac{q n_0}{T} \int \delta\Phi_k \exp(i \mathbf{k} \cdot \mathbf{x}) [1 - \Gamma_0] \frac{d\mathbf{k}}{(2\pi)^3}, \tag{207}$$

where $\Gamma_0 = \exp(-b) I_0(b)$, $b = k_\perp^2 v_t^2 / \Omega^2$, and $I_0(b)$ is the zeroth modified Bessel function of the first kind. This expression is useful for gyrokinetic simulations that use spectral methods in solving the Poisson equation.

These notes were initially written when I visited University of Colorado at Boulder (Sept.-Nov. 2016), where I worked with Dr. Yang Chen, who pointed out that most gyrokinetic simulations essentially employ Frieman-Chen's nonlinear gyrokinetic equation. Therefore a careful re-derivation of the equation to know the gyrokinetic orderings and physics included in the model is highly desirable, which motivates me to write this note.

Appendix A Diamagnetic flow

The perturbed distribution function δF given in Eq. (136) contains two terms. The first term is gyro-phase dependent while the second term is gyro-phase independent. The perpendicular velocity moment of the first term will give rise to the so-called $\delta\mathbf{E} \times \mathbf{B}_0$ flow (seems wrong, need checking) and the second term will give rise to the so-called diamagnetic flow. Let us discuss the diamagnetic flow first. For this case, it is crucial to distinguish between the distribution function in terms of the guiding-center variables, $f_g(\mathbf{X}, \mathbf{v})$, and that in terms of the particle variables, $f_p(\mathbf{x}, \mathbf{v})$. In terms of these denotations, equation (136) is written as

$$\delta F_g = \frac{q}{m} (\delta\Phi - \langle \delta\Phi \rangle_\alpha) \frac{\partial F_{0g}}{\partial \varepsilon} + \delta f_g. \tag{208}$$

Next, consider the perpendicular flow $\mathbf{U}_\perp$ carried by δf_g. This flow is defined by the corresponding distribution function in terms of the particle variables, δf_p, via,

$$n \mathbf{U}_\perp = \int \mathbf{v}_\perp \delta f_p(\mathbf{x}, \mathbf{v}) d\mathbf{v}, \tag{209}$$

where n is the number density defined by $n = \int \delta f_p d\mathbf{v}$. Using the relation between the particle distribution function and guiding-center distribution function given by Eq. (7), i.e.,

$$\delta f_p(\mathbf{x}, \mathbf{v}) = \delta f_g(\mathbf{x} - \boldsymbol{\rho}, \mathbf{v}), \tag{210}$$

equation (209) is written as

$$n\mathbf{U}_{\perp}=\int \mathbf{v}_{\perp}\delta f_g(\mathbf{x}-\boldsymbol{\rho},\mathbf{v})d\mathbf{v}. \tag{211}$$

Using the Taylor expansion near $\mathbf{x}$, $\delta f_g(\mathbf{x}-\boldsymbol{\rho},\mathbf{v})$ can be approximated as

$$\delta f_g(\mathbf{x}-\boldsymbol{\rho},\mathbf{v})\approx\delta f_g(\mathbf{x},\mathbf{v})-\boldsymbol{\rho}\cdot\nabla\delta f_g(\mathbf{x},\mathbf{v}). \tag{212}$$

Plugging this expression into Eq. (211), we obtain

$$n\mathbf{U}_{\perp}\approx\int \mathbf{v}_{\perp}\delta f_g(\mathbf{x},\mathbf{v})d\mathbf{v}-\int \mathbf{v}_{\perp}\boldsymbol{\rho}\cdot\nabla\delta f_g(\mathbf{x},\mathbf{v})d\mathbf{v} \tag{213}$$

As mentioned above, $\delta f_g(\mathbf{x},\mathbf{v})$ is independent of the gyro-angle α. It is obvious that the first integration is zero and thus Eq. (213) is reduced to

$$n\mathbf{U}_{\perp}=-\int \mathbf{v}_{\perp}\boldsymbol{\rho}\cdot\nabla\delta f_g(\mathbf{x},\mathbf{v})d\mathbf{v} \tag{214}$$

Using the definition $\boldsymbol{\rho}=-\mathbf{v}\times\mathbf{e}_{\parallel}/\Omega$, the above equation is written

$$\begin{aligned} n\mathbf{U}_{\perp} &= \int \mathbf{v}_{\perp}\frac{\mathbf{v}\times\mathbf{e}_{\parallel}}{\Omega}\cdot\nabla\delta f_g(\mathbf{x},\mathbf{v})d\mathbf{v} \\ &= \int \mathbf{v}_{\perp}\Big(\frac{\mathbf{e}_{\parallel}}{\Omega}\times\nabla\delta f_g(\mathbf{x},\mathbf{v})\Big)\cdot\mathbf{v}_{\perp}d\mathbf{v}. \\ &= \int \mathbf{v}_{\perp}\mathbf{H}\cdot\mathbf{v}_{\perp}d\mathbf{v}, \end{aligned} \tag{215}$$

where $\mathbf{H}=\frac{\mathbf{e}_{\parallel}}{\Omega}\times\nabla\delta f_g(\mathbf{x},\mathbf{v})$, which is independent of the gyro-angle α because both $\mathbf{e}_{\parallel}(\mathbf{x})/\Omega(\mathbf{x})$ and $\delta f_g(\mathbf{x},\mathbf{v})$ are independent of α. Next, we try to perform the integration over α in Eq. (215). In terms of velocity space cylindrical coordinates $(v_{\parallel},v_{\perp},\alpha)$, $\mathbf{v}_{\perp}$ is written as

$$\mathbf{v}_{\perp}=v_{\perp}(\hat{\mathbf{x}}\cos\alpha+\hat{\mathbf{y}}\sin\alpha), \tag{216}$$

where $\hat{\mathbf{x}}$ and $\hat{\mathbf{y}}$ are two arbitrary unit vectors perpendicular each other and both perpendicular to $\mathbf{B}_0(\mathbf{x})$. $\mathbf{H}$ can be written as

$$\mathbf{H}=H_x\hat{\mathbf{x}}+H_y\hat{\mathbf{y}}, \tag{217}$$

where H_x and H_y are independent of α. Using these in Eq. (215), we obtain

$$\begin{aligned} n\mathbf{U}_{\perp} &= \int v_{\perp}(\hat{\mathbf{x}}\cos\alpha+\hat{\mathbf{y}}\sin\alpha)\,v_{\perp}(H_x\cos\alpha+H_y\sin\alpha)d\mathbf{v} \\ &= \int v_{\perp}^2[\hat{\mathbf{x}}(H_x\cos^2\alpha+H_y\sin\alpha\cos\alpha)+\hat{\mathbf{y}}(H_x\cos\alpha\sin\alpha+H_y\sin^2\alpha)]\,d\mathbf{v}. \end{aligned} \tag{218}$$

Using $d\mathbf{v}=v_{\perp}dv_{\parallel}dv_{\perp}d\alpha$, the above equation is written as

$$\begin{aligned} n\mathbf{U}_{\perp} &= \int_{-\infty}^{\infty}dv_{\parallel}\int_0^{\infty}v_{\perp}dv_{\perp}\int_0^{2\pi}v_{\perp}^2[\hat{\mathbf{x}}(H_x\cos^2\alpha+H_y\sin\alpha\cos\alpha)+\hat{\mathbf{y}}(H_x\cos\alpha\sin\alpha+H_y\sin^2\alpha)]d\alpha \\ &= \int_{-\infty}^{\infty}dv_{\parallel}\int_0^{\infty}v_{\perp}dv_{\perp}\int_0^{2\pi}v_{\perp}^2(\hat{\mathbf{x}}H_x\cos^2\alpha+\hat{\mathbf{y}}H_y\sin^2\alpha)d\alpha \\ &= \int_{-\infty}^{\infty}dv_{\parallel}\int_0^{\infty}v_{\perp}dv_{\perp}[v_{\perp}^2(\hat{\mathbf{x}}H_x\pi+\hat{\mathbf{y}}H_y\pi)] \\ &= \int_{-\infty}^{\infty}dv_{\parallel}\int_0^{\infty}v_{\perp}dv_{\perp}[v_{\perp}^2\mathbf{H}\pi] \\ &= \int_{-\infty}^{\infty}dv_{\parallel}\int_0^{\infty}v_{\perp}dv_{\perp}[v_{\perp}^2\frac{\mathbf{e}_{\parallel}}{\Omega}\times\nabla\delta f_g(\mathbf{x},\mathbf{v})\pi] \\ &= \frac{\mathbf{e}_{\parallel}}{\Omega}\times\nabla\int_{-\infty}^{\infty}dv_{\parallel}\int_0^{\infty}v_{\perp}dv_{\perp}\delta f_g(\mathbf{x},\mathbf{v})\frac{v_{\perp}^2}{2}2\pi \\ &= \frac{\mathbf{e}_{\parallel}}{m\Omega}\times\nabla\delta p_{\perp}, \end{aligned} \tag{219}$$

where

$$\begin{aligned}\delta p_\perp &\equiv \int_{-\infty}^{\infty} dv_\parallel \int_0^\infty v_\perp dv_\perp \delta f_g(\mathbf{x},\mathbf{v}) \frac{m v_\perp^2}{2} 2\pi \\ &= \int \delta f_g(\mathbf{x},\mathbf{v}) \frac{m v_\perp^2}{2} d\mathbf{v},\end{aligned} \tag{220}$$

is the perpendicular pressure carried by $\delta f_g(\mathbf{x},\mathbf{v})$. The flow given by Eq. (219) is called the diamagnetic flow.

Appendix B Transform gyrokinetic equation from $(\mathbf{X}, \boldsymbol{\mu}, \boldsymbol{\varepsilon}, \boldsymbol{\alpha})$ to $(\mathbf{X}, \boldsymbol{\mu}, \boldsymbol{v}_\parallel, \boldsymbol{\alpha})$ coordinates

The gyrokinetic equation given above is written in terms of variables $(\mathbf{X}, \mu, \varepsilon, \alpha)$, where α is the gyro-phase. Next, we transform it into coordinates $(\mathbf{X}', \mu', v_\parallel, \alpha')$ which is defined by

$$\begin{cases} \mathbf{X}'(\mathbf{X}, \mu, \varepsilon, \alpha) = \mathbf{X} \\ \mu'(\mathbf{X}, \mu, \varepsilon, \alpha) = \mu \\ \alpha'(\mathbf{X}, \mu, \varepsilon, \alpha) = \alpha \\ v_\parallel(\mathbf{X}, \mu, \varepsilon, \alpha) = \pm\sqrt{2(\varepsilon - \mu B_0(\mathbf{X}))} \end{cases} \tag{221}$$

Use this definition and the chain rule, the gradient operators in $(\mathbf{X}, \mu, \varepsilon, \alpha)$ variables are written, in terms of $(\mathbf{X}', \mu', v_\parallel, \alpha')$ variables, as

$$\begin{aligned}\left.\frac{\partial}{\partial \mathbf{X}}\right|_{\mu,\varepsilon,\alpha} &= \frac{\partial \mathbf{X}'}{\partial \mathbf{X}} \cdot \left.\frac{\partial}{\partial \mathbf{X}'}\right|_{\mu', v_\parallel, \alpha'} + \frac{\partial \mu'}{\partial \mathbf{X}} \left.\frac{\partial}{\partial \mu'}\right|_{\mathbf{X}', v_\parallel, \alpha'} + \frac{\partial v_\parallel}{\partial \mathbf{X}} \left.\frac{\partial}{\partial v_\parallel}\right|_{\mathbf{X}', \mu', \alpha'} + \frac{\partial \alpha'}{\partial \mathbf{X}} \left.\frac{\partial}{\partial \alpha'}\right|_{\mathbf{X}', \mu', v_\parallel} \\ &= \frac{\partial}{\partial \mathbf{X}'} + 0 \frac{\partial}{\partial \mu'} - \frac{\mu}{v_\parallel} \frac{\partial B_0}{\partial \mathbf{X}} \frac{\partial}{\partial v_\parallel} + 0 \frac{\partial}{\partial \alpha'}\end{aligned} \tag{222}$$

and

$$\begin{aligned}\left.\frac{\partial}{\partial \varepsilon}\right|_{\mathbf{X},\mu,\alpha} &= \frac{\partial \mathbf{X}'}{\partial \varepsilon} \frac{\partial}{\partial \mathbf{X}'} + \frac{\partial \mu'}{\partial \varepsilon} \frac{\partial}{\partial \mu'} + \frac{\partial v_\parallel}{\partial \varepsilon} \frac{\partial}{\partial v_\parallel} + \frac{\partial \alpha'}{\partial \varepsilon} \frac{\partial}{\partial \alpha'} \\ &= 0 \frac{\partial}{\partial \mathbf{X}'} + 0 \frac{\partial}{\partial \mu'} + \frac{1}{v_\parallel} \frac{\partial}{\partial v_\parallel} + 0 \frac{\partial}{\partial \alpha'}\end{aligned} \tag{223}$$

Then, in terms of independent variable $(\mathbf{X}', \mu', v_\parallel, \alpha')$, equation (133) is written

$$\begin{aligned}&\left[\frac{\partial}{\partial t} + (v_\parallel \mathbf{e}_\parallel + \mathbf{V}_D + \delta \mathbf{V}_D) \cdot \frac{\partial}{\partial \mathbf{X}'}\right] \delta G_0 - (v_\parallel \mathbf{e}_\parallel + \mathbf{V}_D + \delta \mathbf{V}_D) \cdot \frac{\mu}{v_\parallel} \frac{\partial B_0}{d\mathbf{X}} \frac{\partial \delta G_0}{\partial v_\parallel} \\ &= -\delta \mathbf{V}_D \cdot \left(\frac{\partial F_0}{\partial \mathbf{X}'} - \frac{\mu}{v_\parallel} \frac{\partial B_0}{d\mathbf{X}} \frac{\partial F_0}{\partial v_\parallel}\right) - \frac{q}{m} \frac{\partial \langle \delta L \rangle_\alpha}{\partial t} \frac{\partial F_0}{\partial v_\parallel} \frac{1}{v_\parallel},\end{aligned} \tag{224}$$

where $\delta \mathbf{V}_D$ and $\langle \delta L \rangle_\alpha$ involve the gyro-averaging operator $\langle ... \rangle_\alpha$. The gyro-averaging operator in $(\mathbf{X}', \mu', v_\parallel, \alpha')$ coordinates is similar to that in the old coordinates since the perpendicular velocity variable μ is identical between the two coordinate systems. Also note that the perturbed guiding-center velocity $\delta \mathbf{V}_D$ is given by

$$\delta \mathbf{V}_D = \frac{\mathbf{e}_\parallel \times \nabla_X \langle \delta \phi \rangle_\alpha}{B_0} + v_\parallel \frac{\langle \delta \mathbf{B}_\perp \rangle_\alpha}{B_0}, \tag{225}$$

where $\partial / \partial \mathbf{X}$ (rather than $\partial / \partial \mathbf{X}'$) is used. Since $\delta \phi(\mathbf{x}) = \delta \phi_g(\mathbf{X}, \mu', \alpha')$, which is independent of $v_\parallel$, then Eq. (222) indicates that $\partial \delta \phi / \partial \mathbf{X} = \partial \delta \phi / \partial \mathbf{X}'$.

Dropping terms of order higher than $O(\lambda^2)$, equation (224) is written as

$$\begin{aligned}&\left[\frac{\partial}{\partial t} + (v_\parallel \mathbf{e}_\parallel + \mathbf{V}_D + \delta \mathbf{V}_D) \cdot \frac{\partial}{\partial \mathbf{X}'}\right] \delta G_0 - \mathbf{e}_\parallel \cdot \mu \nabla B_0 \frac{\partial \delta G_0}{\partial v_\parallel} \\ &= -\delta \mathbf{V}_D \cdot \left(\frac{\partial F_0}{\partial \mathbf{X}'}\right) + \left(\delta \mathbf{V}_D \cdot \mu \nabla B_0 - \frac{q}{m} \frac{\partial \langle \delta L \rangle_\alpha}{\partial t}\right) \frac{\partial F_0}{\partial v_\parallel} \frac{1}{v_\parallel},\end{aligned} \tag{226}$$

Similarly, in terms of independent variable $(\mathbf{X}', \mu', v_\parallel, \alpha')$, equation (138) is written as

$$
\begin{aligned}
&\left[\frac{\partial}{\partial t} + (v_\parallel \mathbf{e}_\parallel + \mathbf{V}_D + \delta\mathbf{V}_D) \cdot \frac{\partial}{\partial \mathbf{X}'}\right]\delta f - \mathbf{e}_\parallel \cdot \mu \nabla B_0 \frac{\partial \delta f}{\partial v_\parallel} \\
&= -\delta\mathbf{V}_D \cdot \left(\frac{\partial F_0}{\partial \mathbf{X}'}\right) + \delta\mathbf{V}_D \cdot \left(\frac{\mu}{v_\parallel} \nabla B_0 \frac{\partial F_0}{\partial v_\parallel}\right) \\
&\quad -\frac{q}{m}\left[-\frac{\partial \langle \mathbf{v} \cdot \delta\mathbf{A}\rangle_\alpha}{\partial t} - \left(v_\parallel \mathbf{e}_\parallel + \mathbf{V}_D + v_\parallel \frac{\langle \delta \mathbf{B}_\perp\rangle_\alpha}{B_0}\right) \cdot \nabla_X \langle \delta\phi\rangle_\alpha\right] \frac{\partial F_0}{\partial v_\parallel} \frac{1}{v_\parallel},
\end{aligned} \tag{227}
$$

The guiding-center velocity in the macroscopic (equilibrium) field is given by

$$
v_\parallel \mathbf{e}_\parallel + \mathbf{V}_D = \frac{\mathbf{B}_0^\star}{B_{\parallel 0}^\star} v_\parallel + \frac{\mu}{\Omega B_{\parallel 0}^\star} \mathbf{B}_0 \times \nabla B_0 + \frac{1}{B_0 B_{\parallel 0}^\star} \mathbf{E}_0 \times \mathbf{B}_0 \tag{228}
$$

where

$$
\mathbf{B}_0^\star = \mathbf{B}_0 + B_0 \frac{v_\parallel}{\Omega} \nabla \times \mathbf{b}, \tag{229}
$$

$$
B_\parallel^\star \equiv \mathbf{b} \cdot \mathbf{B}^\star = B\left(1 + \frac{v_\parallel}{\Omega} \mathbf{b} \cdot \nabla \times \mathbf{b}\right), \tag{230}
$$

Using $B_{\parallel 0}^\star \approx B_0$, then expression (228) is written as

$$
v_\parallel \mathbf{e}_\parallel + \mathbf{V}_D = v_\parallel \mathbf{b} + \underbrace{\frac{v_\parallel^2}{\Omega} \nabla \times \mathbf{b}}_{\text{curvature drift}} + \underbrace{\frac{\mu}{\Omega B_0} \mathbf{B}_0 \times \nabla B_0}_{\nabla B \text{ drift}} + \underbrace{\frac{1}{B_0^2} \mathbf{E}_0 \times \mathbf{B}_0}_{E \times B \text{ drift}}, \tag{231}
$$

where the curvature drift, ∇B drift, and $\mathbf{E}_0 \times \mathbf{B}_0$ drift can be identified. Note that the perturbed guiding-center velocity $\delta\mathbf{V}_D$ is given by

$$
\delta\mathbf{V}_D = \frac{\mathbf{e}_\parallel \times \nabla_X \langle \delta\phi\rangle_\alpha}{B_0} + v_\parallel \frac{\langle \delta\mathbf{B}_\perp\rangle_\alpha}{B_0}. \tag{232}
$$

Using the above results, equation (227) is written as

$$
\begin{aligned}
&\left[\frac{\partial}{\partial t} + (v_\parallel \mathbf{e}_\parallel + \mathbf{V}_D + \delta\mathbf{V}_D) \cdot \frac{\partial}{\partial \mathbf{X}'}\right]\delta f - \mathbf{e}_\parallel \cdot \mu \nabla B_0 \frac{\partial \delta f}{\partial v_\parallel} \\
&= -\delta\mathbf{V}_D \cdot \left(\frac{\partial F_0}{\partial \mathbf{X}'}\right) + \left(\frac{\mathbf{e}_\parallel \times \nabla_X \langle \delta\phi\rangle_\alpha}{B_0} + v_\parallel \frac{\langle \delta\mathbf{B}_\perp\rangle_\alpha}{B_0}\right) \cdot \left(\frac{\mu}{v_\parallel} \nabla B_0 \frac{\partial F_0}{\partial v_\parallel}\right) \\
&\quad -\frac{q}{m}\left[-\frac{\partial \langle \mathbf{v} \cdot \delta\mathbf{A}\rangle_\alpha}{\partial t} - \left(v_\parallel \mathbf{e}_\parallel + \frac{v_\parallel^2}{\Omega} \nabla \times \mathbf{b} + \frac{\mu}{\Omega B_0} \mathbf{B}_0 \times \nabla B_0 + \frac{1}{B_0^2} \mathbf{E}_0 \times \mathbf{B}_0 + v_\parallel \frac{\langle \delta\mathbf{B}_\perp\rangle_\alpha}{B_0}\right) \cdot \nabla_X \langle \delta\phi\rangle_\alpha\right] \frac{\partial F_0}{\partial v_\parallel} \frac{1}{v_\parallel},
\end{aligned} \tag{233}
$$

Collecting coefficients before $\partial F_0 / \partial v_\parallel$, we find that the two terms involving ∇B_0 (terms in blue and red) cancel each other, yielding

$$
\begin{aligned}
&\left[\frac{\partial}{\partial t} + (v_\parallel \mathbf{e}_\parallel + \mathbf{V}_D + \delta\mathbf{V}_D) \cdot \frac{\partial}{\partial \mathbf{X}'}\right]\delta f - \mathbf{e}_\parallel \cdot \mu \nabla B_0 \frac{\partial \delta f}{\partial v_\parallel} \\
&= -\delta\mathbf{V}_D \cdot \left(\frac{\partial F_0}{\partial \mathbf{X}'}\right) \\
&\quad +\frac{q}{m}\left[\frac{m}{q} v_\parallel \frac{\langle \delta\mathbf{B}_\perp\rangle_\alpha}{B_0} \cdot (\mu \nabla B_0) + \frac{\partial \langle \mathbf{v} \cdot \delta\mathbf{A}\rangle_\alpha}{\partial t} + \left(v_\parallel \mathbf{b} + \frac{v_\parallel^2}{\Omega} \nabla \times \mathbf{b} + \frac{1}{B_0^2} \mathbf{E}_0 \times \mathbf{B}_0 + v_\parallel \frac{\langle \delta\mathbf{B}_\perp\rangle_\alpha}{B_0}\right) \cdot \right. \\
&\quad \left. \nabla_X \langle \delta\phi\rangle_\alpha\right] \frac{\partial F_0}{\partial v_\parallel} \frac{1}{v_\parallel},
\end{aligned} \tag{234}
$$

This equation agrees with Eq. (8) in I. Holod's 2009 pop paper (gyro-averaging is wrongly omitted in that paper) and W. Deng's 2011 NF paper. Equation (234) drops all terms higher than $O(\lambda^2)$ and as a result the coefficient before $\partial \delta f / \partial v_\parallel$ contains only the mirror force, i.e.,

$$
\frac{d v_\parallel}{dt} = -\mathbf{e}_\parallel \cdot \mu \nabla B_0, \tag{235}
$$

which is independent of any perturbations.

Appendix C Transform gyrokinetic equation from ($\delta\Phi$, δA) to (δE, δB)

C.1 Expression of $\delta B_\perp$ in terms of δA

Note that

$$\begin{aligned} \delta\mathbf{B}_\perp &= \nabla\times\delta\mathbf{A} - (\mathbf{e}_\parallel\cdot\nabla\times\delta\mathbf{A})\mathbf{e}_\parallel \\ &= \nabla\times(\delta\mathbf{A}_\perp + \delta A_\parallel\mathbf{e}_\parallel) - [\mathbf{e}_\parallel\cdot\nabla\times(\delta\mathbf{A}_\perp + \delta A_\parallel\mathbf{e}_\parallel)]\mathbf{e}_\parallel \end{aligned} \tag{236}$$

Correct to order $O(\lambda)$, $\delta\mathbf{B}_\perp$ in the above equation is written as ($\mathbf{e}_\parallel$ vector can be considered as constant because its spatial gradient combined with $\delta\mathbf{A}$ will give terms of $O(\lambda^2)$, which are neglected)

$$\delta\mathbf{B}_\perp \approx \nabla\times\delta\mathbf{A}_\perp + \nabla\delta A_\parallel\times\mathbf{e}_\parallel - [\mathbf{e}_\parallel\cdot\nabla\times\delta\mathbf{A}_\perp + \mathbf{e}_\parallel\cdot(\nabla\delta A_\parallel\times\mathbf{e}_\parallel)]\mathbf{e}_\parallel \tag{237}$$

$$= \nabla\times\delta\mathbf{A}_\perp + \nabla\delta A_\parallel\times\mathbf{e}_\parallel - (\mathbf{e}_\parallel\cdot\nabla\times\delta\mathbf{A}_\perp)\mathbf{e}_\parallel \tag{238}$$

Using local cylindrical coordinates (r,ϕ,z) with z being along the local direction of $\mathbf{B}_0$, and two components of $\mathbf{A}_\perp$ being A_r and A_ϕ, then $\nabla\times\mathbf{A}_\perp$ is written as

$$\nabla\times\delta\mathbf{A}_\perp = \left(-\frac{\partial\delta A_\phi}{\partial z}\right)\mathbf{e}_r + \left(\frac{\partial\delta A_r}{\partial z}\right)\mathbf{e}_\phi + \frac{1}{r}\left[\frac{\partial}{\partial r}(r\delta A_\phi) - \frac{\partial\delta A_r}{\partial\phi}\right]\mathbf{e}_\parallel. \tag{239}$$

Note that the parallel gradient operator $\nabla_\parallel\equiv\mathbf{e}_\parallel\cdot\nabla = \partial/\partial z$ acting on the the perturbed quantities will result in quantities of order $O(\lambda^2)$. Retaining terms of order up to $O(\lambda)$, equation (239) is written as

$$\nabla\times\delta\mathbf{A}_\perp \approx \frac{1}{r}\left[\frac{\partial}{\partial r}(r\delta A_\phi) - \frac{\partial\delta A_r}{\partial\phi}\right]\mathbf{e}_\parallel, \tag{240}$$

i.e., only the parallel component survive, which exactly cancels the last term in Eq. (238), i.e., equation (238) is reduced to

$$\delta\mathbf{B}_\perp = \nabla\delta A_\parallel\times\mathbf{e}_\parallel. \tag{241}$$

C.2 Expression of $\delta B_\parallel$ in terms of δA

$$\begin{aligned} \delta B_\parallel &= \mathbf{e}_\parallel\cdot\nabla\times\delta\mathbf{A} \\ &= \mathbf{e}_\parallel\cdot\nabla\times(\delta\mathbf{A}_\perp + \delta A_\parallel\mathbf{e}_\parallel) \end{aligned} \tag{242}$$

Accurate to $O(\lambda^1)$, $\delta B_\parallel$ in the above equation is written as ($\mathbf{e}_\parallel$ vector can be considered as constant because its spatial gradient combined with $\delta\mathbf{A}$ will give $O(\lambda^2)$ terms, which are neglected)

$$\begin{aligned} \delta B_\parallel &\approx \mathbf{e}_\parallel\cdot\nabla\times\delta\mathbf{A}_\perp + \mathbf{e}_\parallel\cdot(\nabla\delta A_\parallel\times\mathbf{e}_\parallel) \\ &= \mathbf{e}_\parallel\cdot\nabla\times\delta\mathbf{A}_\perp \end{aligned} \tag{243}$$

[Using local cylindrical coordinates (r,ϕ,z) with z being along the local direction of $\mathbf{B}_0$, and two components of $\delta\mathbf{A}_\perp$ being δA_r and δA_ϕ, then $\nabla\times\delta\mathbf{A}_\perp$ is written as

$$\nabla\times\delta\mathbf{A}_\perp = \left(-\frac{\partial\delta A_\phi}{\partial z}\right)\mathbf{e}_r + \left(\frac{\partial\delta A_r}{\partial z}\right)\mathbf{e}_r + \frac{1}{r}\left[\frac{\partial}{\partial r}(r\delta A_\phi) - \frac{\partial\delta A_r}{\partial\phi}\right]\mathbf{e}_\parallel \tag{244}$$

Note that the parallel gradient operator $\nabla_\parallel\equiv\mathbf{e}_\parallel\cdot\nabla = \partial/\partial z$ acting on the the perturbed quantities will result in quantities of order $O(\lambda^2)$. Retaining terms of order up to $O(\lambda)$, equation (239) is written as

$$\nabla\times\delta\mathbf{A}_\perp \approx \frac{1}{r}\left[\frac{\partial}{\partial r}(r\delta A_\phi) - \frac{\partial\delta A_r}{\partial\phi}\right]\mathbf{e}_\parallel, \tag{245}$$

Using this, equation (243) is written as

$$\delta B_{\parallel} = \frac{1}{r}\left[\frac{\partial}{\partial r}(r\,\delta A_{\phi}) - \frac{\partial \delta A_r}{\partial \phi}\right]. \tag{246}$$

However, this expression is not useful for GEM because GEM does not use the local coordinates (r, ϕ, z).]

C.3 Expressing the perturbed drift in terms of δE and δB

The perturbed drift $\delta\mathbf{V}_D$ is given by Eq. (122), i.e.,

$$\delta\mathbf{V}_D = -\frac{q}{m}\nabla_X \langle \delta L\rangle_\alpha \times \frac{\mathbf{e}_{\parallel}}{\Omega}. \tag{247}$$

Using $\delta L = \delta\Phi - \mathbf{v}\cdot\delta\mathbf{A}$, the above expression can be further written as

$$\begin{aligned}\delta\mathbf{V}_D &= -\frac{q}{m}\nabla_X \langle \delta\Phi - \mathbf{v}\cdot\delta\mathbf{A}\rangle_\alpha \times \frac{\mathbf{e}_{\parallel}}{\Omega}\\ &= \frac{q}{m}\frac{\mathbf{e}_{\parallel}}{\Omega}\times\nabla_X\langle\delta\Phi\rangle_\alpha - \frac{q}{m}\frac{\mathbf{e}_{\parallel}}{\Omega}\times\nabla_X\langle v_{\parallel}\delta A_{\parallel}\rangle_\alpha - \frac{q}{m}\frac{\mathbf{e}_{\parallel}}{\Omega}\times\nabla_X\langle \mathbf{v}_{\perp}\cdot\delta\mathbf{A}_{\perp}\rangle_\alpha. \end{aligned}\tag{248}$$

Accurate to order $O(\lambda)$, the term involving $\delta\Phi$ is

$$\begin{aligned}\frac{q}{m}\frac{\mathbf{e}_{\parallel}}{\Omega}\times\nabla_X\langle\delta\Phi\rangle_\alpha &= \frac{\mathbf{e}_{\parallel}}{B_0}\times\langle\nabla_X\delta\Phi\rangle_\alpha\\ &\approx \frac{\mathbf{e}_{\parallel}}{B_0}\times\langle\nabla_x\delta\Phi\rangle_\alpha\\ &\approx \frac{\mathbf{e}_{\parallel}}{B_0}\times\left\langle -\delta\mathbf{E} - \frac{\partial\delta\mathbf{A}}{\partial t}\right\rangle_\alpha\\ &\approx \frac{\mathbf{e}_{\parallel}}{B_0}\times\langle -\delta\mathbf{E}\rangle_\alpha\\ &\equiv \mathbf{V}_E, \end{aligned}\tag{249}$$

which is the $\delta\mathbf{E}\times\mathbf{B}_0$ drift. Accurate to $O(\lambda)$, the $\langle v_{\parallel}\delta A_{\parallel}\rangle_\alpha$ term on the right-hand side of Eq. (248) is written

$$\begin{aligned}-\frac{q}{m}\frac{\mathbf{e}_{\parallel}}{\Omega}\times\nabla_X\langle v_{\parallel}\delta A_{\parallel}\rangle_\alpha &\approx -\frac{q}{m}\frac{1}{\Omega}\langle\mathbf{e}_{\parallel}\times\nabla_{\mathbf{X}}(v_{\parallel}\delta A_{\parallel})\rangle_\alpha\\ &\approx -\frac{q}{m}\frac{1}{\Omega}\langle\mathbf{e}_{\parallel}\times\nabla_{\mathbf{x}}(v_{\parallel}\delta A_{\parallel})\rangle_\alpha\\ &\approx -\frac{q}{m}\frac{v_{\parallel}}{\Omega}\langle\mathbf{e}_{\parallel}\times\nabla_{\mathbf{x}}(\delta A_{\parallel})\rangle_\alpha\\ &= v_{\parallel}\frac{\langle\delta\mathbf{B}_{\perp}\rangle_\alpha}{B_0}, \end{aligned}\tag{250}$$

which is due to the magnetic fluttering and this is actually not a real drift. In obtaining the last equality, use has been made of Eq. (241), i.e., $\delta\mathbf{B}_{\perp} = \nabla_{\mathbf{x}}\delta A_{\parallel}\times\mathbf{e}_{\parallel}$.

Accurate to $O(\lambda)$, the last term on the right-hand side of expression (248) is written

$$\begin{aligned}-\frac{q}{m}\frac{\mathbf{e}_{\parallel}}{\Omega}\times\nabla_X\langle\mathbf{v}_{\perp}\cdot\delta\mathbf{A}_{\perp}\rangle_\alpha &\approx -\frac{1}{B_0}\langle\mathbf{e}_{\parallel}\times\nabla_X(\mathbf{v}_{\perp}\cdot\delta\mathbf{A}_{\perp})\rangle_\alpha\\ &\approx -\frac{1}{B_0}\langle\mathbf{e}_{\parallel}\times\nabla_x(\mathbf{v}_{\perp}\cdot\delta\mathbf{A}_{\perp})\rangle_\alpha\\ &= -\frac{1}{B_0}\langle\mathbf{e}_{\parallel}\times(\mathbf{v}_{\perp}\times\nabla_x\times\delta\mathbf{A}_{\perp} + \mathbf{v}_{\perp}\cdot\nabla_x\delta\mathbf{A}_{\perp})\rangle_\alpha\\ &= -\frac{1}{B_0}\langle(\mathbf{e}_{\parallel}\cdot\nabla_x\times\delta\mathbf{A}_{\perp})\mathbf{v}_{\perp} + \mathbf{e}_{\parallel}\times\mathbf{v}_{\perp}\cdot\nabla_x\delta\mathbf{A}_{\perp}\rangle_\alpha \end{aligned}$$

Using equation (243), i.e., $\delta B_{\parallel} = \mathbf{e}_{\parallel}\cdot\nabla\times\delta\mathbf{A}_{\perp}$, the above expression is written as

$$\begin{aligned}-\frac{q}{m}\frac{\mathbf{e}_{\parallel}}{\Omega}\times\nabla_X\langle\mathbf{v}_{\perp}\cdot\delta\mathbf{A}_{\perp}\rangle_\alpha &= -\frac{1}{B_0}\langle\delta B_{\parallel}\mathbf{v}_{\perp} + \mathbf{e}_{\parallel}\times\mathbf{v}_{\perp}\cdot\nabla_x\delta\mathbf{A}_{\perp}\rangle_\alpha\\ &\approx -\frac{1}{B_0}\langle\delta B_{\parallel}\mathbf{v}_{\perp} + \mathbf{e}_{\parallel}\times\mathbf{v}_{\perp}\cdot\nabla_X\delta\mathbf{A}_{\perp}\rangle_\alpha\\ &\approx -\frac{1}{B_0}\langle\delta B_{\parallel}\mathbf{v}_{\perp}\rangle_\alpha - \frac{1}{B_0}\mathbf{e}_{\parallel}\times\langle\mathbf{v}_{\perp}\cdot\nabla_X\delta\mathbf{A}_{\perp}\rangle_\alpha.\\ &\approx -\frac{1}{B_0}\langle\delta B_{\parallel}\mathbf{v}_{\perp}\rangle_\alpha. \end{aligned}\tag{251}$$

where use has been made of $\langle \mathbf{v}_\perp \cdot \nabla_X \delta\mathbf{A}_\perp\rangle_\alpha \approx 0$, where the error is of $O(\lambda)\delta\mathbf{A}_\perp$. The term $\langle \delta B_\parallel \mathbf{v}_\perp\rangle_\alpha / B_0$ is of $O(\lambda^2)$ and thus can be neglected (I need to verify this).

Using Eqs. (249), (250), and (251), expression (248) is finally written as

$$\delta\mathbf{V}_D \equiv -\frac{q}{m}\nabla_X \langle \delta L\rangle_\alpha \times \frac{\mathbf{e}_\parallel}{\Omega} = \frac{\langle \delta\mathbf{E}\rangle_\alpha \times \mathbf{e}_\parallel}{B_0} + v_\parallel \frac{\langle \delta\mathbf{B}_\perp\rangle_\alpha}{B_0}. \tag{252}$$

Using this, the first equation of the characteristics, equation (119), is written

$$\begin{aligned} \frac{d\mathbf{X}}{dt} &= v_\parallel \mathbf{e}_\parallel + \mathbf{V}_D + \delta\mathbf{V}_D \qquad (253)\\ &= v_\parallel \mathbf{e}_\parallel + \mathbf{V}_D + \frac{\langle \delta\mathbf{E}\rangle_\alpha \times \mathbf{e}_\parallel}{B_0} + v_\parallel \frac{\langle \delta\mathbf{B}_\perp\rangle_\alpha}{B_0} \\ &\equiv \mathbf{V}_G \qquad (254) \end{aligned}$$

C.4 Expressing the coefficient before $\partial F_0/\partial\varepsilon$ in terms of δE and δB

[Note that

$$\frac{\partial \delta\mathbf{A}_\perp}{\partial t} = -(\delta\mathbf{E}_\perp + \nabla_\perp \delta\Phi), \tag{255}$$

where $\partial\delta\mathbf{A}_\perp/\partial t$ is of $O(\lambda^2)$. This means that $\delta\mathbf{E}_\perp + \nabla_\perp\delta\phi$ is of $O(\lambda^2)$ although both $\delta\mathbf{E}_\perp$ and $\delta\phi$ are of $O(\lambda)$.]

Note that

$$\begin{aligned} \frac{\partial \langle \mathbf{v}\cdot\delta\mathbf{A}\rangle_\alpha}{\partial t} &= v_\parallel \frac{\partial \langle \delta A_\parallel\rangle_\alpha}{\partial t} + \mathbf{v}_\perp \cdot \frac{\partial \langle \delta\mathbf{A}\rangle_\alpha}{\partial t} \\ &= v_\parallel \frac{\partial \langle \delta A_\parallel\rangle_\alpha}{\partial t} + \langle \mathbf{v}_\perp \cdot (-\delta\mathbf{E} - \nabla\delta\Phi)\rangle_\alpha \\ &\approx v_\parallel \frac{\partial \langle \delta A_\parallel\rangle_\alpha}{\partial t} - \langle \mathbf{v}_\perp\cdot\delta\mathbf{E}\rangle_\alpha \qquad (256) \end{aligned}$$

where use has been made of $\langle \mathbf{v}_\perp \cdot \nabla\delta\phi\rangle \approx 0$, This indicates that $\langle \mathbf{v}_\perp\cdot\delta\mathbf{E}\rangle_\alpha$ is of $O(\lambda^1)\delta\mathbf{E}$. Using Eq. (256), the coefficient before $\partial F_0/\partial\varepsilon$ in Eq. (138) can be further written as

$$\begin{aligned} &-\frac{q}{m}\left[-\frac{\partial\langle \mathbf{v}\cdot\delta\mathbf{A}\rangle_\alpha}{\partial t} - \left(v_\parallel \mathbf{e}_\parallel + \mathbf{V}_D - \frac{q}{m}\frac{\mathbf{e}_\parallel}{\Omega}\times\nabla_X\langle \mathbf{v}\cdot\delta\mathbf{A}\rangle_\alpha\right)\cdot\nabla_X\langle\delta\Phi\rangle_\alpha\right] \\ &= -\frac{q}{m}\left[-v_\parallel\frac{\partial\langle\delta A_\parallel\rangle_\alpha}{\partial t} + \langle \mathbf{v}_\perp\cdot\delta\mathbf{E}\rangle_\alpha - \left(v_\parallel \mathbf{e}_\parallel + \mathbf{V}_D - \frac{q}{m}\frac{\mathbf{e}_\parallel}{\Omega}\times\nabla_X\langle \mathbf{v}\cdot\delta\mathbf{A}\rangle_\alpha\right)\cdot\left\langle -\delta\mathbf{E} - \frac{\partial\delta\mathbf{A}}{\partial t}\right\rangle_\alpha\right] \\ &\approx -\frac{q}{m}\left[-v_\parallel\frac{\partial\langle\delta A_\parallel\rangle_\alpha}{\partial t} + \langle \mathbf{v}_\perp\cdot\delta\mathbf{E}\rangle_\alpha - \left(v_\parallel \mathbf{e}_\parallel + \mathbf{V}_D - \frac{q}{m}\frac{\mathbf{e}_\parallel}{\Omega}\times\nabla_X\langle \mathbf{v}\cdot\delta\mathbf{A}\rangle_\alpha\right)\cdot\langle -\delta\mathbf{E}\rangle_\alpha + v_\parallel\left\langle\frac{\partial A_\parallel}{\partial t}\right\rangle_\alpha\right] \\ &= -\frac{q}{m}\left[\langle \mathbf{v}_\perp\cdot\delta\mathbf{E}\rangle_\alpha + \left(v_\parallel \mathbf{e}_\parallel + \mathbf{V}_D - \frac{q}{m}\frac{\mathbf{e}_\parallel}{\Omega}\times\nabla_X\langle \mathbf{v}\cdot\delta\mathbf{A}\rangle_\alpha\right)\cdot\langle\delta\mathbf{E}\rangle_\alpha\right] \\ &\approx -\frac{q}{m}\left[\langle \mathbf{v}_\perp\cdot\delta\mathbf{E}\rangle_\alpha + \left(v_\parallel \mathbf{e}_\parallel + \mathbf{V}_D + v_\parallel\frac{\langle\delta\mathbf{B}_\perp\rangle}{B_0}\right)\cdot\langle\delta\mathbf{E}\rangle_\alpha\right]. \qquad (257) \end{aligned}$$

Using Eq. (257) and (132), gyrokinetic equation (138) is finally written as

$$\begin{aligned} &\left[\frac{\partial}{\partial t} + \left(v_\parallel\mathbf{e}_\parallel + \mathbf{V}_D + \frac{\langle\delta\mathbf{E}\rangle_\alpha\times\mathbf{e}_\parallel}{B_0} + v_\parallel\frac{\langle\delta\mathbf{B}_\perp\rangle_\alpha}{B_0}\right)\cdot\nabla_X\right]\delta f \\ &= -\left(\frac{\langle\delta\mathbf{E}\rangle_\alpha\times\mathbf{e}_\parallel}{B_0} + v_\parallel\frac{\langle\delta\mathbf{B}_\perp\rangle_\alpha}{B_0}\right)\cdot\nabla_X F_0 \\ &-\frac{q}{m}\left[\langle\mathbf{v}_\perp\cdot\delta\mathbf{E}\rangle_\alpha + \left(v_\parallel\mathbf{e}_\parallel + \mathbf{V}_D + v_\parallel\frac{\langle\delta\mathbf{B}_\perp\rangle_\alpha}{B_0}\right)\cdot\langle\delta\mathbf{E}\rangle_\alpha\right]\frac{\partial F_0}{\partial\varepsilon}. \qquad (258) \end{aligned}$$

Appendix D Drift-kinetic limit

In the drift-kinetic limit, $\langle \mathbf{v}_\perp \cdot \delta \mathbf{E} \rangle_\alpha = 0$, $\langle \delta B_\parallel \mathbf{v}_\perp \rangle_\alpha = 0$, and $\langle \delta h \rangle_\alpha = \delta h$, where δh is an arbitrary field quantity. Using these, gyrokinetic equation (258) is written as

$$\begin{aligned}&\left[\frac{\partial}{\partial t} + \left(v_\parallel \mathbf{e}_\parallel + \mathbf{v}_D + \mathbf{v}_E + v_\parallel \frac{\delta \mathbf{B}_\perp}{B_0}\right) \cdot \nabla_X\right] \delta f \\ &= -\left(\mathbf{v}_E + v_\parallel \frac{\delta \mathbf{B}_\perp}{B_0}\right) \cdot \nabla_X F_0 - \frac{q}{m}\left[\left(v_\parallel \mathbf{e}_\parallel + \mathbf{v}_D + v_\parallel \frac{\delta \mathbf{B}_\perp}{B_0}\right) \cdot \delta \mathbf{E}\right] \frac{\partial F_0}{\partial \varepsilon}. \end{aligned} \tag{259}$$

D.1 Linear case

Neglecting the nonlinear terms, drift-kinetic equation (259) is written

$$\begin{aligned}&\left[\frac{\partial}{\partial t} + (v_\parallel \mathbf{e}_\parallel + \mathbf{v}_D) \cdot \nabla_X\right] \delta f \\ &= -\left(\mathbf{v}_E + v_\parallel \frac{\delta \mathbf{B}_\perp}{B_0}\right) \cdot \nabla_X F_0 - \frac{q}{m}[(v_\parallel \mathbf{e}_\parallel + \mathbf{v}_D) \cdot \delta \mathbf{E}] \frac{\partial F_0}{\partial \varepsilon}. \end{aligned} \tag{260}$$

Next let us derive the parallel momentum equation from the linear drift kinetic equation (this is needed in my simulation). Multiplying the linear drift kinetic equation (260) by $qv_\parallel$ and then integrating over velocity space, we obtain

$$\begin{aligned}\frac{\partial \delta j_\parallel}{\partial t} &= -q\int d\mathbf{v} v_\parallel (v_\parallel \mathbf{e}_\parallel + \mathbf{v}_D) \cdot \nabla_X \delta f \\ &- q\int d\mathbf{v} v_\parallel \left(\mathbf{v}_E + v_\parallel \frac{\delta \mathbf{B}_\perp}{B_0}\right) \cdot \nabla_X F_0 - \frac{q}{m} q\int d\mathbf{v} v_\parallel [(v_\parallel \mathbf{e}_\parallel + \mathbf{v}_D) \cdot \delta \mathbf{E}] \frac{\partial F_0}{\partial \varepsilon}. \end{aligned} \tag{261}$$

Equation (261) involve $\nabla_X \delta f$ and this should be avoided in particle methods whose goal is to avoid directly evaluating the derivatives of δf over phase-space coordinates. On the other hand, the partial derivatives of velocity moment of δf are allowed. Therefore, we would like to make the velocity integration of δf appear. Note that $\nabla_X \delta f$ here is taken by holding (ε, μ) constant and thus $v_\parallel$ is not a constant and thus can not be moved inside ∇_X. Next, to facilitate performing the integration over $v_\parallel$, we transform the linear drift kinetic equation (260) into variable $(\mathbf{X}, \mu, v_\parallel)$.

D.2 Transform from $(\mathbf{X}, \mu, \varepsilon)$ to $(\mathbf{X}, \mu, v_\parallel)$ coordinates

The kinetic equation given above is written in terms of variable $(\mathbf{X},\ \mu,\ \varepsilon)$. Next, we transform it into coordinates $(\mathbf{X}', \mu', v_\parallel)$ which is defined by

$$\mathbf{X}'(\mathbf{X}, \mu, \varepsilon) = \mathbf{X}, \tag{262}$$

$$\mu'(\mathbf{X}, \mu, \varepsilon) = \mu, \tag{263}$$

and

$$v_\parallel(\mathbf{X}, \mu, \varepsilon) = \sqrt{2(\varepsilon - \mu B_0(\mathbf{X}))}. \tag{264}$$

Use this, we have

$$\begin{aligned}\frac{\partial \delta G_0}{\partial \mathbf{X}}|_{\mu,\varepsilon} &= \frac{\partial \mathbf{X}'}{\partial \mathbf{X}} \frac{\partial \delta G_0}{\partial \mathbf{X}'} + \frac{\partial \mu'}{\partial \mathbf{X}} \frac{\partial \delta G_0}{\partial \mu'} + \frac{\partial v_\parallel}{\partial \mathbf{X}} \frac{\partial \delta G_0}{\partial v_\parallel} \\ &= \frac{\partial \delta G_0}{\partial \mathbf{X}'}|_{\mu, v_\parallel} + 0\frac{\partial \delta G_0}{\partial \mu'} - \frac{\mu}{v_\parallel} \frac{\partial B_0}{d\mathbf{X}} \frac{\partial \delta G_0}{\partial v_\parallel}, \end{aligned} \tag{265}$$

and

$$\begin{aligned}\frac{\partial F_0}{\partial \varepsilon} &= \frac{\partial F_0}{\partial \mu'} \frac{\partial \mu'}{\partial \varepsilon} + \frac{\partial F_0}{\partial v_\parallel} \frac{\partial v_\parallel}{\partial \varepsilon} \\ &= 0\frac{\partial F_0}{\partial \mu'} + \frac{\partial F_0}{\partial v_\parallel} \frac{\partial v_\parallel}{\partial \varepsilon} \\ &= \frac{\partial F_0}{\partial v_\parallel} \frac{1}{v_\parallel} \end{aligned} \tag{266}$$

Then, in terms of variable $(\mathbf{X}', \mu, v_\parallel)$, equation (260) is written

$$\begin{aligned}&\frac{\partial \delta f}{\partial t} + (v_\parallel \mathbf{e}_\parallel + \mathbf{v}_D) \cdot \nabla \delta f - \mathbf{e}_\parallel \cdot \mu \nabla B \frac{\partial \delta f}{\partial v_\parallel} \\ &= -\left(\mathbf{v}_E + v_\parallel \frac{\delta \mathbf{B}_\perp}{B_0}\right) \cdot \nabla F_0 + \left(\frac{\mathbf{v}_E}{v_\parallel} + \frac{\delta \mathbf{B}_\perp}{B_0}\right) \cdot \mu \nabla B \frac{\partial F_0}{\partial v_\parallel} - \frac{q}{m}\left[\left(\mathbf{e}_\parallel + \frac{\mathbf{v}_D}{v_\parallel}\right) \cdot \delta \mathbf{E}\right] \frac{\partial F_0}{\partial v_\parallel}, \end{aligned} \tag{267}$$

where $\nabla \equiv \partial / \partial \mathbf{X}'|_{\mu, v_\parallel}$.

D.3 Parallel momentum equation

Multiplying the linear drift kinetic equation (267) by $q v_\parallel$ and then integrating over velocity space, we obtain

$$\begin{aligned}&\frac{\partial \delta j_\parallel}{\partial t} + q \int d\mathbf{v} v_\parallel (v_\parallel \mathbf{e}_\parallel + \mathbf{v}_D) \cdot \nabla_X \delta f - q \int d\mathbf{v} v_\parallel \mathbf{e}_\parallel \cdot \mu \nabla B \frac{\partial \delta f}{\partial v_\parallel} \\ &= -q \int d\mathbf{v} v_\parallel \left(\mathbf{v}_E + v_\parallel \frac{\delta \mathbf{B}_\perp}{B_0}\right) \cdot \nabla_X F_0 + q \int d\mathbf{v} v_\parallel \left(\frac{\mathbf{v}_E}{v_\parallel} + \frac{\delta \mathbf{B}_\perp}{B_0}\right) \cdot \mu \nabla B \frac{\partial F_0}{\partial v_\parallel} \\ &- \frac{q}{m} q \int d\mathbf{v} v_\parallel \left[\left(\mathbf{e}_\parallel + \frac{\mathbf{v}_D}{v_\parallel}\right) \cdot \delta \mathbf{E}\right] \frac{\partial F_0}{\partial v_\parallel}. \end{aligned} \tag{268}$$

Consider the simple case that F_0 does not carry current, i.e., $F_0(\mathbf{X}, \mu, v_\parallel)$ is an even function about $v_\parallel$. Then it is obvious that the integration of the terms in red in Eq. (268) are all zero. Among the rest terms, only the following term

$$-\frac{q}{m} q \int d\mathbf{v} v_\parallel [(v_\parallel \mathbf{e}_\parallel) \cdot \delta \mathbf{E}] \frac{\partial F_0}{\partial v_\parallel} \frac{1}{v_\parallel} \tag{269}$$

explicitly depends on $\delta \mathbf{E}$. Using $d\mathbf{v} = 2\pi B dv_\parallel d\mu$, the integration in the above expression can be analytically performed, giving

$$\begin{aligned}&-\frac{q}{m} q \int d\mathbf{v} v_\parallel [(v_\parallel \mathbf{e}_\parallel) \cdot \delta \mathbf{E}] \frac{\partial F_0}{\partial v_\parallel} \frac{1}{v_\parallel} \\ &= -\frac{q^2}{m} \int 2\pi B dv_\parallel d\mu v_\parallel \delta E_\parallel \frac{\partial F_0}{\partial v_\parallel} \\ &= -\frac{q^2}{m} \int 2\pi B d\mu \delta E_\parallel \int v_\parallel \frac{\partial F_0}{\partial v_\parallel} dv_\parallel \\ &= -\frac{q^2}{m} \int 2\pi B d\mu \delta E_\parallel \left(0 - \int F_0 dv_\parallel\right) \\ &= \frac{q^2}{m} \delta E_\parallel n_0. \end{aligned} \tag{270}$$

Using these results, the parallel momentum equation (268) is written

$$\begin{aligned}\frac{\partial \delta j_\parallel}{\partial t} &= \frac{q^2}{m} \delta E_\parallel n_0 - q \int d\mathbf{v} v_\parallel (v_\parallel \mathbf{e}_\parallel + \mathbf{v}_D) \cdot \nabla_X \delta f + q \int d\mathbf{v} v_\parallel \mathbf{e}_\parallel \cdot \mu \nabla B \frac{\partial \delta f}{\partial v_\parallel} \\ &- q \int d\mathbf{v} v_\parallel \left(v_\parallel \frac{\delta \mathbf{B}_\perp}{B_0}\right) \cdot \nabla F_0 + q \int d\mathbf{v} v_\parallel \left(\frac{\delta \mathbf{B}_\perp}{B_0}\right) \cdot \mu \nabla B \frac{\partial F_0}{\partial v_\parallel}, \end{aligned} \tag{271}$$

where the explicit dependence on $\delta \mathbf{E}$ is via the first term $q^2 n_0 \delta E_\parallel / m$, with all the the other terms being explicitly independent of $\delta \mathbf{E}$ (δf and $\delta \mathbf{B}$ implicitly depend on $\delta \mathbf{E}$).

Equation (271) involve derivatives of δf with respect to space and $v_\parallel$ and these should be avoided in the particle method whose goal is to avoid directly evaluating these derivatives. Using integration by parts, the terms involving $\partial / \partial v_\parallel$ can be simplified, yielding

$$\begin{aligned}\frac{\partial \delta j_\parallel}{\partial t} &= \frac{q^2}{m} \delta E_\parallel n_0 - q \int d\mathbf{v} v_\parallel (v_\parallel \mathbf{e}_\parallel + \mathbf{v}_D) \cdot \nabla_X \delta f - q(\mathbf{e}_\parallel \cdot \nabla B_0) \int \mu \delta f d\mathbf{v} \\ &- q \int d\mathbf{v} v_\parallel \left(v_\parallel \frac{\delta \mathbf{B}_\perp}{B_0}\right) \cdot \nabla F_0 - q \left(\frac{\delta \mathbf{B}_\perp}{B_0}\right) \cdot (\nabla B_0) \int \mu F_0 d\mathbf{v}, \end{aligned} \tag{272}$$

Define $p_{\perp 0} = \int m v_\perp^2 F_0 / 2 d\mathbf{v}$ and $\delta p_\perp = \int m v_\perp^2 \delta f / 2 d\mathbf{v}$, then the above equation is written

$$\begin{aligned}\frac{\partial \delta j_\parallel}{\partial t} &= \frac{q^2}{m}\delta E_\parallel n_0 - q\int d\mathbf{v} v_\parallel (v_\parallel \mathbf{e}_\parallel + \mathbf{v}_D)\cdot\nabla_X \delta f - q(\mathbf{e}_\parallel\cdot\nabla B_0)\frac{\delta p_\perp}{m B_0}\\ &\quad - q\int d\mathbf{v} v_\parallel \left(v_\parallel \frac{\delta \mathbf{B}_\perp}{B_0}\right)\cdot\nabla F_0 - q\left(\frac{\delta \mathbf{B}_\perp}{B_0}\right)\cdot(\nabla B_0)\frac{p_{\perp 0}}{m B_0},\end{aligned} \tag{273}$$

Next, we try to eliminate the spatial gradient of δf by changing the order of integration. The second term on the right-hand side of Eq. (273) is written

$$\begin{aligned}&-q\int d\mathbf{v} v_\parallel (v_\parallel \mathbf{e}_\parallel)\cdot\nabla_X \delta f,\\ &= -q\int 2\pi B_0\, dv_\parallel d\mu v_\parallel^2 \mathbf{e}_\parallel\cdot\nabla_X \delta f\\ &= -q 2\pi B_0\, \mathbf{e}_\parallel\cdot\nabla_X \int v_\parallel^2 \delta f dv_\parallel d\mu\\ &= -q B_0\, \mathbf{e}_\parallel\cdot\nabla_X \left(\frac{1}{m B_0}\int m v_\parallel^2 \delta f d\mathbf{v}\right)\\ &= -q B_0\, \mathbf{e}_\parallel\cdot\nabla_X \left(\frac{\delta p_\parallel}{m B_0}\right),\end{aligned} \tag{274}$$

where $\delta p_\parallel = \int m v_\parallel^2 \delta f d\mathbf{v}$. Similarly, the term $-q\int d\mathbf{v} v_\parallel \left(v_\parallel \frac{\delta \mathbf{B}_\perp}{B_0}\right)\cdot\nabla_X F_0$ is written as

$$\begin{aligned}&-q\int d\mathbf{v} v_\parallel \left(v_\parallel \frac{\delta \mathbf{B}_\perp}{B_0}\right)\cdot\nabla_X F_0\\ &= -q\int 2\pi B_0\, dv_\parallel d\mu \left(v_\parallel^2 \frac{\delta \mathbf{B}_\perp}{B_0}\right)\cdot\nabla_X F_0\\ &= -q\left(\frac{\delta \mathbf{B}_\perp}{B_0}\right)\cdot B_0 \nabla_X \int (v_\parallel^2 F_0 2\pi dv_\parallel d\mu)\\ &= -q\left(\frac{\delta \mathbf{B}_\perp}{B_0}\right)\cdot B_0\, \nabla_X \left[\frac{1}{B}\int (m v_\parallel^2 F_0 d\mathbf{v})\right]\\ &= -q\delta \mathbf{B}_\perp\cdot\nabla_X \left(\frac{p_{\parallel 0}}{m B_0}\right)\end{aligned} \tag{275}$$

where $p_{\parallel 0} = \int m v_\parallel^2 F_0 d\mathbf{v}$. Similarly, the term $-q\int d\mathbf{v} v_\parallel \mathbf{v}_D\cdot\nabla_X \delta f$ can be written as the gradient of moments of δf. Let us work on this. The drift $\mathbf{v}_D$ is given by

$$\mathbf{v}_D = \frac{B_0 \frac{v_\parallel}{\Omega}\nabla\times\mathbf{b}}{B_\parallel^\star} v_\parallel + \frac{\mu}{\Omega B_\parallel^\star}\mathbf{B}_0\times\nabla B_0. \tag{276}$$

where $B_\parallel^\star = B_0\left(1 + \frac{v_\parallel}{\Omega}\mathbf{b}\cdot\nabla\times\mathbf{b}\right)$ (refer to my another notes). Using $\mathbf{b}\cdot\nabla\times\mathbf{b}\approx 0$, we obtain $B_\parallel^\star \approx B$. Then $\mathbf{v}_D$ is written

$$\mathbf{v}_D = \frac{v_\parallel^2}{\Omega}\nabla\times\mathbf{b} + \frac{\mu}{\Omega}\mathbf{b}\times\nabla B_0.$$

Using this and $d\mathbf{v} = 2\pi B_0\, dv_\parallel d\mu$, the term $-q\int d\mathbf{v} v_\parallel \mathbf{v}_D\cdot\nabla_X \delta f$ is written as

$$\begin{aligned}-q\int d\mathbf{v} v_\parallel \mathbf{v}_D\cdot\nabla_X \delta f &= -q\int 2\pi B_0\, dv_\parallel d\mu v_\parallel \left(\frac{v_\parallel^2}{\Omega}\nabla\times\mathbf{b} + \frac{\mu}{\Omega}\mathbf{b}\times\nabla B_0\right)\cdot\nabla_X \delta f\\ &= -q 2\pi B_0 \frac{1}{\Omega}(\nabla\times\mathbf{b})\cdot\nabla_X \int v_\parallel^3 \delta f dv_\parallel d\mu - q 2\pi B_0 \frac{1}{\Omega}(\mathbf{b}\times\nabla B_0)\cdot\nabla_X \int v_\parallel \mu \delta f dv_\parallel d\mu\\ &= -q B_0 \frac{1}{\Omega}(\nabla\times\mathbf{b})\cdot\nabla_X \left(\frac{1}{B_0}\int v_\parallel^3 \delta f d\mathbf{v}\right) - q B_0 \frac{1}{\Omega}(\mathbf{b}\times\nabla B_0)\cdot\nabla_X \left(\frac{1}{B_0}\int v_\parallel \mu \delta f d\mathbf{v}\right),\\ &= -m(\nabla\times\mathbf{b})\cdot\nabla_X \left(\frac{1}{B_0}\int v_\parallel^3 \delta f d\mathbf{v}\right) - m(\mathbf{b}\times\nabla B_0)\cdot\nabla_X \left(\frac{1}{B_0}\int v_\parallel \mu \delta f d\mathbf{v}\right),\end{aligned} \tag{277}$$

which are the third order moments of δf and may be neglect-able (a guess, not verified). Using the above results, the linear parallel momentum equation is finally written

$$\begin{aligned}\frac{\partial\delta j_\parallel}{\partial t} &= \frac{e^2 n_{e0}}{m}\delta E_\parallel + eB_0\,\mathbf{b}\cdot\nabla_X\left(\frac{\delta p_\parallel}{mB_0}\right) + e(\mathbf{b}\cdot\nabla B_0)\frac{\delta p_\perp}{mB_0} \\ &\quad + e\delta\mathbf{B}_\perp\cdot\nabla_X\left(\frac{p_{\parallel 0}}{mB_0}\right) + e\left(\frac{\delta\mathbf{B}_\perp}{B_0}\right)\cdot(\nabla B_0)\frac{p_{\perp 0}}{mB_0}. \\ &\quad - m(\nabla\times\mathbf{b})\cdot\nabla_X\left(\frac{1}{B_0}\int v_\parallel^3\delta f d\mathbf{v}\right) - m(\mathbf{b}\times\nabla B_0)\cdot\nabla_X\left(\frac{1}{B_0}\int v_\parallel\mu\delta f d\mathbf{v}\right)\end{aligned} \tag{278}$$

Define

$$\mathbf{D}_0 = \nabla\left(\frac{p_{\parallel 0}}{mB_0}\right) + \frac{\nabla B_0}{B_0}\frac{p_{\perp 0}}{mB_0}, \tag{279}$$

which, for the isotropic case ($p_{\parallel 0} = p_{\perp 0} = p_0$), is simplified to

$$\mathbf{D}_0 = \frac{\nabla p_0}{mB_0}. \tag{280}$$

then Eq. (278) is written as

$$\begin{aligned}\frac{\partial\delta j_\parallel}{\partial t} &= \frac{e^2 n_0}{m}\delta E_\parallel + e\delta\mathbf{B}_\perp\cdot\mathbf{D}_0 \\ &+ eB_0\,\mathbf{b}\cdot\nabla_X\left(\frac{\delta p_\parallel}{mB_0}\right) + e(\mathbf{b}\cdot\nabla B_0)\frac{\delta p_\perp}{mB_0} \\ &- m(\nabla\times\mathbf{b})\cdot\nabla_X\left(\frac{1}{B_0}\int v_\parallel^3\delta f d\mathbf{v}\right) - m(\mathbf{b}\times\nabla B_0)\cdot\nabla_X\left(\frac{1}{B_0}\int v_\parallel\mu\delta f d\mathbf{v}\right).\end{aligned} \tag{281}$$

D.4 Special case in uniform magnetic field

In the case of uniform magnetic field, the parallel momentum equation (278) is written as

$$\frac{\partial\delta j_\parallel}{\partial t} = \frac{q}{m}qE_\parallel n_{e0} - q\,\mathbf{e}_\parallel\cdot\nabla_X(\delta p_\parallel) - q\frac{\delta\mathbf{B}_\perp}{B_0}\cdot\nabla_X p_{\parallel 0}. \tag{282}$$

D.5 Electron perpendicular flow

Using the gyrokinetic theory and taking the drift-kinetic limit, the perturbed perpendicular electron flow, $\delta\mathbf{V}_{e\perp}$, is written (see Sec. A or Appendix in Yang Chen's paper[2])

$$n_{e0}\delta\mathbf{V}_{e\perp} = \underbrace{\frac{n_{e0}}{B_0}\delta\mathbf{E}\times\mathbf{b}}_{E\times B\text{ flow}} - \underbrace{\frac{1}{eB_0}\mathbf{b}\times\nabla\delta p_{\perp e}}_{\text{diamagnetic flow}} \tag{283}$$

where n_{e0} is the equilibrium electron number density, $\delta p_{e\perp}$ is the perturbed perpendicular pressure of electrons.

D.5.1 Drift kinetic equation

Drift kinetic equation is written

$$\frac{\partial f}{\partial t} + \left(v_\parallel\tilde{\mathbf{b}} + \mathbf{v}_D + \frac{\delta\mathbf{E}\times\mathbf{b}}{B_0}\right)\cdot\nabla f + \left(-\frac{e}{m}\delta E_\parallel - \mu\tilde{\mathbf{b}}\cdot\nabla B\right)\frac{\partial f}{\partial v_\parallel} = 0, \tag{284}$$

where $f = f(\mathbf{x}, \mu, v_\parallel, t)$, $\mu = mv_\perp^2/B_0$ is the magnetic moment, $\tilde{\mathbf{b}} = \mathbf{b} + \delta\mathbf{B}_\perp/B_0$, $\mathbf{b} = \mathbf{B}_0/B_0$ is the unit vector along the equilibrium magnetic field, $\mathbf{v}_D = \mathbf{v}_D(\mathbf{x}, \mu, v_\parallel)$ is the guiding-center drift in the equilibrium magnetic field. $\delta\mathbf{E}$ and $\delta\mathbf{B}$ are the perturbed electric field and magnetic field, respectively.

D.5.2 Parallel momentum equation

Multiplying the drift kinetic equation () by $v_\parallel$ and then integrating over velocity space, we obtain

$$\int \frac{\partial f_e v_\parallel}{\partial t} d\mathbf{v} + \int v_\parallel \left(v_\parallel \tilde{\mathbf{b}} + \mathbf{v}_D + \frac{\delta \mathbf{E} \times \mathbf{b}}{B_0} \right) \cdot \nabla f_{\ e} d\mathbf{v} + \int v_\parallel \left(-\frac{e}{m} \delta E_\parallel - \mu \tilde{\mathbf{b}} \cdot \nabla B \right) \frac{\partial f_e}{\partial v_\parallel} d\mathbf{v} = 0, \tag{285}$$

which can be written as

$$\frac{\partial J_{\parallel e}}{\partial t} + \int v_\parallel \left(v_\parallel \tilde{\mathbf{b}} + \mathbf{v}_D + \frac{\delta \mathbf{E} \times \mathbf{e}_\parallel}{B_0} \right) \cdot \nabla f_e \, d\mathbf{v} + \int v_\parallel \left(-\frac{e}{m} \delta E_\parallel - \mu \tilde{\mathbf{b}} \cdot \nabla B \right) \frac{\partial f_e}{\partial v_\parallel} d\mathbf{v} = 0, \tag{286}$$

Using $d\mathbf{v} = B^{-1} 2\pi m \, dv_\parallel d\mu$, the last term on the RHS of the above equation is written

$$\begin{aligned}
&\iint v_\parallel \left(-\frac{e}{m} \delta E_\parallel - \mu \tilde{\mathbf{b}} \cdot \nabla B \right) \frac{\partial f_e}{\partial v_\parallel} d\mathbf{v} \\
=& \iint v_\parallel \left(-\frac{e}{m} \delta E_\parallel - \mu \tilde{\mathbf{b}} \cdot \nabla B \right) \frac{\partial f_e}{\partial v_\parallel} 2\pi \frac{B}{m} \, dv_\parallel d\mu \\
=& -\frac{e}{m} \delta E_\parallel 2\pi \frac{B}{m} \iint v_\parallel \frac{\partial f_e}{\partial v_\parallel} \, dv_\parallel d\mu - (\tilde{\mathbf{b}} \cdot \nabla B) 2\pi \frac{B}{m} \int \mu \int v_\parallel \frac{\partial f_e}{\partial v_\parallel} \, dv_\parallel d\mu \\
=& -\frac{e}{m} \delta E_\parallel 2\pi \frac{B}{m} \int \left(v_\parallel f_e |_{-\infty}^{+\infty} - \int f_e dv_\parallel \right) d\mu - (\tilde{\mathbf{b}} \cdot \nabla B) 2\pi \frac{B}{m} \int \mu \left(v_\parallel f_e |_{-\infty}^{+\infty} - \int f_e dv_\parallel \right) d\mu \\
=& \frac{e}{m} \delta E_\parallel 2\pi \frac{B}{m} \iint f_e dv_\parallel d\mu + (\tilde{\mathbf{b}} \cdot \nabla B) 2\pi \frac{B}{m} \int \mu \int f_e dv_\parallel d\mu \\
=& \frac{e}{m} \delta E_\parallel n_e + \iint \mu (\tilde{\mathbf{b}} \cdot \nabla B) f_e d\mathbf{v} \\
\approx& \frac{e}{m} \delta E_\parallel n_e
\end{aligned} \tag{287}$$

$$\begin{aligned}
&\iint v_\parallel (v_\parallel \tilde{\mathbf{b}}) \cdot \nabla f_{\ e} d\mathbf{v} \\
=& \iint \tilde{\mathbf{b}} \cdot \nabla (v_\parallel^2 f_{\ e}) d\mathbf{v} \\
=& \iint \tilde{\mathbf{b}} \cdot \nabla (v_\parallel^2 f_{\ e}) 2\pi \frac{B}{m} \, dv_\parallel d\mu \\
=& \tilde{\mathbf{b}} \cdot \nabla \left(\iint v_\parallel^2 f_{\ e} 2\pi \frac{1}{m} \, dv_\parallel d\mu \right) B_0 \\
=& \tilde{\mathbf{b}} \cdot \nabla \left(\frac{p_\parallel}{B_0} \right) B_0 \\
=& \tilde{\mathbf{b}} \cdot \nabla \left(\frac{p_{\parallel 0} + \delta p_\parallel}{B_0} \right) B_0 \\
=& \tilde{\mathbf{b}} \cdot \nabla \left(\frac{p_{\parallel 0}}{B_0} \right) B_0 + \tilde{\mathbf{b}} \cdot \nabla \left(\frac{\delta p_\parallel}{B_0} \right) B_0 \\
\approx& \mathbf{b} \cdot \nabla \left(\frac{p_{\parallel 0}}{B_0} \right) B_0 + \frac{\delta \mathbf{B}_\perp}{B_0} \cdot \nabla \left(\frac{p_{\parallel 0}}{B_0} \right) B_0 + \mathbf{b} \cdot \nabla \left(\frac{\delta p_\parallel}{B_0} \right) B_0 \\
\approx& \mathbf{b} \cdot \nabla (p_{\parallel 0}) + \frac{\delta \mathbf{B}_\perp}{B_0} \cdot \nabla (p_{\parallel 0}) + \mathbf{b} \cdot \nabla (\delta p_\parallel) \\
=& \frac{\delta \mathbf{B}_\perp}{B_0} \cdot \nabla (p_{\parallel 0}) + \mathbf{b} \cdot \nabla (\delta p_\parallel)
\end{aligned} \tag{288}$$

where use has been made of $\mathbf{b} \cdot \nabla p_{\parallel 0} = 0$.

$$\frac{\partial \delta J_{e\parallel}}{\partial t} = -\frac{e}{m} \delta E_\parallel n_e - \frac{\delta \mathbf{B}_\perp}{B_0} \cdot \nabla (p_{\parallel 0}) - \mathbf{b} \cdot \nabla (\delta p_\parallel) \tag{289}$$

Using Eq. () in Eq. (), we obtain

$$\mu_0 e \frac{e}{m} \delta E_{\parallel} n_e + \mathbf{b} \cdot \nabla \times \nabla \times \delta \mathbf{E} = -\mu_0 e \left[\frac{\delta \mathbf{B}_{\perp}}{B_0} \cdot \nabla (p_{e\parallel 0}) + \mathbf{b} \cdot \nabla (\delta p_{e\parallel}) \right] \tag{290}$$

$$-\mathbf{b} \cdot \nabla \times \nabla \times \delta \mathbf{E} = -\mu_0 e \left(-q \frac{\delta \mathbf{B}_{\perp}}{B_0} \cdot \nabla_X p_{\parallel 0} + \frac{q}{m} q E_{\parallel} n_{e0} - q\, \mathbf{e}_{\parallel} \cdot \nabla_X (\delta p_{\parallel}) \right). \tag{291}$$

ddddd

$$\begin{aligned}
&\iint v_{\parallel} \left(v_{\parallel} \tilde{\mathbf{b}} + \mathbf{v}_D + \frac{\delta \mathbf{E} \times \mathbf{b}}{B_0} \right) \cdot \nabla f_{\ e} d\mathbf{v} \\
&= \iint \left(v_{\parallel} \tilde{\mathbf{b}} + \mathbf{v}_D + \frac{\delta \mathbf{E} \times \mathbf{b}}{B_0} \right) \cdot \nabla (v_{\parallel} f_{\ e}) d\mathbf{v} \\
&= \iint v_{\parallel} \tilde{\mathbf{b}} \cdot \nabla (v_{\parallel} f_{\ e}) d\mathbf{v} + \iint \mathbf{v}_D \cdot \nabla (v_{\parallel} f_{\ e}) d\mathbf{v} + \iint \frac{\delta \mathbf{E} \times \mathbf{b}}{B_0} \cdot \nabla (v_{\parallel} f_{\ e}) d\mathbf{v} \\
&= \iint \tilde{\mathbf{b}} \cdot \nabla (v_{\parallel}^2 f_{\ e}) d\mathbf{v} + \iint \mathbf{v}_D \cdot \nabla (v_{\parallel} f_{\ e}) d\mathbf{v} + \iint \frac{\delta \mathbf{E} \times \mathbf{b}}{B_0} \cdot \nabla (v_{\parallel} f_{\ e}) d\mathbf{v} \\
&= \iint \tilde{\mathbf{b}} \cdot \nabla (v_{\parallel}^2 f_{\ e}) 2\pi \frac{B}{m} dv_{\parallel} d\mu + \iint \mathbf{v}_D \cdot \nabla (v_{\parallel} f_{\ e}) 2\pi \frac{B}{m} dv_{\parallel} d\mu + \iint \frac{\delta \mathbf{E} \times \mathbf{b}}{B_0} \cdot \nabla (v_{\parallel} f_{\ e}) 2\pi \frac{B}{m} dv_{\parallel} d\mu \\
&= \int \tilde{\mathbf{b}} \cdot \nabla \left(v_{\parallel}^2 f_{\ e} 2\pi \frac{1}{m} dv_{\parallel} d\mu \right) B + \iint \mathbf{v}_D \cdot \nabla (v_{\parallel} f_{\ e}) 2\pi \frac{B}{m} dv_{\parallel} d\mu + \iint \frac{\delta \mathbf{E} \times \mathbf{b}}{B_0} \cdot \nabla (v_{\parallel} f_{\ e}) 2\pi \frac{B}{m} dv_{\parallel} d\mu \\
&= \tilde{\mathbf{b}} \cdot \nabla \left(\frac{p_{\parallel}}{B} \right) B + \iint \mathbf{v}_D \cdot \nabla (v_{\parallel} f_{\ e}) 2\pi \frac{B}{m} dv_{\parallel} d\mu + \iint \frac{\delta \mathbf{E} \times \mathbf{b}}{B_0} \cdot \nabla (v_{\parallel} f_{\ e}) 2\pi \frac{B}{m} dv_{\parallel} d\mu \\
&= \tilde{\mathbf{b}} \cdot \nabla \left(\frac{p_{\parallel}}{B} \right) B + \iint \frac{1}{m\Omega} \mathbf{b} \times (\mu \nabla B) \cdot \nabla (v_{\parallel} f_{\ e}) 2\pi \frac{B}{m} dv_{\parallel} d\mu + \iint \frac{1}{m\Omega} \mathbf{b} \times (m v_{\parallel}^2 \boldsymbol{\kappa}) \cdot \nabla (v_{\parallel} f_{\ e}) 2\pi \frac{B}{m} dv_{\parallel} d\mu \\
&+ \iint \frac{\delta \mathbf{E} \times \mathbf{b}}{B_0} \cdot \nabla (v_{\parallel} f_{\ e}) 2\pi \frac{B}{m} dv_{\parallel} d\mu
\end{aligned}$$

Appendix E Modern view of gyrokinetic equation

The modern form of the nonlinear gyrokinetic equation is in the total-f form. The modern way of deriving the gyrokinetic equation is to use transformation methods to eliminate gyro-phase dependence of the total distribution function and thus obtain an equation for the resulting gyro-phase independent distribution function (called gyro-center distribution function).

The resulting equation for the gyro-center distribution function is given by (see Baojian's paper)

$$\left(\frac{\partial}{\partial t} + \dot{\mathbf{X}} \cdot \nabla + \dot{v}_{\parallel} \frac{\partial}{\partial v_{\parallel}} \right) f(\mathbf{X}, v_{\parallel}, \mu, t) = 0, \tag{292}$$

where

$$\dot{\mathbf{R}} = \mathbf{V}_D + \frac{\mathbf{e}_{\parallel}}{B_0} \times \langle \nabla \delta \Phi \rangle_{\alpha} + v_{\parallel} \frac{\langle \delta \mathbf{B}_{\perp} \rangle_{\alpha}}{B_0} \tag{293}$$

$$\dot{v}_{\parallel} = -\frac{1}{m} \frac{\mathbf{B}^{\star}}{B_{\parallel}^{\star}} \cdot (q \nabla \langle \delta \Phi \rangle + \mu \nabla B_0) - \frac{q}{m} \frac{\partial \langle \delta A_{\parallel} \rangle_{\alpha}}{\partial t}. \tag{294}$$

Here the independent variables are gyro-center position $\mathbf{X}$, magnetic moment μ and parallel velocity $v_{\parallel}$.

The gyro-phase dependence of the particle distribution can be recovered by the inverse transformation of the transformation used before. The pull-back transformation gives rise to the polarization density term. (phase-space-Lagrangian Lie perturbation method (Littlejohn, 1982a, 1983), I need to read these two papers.).

In the traditional iterative method of deriving the gyrokinetic equation, the perturbed distribution function has two terms, one of which depends on the gyro-angle, the other of which does not. The former will give rise to the the polarization density and the latter is described by the gyrokinetic equation.

Bibliography

[1] P J Catto. Linearized gyro-kinetics. *Plasma Physics*, 20(7):719, 1978.

[2] Yang Chen and Scott E. Parker. Particle-in-cell simulation with vlasov ions and drift kinetic electrons. *Phys. Plasmas*, 16:52305, 2009.

[3] E. A. Frieman and Liu Chen. Nonlinear gyrokinetic equations for low-frequency electromagnetic waves in general plasma equilibria. *Physics of Fluids*, 25(3):502–508, 1982.